Profiling of Embryonic Stem Cell Differentiation

Nobuaki Shiraki, Soichiro Ogaki, and Shoen Kume

Department of Stem Cell Biology, Institute of Molecular Embryology and Genetics, Kumamoto University, Honjo 2-2-1, Kumamoto 860-0811, Japan. Address correspondence to: Shoen Kume, e-mail: skume@kumamoto-u.ac.jp

Manuscript submitted March 22, 2013; resubmitted May 6, 2013; accepted July 3, 2013

Abstract

Embryonic stem (ES) cells have been shown to recapitulate normal developmental stages. They are therefore a highly useful tool in the study of developmental biology. Profiling of ES cell-derived cells has yielded important information about the characteristics of differentiated cells, and allowed the identification of novel marker genes and pathways of differentiation. In this review, we focus on recent results from profiling studies of mouse embryos, human islets, and human ES cell-derived differentiated cells from several research groups. Global gene expression data from mouse embryos have been used to identify novel genes or pathways involved in the developmental process, and to search for transcription factors that regulate direct reprogramming. We introduce gene expression databases of human pancreas cells (Beta Cell Gene Atlas, EuroDia database), and summarize profiling studies of islet- or human ES cell-derived pancreatic cells, with a focus on gene expression, microRNAs, epigenetics, and protein expression. Then, we describe our gene expression profile analyses and our search for novel endoderm, or pancreatic, progenitor marker genes. We differentiated mouse ES cells into mesendoderm, definitive endoderm (DE), mesoderm, ectoderm, and Pdx1-expressing pancreatic lineages, and performed DNA microarray analyses. Genes specifically expressed in DE, and/or in Pdx1-expressing cells, were extracted and their expression patterns in normal embryonic development were studied by in situ hybridization. Out of 54 genes examined, 27 were expressed in the DE of E8.5 mouse embryos, and 15 genes were expressed in distinct domains in the pancreatic buds of E14.5 mouse embryos. Akr1c19, Adip2, Pdx1p1, and Creb3l1 were all novel, and none has been described as being expressed, either in the DE, or in the pancreas. By introducing the profiling results of ES cell-derived cells, the benefits of using ES cells to study early embryonic development will be discussed.

Keywords: diabetes · embryonic stem cell · differentiation · beta-cell · Pdx1 · Ngn3 · Sox · gene profiling · microRNA

1. Introduction

Endoderm gives rise to respiratory and digestive organs, such as pancreas, liver, lung, stomach, and intestine. Multipotent endoderm has the potential to be used in tissue repair. However, despite the importance of definitive endoderm (DE)-derived tissues, not much is known about how they emerge from the primary gut tube. Fate mapping studies suggest that the DE fate begins to segregate at embryonic day 6-6.5 (E6-E6.5), and that the progenitors fated to become specific tissues of the gut tube appear shortly after the completion of gastrulation [1, 2]. The expression of the region-specific transcription factors has provided clues as to how the endoderm is patterned into different organ domains. Pancreatic and duodenal homeobox gene 1 (Pdx1) expression is the first clear sign of pancreatic differentiation, and is detected at E8.5 in the dorsal endoderm of the gut. Pdx1 is expressed before the buds become evident, and is required for the progression of pancreatic and rostral duodenal development [3]. Genetic lineage tracing studies have shown that Pdx1-expressing cells give rise to all three cell lineages in the pancreas: endocrine, exocrine, and duct cells [4].

Recent advances in the analysis and identification of early endodermal or pancreatic genes has been remarkable [5-9]. Several reports have dem-
onstrated the identification of novel endodermal genes using early embryos. Progress in embryonic stem (ES) cell studies has demonstrated that ES cells provide a good system for studying developmental biology. In particular, the human ES cell differentiation system is a useful tool to study the molecular mechanisms of human embryonic pancreas development, most notably from germ layer specification to pancreatic endocrine cell differentiation [10, 11].

Here, we first summarize gene profiling studies using mouse embryo and mouse/human ES cell-derived cells. Then, we describe our in vitro differentiation method and gene expression profile analyses of mouse ES cell-derived DE and Pdx1-expressing cells.

2. Profiling studies of mouse early embryos

Analyses of individual genes have defined critical stages in the development of the endocrine pancreas. Global gene expression analyses provide fundamental information on the processes that regulate the normal development of the endocrine pancreas.

2.1 Finding genes that regulate pancreatic development

Gene profiling analyses of E7.5 endoderm and mesoderm, E10.5 GFP+ and GFP- cells from Pdx1/GFP-transgenic mice, E13.5 GFP+ and GFP-cells from Neurogenin3 (Ngn3)/GFP-transgenic mice, and adult islets, have been described [12]. In this study, Myelin transcription factor 1 (Myt1) was identified as a candidate gene expressed in E13.5 Ngn3/GFP-positive cells, and an investigation of its loss-of-function revealed that Myt1 is a downstream effector of Ngn3 [12]. The function of Myt1 was further confirmed by a mutant mouse study [13]. Sherwood and coworkers carried out gene expression analysis of the E8.5 DE and visceral endoderm using Sox17/GFP-transgenic mice and cell surface markers, such as epithelial cell adhesion molecule (EpCAM) and dolichos biflorus agglutinin (DBA) [7]. By developing an early endoderm gene expression signature, they characterized the transcriptional similarities and differences between DE and visceral endoderm [7]. Also, they performed profiling analyses on several early endodermal organ domains, such as the mouse E11.5 esophageal, lung, distal tracheal, stomach, hepatic, and pancreatic regions, using cell surface markers, such as EpCAM, Liv2, and Rae [8]. An

### Abbreviations:

- Aebp2 - AE-binding protein 2
- AFP - alpha-fetoprotein
- Akrlc19 - aldol-keto reductase family 1 member C19
- bFGF - basic fibroblast growth factor
- BIO - 6-bromoindirubin-3'-oxime
- BMP - bone morphogenetic protein
- CALB1 - calbindin 1
- cAMP - cyclic adenosine monophosphate
- CHIP-Seq - chromatin immunoprecipitation DNA-sequencing
- Creb3-1 - AMP responsive element binding protein 3-like 1
- CTCF - CCCTC-binding factor
- Cxcr4 - chemokine (C-X-C motif) receptor 4
- DBA - dolichos biflorus agglutinin
- DE - definitive endoderm
- Dex - dexamethasone
- E - embryonic day
- ECT - ectoderm
- EpCAM - epithelial cell adhesion molecule
- ES - embryonic stem
- FAIRE - formaldehyde-assisted isolation of regulatory elements
- GCNT2 - glucosaminyl (N-acetyl) transferase 2
- GFP - green fluorescence protein
- GLP1 - glucagon-like peptide 1
- GPR50 - G protein-coupled receptor 50
- GSK - glycogen synthase kinase
- H3K4me3 - histone 3 lysine 4 trimethylation
- HGF - hepatocyte growth factor
- Hnf1beta - hepatocyte nuclear factor 1beta
- IL-6 - interleukin 6
- IPS - induced pluripotent stem
- KDM5B - lysine (K)-specific demethylase 5B
- KDM3B - lysine (K)-specific demethylase 5B
- LPM - lateral plate mesoderm
- Mafa - v-maf musculoaponeurotic fibrosarcoma oncogene homolog a
- MES - mesendoderm
- MPSS - massively parallel signature sequencing
- mRNA - messenger ribonucleic acid
- Myt1 - myelin transcription factor 1
- NEUROG3 - neurogenin 3
- Ngf - neurogenin 3
- PAM - paraxial mesoderm
- PAK6 - p21-activated kinase 6
- Pbxip1 - pre B cell leukemia transcription factor interacting protein 1
- PcG - polycomb
- Pdx1 - pancreatic and duodenal homeobox 1
- Plekhh1 - pleckstrin homology domain-containing h
- PRDM1 - positive regulatory domain I-binding factor 1
- Ptf1a - pancreas transcription factor 1 subunit alpha
- Rae - RNA export factor
- Rbm47 - RNA-binding motif protein 47
- RBPMS2 - RNA-binding protein with multiple splicing 2
- SAGE - serial analysis of gene expression
- SCID - severe combined immunodeficiency
- SCG3 - suppressor of cytokine signaling 3
- Sox17 - sex-determining region Y (Sry) box 17
- SSEA1 - stage-specific embryonic antigen 1
- STAT3 - signal transducer and activator of transcription 3
- TROP2 - tumor-associated calcium signal transducer 2
- UCN3 - urocortrin 3
- Wnt - wingless-type MMTV integration site family

www.The-RDS.org

endoderm transcription factor map at E9.5 was constructed, and anterior-posterior patterning dynamics were revealed [8]. Hoffman et al. performed serial analysis of gene expression (SAGE) of E10.5-E18.5 pancreas, adult duct, and islets. After extracting candidate genes by SAGE, the GenePaint database was used to validate their results [9]. Selective isolation of cells is necessary for profiling analysis of specific cell types. Fagman et al. employed laser capture microdissection and microarray analysis, to define genes expressed in the mouse E10.5 thyroid and lung. They found a regulatory pathway involving the anti-apoptotic gene Bcl2 that controls cell survival in early thyroid development [14]. These studies indicated that global gene expression analyses of the mouse embryo are useful at the molecular level to characterize the similarities and differences between the various developmental domains, stages, or lineages, and to identify novel genes or pathways involved in developmental processes.

2.2 Finding genes related to reprogramming

Gene expression profiling is also useful to identify candidate genes that regulate reprogramming. Zhou et al. performed a genome-wide transcription-factor expression analysis of mouse E14.5 pancreas. The expression pattern of 1,100 mouse transcription factors was confirmed by whole mouse in situ hybridization. There are at least 20 transcription factors expressed in mature β-cells and their precursors or endocrine progenitors. Mutagenesis of 9 of these genes resulted in β-cell developmental phenotypes [5]. Reprogramming from exocrine cells to pancreatic β-cells was attempted by overexpressing these 9 genes, including 3 transcription factors (Ngn3, Pdx1, and Mafa) found to reprogram pancreatic exocrine cells into cells closely resembling pancreatic β-cells [15]. This report demonstrated the possibility of direct cell reprogramming into other lineages [16].

3. Profiling studies of human islets

Profiling studies of mature human pancreatic β-cells and islets were performed, with a focus on gene expression [17, 18], microRNAs [19-21], epigenetics [22-24], and protein expression [25].

3.1 Gene expression profiling of human islets

The Beta Cell Gene Atlas (http://www.t1dbase.org/page/AtlasHome) is a useful resource that contains detailed information on the gene expression profiles of pancreatic β-cells, islets, and insulin-producing cell lines. A ‘massively parallel signature sequencing (MPSS) analysis’ of human pancreatic islet samples and microarray analyses were performed with purified rat pancreatic β-cells, α-cells, and INS-1 cells. The results were compared with array data available in literature [17]. Another database, EuroDia database (http://eurodia.vital-it.ch), was established to build a unique collection of gene expression analyses performed on β-cells of the human, mouse, and rat. The EuroDia database is now available to the entire diabetes research community to ensure continuous access to this valuable data collection after the formal end of the project [18].

3.2 microRNA profiling of human islets

Joglekar and coworkers performed microRNA profiling of human pancreatic islet cells [20]. Islets were isolated from 55 human fetal pancreata during 8-37 weeks of gestation, from neonatal pancreas and human fetal liver. Skin and muscle tissues were used for comparison. It showed that miR-375 might be involved in human pancreatic islet development. Cell sorting based on immunostaining with antibodies against intracellular molecules is possible because microRNAs remain stable after fixation. Based on this technique, microRNA profiling of glucagon+ α-cells, and C-peptide+ β-cells from a human pancreas, were performed [21].

3.3 Epigenetics of human islets

Analyses, using human pancreatic islets, were performed to characterize epigenetic regulation. DNase I hypersensitive sites, histone H3 lysine methylation modifications (H3K4me1, H3K4me3, and H3K79me2), and CCCTC factor (CTCF) binding in human islet cells were examined [24]. Another group mapped the genome-wide location of 4 histone marks (H3K4me1, H3K4me2, H3K4me3, and H3K27me3) [23]. Gaulton and colleagues profiled chromatin, using formaldehyde-assisted isolation of regulatory elements coupled with high-throughput sequencing (FAIRE-seq), to identify regulatory DNA sites active in human pancreatic islets. A comparison of FAIRE-seq data from human islets with data from 5 non-islet cell lines revealed ~3,300 physically linked clusters of islet-selective open chromatin sites, which typically encompassed genes that showed islet-specific expression [22].
These data provided insight into pancreatic β-cell function and the molecular mechanisms causing diabetes. The results are also useful for the validation of ES cell-derived pancreatic β-cells.

4. Human ES cells

Profiling studies of human ES cell-derived cells have been performed to examine the characteristics of the differentiated cells, identify novel marker genes, and understand embryonic development.

4.1 Gene profiling of human ES cell-derived cells

By microarray analysis, Cxcr4 was identified as a gene that encodes CXCR4, which can be used as a cell surface marker specifically expressed in DE, but not in visceral endoderm [11]. Wang et al. established a Sox17/GFP knock-in human ES cell line, and carried out gene expression analysis of Sox17/GFP+ cells that were differentiated based on the procedure established by the D’Amour group. The results of their gene expression analysis, in vitro differentiation, and transplantation-based assays showed that CD49e+CD141+CD238+ cells are primitive gut tube endoderm cells [26]. Human ES cell lines were established, with a Sox17/GFP or Pdx1/GFP transgene introduced via BAC vectors. These cells are useful resources for the identification of novel cell surface markers. G protein-coupled receptor 50 (GPR50) and tumor-associated calcium signal transducer 2 (TROP2) were identified as cell surface proteins that were highly enriched in pancreatic progenitor cells [27]. The identification of cell surface marker genes enabled the isolation of DE [11, 28], primitive gut tube endoderm [26], and pancreatic progenitor cells [27], without genetic manipulation of ES cells. This method represents a powerful tool for future characterization of similar cell populations.

Recently, a pancreatic differentiation protocol was developed by sequentially exposing human ES cells to different growth factors and small molecules. However, the resultant differentiated cells are immature, and are mostly polyhormonal cells [10, 29]. Using a modified procedure, Basford et al. established Insulin/GFP knock-in human ES cells for prospective isolation and the study of gene expression profiles by microarray analysis to characterize human ES cell-derived pancreatic cells, both functionally and molecularly [30]. Genomic analyses revealed that Insulin/GFP+ cells collectively resemble immature endocrine cells [31]. These findings suggest that additional effort is required to derive fully mature β-cells from human ES cells.

Gene expression profiling of ES cell-derived cells was also performed for other lineages, including neural cells [32], intestinal cells [33], adipocytes [34], or myoblasts [35]. For hepatocyte differentiation, expression profiling was performed to estimate the maturation state of ES cell-derived hepatic cells in comparison with adult hepatocytes [36, 37]. Similarly, ES cells bearing fluorescent reporter genes were used for microarray analyses of hepatic differentiation. Chiao et al. used a lentiviral vector containing the alpha fetoprotein promoter to drive enhanced green fluorescent protein expression (AFP:eGFP) [38], and our group established albumin/mKO1 knock-in human ES/iPS cells [39].

4.2 MicroRNA profiling of human ES cell-derived cells

MicroRNAs are endogenous small non-coding RNAs that play important roles in embryogenesis, cell fate, growth control, and apoptosis, and are also targets of profiling studies. Human ES cell-derived pancreatic islet-like clusters showed very high expression of the microRNAs miR-186, miR-199a, and miR-339, which downregulate the expression of LIN28, PRDM1, CALB1, GCNT2, RBM47, PLEKHH1, RBPMS2, and PAK6 [40]. Wei et al. also reported that miR-34a was expressed during pancreatic progenitor differentiation from endoderm cells, and that miR-146a, miR-7, and miR-375 were specifically expressed during differentiation from pancreatic progenitors to insulin-positive cells. Overexpression of miR-375 downregulated Hnf1β and Sox9 expression [41]. Although these microRNA approaches are promising, further research is required to utilize microRNA for in vitro maturation of ES cell-derived pancreatic cells.

4.3 Epigenetics of human ES cell-derived cells

As described above, embryonic development and ES cell differentiation are characterized by dynamic changes in genome-wide gene expression. Yet, the roles of epigenetic modifications remain elusive in these events. Recently, two groups reported the profiling of histone modifications remain elusive in these events. Recently, two groups reported the profiling of histone modifications. Gutierrez performed 3 types of genome-wide profiling (mRNA expression, microRNA expression, and histone 3 lysine 4 trimethylation (H3K4me3)), to identify novel pancreatic endocrine maturation pathways.
H3K4me3 is found at all active transcriptional start sites. Undifferentiated ES (day 0), mesendoderm (day 1), DE (day 2), primitive foregut (day 5), pancreatic progenitor (day 8), and pancreatic endocrine (day 11) cells were used for this profiling study. Data analysis suggested the involvement of novel gene networks, such as NEUROG3/E2F1/KDM5B and SOCS3/STAT3/IL-6, in endocrine cell differentiation. Finally, they showed that the addition of IL-6 increased Nkx2.2 and NEUROG3 expression [42].

Other groups performed RNA-seq and CHIP-seq profiling to identify the gene targets for H3K27me3 and H3K4me3 in ES cell-derived cells. H3K27me3 is enriched in genes that are repressed by polycomb (PcG) proteins. Cells differentiated in vitro (gut tube, posterior foregut, pancreatic endoderm, and polyhormonal cells) and functional endocrine cells produced by further differentiation in vivo in mice were used for these analyses. They demonstrated that in vivo, but not in vitro, differentiated endocrine cells exhibit close similarity to human islet and endocrine cells produced in vitro, but that they do not fully eliminate the PcG-mediated repression of endocrine-specific genes, such as insulin, GLP1, and UCN3, which are thought to contribute to maturation [43]. Epigenetic profiling of ES cell-derived cells to date has suggested future strategies for manipulating epigenetic signatures to improve cell differentiation in vitro.

5. Mouse ES cells

In the next sections, we describe the in vitro differentiation method and gene expression profile analysis of mouse ES cell-derived DE and Pdx1-expressing cells performed by our group.

5.1 Pancreatic differentiation of mouse ES cells using M15 cells

The embryonic endoderm requires signals from the adjacent germ layers for subsequent regionalization into specific endoderm organs [44]. The requirement to induce signals from the mesoderm led to the idea that coculture of ES cells with a feeder cell line would induce the ES cells to differentiate into DE cells. This led to the discovery of M15, a mesonephros-derived cell line, which has been shown to be an excellent endoderm inductive source [45]. The M15 system efficiently and reproducibly supports ES cells to give rise to the DE and Pdx1-expressing cells. The use of a Pdx1/GFP-expressing ES cell line (SK7), cultured on M15 cells, has allowed a closer examination of the differentiation processes. The differentiation of ES cells to Pdx1/GFP-expressing cells is a multistep process. In the early phase, ES cells are first differentiated into mesendoderm (MES) or ectoderm (ECT) cells. In the next phase, the bipotential mesendoderm differentiates into mesoderm or DE. Finally, in the late phase, DE gives rise to region-specific tissue of the endoderm. The molecular bases of the signaling events involved in each step of the process are summarized in Figure 1.

Activin and basic fibroblast growth factor (bFGF) both promote ES cell differentiation at all phases of induction. Therefore, activin and/or

**Figure 1. A schematic drawing of M15 cell-mediated signaling events.** Signaling molecules involved in the in vitro differentiation process mediated by M15 are shown. Abbreviations: BIO - 6-bromoirindirubin-3'-oxime; BMP7 - bone morphogenetic protein 7; DAPT - N-(N-3,5-difluorophenacetyl)-L-alanyl)-S-phenylglycine t-butyl ester; Dex - dexamethasone; ES - embryonic stem; FGF - fibroblast growth factor; HGF - hepatocyte growth factor; Pdx1 - pancreatic and duodenal homeobox 1.
bFGF were added throughout the entire process of ES differentiation. The simultaneous treatment of activin and bFGF resulted in a dramatic increase of Pdx1/GFP+ cells, from 2% to 31%. When ES cell-derived Pdx1/GFP cells were grafted under the kidney capsule of mice with SCID (severe combined immunodeficiency), they differentiated into all 3 pancreatic lineages: endocrine, exocrine, and duct cells. Therefore, the ES cell-derived Pdx1/GFP+ cells we obtained had the potential to differentiate similarly into embryonic Pdx1/GFP cells.

5.2 Non-pancreatic endoderm, ectoderm, and mesoderm differentiation of mouse ES cells using M15 cells

The M15 cell line was later proved as an inducing source for pancreatic differentiation, for hepatic and intestinal differentiation [46, 47], and for induction of the ectoderm and mesoderm cell lineages [48]. Pancreatic differentiation was at the expense of hepatic differentiation. The withdrawal of activin and bFGF induced alpha-fetoprotein (AFP) expression. The addition of hepatocyte growth factor (HGF) and dexamethasone (Dex) promoted hepatic differentiation [46]. ES cells were differentiated into DE and challenged with various growth factors or chemicals that affect certain signaling pathways at a late stage (Figure 1) to establish optimal conditions for differentiation into intestinal cell lineages. Among these tested growth factors and chemicals, we found that intestinal differentiation was efficiently induced through:

1. Activation of the Wnt/β-catenin and inhibition of the Notch signaling pathways.

2. Simultaneous application of 6-bromoindirubin-3'-oxime (BIO), a glycogen synthase kinase (GSK)-3β inhibitor, and DAPT, a known γ-secretase inhibitor [47].

SB 203580, a p38 MAPK inhibitor, increased the neuroectodermal population (Figure 1). These SB 203580-treated cells were multipotent neuronal progenitors able to give rise to astrocytes, oligodendrocytes, neurons, and dopaminergic neurons [48]. Bone morphogenetic protein (BMP) antagonized activin and resulted in the potentiation of mesodermal differentiation (Figure 1). Further differentiation into lineage-specific cells was achieved by subjecting ES cell-derived mesodermal cells to adipogenic or osteogenic differentiation conditions. Differentiation into Alizarin red S-
positive osteogenic cells or oil red O-positive adipogenic cells was observed at day 20 [48].

5.3 Microarray analysis of mouse ES cell-derived cells

We isolated mouse ES cell-derived differentiated cells for microarray analysis using the procedure described above, by tracking the expression of specific cell surface antigens using flow cytometry. The cell types and cell surface markers (or GFP) used for prospective cell isolation were: SSEA1+Flk1+PDGFRα- (ES cell-derived ectoderm (ECT)), E-cadherin+PDGFRα+ (mesendoderm (MES)), E-cadherin-PDGFRα+Flk1+ (paraxial mesoderm (PAM)), E-cadherin-PDGFRα-Flk1+ (lateral plate mesoderm, LPM), and E-cadherin+CXCR4+ (DE) populations (Figure 2). DE cells at D5, D7, or D8 were collected. DE at D8 was further subdivided into Pdx1/GFP-negative and -positive populations (D8 DE Pdx1- and Pdx1+). A remarkable transition in the gene expression profile was observed from D5 to D7 DE, and thereafter (Figure 2B). Comparison between ES cell-derived cells and embryonic tissue, such as E7.5 embryonic endoderm [12], E8.25 endoderm [7], and E10.5 Pdx1+ cells [12], suggested that gene expression profiles in D5, D7, and D8 DE were similar to E7.5, E8.25, or E10.5 embryonic Pdx1+ cells, respectively. These analyses show that ES cell-derived DE cells or Pdx1+ cells mimic cells in normal developmental processes.

5.4 Identification of DE-specific genes in ES cell-differentiation

Gene expression profiles of undifferentiated ES cells and ES cell-derived differentiated cells of the 3 germ layers (ECT, LPM, PAM, MES, D5 DE, D7 DE, Pdx1+D8 DE, and Pdx1+D8 DE) were compared. Figure 3 and Tables 1 and 2 show the summary of the numbers of genes analyzed and indicate the genes expressed in the gut endoderm at E8.5 and/or the pancreatic bud at E14.5. These analyses show that ES cell-derived differentiated cells served as a good model cellular system for studying the gene expression of normal developmental stages.

Decay accelerating factor (DAF1/CD55), a gene found to be highly expressed in ES cell-derived DE, was detected in the DE and mesoderm in early embryos at E8.5 [49]. Flow cytometry analysis of ES cell-derived differentiated cells revealed that DAF1+ cells also expressed CXCR4 on the cell surface. Moreover, DAF1 expression is maintained until differentiation day 12 in ES cell-derived DE cells. Analysis of the Pdx1+GFP+ cells in E9.5 embryos and ES cell-derived cells with anti-DAF1 revealed that most Pdx1+GFP+ cells expressed DAF1. These results suggest that DAF1, when
used in combination with E-cadherin, is useful for the prospective identification of DE cells.

Among the genes whose expression is increased in the ES cell-derived DE population, Foxq1 [50], Cpm [51-53], Foxp4 [54, 55], Pcdh1 [56], and Zmiz1 [57] were found to be expressed in the foregut, hindgut, or whole gut at E8.5. Parm1 [6], Tmem184 [58], Hipk2 [59], Nptx2 [60, 61], Tcf7l2 [62-65], C2Cd4b [66], Sox4 [67-69], and Kiss1r [70-72] were revealed for the first time to be expressed at this early stage of E8.5 and E14.5. Hipk2 was co-expressed with glucagon, but not insulin, implicating that it might be associated with β-cell differentiation (Figure 4). C2Cd4b, a gene expressed in the trunk, was co-expressed with insulin, but not glucagon, implicating its function in endocrine β-cell differentiation (Figure 4). It is of interest that genes responsible for β-cell maturation are expressed at early stages of development. Future studies examining the functions of these genes should reveal their role in β-cell replication or differentiation of the pancreas.

We found for the first time that the following 4 genes are expressed in the E8.5 endoderm or E14.5 pancreas:

1. Aldo-keto reductase family 1 member C19 (Akr1c19)
2. AE binding protein 2 (Aebp2)
3. Pre B cell leukemia transcription factor interacting protein 1 (Pbxip1)
4. cAMP responsive element binding protein 3-like 1 (Creb3l1)

Akr1c19 was reported to be highly expressed in the liver and gastrointestinal tract [73]. Aebp2 encodes a zinc finger protein that interacts with the mammalian polycomb repression complex 2 (PRC2) [74]. Its Drosophila homolog, jing, is a zinc-finger transcription factor that interacts with the fly polycomb group (PcG) protein complexes, and plays an essential role in controlling CNS midline and tracheal cell differentiation [75]. Pbxip1 is a PBX interacting protein, also known as HPIP,
which inhibits the binding of Pbx1-Hox complexes to DNA [76]. Creb3l1, also known as OASIS, is a ZIP (basic leucine zipper) transcription factor, which is a member of the CREB/ATF family and has been identified as an ER stress transducer [77].

There are genes whose expression we could not detect in Pdx1+ cells during normal pancreatic development. This might be due to their low expression levels and/or technical limitations of our experimental setup. In addition, some of the genes show expression patterns that are difficult to be catalogued at E14.5, since pancreatic differentiation undergoes a secondary transition at this stage, and many genes show a dramatic change in their expression patterns after this transition.

6. Conclusions

We reviewed gene expression profiling studies using mouse embryo, islets, and ES cell-derived cells, and described our in vitro differentiation method that used feeder cells and growth factors. Then, we described our gene expression profile
analyses. These analyses revealed that ES cell-derived cells mimic cells that arise during normal development. Profiling of ES cell-derived cells yielded important information about the characteristics of differentiated cells, identified novel marker genes, and revealed novel pathways of differentiation. Currently, several groups have reported the generation of pancreatic β-cell like cells. Although these cells were immature human pancreatic progenitor cells, they matured into functional β-cell after transplantation [10, 29].

Multi-level genome-wide profiling assessing gene expression, microRNAs expression, proteome composition, metabolome makeup, DNA methylation patterns, and histone modifications might provide us with useful information to induce in vitro maturation of ES cell-derived pancreatic cells. During the last decade, high-throughput techniques have been developed, including microarray and next-generation sequencing, together with public databases, such as Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/), the RIKEN FANTOM project (http://fantom.gsc.riken.jp/), GenePaint [78, 79], and the Mouse Atlas Website (http://www.mouseatlas.org/). In the near future, profiling studies using the aforementioned new technologies will lead to the identification of novel signaling molecules which may promote pancreatic development, and which may offer novel targets for the treatment of diabetes.

Disclosures: The authors do not have any conflict of interests. There are no financial, institutional, corporate, or patent-licensing arrangements associated with this paper.

References

12. Gu G, Wells JM, Dombkowski D, Preffer F, Aronow B, Melton DA. Global expression analysis of gene regu-


