The effect of neurogenin3 deficiency on pancreatic gene expression in embryonic mice

Andreas Petri*1,2, Jonas Ahnfelt-Rønne*2, Klaus Stensgaard Frederiksen1, David George Edwards3, Dennis Madsen4, Palle Serup2, Jan Fleckner1 and R Scott Heller2

1Department of Molecular Genetics, Novo Nordisk A/S, DK-2880 Bagsvaerd, Denmark
2Department of Developmental Biology, Hagedorn Research Institute, DK-2820 Gentofte, Denmark
3Department of Biostatistics, Novo Nordisk A/S, DK-2880 Bagsvaerd, Denmark
4Department of Scientific Computing, Novo Nordisk A/S, DK-2760 Målev, Denmark

(Requests for offprints should be addressed to R S Heller; Email: shll@hagedorn.dk)

*(A Petri and J Ahnfelt-Ronne contributed equally to this work)

Abstract

To understand the molecular mechanisms regulating pancreatic endocrine development and function, pancreatic gene expression was compared between Ngn3-deficient mice and littermate controls on embryonic days 13 and 15. Microarray analysis identified 504 genes with significant differences in expression. Fifty-two of these showed at least twofold reduction in Ngn3 knockouts compared to controls. Many of them were previously described to be involved in endocrine development and function. Among the genes not previously characterized were Rhomboid veinlet-like 4, genes involved in tetrahydrobiopterin biosynthesis and the Iroquois-type homeobox gene Irx1, the latter was selected for further investigation. In situ hybridisation demonstrated that two Iroquois genes, Irx1 and Irx2, were expressed in pancreatic endoderm of wild-type, but not Ngn3 mutant embryos. Furthermore, ectopic Ngn3 induced prominent Irx2 expression in chicken endoderm. Co-labelling established that Irx1 and Irx2 mRNA is located to glucagon-, but not insulin- or somatostatin-producing cells in mice and chicken. These data suggest that Irx1 and Irx2 serve an evolutionary conserved role in the regulation of α-cell-specific gene expression.

Journal of Molecular Endocrinology (2006) 37, 301–316

Introduction

The pancreas is an organ located next to the duodenal part of the intestine and serves two main functions. It produces and secretes enzymes required for the digestion of proteins, carbohydrates, nucleic acids, as well as fats in the small intestine, and hormones including insulin and glucagon required for the regulation of blood glucose levels (Slack 1995). Three distinct populations of cells make up the adult pancreas. The acinar cells produce and secrete enzymes, the duct cells form a highly branched tube system facilitating delivery of enzymes to the intestine and the endocrine cells produce hormones. In adult pancreas, the endocrine cells constitute approximately 2% of the pancreas and are grouped into cell clusters known as islets of Langerhans. Five distinct cell types are discernible in the endocrine compartment. The majority of cells are insulin-producing β-cells, whereas glucagon-producing α-cells, somatostatin-producing δ-cells, pancreatic polypeptide-producing PP-cells and ghrelin-producing ε-cells constitute a minor fraction (Slack 1995, Wierup et al. 2002, Prado et al. 2004, Heller et al. 2005).

At embryonic day 9 (e9), distinct patches of primitive gut epithelium start to evaginate and give rise to the dorsal and the ventral pancreatic buds (Slack 1995). These buds are characterized by homogeneous expression of pancreatic-duodenal homeobox 1 (Pdx1; Ohlsson et al. 1993) and these Pdx1-positive cells give rise to all cell types of the adult pancreas (Gu et al. 2003). As the buds grow, they invade the surrounding mesenchyme and the resulting epithelial–mesenchymal interactions facilitate further proliferation and differentiation (Kim & Hebrok 2001). Several cell fate choices are required for the development of a complex organ like the pancreas from initially identical cells, and research done during the last decade has revealed several important steps of pancreas development (Murtaugh & Melton 2003, Wilson 2003).
requirement of this transcription factor for endocrine development was demonstrated by Gradwohl et al. (2000) who showed that mice lacking functional Ngn3 fail to develop endocrine cells, but develop an apparently normal exocrine pancreas although accumulation of secretory granules was observed at postnatal day 1 (Gradwohl et al. 2000). Transgenic mice expressing Ngn3 under the Pdx1 promoter have small under-developed pancreata consisting primarily of endocrine cells, showing that ectopic expression of Ngn3 is sufficient to drive Pdx1+ pancreatic precursor cells to an endocrine fate, thereby depleting the pool of pancreatic precursors capable of differentiating into exocrine tissue (Apelqvist et al. 1999, Jensen et al. 2000b, Schwitzgebel et al. 2000). Further evidence for the pro-endocrine nature of Ngn3 was provided by Grapin-Botton et al. (2001) by showing that in ovo electroporation of chicken endoderm with Ngn3-expressing plasmids promoted endocrine differentiation not only in the pancreas region but also in the entire primitive gut tube. However, both studies found that endocrine cells with an α-cell fate was the predominant cell type produced by forced expression of Ngn3, indicating that additional factors are required for the development of β-cells. Ngn3 expression is turned off prior to terminal differentiation (Jensen et al. 2000a) and is thus not co-expressed with hormone-producing islet cells. Functional studies have identified several transcription factors acting downstream of Ngn3, which are necessary for proper development of the endocrine cells.

Here we present the microarray analysis of Ngn3-deficient mice. A cDNA chip containing 20 188 probes was used to compare a major part of the transcriptome in dorsal pancreas buds from wild-type and Ngn3-deficient embryos isolated at e13 and e15. In wild-type mice, these days mark the beginning and the end of the secondary transition, a stage of pancreas development where the majority of endocrine cells differentiate (Murtaugh & Melton 2003). The aim of this study is to identify genes expressed in endocrine cells and their precursors and to map important pathways that operate during the formation of the endocrine cells.

Genotyping

Genotyping of the embryos was done by PCR analysis on genomic DNA isolated from forelimbs, as described by Gradwohl et al. (2000), using BioTaq polymerase (Bioline GmbH, Luckenwalde, Germany) according to the manufacturer’s recommendations. Genomic DNA was extracted by incubating 50 mM Tris (pH 8), 100 mM EDTA and 0-5% SDS, in 10 mg/ml protease K, at 55 °C overnight. Proteins were precipitated by the addition of NaCl to a final concentration of 1-3 M and the lysate was cleared by centrifugation at 10 000 g for 5 min. Supernatants were transferred to a clean tube and nucleic acids precipitated by the addition of 0-5 vol isopropanol, followed by gentle mixing and centrifugation at 10 000 g for 10 min. The pellet was washed in 70% EtOH followed by 96% EtOH and finally resuspended in 100 µl Milli-Q water. Thermal cycling was done in a PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA, USA) for 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min. PCR products were visualized by agarose gel electrophoresis in the presence of EtBr.

Total RNA isolation

Mini RNA isolation kit (Zymo Research, Orange, CA, USA) was used to isolate total RNA from pancreas tissue using manufacturer’s instructions with a few changes. After addition of lysis buffer, the tissue was homogenized using disposable pistils for 1-5 ml Eppendorf tubes (Buch&Holm, Herlev, Denmark). The wash buffer was substituted with 80% EtOH and the total RNA was eluted with Milli-Q water. Subsequently, the integrity and the quantity of the purified RNA were determined on a Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA) using the Eukaryotic Total RNA Nano chip (Agilent) as described by the manufacturer.

Sample labelling

The total RNA from dorsal pancreata was amplified using a combination of the RiboAmp kit (Arcturus, Mountain View, CA, USA) and T7 MEGAscript kit (Ambion). Briefly, generation and purification of double-stranded cDNA was done with components from the Arcturus kit following the manufacturer’s recommendations. The resulting cDNA was in vitro transcribed using components from the MEGAscript kit and aminoallyl-modified UTP (aa-UTP; Sigma). Briefly, the generated double-stranded cDNA was in vitro transcribed in 40 ml reactions (1 X T7 reaction buffer, 7-5 mM ATP, 7-5 mM CTP, 7-5 mM GTP, 3-75 mM UTP, 3-75 mM aa-UTP and 4 µl T7 Enzyme Mix) at 37 °C for 5 h. Subsequently, the cDNA template was digested with 4 U DNaseI for 15 min at 37 °C and amino allyl cRNA

Materials and methods

Tissue isolation

Time-mated Ngn3-mutant mice (Gradwohl et al. 2000) were killed by cervical dislocation at e13 and e15. The dorsal pancreatic bud was isolated from embryos and stored in RNAlater (Ambion, Austin, TX, USA) at 4 °C overnight. The next day, RNAlater was removed and the tissue was stored at −80 °C.

Journal of Molecular Endocrinology (2006) 37, 301–316

www.endocrinology-journals.org

Downloaded from Bioscientifica.com at 08/25/2019 01:40:25AM
via free access
was purified on RNeasy columns (Qiagen). Manufacturer’s recommendations were modified by replacing the RPE wash buffer with 80% EtOH, adding an additional wash step and eluting the aa-cRNA from the columns in two times 30μl Milli-Q water. The purified aa-cRNA was split into 2μg aliquots, completely dried in a SpeedVac (Holm and Halby, Brøndby, Denmark) and stored at −20°C. Coupling of fluorescent dyes to aa-cRNA was done by mixing 2μg aa-cRNA resuspended in 4·5μl freshly made 0·1 M sodium carbonate buffer (pH 9·0) with Alexa Fluor 555 or Alexa Fluor 647 succinimidyl ester (Molecular Probes, Eugene, OR, USA) resuspended in 4·5μl dimethyl sulfoxide (DMSO) (Sigma) and incubating at 25°C for 1 h. Subsequently, uncoupled succinimidyl ester dyes were quenched by the addition of 4·5μl, 4 M hydroxyamine and incubating for 15 min at 25°C. Labelled cRNA was purified on RNeasy columns with the following modifications to manufacturer’s protocol. An additional wash step with 500μl RPE buffer was added and the dye-coupled cRNA was eluted two times in 40μl Milli-Q water. Labelling efficiency and concentration of nucleic acid were determined by measuring absorbance of the cRNA-dye conjugate at λ-max of the dye and 260 nm respectively. All incubation steps were done in a MJ Research PTC-200 Peltier Thermal Cycler.

Microarray
The ArrayTAG 20k murine gene collection (LION Bioscience AG, Heidelberg, Germany) was spotted in duplicate on CodeLink slides (GE Healthcare, Hølserød, Denmark) using a BioRobotics Microgrid II spotter (Genomic Solutions Ltd, Cambridgeshire, UK). Probes contained an amino modification at the 5’ end of the sense strand allowing specific covalent coupling of the sense strand to N-hydroxysuccinimide ester groups on the slide. Prior to hybridisation, residual reactive groups on the slide were blocked by incubating, 0·1 M Tris (pH 9), in 50 mM ethanolamine, at 50°C for 30 min. Slides were rinsed twice with deionised water and incubated in 4×SSC, 0·1% SDS for 30 min. Subsequently, slides were rinsed in deionised water and placed in boiling water for 2 min to render the covalently attached probes single stranded. Following two additional rinses with water, the slides were prehybridised in digoxigenin (DIG) Easy Hyb buffer (Roche) for 30 min at 42°C. Finally, slides were rinsed twice in water and dried using compressed air. Dye-conjugated cRNA corresponding to 30 pmol of each fluorescent dye was mixed with 7μg mouse Cot-1 DNA (Invitrogen) and 1μg poly dA (DNA technology, Århus, Denmark). The hybridisation mix was dried in a SpeedVac and resuspended in 80μl digoxigenin (DIG) Easy Hyb buffer. Nucleic acids of the hybridisation mix were denatured at 95°C for 2 min immediately prior to setting up the hybridisation. Labelled cRNA was hybridised to the arrays under LifterSlips (Erie Scientific Company, Portsmouth, NH, USA) in GeneMachines HybChambers (Genomic Solutions Ltd, Cambridgeshire, UK) immersed in a 42°C water bath for 16 h.

The LifterSlips were removed by immersing slides in 2×SSC, 0·1% SDS, and the slides were placed subsequently in a slide rack immersed in 2×SSC, 0·1% SDS. All post-hybridisation washes were performed on a shaker to ensure gentle agitation of the wash solution. Slides were washed twice in 2×SSC, 0·1% SDS at room temperature for 5 min, twice in 0·5× SSC, 0·1% SDS at 55°C for 15 min, and finally a quick rinse in water immediately prior to drying with compressed air. Slides were scanned in an Axon 4000B scanner at a resolution of 5μm/pixel and image analysis was done in GenePix Pro 5 (Molecular Devices, Sunnyvale, CA, USA).

Experimental setup
The experiment was designed to analyse the expression of genes in dorsal pancreatic cells at two temporally separated stages of pancreas development. This was accomplished by comparing expression profiles of embryonic dorsal pancreas tissue from Ngn3 null mice with wild-type littermates at days 13 and 15 of embryonic development. The comparison of gene expression in mutant and wild-type pancreas was primarily expected to show genes that are lower expressed/missing in the mutant, as Ngn3 null mice have no endocrine pancreas tissue. From each developmental stage, five wild-type and five mutant samples were chosen, representing embryos from at least three different litters. Wild-type and mutant samples from the common stage of development were paired randomly and analysed in flipped colour. Probes were spotted in duplicate on each slide in a randomised (fixed) layout, effectively distributing the duplicate spots randomly over the slide.

Data analysis
Analysis of the microarray data was done using R-2.2.0 (R Development Core Team 2005) and Bioconductor 1·7 packages (Gentleman et al. 2004). The limma package (Smyth et al. 2005) was used for the normalisation and identification of differentially expressed transcripts essentially as described in the Limma User’s Guide. The arrays were normalised by print-tip loess normalisation followed by between slide-scale normalisation. After averaging within slide duplicates, a linear model was fitted to the data and empirical Bayesian smoothing was applied to the standard errors and
finally, the Benjamini and Hochberg procedure of controlling false-discovery rate (Benjamini & Hochberg 1995) was used to adjust \( P \) values and a false-discovery rate of 0.01 was chosen. The dye-swap replicates were treated as described in the user’s guide. The AnnBuilder package (Zhang et al. 2003) was used to build annotation database for the LION array; GO, GOstats and Rgraphviz packages were used for GO analysis and construction of the GO graph.

**DNA constructs**

The mouse cDNA for Ngn3 was a gift from F Guillemot. Mouse cDNA for Irx1 was a gift from Dr Vincent M Christoffel, University of Amsterdam, Holland. The full-length mouse Irx2 was purchased as an IMAGE EST clone. The chicken Irx genes were purchased from MRC Gene Service, Cambridge, UK. For electroporation, the Ngn3 cDNA was subcloned into the pCAGGS-IRE-EGFP vector (a gift from Dr Anne Grapin-Botton). Restriction digests and sequencing confirmed the correct orientation and reading frames of the constructs. Plasmids were purified using Endofree Maxiprep system (Qiagen) according to manufacturer’s instructions.

**In ovo endoderm electroporation**

This method was performed as described earlier (Grapin-Botton et al. 2001, Pedersen & Heller 2005). After 72 h, surviving embryos were recovered and fixed in fresh 4% paraformaldehyde (PFA) in PBS overnight. Only highly electroporated embryos, as determined by the abundance of green fluorescent protein (GFP\(^{+}\)) cells in the embryos, were further processed. Embryos were embedded in OCT compound (Sakura Fineteck Europe, Zoeterwoude, The Netherlands) and frozen on dry ice and stored at −80°C until sectioned.

**Immunocytochemistry and in situ hybridisation**

*In situ* hybridisations were performed on 10 \( \mu \)m frozen sections with DIG-labelled antisense RNA probes for Irx1-5 created using the DIG RNA synthesis kit from Roche, following the manufacturer’s instructions. Mouse cDNAs for Irx1-5 were a gift from Dr Vincent M Christoffel, University of Amsterdam, Holland. The chicken Irx genes were purchased from MRC Gene Service, Cambridge, UK. Probes were hybridised to the tissues at 68°C followed by an AP-conjugated anti-DIG antibody (Roche). Deposition of the Nitro-Blue Tetrazolium/5-Bromo-4-chloro-3-indoly phosphate (NBT/BCIP) chromophore (Sigma) was catalysed in 100 mM NaCl 1 ml 1 M 100 mM Tris, pH9, 51 ml 1 M 50 mM MgCl\(_2\) 1ml 0.5 M 0.1% Tween 20 10 \( \mu \)l 100% (NTMT) buffer at pH 9.5 and fast red (Sigma) signals for the fluorescent *in situ* were developed in 0.1 M Tris (pH 8.2). Immunostaining reactions were performed after antigen retrieval by microwave treatment in 10 mM sodium citrate (pH 6.0). The primary antiserum was incubated overnight at 4°C. Immunoreactivity was visualised with a Tyramide Signal Amplification (TSA) kit (Perkin-Elmer Life Science, Wellesley, MA, USA) or with the Alexa-series of secondary antibodies (Molecular Probes).

### RT-PCR

Quantitative reverse transcriptase (RT)-PCR was performed on an ABI Prism 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA) using Assays-on-Demand Gene Expression Products (Applied Biosystems). Assays for verification of microarray results were delivered in low-density arrays. Total RNA was reverse-transcribed using Reverse Transcriptase Reagents (Applied Biosystems) according to manufacturer’s recommendations. Relative quantification of gene expression measured on low-density arrays was performed using the \( \Delta \Delta C_i \) method. Critical threshold \( (C_i) \) cycle is defined as the cycle in which the reporter signal crosses a given threshold. The \( C_i \) values were normalized to 18S levels by subtracting the 18S \( C_i \) values resulting in \( \Delta C_i \) values. The \( \Delta \Delta C_i \) values were

### Table 1. DNA constructs

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Species</th>
<th>Dilution</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pdx1 253</td>
<td>Rabbit</td>
<td>1:5000</td>
<td>Joel Habener, MGH, Boston, MA</td>
</tr>
<tr>
<td>Pax6</td>
<td>Rabbit</td>
<td>1:2000</td>
<td>Babco, Richmond, CA.</td>
</tr>
<tr>
<td>Nkx6-1</td>
<td>Rabbit</td>
<td>1:5000</td>
<td>Jensen et al. (1996)</td>
</tr>
<tr>
<td>Ngn3</td>
<td>Rabbit</td>
<td>1:1000</td>
<td>Michael German, UCSF</td>
</tr>
<tr>
<td>ISL-1</td>
<td>Rabbit</td>
<td>1:1000</td>
<td>Helena Edlund, Umeå, Sweden</td>
</tr>
<tr>
<td>Insulin</td>
<td>Mouse</td>
<td>1:75</td>
<td>Novo Nordisk A/S, Denmark</td>
</tr>
<tr>
<td>Glucagon</td>
<td>Mouse</td>
<td>1:50</td>
<td>DAKO, Glostrup, Denmark</td>
</tr>
<tr>
<td>PP</td>
<td>Rabbit</td>
<td>1:50</td>
<td>Novo Nordisk A/S, Denmark</td>
</tr>
<tr>
<td>Somatostatin</td>
<td>Mouse</td>
<td>1:10</td>
<td>DAKO, Glostrup, Denmark</td>
</tr>
<tr>
<td>Somatostatin</td>
<td>Rabbit</td>
<td>1:100</td>
<td>DAKO, Glostrup, Denmark</td>
</tr>
<tr>
<td>Ghrelin</td>
<td>Goat</td>
<td>1:500</td>
<td>Santa Cruz Biotechnology, CA</td>
</tr>
<tr>
<td>cMYC (A14)</td>
<td>Rabbit</td>
<td>1:80</td>
<td>Santa Cruz Biotechnology, CA</td>
</tr>
<tr>
<td>cMYC (9E10)</td>
<td>Mouse</td>
<td>1:50</td>
<td>Santa Cruz Biotechnology, CA</td>
</tr>
</tbody>
</table>

**Journal of Molecular Endocrinology** (2006) 37, 301–316

www.endocrinology-journals.org

Downloaded from Bioscientifica.com at 08/25/2019 01:40:45AM via free access
obtained by subtracting ΔCt values from mutant and wild type, resulting in a log2 ratio of target gene expression in mutant over wild type (equivalent to the M value determined in the microarray data analysis). The following assays were ordered as low-density arrays: Mm00514341_m1 (Chga), Mm00517617_m1 (Cryba2), Mm00432448_m1 (Cdkn1a), Mm00484138_m1 (Cyp4b1), Mm00456190_m1 (Mvt1), Mm00440465_g1 (Neurod2), Mm00440480_m1 (Nnat), Mm00499982_m1 (Pcdh21), Mm00485697_m1 (Resp18), Mm00803260_m1 (Rbp4), Mm00460511_m1 (Rhbdl4), Mm00460679_m1 (Sesn2), Mm00600526_m1 (Syt13), Mm00443220_m1 (Tcp11), Mm00451001_m1 (Tm4sf2), Mm00491265_m1 (2510004L01Rik), Mm00549967_m1 (C81234), Mm00731595_gH (Ins2), Mm00437606_s1 (Neurog3).

Results

Microarray analysis of Ngn3-deficient mice

To examine the difference in transcriptomes of dorsal pancreas from wild-type and Ngn3-deficient mice, a spotted cDNA microarray comprising 20,188 probes covering 12,140 genes plus ESTs was used. The comparison was made at two different stages of pancreas development, allowing monitoring of large-scale parallel gene expression at the beginning and end of the secondary transition. Total RNA was isolated from homozygous wild-type and mutant pancreata, and a single round of RNA amplification yielded sufficient material for microarray analysis of individual pancreas.

Spatial analysis of the foreground and background intensities obtained after image analysis of the hybridised arrays did not reveal any major flaws (data not shown; Petri et al. 2004). Data analysis was performed as described in ‘Materials and methods’ and resulted in identification of 504 transcripts showing differential expression between Ngn3-deficient and wild-type mice at e13 and/or e15 (see Fig. 1). The MA plots show that many of the identified genes show only relatively small differences in expression levels (|M|<1 corresponding to less than twofold). Furthermore, the majority of genes show decreased expression levels in mutant mice compared to wild type. To identify genes that exhibited a robust difference between wild type and mutant, we defined an M value cut-off corresponding to at least twofold difference. This filtering strategy reduced the list of genes to 52 showing significant difference at e13 and/or e15 (Table 1), all with decreased expression in Ngn3-deficient mice. Expression of several neuropeptide genes (Ghrl, Gip, Gal, Cck and Npy) as well as genes whose products are involved in hormone processing (Pcsk2, Pcsk1n, and 7B2) are lacking in the mutant mice. Multiple members of the granin family (Chga, Chgb, Scg3), which are known to be expressed in endocrine cells of the pancreas, displayed large differences between mutant and wild-type mice (M<−2). Two members (Syt7 and Syt13) of the synaptotagmin family of vesicle proteins, which are believed to mediate calcium-dependent regulation of membrane trafficking (Yoshihara & Montana 2004) and transthyretin, which has an important role in the plasma transport of thyroid hormone T4 and retinol
### Table 1: List of genes showing robust difference in expression levels between Ngn3-deficient and wild-type mice at e13 or e15

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hormone, hormone processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-2n7</td>
<td>Cholecystokinin</td>
<td>Cck</td>
<td>-2.18</td>
<td>$3.19 \times 10^{-9}$</td>
<td>-1.99</td>
<td>$3.21 \times 10^{-7}$</td>
</tr>
<tr>
<td>MG-8-92o10</td>
<td>Galectin</td>
<td>Gal</td>
<td>-1.40</td>
<td>$5.34 \times 10^{-7}$</td>
<td>-0.57</td>
<td>$9.32 \times 10^{-6}$</td>
</tr>
<tr>
<td>MG-12-168g11</td>
<td>Gastric inhibitory polypeptide</td>
<td>Gip</td>
<td>-1.26</td>
<td>$1.28 \times 10^{-4}$</td>
<td>-0.55</td>
<td>$3.10 \times 10^{-4}$</td>
</tr>
<tr>
<td>MG-3-3o2</td>
<td>Ghrelin</td>
<td>MGI:1930008</td>
<td>-2.56</td>
<td>$1.23 \times 10^{-7}$</td>
<td>-2.23</td>
<td>$1.96 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-6-55p2</td>
<td>Neurotensin Y</td>
<td>Npy</td>
<td>-0.94</td>
<td>$3.45 \times 10^{-4}$</td>
<td>-2.26</td>
<td>$1.16 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-6-22i2</td>
<td>Proprotein convertase subtilisin/kexin type 1 inhibitor</td>
<td>Pcsk1n</td>
<td>-1.84</td>
<td>$7.25 \times 10^{-7}$</td>
<td>-1.37</td>
<td>$4.66 \times 10^{-9}$</td>
</tr>
<tr>
<td>MG-3-7en17</td>
<td>Proprotein convertase subtilisin/kexin type 2</td>
<td>Pcsk2</td>
<td>-2.22</td>
<td>$4.34 \times 10^{-7}$</td>
<td>-2.22</td>
<td>$9.20 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-14-4o13</td>
<td>Secretory granule neuroendocrine protein 1, 7B2 protein</td>
<td>Sgne1</td>
<td>-1.26</td>
<td>$2.52 \times 10^{-7}$</td>
<td>-1.69</td>
<td>$2.21 \times 10^{-8}$</td>
</tr>
<tr>
<td>Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-3j21</td>
<td>Chromogranin A</td>
<td>Chga</td>
<td>-3.02</td>
<td>$1.07 \times 10^{-9}$</td>
<td>-2.50</td>
<td>$2.21 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-14-45h10</td>
<td>Chromogranin B</td>
<td>Chgb</td>
<td>-3.17</td>
<td>$7.47 \times 10^{-8}$</td>
<td>-2.38</td>
<td>$1.18 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-6-55a14</td>
<td>Secretogranin III</td>
<td>Scg3</td>
<td>-1.98</td>
<td>$3.72 \times 10^{-7}$</td>
<td>-2.03</td>
<td>$6.61 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-8-10i4</td>
<td>Retinol binding protein 4, plasma</td>
<td>Rbp4</td>
<td>-1.62</td>
<td>$6.37 \times 10^{-7}$</td>
<td>-1.51</td>
<td>$4.00 \times 10^{-9}$</td>
</tr>
<tr>
<td>MG-6-25p16</td>
<td>Secretagogin, EF-hand calcium binding protein</td>
<td>Scgn</td>
<td>-1.99</td>
<td>$4.07 \times 10^{-6}$</td>
<td>-2.15</td>
<td>$2.45 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-16-11p10</td>
<td>Protocadherin 21</td>
<td>Pcdh21</td>
<td>-0.97</td>
<td>$2.77 \times 10^{-4}$</td>
<td>-1.23</td>
<td>$2.41 \times 10^{-7}$</td>
</tr>
<tr>
<td>Transcription, translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-4c22</td>
<td>Iroquois related homeobox 1 (Drosophila)</td>
<td>Irx1</td>
<td>-1.35</td>
<td>$2.39 \times 10^{-7}$</td>
<td>-0.78</td>
<td>$2.28 \times 10^{-5}$</td>
</tr>
<tr>
<td>MG-6-27f14</td>
<td>Myelin transcription factor 1</td>
<td>Myt1</td>
<td>-1.03</td>
<td>$1.32 \times 10^{-5}$</td>
<td>-1.21</td>
<td>$2.21 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-6-2g18</td>
<td>Ribosomal protein S12</td>
<td>Rps12</td>
<td>-1.04</td>
<td>$3.31 \times 10^{-8}$</td>
<td>-1.38</td>
<td>$4.37 \times 10^{-8}$</td>
</tr>
<tr>
<td>Transport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-64j8</td>
<td>Synaptotagmin 7</td>
<td>Syt7</td>
<td>-0.67</td>
<td>$5.84 \times 10^{-5}$</td>
<td>-1.25</td>
<td>$5.59 \times 10^{-10}$</td>
</tr>
<tr>
<td>MG-6-64b7</td>
<td>Synaptotagmin 7</td>
<td>Syt7</td>
<td>-0.90</td>
<td>$5.34 \times 10^{-7}$</td>
<td>-1.40</td>
<td>$5.59 \times 10^{-10}$</td>
</tr>
<tr>
<td>MG-8-12f2</td>
<td>Synaptotagmin 13</td>
<td>Syt13</td>
<td>-1.65</td>
<td>$3.68 \times 10^{-6}$</td>
<td>-1.71</td>
<td>$8.88 \times 10^{-9}$</td>
</tr>
<tr>
<td>MG-14-11n19</td>
<td>Transnethrin</td>
<td>Trt</td>
<td>-1.73</td>
<td>$6.57 \times 10^{-7}$</td>
<td>-0.40</td>
<td>$7.86 \times 10^{-4}$</td>
</tr>
<tr>
<td>MG-11-1e9</td>
<td>Transnethrin</td>
<td>Trt</td>
<td>-1.74</td>
<td>$4.07 \times 10^{-6}$</td>
<td>-0.41</td>
<td>$4.44 \times 10^{-4}$</td>
</tr>
<tr>
<td>Signalling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-36e16</td>
<td>Rhomboid, veinlet-like 4 (Drosophila)</td>
<td>Rhibd4</td>
<td>-0.98</td>
<td>$2.35 \times 10^{-4}$</td>
<td>-1.00</td>
<td>$2.55 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-8-89n12</td>
<td>GNAS (guanine nucleotide binding protein, alpha stimulating) complex locus</td>
<td>Gnas</td>
<td>-1.90</td>
<td>$1.15 \times 10^{-7}$</td>
<td>-2.18</td>
<td>$2.21 \times 10^{-8}$</td>
</tr>
<tr>
<td>Kinase/phosphatase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-30c20</td>
<td>Protein phosphatase 1, regulatory (inhibitor) subunit 1A</td>
<td>Ppp1r1a</td>
<td>-0.83</td>
<td>$3.72 \times 10^{-7}$</td>
<td>-1.60</td>
<td>$4.23 \times 10^{-8}$</td>
</tr>
<tr>
<td>Not classified</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-19-1k5</td>
<td>Ureidopropionase, beta</td>
<td>Upb1</td>
<td>-1.16</td>
<td>$1.53 \times 10^{-6}$</td>
<td>-0.52</td>
<td>$4.92 \times 10^{-5}$</td>
</tr>
<tr>
<td>MG-6-18n16</td>
<td>Nucleosome assembly protein 1-like 5</td>
<td>Nap115</td>
<td>-1.75</td>
<td>$1.79 \times 10^{-6}$</td>
<td>-0.69</td>
<td>$5.24 \times 10^{-6}$</td>
</tr>
<tr>
<td>MG-12-14o7</td>
<td>Serine protease inhibitor, Kazal type 3</td>
<td>Spink8</td>
<td>-1.53</td>
<td>$1.76 \times 10^{-4}$</td>
<td>-0.53</td>
<td>$1.76 \times 10^{-1}$</td>
</tr>
<tr>
<td>MG-8-72n2</td>
<td>Fibrinogen, B beta polypeptide</td>
<td>Fgb</td>
<td>-1.33</td>
<td>$3.20 \times 10^{-3}$</td>
<td>-0.28</td>
<td>$1.03 \times 10^{-1}$</td>
</tr>
<tr>
<td>MG-60-1p11</td>
<td>Crystallin, beta A2</td>
<td>Cryba2</td>
<td>-0.74</td>
<td>$1.76 \times 10^{-4}$</td>
<td>-1.61</td>
<td>$1.89 \times 10^{-8}$</td>
</tr>
<tr>
<td>MG-6-38n22</td>
<td>Regulated endocrine-specific protein 18</td>
<td>Resp18</td>
<td>-1.59</td>
<td>$1.37 \times 10^{-7}$</td>
<td>-1.23</td>
<td>$2.16 \times 10^{-7}$</td>
</tr>
<tr>
<td>MG-3-8vb2</td>
<td>Solute carrier family 38, member 5</td>
<td>Sli38a5</td>
<td>-3.84</td>
<td>$1.15 \times 10^{-7}$</td>
<td>-2.21</td>
<td>$2.79 \times 10^{-9}$</td>
</tr>
<tr>
<td>EST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG-6-60f22</td>
<td>RIKEN cDNA 1700012H17 gene</td>
<td>1700012H17Rik</td>
<td>-1.86</td>
<td>$7.72 \times 10^{-7}$</td>
<td>-0.43</td>
<td>$1.08 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
Microarray analysis of NGN3 mutants

A. Petri, J. Ahnfelt-Rönn and others

Microarray analysis of NGN3 mutants

found in Supplementary Tables 2 and 3 (e13 and e15 data type at e13 and/or e15). The complete list of genes can be presented in Table 1 represent only a small subset (52 out of 504 genes) of the transcripts shown by microarray analysis to differ significantly between mutant and wild type. However, the genes showing significant and robust difference in expression between mutant and wild type. Expression of genes involved in signalling pathways is also affected by Ngf5 deficiency. Rhomboid venlet-like 4, a vertebrate homologue of Drosophila Rhomboid (Jaszai & Brand 2002) encoding a protein involved in regulated intramembrane proteolysis, as well as Stimulatory G protein alpha subunit, are expressed at lower levels in mutant mice. Several ESTs and transcripts of genes with unknown function (RIKEN clones) are found at lower levels in dorsal pancreas from mutant mice, some of which show large differences. Interestingly, in this list of robustly differentially expressed transcripts, we detect two transcription factors, namely Mtv1L, which was recently reported to be expressed in endocrine precursor cells (Gu et al. 2004) and Iroquois related homeobox 1 (Irx1). Irx1 belongs to a family of homeodomain proteins, which is conserved from nematodes to humans (Cavodeassi et al. 2001). Six different Irx genes have been identified in mouse, of which Irx1, Irx3 and Irx5 are present on the array, but differential expression between mutant and wild type is only detected for Irx1.

Differential expression of several transcripts showing lower expression in mutant mice was verified by RT-PCR on total RNA isolated from another batch of mice (n = 3). We verified the expression of a few genes showing more than twofold difference as well as some genes showing less than twofold difference in expression levels at either e13 or e15 (see Fig. 2). The data from these experiments have been summarised in Fig. 2 and the actual values are shown in Supplementary Table 1 (See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). The RT-PCR results corroborated the microarray data. Only Cyp4b1 shows disagreement between the two platforms at both e13 and e15, whereas Cdkn1a does so at e15.

Further microarray data analysis

The filtration of microarray data described in the previous section resulted in identification of genes showing significant and robust difference in expression between mutant and wild type. However, the genes presented in Table 1 represent only a small subset (52 out of 504 genes) of the transcripts shown by microarray analysis to differ significantly between mutant and wild type at e13 and/or e15. The complete list of genes can be found in Supplementary Tables 2 and 3 (e13 and e15 data respectively: See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). In particular, the additional filtration results in omission of a substantial fraction of genes showing increased expression, albeit small, at e13 (Fig. 1, red dots). The full sets of genes identified at e13 and e15 were divided

(Soares et al. 2003) are expressed at decreased levels in the mutant. Several genes encoding proteins involved in phosphorylation reactions are showing robust differences between mutant and wild type. Expression of genes involved in signalling pathways is also affected by Ngf5 deficiency. Rhomboid venlet-like 4, a vertebrate homologue of Drosophila Rhomboid (Jaszai & Brand 2002) encoding a protein involved in regulated intramembrane proteolysis, as well as Stimulatory G protein alpha subunit, are expressed at lower levels in mutant mice. Several ESTs and transcripts of genes with unknown function (RIKEN clones) are found at lower levels in dorsal pancreas from mutant mice, some of which show large differences. Interestingly, in this list of robustly differentially expressed transcripts, we detect two transcription factors, namely Mtv1L, which was recently reported to be expressed in endocrine precursor cells (Gu et al. 2004) and Iroquois related homeobox 1 (Irx1). Irx1 belongs to a family of homeodomain proteins, which is conserved from nematodes to humans (Cavodeassi et al. 2001). Six different Irx genes have been identified in mouse, of which Irx1, Irx3 and Irx5 are present on the array, but differential expression between mutant and wild type is only detected for Irx1.

Differential expression of several transcripts showing lower expression in mutant mice was verified by RT-PCR on total RNA isolated from another batch of mice (n = 3). We verified the expression of a few genes showing more than twofold difference as well as some genes showing less than twofold difference in expression levels at either e13 or e15 (see Fig. 2). The data from these experiments have been summarised in Fig. 2 and the actual values are shown in Supplementary Table 1 (See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). The RT-PCR results corroborated the microarray data. Only Cyp4b1 shows disagreement between the two platforms at both e13 and e15, whereas Cdkn1a does so at e15.

Further microarray data analysis

The filtration of microarray data described in the previous section resulted in identification of genes showing significant and robust difference in expression between mutant and wild type. However, the genes presented in Table 1 represent only a small subset (52 out of 504 genes) of the transcripts shown by microarray analysis to differ significantly between mutant and wild type at e13 and/or e15. The complete list of genes can be found in Supplementary Tables 2 and 3 (e13 and e15 data respectively: See Supplementary data at http://jme. endocrinology-journals.org/content/vol37/issue2). In particular, the additional filtration results in omission of a substantial fraction of genes showing increased expression, albeit small, at e13 (Fig. 1, red dots). The full sets of genes identified at e13 and e15 were divided

(Soares et al. 2003) are expressed at decreased levels in the mutant. Several genes encoding proteins involved in phosphorylation reactions are showing robust differences between mutant and wild type. Expression of genes involved in signalling pathways is also affected by Ngf5 deficiency. Rhomboid venlet-like 4, a vertebrate homologue of Drosophila Rhomboid (Jaszai & Brand 2002) encoding a protein involved in regulated intramembrane proteolysis, as well as Stimulatory G protein alpha subunit, are expressed at lower levels in mutant mice. Several ESTs and transcripts of genes with unknown function (RIKEN clones) are found at lower levels in dorsal pancreas from mutant mice, some of which show large differences. Interestingly, in this list of robustly differentially expressed transcripts, we detect two transcription factors, namely Mtv1L, which was recently reported to be expressed in endocrine precursor cells (Gu et al. 2004) and Iroquois related homeobox 1 (Irx1). Irx1 belongs to a family of homeodomain proteins, which is conserved from nematodes to humans (Cavodeassi et al. 2001). Six different Irx genes have been identified in mouse, of which Irx1, Irx3 and Irx5 are present on the array, but differential expression between mutant and wild type is only detected for Irx1.

Differential expression of several transcripts showing lower expression in mutant mice was verified by RT-PCR on total RNA isolated from another batch of mice (n = 3). We verified the expression of a few genes showing more than twofold difference as well as some genes showing less than twofold difference in expression levels at either e13 or e15 (see Fig. 2). The data from these experiments have been summarised in Fig. 2 and the actual values are shown in Supplementary Table 1 (See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). The RT-PCR results corroborated the microarray data. Only Cyp4b1 shows disagreement between the two platforms at both e13 and e15, whereas Cdkn1a does so at e15.

Further microarray data analysis

The filtration of microarray data described in the previous section resulted in identification of genes showing significant and robust difference in expression between mutant and wild type. However, the genes presented in Table 1 represent only a small subset (52 out of 504 genes) of the transcripts shown by microarray analysis to differ significantly between mutant and wild type at e13 and/or e15. The complete list of genes can be found in Supplementary Tables 2 and 3 (e13 and e15 data respectively: See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). In particular, the additional filtration results in omission of a substantial fraction of genes showing increased expression, albeit small, at e13 (Fig. 1, red dots). The full sets of genes identified at e13 and e15 were divided
into groups of genes showing increased or decreased expression in the mutant. Subsequently, these gene lists were analysed for over-representation of gene ontology (GO) terms in the biological process (BP) branch of the ontology (see Table 2). The number of genes showing increased expression at e15 is quite small, not many terms are identified and those identified contain few genes. However, the other groups of genes show several notable terms. Not surprisingly, we find ‘peptide hormone processing’ and ‘neuropeptide-signalling pathway’ represented by genes showing decreased expression in the mutant at both e13 and e15. At e15, we further identify genes involved in ‘transport’ and ‘G-protein-coupled receptor protein-signalling pathway’ in the group of genes showing reduced expression. The most significant finding at e13 among decreased transcripts is terms related to tetrahydrobiopterin biosynthesis. Tetrahydrobiopterin is an obligatory co-factor of nitric oxide synthases (Kwon et al. 1989, Tayeh & Marletta 1989) and in vitro experiments in INS-1 cells have shown that tetrahydrobiopterin synthesis precedes nitric oxide-dependent inhibition of insulin secretion (Laffranchi et al. 1997). The terms that are associated with the largest number of genes showing reduced expression at e13 are related to apoptosis. Interestingly, genes involved in ‘cell communication’ and ‘cell adhesion’ show increased expression in the mutant at e13 and there is also an over-representation of genes involved in ‘neurogenesis’. To further compare the difference between increased and decreased genes at e13, we computed the induced GO graph of genes identified at e13 and used a green–black–red colour gradient to indicate whether genes annotated at a particular node (either directly or by inheritance) showed increased expression, decreased expression, or a mix of both (Fig. 3 and Supplementary Figure S1 See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). GO terms that were identified as significantly over-represented (Table 2) have been emphasized in Fig. 3. While only few terms are identified as significantly over-represented, this graph identifies processes, which are only affected by decreasing gene expression. Some of these terms have been labelled in Fig. 3.

**α-Cell-specific Irx1/2 expression depend on and is induced by NGN3**

Irx1 and Irx2 have previously been shown to be expressed in the pancreas (Houweling et al. 2001) but their function and cell-specific expression pattern are unknown. Our microarray data suggest that Irx1 is expressed in the endocrine lineage. To determine whether Irx1 and Irx2 were expressed in mature endocrine cells or their precursors, we performed fluorescent in situ hybridisations for the Irx1/2 genes followed by immunofluorescent stainings with a marker specific for these two populations. Irx1- and Irx2-expressing cells were found to be ISL1+ (Fig. 4A and C), a marker of mature endocrine cells (Thor et al. 1991, Ahlgren et al. 1997), but negative for the precursor marker NGN3 (Gradwohl et al. 2000, Jensen et al. 2000a, Schwitzgebel et al. 2000, Gu et al. 2002) although Irx1/2-expressing cells were often found in close association with NGN3+ cells (Fig. 4B and D).
To determine if Irx1/2 gene expression was restricted to a particular endocrine subtype, we performed in situ hybridisation for Irx1/2 (Fig. 5A and F) followed by fluorescent antibody staining for the major pancreatic hormones. We found Irx1 and Irx2 to be co-expressed with glucagon (Fig. 5D and I), but not insulin (Fig. 5C and H) and somatostatin at e18 (Fig. 5E and J). Irx1/2 did co-localize with a subpopulation of PP cells, most of which also expressed glucagon (data not shown). The α-cell-specific expression of Irx1/2 was also seen in E4 chicken pancreas (Fig. 5A–D). Additionally, RT-PCR analysis demonstrated prominent Irx1/2 expression in the mouse glucagonoma cell line, αTC1.9 (Hamaguchi & Leiter 1990) while the insulinoma cell line βTC3 (Efrat et al. 1988) had barely detectable levels (data not shown).

Consistent with our microarray data, expression of Irx1/2 was lost in mice deficient in Ngn3 (Fig. 6A–D). Furthermore, forced expression of Ngn3 in chick endoderm induced the expression of both Irx1 (data

Table 2 Over-represented GO BP terms

<table>
<thead>
<tr>
<th>GO id</th>
<th>Term</th>
<th>Genes</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0046146</td>
<td>Tetrahydriobiopterin metabolism</td>
<td>3</td>
<td>7.02×10⁻⁵</td>
</tr>
<tr>
<td>GO:0006729</td>
<td>Tetrahydriobiopterin biosynthesis</td>
<td>3</td>
<td>7.02×10⁻⁵</td>
</tr>
<tr>
<td>GO:0042559</td>
<td>Pteridine and derivative biosynthesis</td>
<td>3</td>
<td>5.56×10⁻⁴</td>
</tr>
<tr>
<td>GO:0042558</td>
<td>Pteridine and derivative metabolism</td>
<td>3</td>
<td>7.84×10⁻⁴</td>
</tr>
<tr>
<td>GO:0019438</td>
<td>Aromatic compound biosynthesis</td>
<td>3</td>
<td>1.06×10⁻³</td>
</tr>
<tr>
<td>GO:0016486</td>
<td>Peptide hormone processing</td>
<td>2</td>
<td>3.62×10⁻³</td>
</tr>
<tr>
<td>GO:0008219</td>
<td>Cell death</td>
<td>12</td>
<td>4.15×10⁻³</td>
</tr>
<tr>
<td>GO:0016265</td>
<td>Death</td>
<td>12</td>
<td>4.69×10⁻³</td>
</tr>
<tr>
<td>GO:0043064</td>
<td>Flagellum organization and biogenesis</td>
<td>2</td>
<td>5.36×10⁻³</td>
</tr>
<tr>
<td>GO:0007218</td>
<td>Neuropeptide-signalling pathway</td>
<td>4</td>
<td>5.52×10⁻³</td>
</tr>
<tr>
<td>GO:0000082</td>
<td>G1/S transition of mitotic cell cycle</td>
<td>3</td>
<td>5.57×10⁻³</td>
</tr>
<tr>
<td>GO:0006915</td>
<td>Apoptosis</td>
<td>11</td>
<td>6.68×10⁻³</td>
</tr>
<tr>
<td>GO:0012501</td>
<td>Programmed cell death</td>
<td>11</td>
<td>7.10×10⁻³</td>
</tr>
<tr>
<td>GO:0007585</td>
<td>Respiratory gaseous exchange</td>
<td>2</td>
<td>9.76×10⁻³</td>
</tr>
<tr>
<td>GO:0006558</td>
<td>l-phenylalanine metabolism</td>
<td>2</td>
<td>9.76×10⁻³</td>
</tr>
</tbody>
</table>

Over-represented GO BP terms, e13 decreased expression

<table>
<thead>
<tr>
<th>GO id</th>
<th>Term</th>
<th>Genes</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0009880</td>
<td>Embryonic pattern specification</td>
<td>3</td>
<td>3.01×10⁻⁴</td>
</tr>
<tr>
<td>GO:0007224</td>
<td>Smoothed receptor-signalling pathway</td>
<td>2</td>
<td>1.61×10⁻³</td>
</tr>
<tr>
<td>GO:0009953</td>
<td>Dorsal/ventral pattern formation</td>
<td>2</td>
<td>1.61×10⁻³</td>
</tr>
<tr>
<td>GO:0007389</td>
<td>Pattern specification</td>
<td>4</td>
<td>1.83×10⁻³</td>
</tr>
<tr>
<td>GO:0007389</td>
<td>Neurogenesis</td>
<td>7</td>
<td>3.23×10⁻³</td>
</tr>
<tr>
<td>GO:0007420</td>
<td>Brain development</td>
<td>3</td>
<td>4.52×10⁻³</td>
</tr>
<tr>
<td>GO:0009799</td>
<td>Determination of symmetry</td>
<td>2</td>
<td>4.70×10⁻³</td>
</tr>
<tr>
<td>GO:0009855</td>
<td>Determination of bilateral symmetry</td>
<td>2</td>
<td>4.70×10⁻³</td>
</tr>
<tr>
<td>GO:0007368</td>
<td>Determination of left/right symmetry</td>
<td>2</td>
<td>4.70×10⁻³</td>
</tr>
<tr>
<td>GO:0000059</td>
<td>Protein-nucleus import, docking</td>
<td>2</td>
<td>5.71×10⁻³</td>
</tr>
<tr>
<td>GO:0007155</td>
<td>Cell adhesion</td>
<td>8</td>
<td>5.97×10⁻³</td>
</tr>
<tr>
<td>GO:0007417</td>
<td>Central nervous system development</td>
<td>3</td>
<td>6.31×10⁻³</td>
</tr>
<tr>
<td>GO:0007154</td>
<td>Cell communication</td>
<td>22</td>
<td>7.51×10⁻³</td>
</tr>
</tbody>
</table>

Over-represented GO BP terms, e13 increased expression

<table>
<thead>
<tr>
<th>GO id</th>
<th>Term</th>
<th>Genes</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0008816</td>
<td>Calcium ion transport</td>
<td>8</td>
<td>3.02×10⁻⁶</td>
</tr>
<tr>
<td>GO:0016486</td>
<td>Peptide hormone processing</td>
<td>3</td>
<td>5.63×10⁻⁵</td>
</tr>
<tr>
<td>GO:0007017</td>
<td>Microtubule-based process</td>
<td>9</td>
<td>7.93×10⁻⁵</td>
</tr>
<tr>
<td>GO:0015674</td>
<td>Dl-, tri-valent inorganic cation transport</td>
<td>8</td>
<td>8.10×10⁻⁵</td>
</tr>
<tr>
<td>GO:0006810</td>
<td>Transport</td>
<td>38</td>
<td>4.44×10⁻⁴</td>
</tr>
<tr>
<td>GO:0042445</td>
<td>Hormone metabolism</td>
<td>4</td>
<td>1.49×10⁻³</td>
</tr>
<tr>
<td>GO:0006812</td>
<td>Cation transport</td>
<td>11</td>
<td>1.62×10⁻³</td>
</tr>
<tr>
<td>GO:0016485</td>
<td>Protein processing</td>
<td>3</td>
<td>2.24×10⁻³</td>
</tr>
<tr>
<td>GO:0030001</td>
<td>Metal ion transport</td>
<td>9</td>
<td>2.55×10⁻³</td>
</tr>
<tr>
<td>GO:0007186</td>
<td>G-protein-coupled receptor protein-signalling pathway</td>
<td>10</td>
<td>3.34×10⁻³</td>
</tr>
<tr>
<td>GO:0006811</td>
<td>Ion transport</td>
<td>13</td>
<td>5.09×10⁻³</td>
</tr>
<tr>
<td>GO:0007026</td>
<td>Microtubule stabilisation</td>
<td>2</td>
<td>8.47×10⁻³</td>
</tr>
</tbody>
</table>

Over-represented GO BP terms, e15 increased expression

<table>
<thead>
<tr>
<th>GO id</th>
<th>Term</th>
<th>Genes</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0016049</td>
<td>Cell growth</td>
<td>2</td>
<td>2.94×10⁻³</td>
</tr>
<tr>
<td>GO:0008361</td>
<td>Regulation of cell size</td>
<td>2</td>
<td>3.21×10⁻³</td>
</tr>
<tr>
<td>GO:0040007</td>
<td>Growth</td>
<td>2</td>
<td>5.88×10⁻³</td>
</tr>
</tbody>
</table>
Only decreasing

Only increasing

Figure 3 Induced GO tree of all genes showing significant difference between Ngn3-deficient and wild-type mice. Each node in the tree corresponds to a particular GO term, which might be child to one or more GO terms and in turn be parent to the same. The root of the GO tree corresponds to very generic terms and progressing upwards from parent term to child term results in increasingly more specific GO terms. Some nodes in this induced GO graph are not represented by probes on the chip and these have been coloured grey.

Figure 4 Irx1- and Irx2-expressing cells belong to the endocrine lineage (A–D) Localisation of Irx1 and Irx2 with ISL1 and NGN3 at e16 visualised by fluorescent in situ hybridisations for Irx1 and Irx2 (in red) followed by immunofluorescent stainings for NGN3 and ISL1 (in green). ISL1 is detectable in Irx1- (A) and Irx2- (C) positive cell clusters showing their endocrine nature. (B, D) Some NGN3-positive endocrine precursors lie close to Irx1- and Irx2-expressing cells. A and C are 100 × magnification. B and D are 250 × magnification.
and Irx2 and a subset of the Ngn3/Irx positive cell clusters were also glucagon immunoreactive (Fig. 6E–H).

Discussion

The microarray experiment described here was set up to analyse changes in gene expression in the pancreas of Ngn3-deficient mice. Owing to the absence of Ngn3, cells of the developing pancreas cannot adopt an endocrine cell fate resulting in mice lacking endocrine pancreatic cells (Gradwohl et al. 2000). The transcription profiling was done on spotted cDNA microarrays containing 20,188 probes from LION Bioscience’s murine arrayTAG set. A large fraction of the clones in the murine LION clone set were derived from tissues such as brain, testis, embryos and embryonic stem cells, but not pancreas. Therefore, several pancreas-related genes including insulin, glucagon, Ngn3, Ptx1, NeuroD, Isl1 and Pax4 are not present on the array. Nonetheless, due to the many similarities between pancreatic endocrine cells and neuronal cells, both in terms of developmental mechanisms and expressed genes (Slack 1995, Wilson 2003), the murine arrayTAG clone collection is well suited for the analysis of gene expression in pancreas lacking endocrine cells. Initially, we analysed the data to find genes showing significant and more than twofold difference between mutant and wild type at e13 and/or e15. This strict filtering
approach resulted in identification of a small set of 52 genes, all showing lower transcription levels in mutant compared to wild-type mice (see Fig. 1 and Table 1). The simplest explanation for the reduced level of transcription of these genes in the mutant is that they are normally expressed in the endocrine lineage of the pancreas and due to the lack of this lineage in Ngn3-deficient mice, the transcript levels compared to wild

Figure 6 Irx1- and Irx2-expressing cells belong to the endocrine lineage and can be induced by exogenous Ngtn3. (A–D) Irx1 and Irx2 expression in e14 wild-type (WT) and Ngn3-deficient (Ngn3<sup>KO/KO</sup>) pancreata visualised by in situ hybridisation. (A–B) Irx1 and Irx2 are expressed in a characteristic pattern resembling endocrine cell clusters, which mainly consist of α-cells at this time point and specific expression is denoted by arrows (C and D). The expression of Irx1 and Irx2 is lost in the Ngn3 KO demonstrating that these genes are expressed in the endocrine lineage. Black arrows denote areas of Irx expression, while white arrows depict areas with glucagon-expressing cells. Asterisks denote autofluorescent red blood cells. (E–H) Expression of Ngn3 in chicken endoderm induces the expression of Irx2. (E) Expression of the pEMPty-IRES-EGFP plasmid does not induce Irx2 expression (F), while expression of pNgn-3-IRES-EGFP (G) induces glucagon and Irx2 expression (H). White and black arrows denote the areas on the section where glucagon and Irx2 are expressed in the same area. A–D, F, and H are 100× magnification. E and G are 250× magnification.
type appear lower. Indeed, many of the genes in Table 1 are known from the literature to be expressed in the endocrine lineage (Npy (Ding et al. 1997), Ghr (Wierup et al. 2002, Prado et al. 2004, Heller et al. 2005), Pskln (Guest et al. 2002), Psk2 (Furuta et al. 1997), Myt1 (Gu et al. 2004), Chga, Chgb (Karlssoon 2001), Scg3 (Sakai et al. 2004), Cdkn1a (Kaneto et al. 1999)). However, several genes in Table 1 have not yet been described in pancreas biology. Rhomboid, veinlet-like 4 is a vertebrate homolog of Drosophila Rhomboid. In Drosophila, Rhomboid-1 is a key regulator of EGF-receptor signalling (Bier et al. 1990), controlling cleavage of the membrane-bound EGF-receptor ligand Spitz. The components of EGFR signalling are ubiquitously expressed during Drosophila development, whereas the rhomboid proteases are not and the regulated intramembrane cleavage of Spitz by Rhomboid is thus the primary mechanism for controlling EGFR signalling in developing Drosophila embryos (Rutledge et al. 1992, Urban et al. 2002). Interestingly, EGF pathway components are expressed in the developing pancreas in mice (Miettinen et al. 2000) and Erbb1 as well as Erbb2 is required for normal endocrine pancreas development (Miettinen et al. 2000, Huotari et al. 2002). Further studies are needed to determine the location and function of Rhomboid, veinlet-like 4 in endocrine pancreas development. Additionally, many ESTs and genes encoding proteins with unknown functions were identified in our analysis. Some of the ESTs identified in this experiment do not have sequence homology to any sequences in GenBank and since the sequences on this platform are derived from the 3’UTR, homology search might not reveal homology to known families of proteins with conserved motifs. RT-PCR verifications of microarray data showed good agreement between the two platforms (Fig. 2 and Supplementary Table 1; See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2). Slight discrepancies were only observed for Cyp4b1 and Cdkn1a. BLAST analysis of the probe sequences for Cyp4b1 revealed that, in addition to Cyp4b1, the probe showed homology to a predicted gene similar to Cyp4b1 (CYP1B1, XM_983069; data not shown). This suggests that the Cyp4b1 discrepancy between microarray and RT-PCR could be caused by cross-hybridisation. Cdkn1a data only show discrepancy at e15, where microarray data shows 0–8-fold reduction, while RT-PCR shows 1–2-fold increase in mutants. BLAST analysis of the Cdkn1a probe sequence did not show any unexpected hybridisation partners and discrepancy at only e15 suggests minor inaccuracies in one of the platforms. In a previous study of gene expression in pancreas, it was shown that Cdkn1a is enriched in Ngn3+ cells versus pancreatic tissue of other stages (Gu et al. 2004) which is in line with the microarray data presented here.

Several transcription factors (Myt1, Hes1, Hes6, MafB, Tcf2, Pax6 and Irx1) show reduced expression in Ngn3 mutants compared to wild type and identification of some of these serve as positive control for the analysis. Previous studies have shown that Pax6 is expressed in all cells of the endocrine lineage and loss of this transcription factor leads to arrest of endocrine development prior to final differentiation (Sander et al. 1997, St Onge et al. 1997). Myt1 was recently identified by Gu et al. (2004) in a microarray experiment involving distinct populations of cells from different stages of pancreatic development. Myt1 was shown to be expressed in endocrine progenitor cells and the ability of Ngn3 to promote α- and β-cell development likely depends on Myt1 (Gu et al. 2004). The transcription factor, Irx1, belongs to a family of homeodomain proteins known as Iroquois and includes six members. The Irx family of transcription factors have been shown to be required early during development to define large territories, but also play a role later on for subdividing these domains (Cavodeassi et al. 2001). We confirmed and extended previous observations (Houweling et al. 2001, Lebel et al. 2003) showing that Irx1 and 2 but not 3, 4 and 5 (data not shown) are expressed in the pancreas. We demonstrate here that Irx1 and Irx2 are specifically expressed in α-cells. Furthermore, our data suggest that Irx1 and Irx2 could be operating at early stages of endocrine subtype specification soon after commitment to the endocrine lineage as Irx1/2 was induced by Ngn3 in many cells that did not yet express glucagon. The identical expression of the two Irx genes and their highly related homeodomains suggest that they might serve redundant functions during α-cell development. It was recently demonstrated that MAP kinase-mediated phosphorylation can change Irx2 from a repressor to an activator in developing midbrain tissue (Matsumoto et al. 2004), but the transcriptional activity and function of Irx1/2 proteins in α-cells is unknown. Further, gain- and loss-of-function experiments will be required to dissect the precise role of this family of transcription factors in the pancreas.

The MA plots in Fig. 1 show that genes exhibiting a significant difference in expression level of more than twofold between wild type and mutant constitute only a minor fraction of the significant results. In an attempt to provide an unbiased analysis of the microarray data, we carried out additional analysis based on all genes found to show statistically significant difference between the wild type and the mutant. The genes identified at e13 and e15 were divided into groups of genes showing increased expression and those showing decreased expression. For each of these groups of genes, we used the Gostats package to carry out a hypergeometric test to determine whether there was over-representation of genes at particular GO terms in the Biological Process branch of the ontology (Table 2). The genes showing decreased expression levels in the
mutant might primarily be expected to be expressed in endocrine cells. While ‘peptide-hormone processing’ and ‘neuropeptide signalling’ are common to both e13 and e15, there are differences in the terms that are over-represented at the two stages of development. It is tempting to speculate that the difference in BP terms reflects that Ngn3 lineage cells at e13 and e15 are at different developmental stages. At e13, a large group of apoptosis-related genes are identified (Rnf130, Gadd45g, Cdkn1a, Btg2, Txin1, Sphk1, 1700020C11Rik, Smndc1, Pim2, Unc5a and E2f2). While cell death has been demonstrated to be a common event in the development of several organs (Barres et al. 1992, Coles et al. 1993, Coucouvanis & Martin 1995), nothing is known about the role of apoptosis in formation of the pancreas. Several negative regulators of apoptosis are identified (Pim2, Sphk1, Btg2, and Cdkn1a), which might indicate that suppression of apoptotic signals is necessary for proper development of endocrine cells. At e15, many of the genes showing decreased expression encode proteins involved in transport or ‘G-protein-coupled receptor protein-signalling pathway’. It is well known that insulin secretion is triggered by Ca\(^{2+}\) influx through voltage-dependent Ca\(^{2+}\) channels (reviewed in (Mears 2004)) and recently G-protein-coupled receptor signalling has been shown to influence insulin secretion (Briscoe et al. 2003, Itoh et al. 2003). Thus both, ‘transport’ and offspring terms as well as ‘G-protein-coupled receptor protein-signalling pathway’ are reconcilable with a predominantly mature phenotype of the Ngn3-lineage cell at e15.

Interestingly, 116 probes on the array detect slightly increased expression levels in Ngn3-deficient mice at e13. It has been shown that Ngn3 is both required and sufficient for inducing endocrine differentiation in pancreatic cells (Gradwohl et al. 2000, Grapin-Botton et al. 2001); thus many genes showing increased expression in the mutant could suggest that the absence of endocrine and progenitor cells has an effect on parts of the remaining pancreatic cells. The fact that only 21 genes with increased expression are found at e15 could be caused by a dilution effect where cells that are potentially affected by the lack of endocrine progenitor or mature endocrine cells is outnumbered by non-endocrine cells which are more abundant at e15 than at e13. Among the genes showing increased expression in the mutant at e13, we identified GO terms ‘neurogenesis’, ‘cell adhesion’, and ‘cell communication’ as significantly over-represented. Whether these GO terms reflect some sort of compensatory mechanism or a response to lacking endocrine and precursor cells, requires further study.

The entire list of genes identified at e13 was used to compute the induced GO graph in the BP branch. Colour coding of this graph shows that there are biological processes that appear to be represented only in the list of increasing or decreasing genes. The group of genes showing decreased expression in the mutant is larger than those showing increased expression and consequently, there is a clear overweight of different shades of green. ‘RNA metabolism’ and offspring terms like ‘nuclear mRNA splicing, via spliceosome’ are only represented by genes showing decreased expression. In a recent paper, Gunton et al. (2005) presented microarray data showing that a group of RNA processing genes are significantly altered in islets from type-2 diabetic patients. Whether RNA processing plays a crucial role in endocrine cell development or function requires further analysis. The interactive version of the GO graph presented in the Supplementary Data (Figure S1; See Supplementary data at http://jme.endocrinology-journals.org/content/vol37/issue2) can be used to browse the data and give more detailed GO annotation of the genes identified at e13.

A few issues should be kept in mind when using GO annotations to summarise microarray results, as done in this paper. First, only 50% of the genes in the LION murine arrayTAG set have GO annotations and some of these are solely inferred from electronic annotation, meaning that no human judgement was used in the process. Secondly, the GO database is subject to constant change as our understanding of biology and biological processes changes meaning that redoing the analysis in the future will probably give slightly different results. Thirdly, the fact that development of the pancreas involves differentiation of several different cell types differentiating asynchronously, makes analysis of large gene lists difficult, and identification of multiple genes with similar GO annotations does not imply that the encoded proteins interact physically. Nevertheless, the use of GO terms provides an unbiased way of analysing large data sets like microarray data and importantly, brings together the combined knowledge from several different model organisms.

In conclusion, using advanced microarray technology and data mining, we have revealed Irx-1/2, as new pancreatic glucagon-producing cell-specific genes, whose function in the pancreas remains unknown. Additionally, our microarray data strongly suggest that the absence of Ngn3 positive and mature endocrine cells results in changes in gene expression in embryonic exocrine/ductal cells. GO-based analysis of genes showing reduced expression in the mutant and genes showing increased expression revealed multiple biological processes that are affected solely by increased or decreased gene expression, suggesting that these processes are involved in the development or the function of pancreatic cells or are important in the cross-talk between the formation of endocrine tissue and the other pancreatic cell types during development.
References


Diabetes 39 415–425.


Gu G, Dubaukaita J & Melton DA 2002 Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development 129 2447–2457.


Acknowledgements

The authors would like to thank Dr Nanin Din as well as Kirsten Brunstedt, Tine Pedersen and Anette Seneca for amplifying, purifying and printing the LION clone collection. We would like to thank Dr Gerard Gradwoh for valuable discussions. Furthermore, we would like to thank the R project and Bioconductor Consortium for contributing a splendid software solution for microarray data analysis. This work was made possible through support from JDRF and NIH (U19 DK072495-01, PS) and the Danish Research Council (AP). The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.


Mears D 2004 Regulation of insulin secretion in islets of Langerhans by Ca(2+) channels. *Journal of Membrane Biology* 200 57–66.


Received 23 May 2006

Accepted 9 June 2006