Exploring the Use of Human Pluripotent Stem Cells to Create Functional Pancreatic \(\beta\) Cells

Professor Douglas Melton Siniša Hrvatin

Exploring the use of human pluripotent stem cells to create functional pancreatic β cells

Abstract

Directed differentiation of human pluripotent stem cells (hPSCs) has the potential to produce human cell types that can be used for disease modeling and cell transplantation. Two key challenges in the differentiation from hPSCs to β cells are the specification from pancreatic progenitors to insulin-expressing (INS^+) cells and the maturation of INS^+ cells into glucose responsive β cells.

To address the first, two high-content chemical screens identified PKC inhibitors as inducers of INS^+ cells from pancreatic progenitors. PKC inhibition generated up to ten-fold more INS^+ cells while PKC agonists blocked differentiation into INS^+ cells. Transplantation of PKCβ inhibitor-treated pancreatic progenitors, containing higher proportions of endocrine progenitors and endocrine cells, resulted in mature β cells showing higher levels of glucose-stimulated human c-peptide production *in vivo*. This indicates that *in vitro* derived INS^+ cells might be competent to mature into functional β cells.

To address the second challenge, we first studied mouse and human β cell maturation *in vivo*. Postnatal mouse β cell maturation was marked by an increase in the glucose threshold for insulin secretion and by expression of the gene urocortin 3. To study human β cell maturation, a Method for Analyzing RNA following Intracellular Sorting (MARIS) was developed and used for transcriptional profiling of sorted human fetal and adult β cells. Surprisingly, transcriptional differences between human fetal and adult β cells did not resemble differences between mouse fetal and adult β cells, calling into question inter-species homology at the late stages of development.

A direct comparison between hPSC-derived INS^+ cells, and β cells produced during human development is essential to validate directed differentiation and provide a roadmap for maturation of hPSC-derived INS^+ cells. Genome-wide transcriptional analysis of sorted INS^+ cells derived from three hPSC-lines suggest that different lines produce highly similar INS^+ cells, confirming robustness of directed differentiation protocols. Furthermore, non-functional hPSC-derived INS^+ cells resemble human fetal β cells, which are distinct from adult β cells. We therefore suggest that *in vitro* directed differentiation mimics normal human development and reveal differences in gene expression that may account for the functional differences between hPSC-derived INS^+ cells and true β cells.

Table of contents

List of figures

List of tables

Acknowledgements

Throughout the course of my graduate study, I have greatly benefited from the guidance and encouragement of many individuals. Most notably, I would like to thank Professor Doug Melton for the opportunity to work with him in his lab and for his steadfast mentorship, support and inspiration. My desire to study stem cell biology arose when I took Doug's MCB125 course as a Harvard College senior. Doug had a unique and inspiring way to get across to students, to distill and explain scientific discoveries with amazing clarity and simplicity. Since joining his lab as a graduate student, he has always encouraged me to ask the most interesting and most important questions and to strive for the most elegant and informative experiment to answer them. I am forever grateful for the privilege of working with and learning from him.

I'd like to thank all my collaborators and all the members of the Melton Lab for their contributions, advice, helpful discussions, technical help and reagents. In particular I am grateful to Dr. Rene Maehr, my lab rotation mentor, and Francis Deng, a Harvard College undergraduate student who worked closely with me.

I'd also like to thank my early scientific mentors, Prof. Dr. Joern Piel in whose laboratory in Bonn, Germany I did my first independent research and Dr. Chris T. Walsh at Harvard Medical School who was my undergraduate thesis mentor. I quickly discovered working with them and in their laboratories the exhilaration and fulfilling satisfaction of

scientific discovery, which solidified my decision to pursue graduate school and a career in science.

I want to thank my family, especially my mom Gordana, for love and support, for smiles and tears throughout my studies and the many years that I have been away from home. My late father Milan, a true Renaissance man and teacher, instilled in me as a child the love for logic, mathematics and science. Without them, I would not be where I am today.

In closing, I'd like to thank my Dissertation Advisory Committee members Dr. Andy McMahon, Dr. Len Zon, Dr. Sean Wu, Dr. Chad Cowan and Dr. Kiran Musunuru for helpful and critical advice throughout my dissertation work. I am truly honored to have Dr. Len Zon, Dr. Kevin Eggan, Dr. Gordon Weir and Dr. Dale Greiner on my Dissertation Defense committee. I am thankful for their time and the opportunity to present and discuss my work with such remarkable scientists and thought leaders I can only aspire to one day become.

Statement of contribution

This dissertation presents three projects that I have worked on in the Melton Lab.

Work presented in Chapter 2 is being submitted to the journal Stem Cells with equal contribution from Shuibing Chen and me. The author list is: Siniša Hrvatin, Shuibing Chen, Kelvin Lam, Anthony Arvanites, Alireza Rezania, Lee L. Rubin, Douglas A. Melton. Dr. Chen and I designed and worked entirely independently on two separate projects that were merged into one story. Dr. Chen carried out Screen 1 and all the work using the PKCβ inhibitor including the *in vivo* transplantation study. I carried out Screen 2 and the work using the BisI inhibitor. The manuscript was primarily written by me, with contributions from Dr. Chen and Dr. Melton. Other collaborators helped us with the chemical screens, data analysis and reagents.

Work presented in Chapter 3 represents a manuscript published in Nature Biotechnology. Dr. Barak Blum was the first author and I was the second author. Dr. Blum and I collaborated on this project at the beginning of my dissertation work. Together we designed the study and carried out the initial experiments including the static GSIS tests during early postnatal development and the microarray analysis. Subsequently Dr. Blum analyzed the microarray data and identified *Ucn3* as a marker of β cell maturation. He described the pattern of expression of *Ucn3* in mouse and functionally characterized

early postnatal mouse maturation. I carried out directed differentiation of human embryonic stem cells (hESCs) and *in vivo* transplantation to generate glucose responsive grafts. Dr. Blum stained these hESC-derived cells and grafts for UCN3 expression and wrote the paper. Other collaborators contributed materials and carried out immunofluorescent staining.

Work presented in Chapter 4 has not yet been submitted for publication. I am responsible for the design and execution of this project. This project started as a fruitful collaboration with Francis Deng, an undergraduate at Harvard College, who was mentored by me. Together we developed and optimized the methods and carried out the initial analysis of hESC-derived INS^{+} cells and human adult β cells. Subsequently, this project benefited from a collaboration with Dr. Charlie O'Donnell who analyzed the raw RNA-Seq data as well as generated correlation coefficients and hierarchical clusters. Dr. Phil DiIorio provided human fetal samples. Dr. Ali Rezania and several member of the Melton Lab generously provided reagents, cells, technical help and advice. I carried out gene expression analysis, interpreted the data and wrote the manuscript.

Chapter 1

Introduction

Regenerative medicine holds the promise of replacing, repairing or regenerating human cells, tissues and organs to restore normal function. Many diseases, including diabetes and neurodegeneration, result from loss or malfunction of specific cell types in the human body. Through cell replacement therapy, regenerative medicine could provide a cure for these and other conditions. However, cell replacement therapies are limited by a shortage of cell sources, usually cadaveric donors. Stem cell biology has the potential to create a renewable and accessible source of cells for cell replacement transplantation therapy.

1.1. Pluripotent Stem Cells

Stem cells are defined by the ability to self-renew and differentiate into other cells. Self-renewal is the ability to divide and generate at least one daughter cell equivalent to the mother stem cell. Differentiation is the ability to generate other cell types. Pluripotent stem cells can differentiate any cell type in the organism. Three types of pluripotent cells have been identified: 1) embryonic germ (EG) cells, derived from primordial germ cells[1-3]; 2) embryonic carcinoma (EC) cells, derived from germ cell tumors[4-6] and 3) embryonic stem (ES) cells, derived from the inner cell mass (ICM) of the blastocyst. Of the three, embryonic stem cells are the most common pluripotent cell type used in research because they are generally chromosomally normal (unlike EC cells) and can be self-renewed and expanded *in vitro* (unlike EG cells).

1.1.1. Embryonic Stem Cells

The blastocyst represents an early pre-implantation stage of development still containing pluripotent cells. Isolation and *in vitro* culture of the inner cell mass of the mouse blastocyst created embryonic stem cells capable of extensive, perhaps unlimited, self-renewal and differentiation into any cell type in the mouse embryo [7,8]. Isolation of stable human embryonic stem cells (hESCs) from donated, otherwise discarded, human blastocysts therefore created a potentially unlimited source of any human cell type[9-12].

1.1.2. Induced pluripotent stem cells

The ability to convert a mature, adult cell into a pluripotent cell received the 2012 Medicine and Physiology Nobel Prize. It was first achieved in *Xenopus laevis* by somatic cell nuclear transfer (SCNT), a method in which a nucleus of one cell is reprogrammed into a pluripotent cell by injection into the enucleated egg [13]. In mammals, SCNT [14,15]; altered nuclear transfer[16]; the addition of a somatic cell nucleus to an oocyte[17]; and fusion of somatic cells with ES cells [18] were all successful at generating embryonic stem cells. Recently, a breakthrough technology, co-awarded with the 2012 Medicine and Physiology Nobel Prize, described the creation of induced pluripotent stem cells (iPSCs) from the skin of adult mice or humans by viral introduction of defined genetic factors[19-22]. With iPSC technology, patient's skin cells can be used to generate genetically identical pluripotent stem cells that can be differentiated into cells for autologous cell replacement therapy. Additionally, iPSCs derived from patients carrying genetic diseases can been used to study and model disease in otherwise inaccessible cell types [23].

1.1.3. Differentiation of pluripotent stem cells

Mouse embryonic and induced pluripotent stem cells, together referred to as pluripotent stem cells, can generate a live mouse when injected into tetraploid blastocysts (tetraploid complementation) [24]. For obvious ethical reasons, the differentiation potential of human pluripotent stem cells (hPSCs) cannot be tested using blastocyst

complementation. Instead, hPSCs are injected into adult immunodeficient mice to form teratomas containing derivatives of all three germ layers[25]. *In vitro*, cultured hPSCs can spontaneously differentiate into cells of all three germ layers [12].

Lineage-specific hPSC differentiation attempts to recapitulate normal human development *in vitro*. To this end, directed differentiation protocols rely on the exogenous addition of growth factors or small molecules that mimic key developmental transitions. For instance, activin A (TGFβ family ligand) and BMP4 have been found to induce formation of mesoderm from pluripotent stem cells[26].

Recently, forced expression of defined transcription factors using viral vectors, transposable genetic elements or modified RNAs successfully differentiated hPSCs towards specific lineages [27-29].

1.2. Diabetes mellitus and pancreatic β cells

1.2.1. Glucose homeostasis and pancreatic β cells

Glucose homeostasis is critical for normal body function. Following a meal, food is processed and complex sugars metabolized into glucose, which is absorbed by the intestine. Glucose is distributed to and used by most cells in the body as a critical source of energy. Importantly, when abundant, glucose is taken up from the blood by liver and muscle for storage in the form of glycogen. Between meals, when serum glucose levels are low, glycogen can be converted into glucose and secreted back into the blood. Two cell types

residing in pancreatic islets of Langerhans, pancreatic βcells and αcells specialize in monitoring blood glucose levels. Specifically, increased glucose levels trigger secretion of the hormone insulin by pancreatic β cells. Insulin signals to the liver and muscles to take up glucose from the blood. Conversely, decreased glucose levels trigger secretion of glucagon by pancreatic α cells. Glucagon signals to the liver and muscle not to take up but rather generate more glucose from stored glycogen.

1.2.2. Diabetes mellitus

Diabetes mellitus is a condition characterized by hyperglycemia, which is in turn caused by either insufficient production or resistance to insulin. Type I diabetes (T1DM) is an early onset autoimmune disease affecting 0.4% of the population[30]. It results from the autoimmune destruction of pancreatic β cells, which are responsible for glucose-dependent insulin secretion [31]. Type II diabetes is characterized by insulin resistance and insufficient insulin production from the pancreatic βcells. It affects a very large and growing population worldwide [30]. In later stages of disease progression, type II diabetics cannot control blood glucose levels and depend on insulin injections. Patients with diabetes experience glucose deregulation and are at risk for complications such as infections, retinopathy, nephropathy, peripheral neuropathy, and macrovascular disease [32].

Constant glucose monitoring and insulin replacement therapy can successfully treat most forms of diabetes, however, long-term complications are still common. Moreover, a

subset of patients, despite careful medical treatment, suffer from potentially life-threatening episodes of hypoglycemia and/or ketoacidosis [33].

1.2.3. Cell replacement therapy for diabetes

Increasing functional β cell mass is a commonly accepted strategy for curing diabetes. The Edmonton Protocol described the first effective cell replacement therapy for T1DM [34]. Cadaveric donor islets were isolated and infused into the liver of diabetic patients via the portal vein, coupled with immunosuppression. Several patients achieved normal glucose tolerance [35]. Recent data suggest that independence from exogenous insulin injection can be achieved for more than five years in half of islet recipients [36]. The remaining patients, although not free from insulin replacement therapy, are protected against severe hypoglycemic episodes and diabetes-related complications [36]. The Edmonton protocol presents great potential as a treatment and cure for diabetes.

The potential of cell based therapy is greatly limited by the scarce supply of organs from cadaveric donors [37]. For islet and β cell transplantation to become a widespread cure for T1DM, novel sources of pancreatic β cell must be considered. Additionally, as islets are currently derived from cadavers, host-graft immune rejection is a significant limitation to transplantation therapy [34]. hPSCs differentiated to pancreatic β cells could provide a renewable source of β cell for transplantation therapy. Moreover, patient-specific iPSCs could generate autologous functional islets for disease modeling or rejection-free transplantation.

1.3. Directed differentiation of hPSCs toward pancreatic β cells

Initial studies aimed at producing β cells from hESCs relied on spontaneous embryoid body (EB) differentiation and generated very small numbers of insulin-expressing cells[38]. Subsequent studies attempted to increase the efficiency of insulin-expressing cells using a single step induction process from embryonic stem cells. One method reported the generation of 10-30% insulin-positive cells using a modified neuro-ectoderm induction protocol[39]. However, careful investigation demonstrated that insulin staining resulted from insulin uptake from the media by apoptotic cells and not insulin production[40]. Failure of these methods to efficiently produce insulin-expressing cells fueled the creation of directed differentiation protocols *t*hat recapitulated normal, stepwise development. To generate human β cells from hESCs, we therefore must understand the molecular mechanisms that guide normal pancreatic development (Figure 1-1).

1.3.1. Murine Pancreatic development informs directed differentiation

The first stage of β cell development is the TGFβ mediated specification of the definitive endoderm during gastrulation [41]. Definitive endoderm is marked by co-expression of SOX17 and FOXA2 transcription factors. In mice, the pancreas is subsequently specified around E8.5 in a region of the foregut endoderm. This region is marked by expression of the pancreatic and duodenal homeobox 1 (PDX1) and pancreas

specific transcription factor 1a (PTF1a), and is referred to as the pancreatic bud [42,43]. Genetic lineage tracing demonstrated that $Pdx1^+$ cells give rise to all cell lineages of the pancreas[43,44]. The pancreatic bud subsequently thickens and evaginates into the surrounding mesenchyme (E9-11.5) [45]. The first endocrine progenitors marked by neurogenin 3 (NGN3), a bHLH transcription factor, are detectable in the pancreatic epithelium at E9. Their number reaches a maximum around E15.5 and declines towards birth [46,47]. Neurogenin 3-expressing ($NGN3^+$) cells differentiate into five different types of endocrine cells: α cells secreting glucagon (GCG), β cells secreting insulin (INS), δ cells producing somatostatin (SST), PP cells secreting pancreatic polypeptide (PP) and ε cells producing ghrelin (GHRL)[43]. Endocrine cells migrate into the mesenchyme and aggregate into islets of Langerhans. NGN3 is required for the formation of all pancreatic endocrine cells[46,48].

Several signaling pathways are critical for pancreatic development: a) TGFβ signaling (activin and nodal family) is required for early endoderm formation [49], b) retinoic acid (RA) is a mediator of anteroposterior patterning and required for specification of the dorsal pancreas [50], c) manipulation of Fgf4 levels influences the size of the pre-pancreatic domain [51], d) Notch signaling is believed to delay the differentiation into the endocrine lineage until E13.5 when the $Ngn3^+$ endocrine progenitors preferentially differentiate into β cells (secondary transition) [52] and e) mesenchymal secretion of Fgf10 promotes the proliferation of Pdx1 progenitors, induces exocrine differentiation, and maintains Notch signaling.

In contrast to early pancreatic development, little is understood about the signals that induce Ngn3^{+} progenitors and specify them into five different endocrine lineages[53-55]. Even less understood is the process by which embryonic insulin-producing cells mature into functional β cells. Understanding the mechanisms responsible for the generation and maturation of β cells is critical for the successful differentiation of hPSCs towards human β cells.

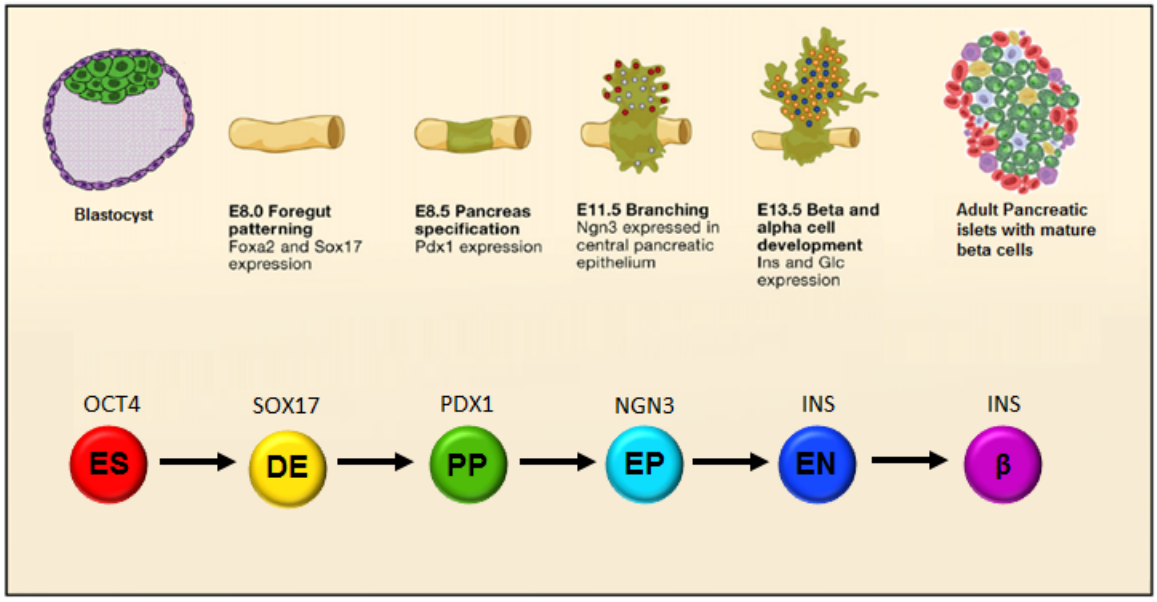


Figure 1-1. Pancreatic development and directed differentiation

TOP: Schematic representation of mouse development with key genes marking different stages. BOTTOM: Schematic representation of hPSC- directed differentiation. Each ball indicates one differentiation stage approximately matched to mouse development. Embryonic stem cells (ES), definitive endoderm (DE), pancreatic progenitors (PP), endocrine progenitors (EP), endocrine cells (EN) and β cells (β). Key marker genes are indicated above each stage.

Images adopted from[56-58]

1.3.2. hPSC-derived insulin-producing cells

The first development of a step-wise directed differentiation protocol to generate immature insulin-expressing cells from hPSCs generated tremendous hope for the generation of functional human β cells [59]. Since, many groups have used modifications to this original protocol to generate similar insulin-producing cells [59-71]. Although the hPSC-derived insulin-expressing cells (hPSC-INS$^+$) expressed insulin and several other β cell markers, they also frequently expressed multiple hormones (poly-hormonal), lacked the expression of certain mature β cell markers (NKX6-1, MAFA) and did not show glucose stimulated insulin secretion (GSIS). It has been therefore hypothesized that the hPSC-INS$^+$ cells resemble immature human β cells [59]. It is unclear whether hPSC-INS$^+$ cells can further differentiate into mature β cells and functionally maintain glucose homeostasis. A recent study produced hPSC-INS$^+$ cells that resembled mono-hormonal, glucose responsive human β cells following long-term passaging and subsequent differentiation of definitive endoderm cells [72]. It is unclear whether the produced β-like cells have the ability to rescue diabetes following transplantation.

1.3.3. hESC-derived pancreatic progenitors rescue diabetes in mice

A landmark study by Novocell (now Viacyte) showed rescue of chemically induced diabetes in mice transplanted with hESC-derived pancreatic progenitors. Kroon et al. transplanted a mixed population of pancreatic progenitors under the kidney capsule or the fat pad of immunodeficient (SCID)-beige (Bg) mice [61]. 110 days following transplant the grafts contained human β–like cells that were mono-hormonal and expressed several

mature β cell markers. The animals were treated with streptozotocin (STZ) to ablate their endogenous β cells. Animals engrafted with hESC-derived pancreatic progenitors maintained normal serum glucose levels and contained high concentrations of human c-peptide (a byproduct of human insulin processing) in the serum. The study was the first pre-clinical proof-of-concept disease rescue suggesting that hESC-derived pancreatic progenitors can differentiate into functional β cells following *in vivo* transplantation. A different group using a modified directed differentiation protocol and a different hESC-line confirmed these findings [63]. Together, this data implies that hESC-derived pancreatic progenitors can differentiate into functional β cells and, importantly, that transplantation of hESC-derived pancreatic progenitors may be a feasible diabetes therapy.

1.4. Key challenges of directed differentiation

As we discussed, transplantation of human islets from cadaveric donors into T1DM patients can achieve insulin independence shortly after transplantation[73,74]. In one remarkable case a diabetic patient achieved normoglycemia one day following islet transplantation[75]. Hence, the goal of hPSC directed differentiation is to generate human β cells, or β-like cells, that can approach the therapeutic efficacy of human islets. Several unanswered questions and challenges remain. 1) Can the immature and non-functional hPSC-INS$^+$ cells further differentiate into mature β cells following transplantation? 2) How can we generate more hESC-INS$^+$ cells? Higher efficiencies will be necessary to achieve better efficacy following transplantation into patients. 3) How can we generate mature and

functional hPSC-INS$^+$ cells *in vitro*? 4) How similar are the *in vitro* hESC-INS$^+$ cells to human fetal or adult β cells? These challenges are described in detail below (Figure 1-2).

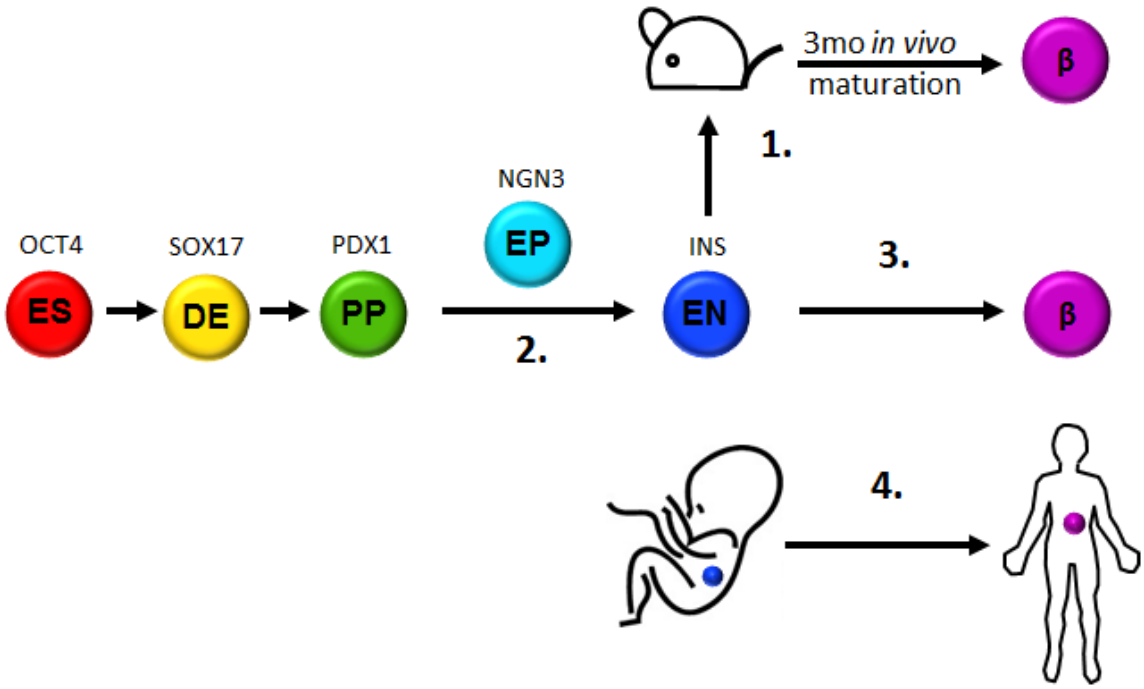


Figure 1-2. Challenges in directed differentiation to functional β cells

Schematic of hPSC directed differentiation towards pancreatic β cells. Each colored ball represents a distinct differentiation stage. Numbers 1-4 indicate specific challenges which are addressed in this dissertation. 1) Can the immature hPSC-INS$^+$ cells further differentiate into mature β cells following transplantation? 2) How to generate more hESC-INS$^+$ cells? 3) How to generate mature and functional hPSC-INS$^+$ cells *in vitro*? 4) How similar are the *in vitro* hESC-INS$^+$ cells to human fetal or adult β cells?

Fetal image adopted from[76]

1.4.1. Competence of in vitro hPSC-INS$^+$ cells to generate functional β cells

Lineage tracing experiments in mice suggest that Neurogenin3-expressing endocrine progenitors present during E8.5 and E12.5 give rise to adult β cells [43]. Additionally, lineage tracing suggests that insulin-expressing cells and not glucagon-expressing cells give rise to adult β cells [77]. Taken together, these findings indicate that early embryonic insulin-expressing cells differentiate into mature adult β cells.

Mouse and human adult islets contain endocrine cells expressing only one hormone. Early stages of mouse pancreatic development indicate the presence of a small number of poly-hormonal cells which do not appear to differentiate into single-hormonal adult endocrine cells[77]. Human fetal pancreas, unlike mouse embryonic pancreas contains a large number of poly-hormonal cells, up to 30% of all endocrine cells during gestational week 9-16 [78]. Whether human fetal poly-hormonal cells contribute directly to adult lineages and whether fetal insulin-expressing cells differentiate into adult β cells remains unknown.

Insulin-expressing cells derived from hPSCs contain both single-hormonal and poly-hormonal cells. Recently, transplantation of purified hESC-INS$^+$ cells using a genetically modified insulin:GFP hESC-reporter line gave rise to grafts containing only α cells and not β cells [64]. Two additional studies have transplanted enriched hESC-derived endocrine cells and reported the formation of α cell enriched grafts [62,79]. Together, these observations lead to a conclusion that hESC-INS$^+$ cells are human α cell precursors. Consistent with mouse pancreatic development these cells may represent the first wave of endocrine induction, whose developmental competence is restricted to α cells[80].

These experiments suggest that *in vitro* generation of endocrine cells is not beneficial for the ultimate goal of creating functional human β cells capable of rescuing diabetes. Instead, it has been proposed that only the undifferentiated pancreatic progenitors are competent to differentiate into mature β cells.

Formally, we cannot exclude the possibility that purified endocrine and INS^+ cells require the presence of surrounding cell types to differentiate into β cells following *in vivo* transplantation. This hypothesis is supported by evidence from two studies showing that the presence of hESC-derived endocrine cells at the time of transplantation significantly increases circulating human c-peptide levels 11, 12 and 16 weeks following transplant [62,63]. Graft analysis at week 17 suggests that the large majority of insulin-expressing cells is single-hormonal and expressing all the molecular markers of mature β cells[63]. It can be therefore concluded that hESC-derived endocrine cells present at transplantation contribute beneficially to the formation of mature β cells following transplantation. Lineage tracing experiments of insulin-expressing cell or endocrine cells transplanted together with pancreatic progenitors are required to determine whether these can directly contribute to mature β cells. Alternatively, the presence of endocrine cells may facilitate the differentiation of pancreatic progenitors into β cells.

Independent of the mechanism, generating hESC-derived endocrine cells with higher efficiency appears to be beneficial for *in vivo* β cell maturation.

1.4.2. Efficiency of production of hPSC-INS^+ cells

Directed differentiation protocols generate insulin$^+$ cells in a stepwise manner. The efficiency of differentiation at each step/stage is measured by the percentage of cells at that stage expressing defined molecular markers. Developmental signals during early stages of pancreatic endoderm development have been well described in model organisms. Directed differentiation protocols apply these signals in a step-wise manner to achieve high efficiencies of differentiation. For example, pancreatic progenitors, marked by expression of PDX1 can be generated with efficiencies >90% [63]. Subsequent differentiation into endocrine progenitors and endocrine cells has not been extensively studied. It is therefore not surprising that the efficiency of generating insulin-expressing cells from pancreatic progenitors is relatively low, with the best differentiations approaching 25-30% of insulin$^+$ cells [60,68,71]. Importantly, these higher efficiencies achieved by protocol optimization using a particular hPSC-line often cannot be replicated using a different hPSC-line[60]. Most hESC-lines routinely produce less than 1-10% of insulin-expressing cells.

High content chemical screening has been successful at identifying novel molecules and pathways that enhance the efficiency of directed differentiation protocols [65,66]. In the absence of information from model organisms, unbiased screening approaches are best suited to identify novel pathways implicated in pancreatic endocrine induction.

1.4.3. Maturation of hPSC-INS$^+$ cells into functional β cells

The goal of directed differentiation of hPSCs towards β cells is to create functional, mature β cells. However, functionality and therefore maturation may be defined in many

different ways. A major goal of developing new sources of human β cells is cell replacement therapy. We therefore propose the following operational definition for mature β cells: mature β cells can autonomously maintain normal serum glucose levels in an organism with a similar efficacy to adult β cells. In a standard assay, transplantation of 3,000 to 5,000 human islet equivalents (IEQ) containing approximately 0.5-1 million human β cells normalizes serum glucose levels in diabetic mice shortly following transplantation[61,63]. hESC-INS$^+$ cells, on the other hand, lack expression of several mature β cell markers, do not exhibit glucose stimulated insulin secretion and cannot maintain normal glucose levels immediately following transplantation into diabetic mice.

Instead, hPSC-derived pancreatic progenitors, endocrine progenitors and endocrine cells transplanted under the kidney capsule of diabetic mice differentiate into mature β cells and maintain normoglycemia after approximately 3-4 months [63]. Following *in vivo* maturation, INS$^+$ cells express markers of adult human β cells and secrete high levels of insulin in response to glucose. As previously discussed, it is unclear whether hESC-INS$^+$ cells present at the time of transplant differentiate into functional β cells *in vivo* or whether all the functional β cells arise from undifferentiated pancreatic progenitors. The same uncertainty applies to the lineage relationship between human fetal and adult β cells.

hPSC directed differentiation has greatly benefited from studies of pancreatic development in model organisms. In mouse, early embryonic β cells differentiate into adult β cells[43,55]. However, little is known about the molecular and functional changes during mouse β cell maturation. Certain molecular markers, such as MAFB, MAFA and PAX4

have been implicated at different stages during maturation [81,82]. Analysis of insulin-expressing cells at different stages of mouse development could better our understanding of mouse β cell maturation.

However, it still remains uncertain whether human and mouse β cell maturation share molecular and functional similarities. Using current methods it is not possible to purify human fetal or adult β cells for comprehensive expression profiling. Moreover, it remains unclear whether understanding human and mouse *in vivo* development will inform the process of β cell maturation from hESC culture. Although hESC-derived progenitors have the developmental potential to create mature β cells, the process of maturation may be distinct from normal mouse or human development. In the absence of tools to purify hESC-INS$^+$ cells, it is not possible to molecularly compare them to developing mouse or human β cells. This issue is further discussed below in the context of a question: Does directed differentiation recapitulate human development?

To achieve the goal of generating mature β cells from hPSCs, multiple parallel approaches may be required: 1) The study of mouse β cell maturation; 2) molecular and functional characterization of human fetal and adult β cells; 3) characterization of hPSC-INS$^+$ cells and 4) characterization of changes during *in vivo* maturation of hPSC-derived pancreatic progenitors transplanted into immuno-compromised mice.

1.4.4. Comparison between directed differentiation and human development

A defining feature of embryonic stem cells is their ability to differentiate into all cell types of the organism. As a proof of concept mouse ES cells and iPS cells can generate an entire organisms when injected into tetraploid blastocysts[24]. hESCs for ethical reasons cannot be subjected to the same test. Instead, hESCs generate teratomas and can be differentiated *in vitro* into cell types of all three germ layers [12]. However, *in vitro* differentiation of hESCs has met several challenges questioning the statement that *in vitro* differentiation recapitulates *in vivo* development.

The first challenge concerns variation between hPSCs. Despite the common understanding that human development proceeds consistently in all individuals, there are remarkable variations between hESC-lines regarding their propensity to differentiate into different tissues[83]. These variations may be explained by random epigenetic differences resulting in differential gene expression, differences in the rate of replication and DNA methylation patterns[60,83]. Additionally, long-term *in vitro* culture may result in the accumulation of genetic mutations, translocations and chromosomal abnormalities, which may affect the propensity of differentiation [84]. As a result of these differences, researchers frequently develop and optimize directed differentiation protocols using only one hESC-line.

hiPSCs add an additional layer of variation. They retain epigenetic memory of their previous mature cells state, undergo slow erosion of X-chromosome inactivation and often contain partially silenced genetic viral integrations [85-89].

A fundamental question, underpinning all directed differentiation efforts remains unanswered: Despite aforementioned differences, can different hESC-lines and hiPSC-lines produce the same cells types? For example, how similar are hPSC-INS$^+$ cells from different hESC-lines and hiPSC-lines. Large differences between cells of the same cell type derived using different hPSC-lines would question the reproducibility, utility and biological relevance of *in vitro* directed differentiation.

A second issue concerns the realization that hPSC directed differentiation has not produced functional adult-like β cells *in vitro.* Instead, hPSC-INS$^+$ cells appear to share characteristics, such as poly-hormone expression, with fetal β cells[59]. However, the inability to sort and transcriptionally analyze hPSC-INS$^+$ cells and human fetal and adult β cells limits the comparison between these cell types to a few markers. It is therefore unknown whether hPSC-INS$^+$ cells resemble human fetal β cells or are instead a culture artifact dissimilar to any cell during human development.

Beyond pancreatic β cells, it is critical for the field of hESC directed differentiation to evaluate whether *in vitro* hPSC differentiation has the potential to recapitulate *in vivo* development. In the absence of cell surface markers and genetic reporter lines to isolate specific cell types, it has not been possible to evaluate the degree to which these *in vitro* derived cells resemble their *in vivo* counterparts.

1.5. Topics addressed in this dissertation

First, using a high-content chemical screen our work discovers small molecules modulating novel pathways involved in pancreatic endocrine induction and proposes a dynamic role for PKC signaling in pancreatic development. We improve the differentiation efficiency of hPSC-INS$^+$ cells and improve human c-peptide secretion following *in vivo* differentiation into mature β cells. Additionally this work directly challenges the common view that hESC-INS$^+$ cells cannot differentiate *in vivo* into mature β cells.

Second, we present a study of early postnatal mouse β cell maturation that identifies UCN3 as a marker of mouse and human mature β cells. UCN3 is a potential marker for maturation of hESC-INS$^+$ cells as it is expressed in *in vivo* matured hESC-derived β cells but not in *in vitro* derived hESC-INS$^+$ cells.

Third, we develop a method to transcriptionally analyze cells following fixation, intracellular immunofluorescent staining and FACS. Using this method, we determine a high degree of similarity in global gene expression of sorted INS$^+$ cells from two different hESC-lines and one hiPSC-line. This suggests that despite differences between hPSC-lines, *in vitro* differentiation reliably generates the same cell types. Next, we obtain the first gene expression profile of sorted human fetal and adult β cells. Comparison of human and mouse maturation suggests significant differences between species and points to challenges involved in studying human development and disease through animal models. Finally, transcriptome analysis identifies that hPSC-INS$^+$ closely resemble human fetal not adult β

cells suggesting broad similarities between directed differentiation and human development.

Work is currently underway using gene expression patterns identified in this study to inform the hPSC differentiation into mature human β cells.

1.6. References

1 Matsui, Y. *et al.* Effect of Steel factor and leukaemia inhibitory factor on murine primordial germ cells in culture. *Nature* **353**, 750-752, doi:10.1038/353750a0 (1991).

2 Resnick, J., Bixler, L., Cheng, L. & Donovan, P. Long-term proliferation of mouse primordial germ cells in culture. *Nature* **359**, 550-551, doi:10.1038/359550a0 (1992).

3 Shamblott, M. *et al.* Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 13726-13731 (1998).

4 Kleinsmith, L. & Pierce, G. MULTIPOTENTIALITY OF SINGLE EMBRYONAL CARCINOMA CELLS. *Cancer research* **24**, 1544-1551 (1964).

5 Andrews, P. Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line *in vitro*. *Developmental biology* **103**, 285-293 (1984).

6 Pera, M., Cooper, S., Mills, J. & Parrington, J. Isolation and characterization of a multipotent clone of human embryonal carcinoma cells. *Differentiation; research in biological diversity* **42**, 10-23 (1989).

7 Evans, M. & Kaufman, M. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156, doi:10.1038/292154a0 (1981).

8 Martin, G. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7634-7638 (1981).

9 Amit, M. *et al.* Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Developmental biology* **227**, 271-278, doi:10.1006/dbio.2000.9912 (2000).

10 Cowan, C. *et al.* Derivation of embryonic stem-cell lines from human blastocysts. *The New England journal of medicine* **350**, 1353-1356, doi:10.1056/NEJMsr040330 (2004).

11 Lerou, P. *et al.* Human embryonic stem cell derivation from poor-quality embryos. *Nature biotechnology* **26**, 212-214, doi:10.1038/nbt1378 (2008).
12 Thomson, J. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)* **282**, 1145-1147 (1998).
13 Gurdon, J. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of embryology and experimental morphology* **10**, 622-640 (1962).
14 Byrne, J. *et al.* Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* **450**, 497-502, doi:10.1038/nature06357 (2007).
15 French, A. *et al.* Development of human cloned blastocysts following somatic cell nuclear transfer with adult fibroblasts. *Stem cells (Dayton, Ohio)* **26**, 485-493, doi:10.1634/stemcells.2007-0252 (2008).
16 Meissner, A. & Jaenisch, R. Generation of nuclear transfer-derived pluripotent ES cells from cloned Cdx2-deficient blastocysts. *Nature* **439**, 212-215, doi:10.1038/nature04257 (2006).
17 Noggle, S. *et al.* Human oocytes reprogram somatic cells to a pluripotent state. *Nature* **478**, 70-75, doi:10.1038/nature10397 (2011).
18 Cowan, C., Atienza, J., Melton, D. & Eggan, K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science (New York, N.Y.)* **309**, 1369-1373, doi:10.1126/science.1116447 (2005).
19 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.)* **318**, 1917-1920, doi:10.1126/science.1151526 (2007).
20 Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141-146, doi:10.1038/nature06534 (2008).
21 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
22 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872, doi:10.1016/j.cell.2007.11.019 (2007).
23 Kiskinis, E. & Eggan, K. Progress toward the clinical application of patient-specific pluripotent stem cells. *The Journal of clinical investigation* **120**, 51-59, doi:10.1172/jci40553 (2010).
24 Zhao, X.-y. *et al.* iPS cells produce viable mice through tetraploid complementation. *Nature* **461**, 86-90, doi:10.1038/nature08267 (2009).
25 Reubinoff, B., Pera, M., Fong, C., Trounson, A. & Bongso, A. Embryonic stem cell lines from human blastocysts: somatic differentiation *in vitro*. *Nature biotechnology* **18**, 399-404, doi:10.1038/74447 (2000).
26 Yang, L. *et al.* Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. *Nature* **453**, 524-528, doi:10.1038/nature06894 (2008).

27 Lacoste, A., Berenshteyn, F. & Brivanlou, A. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. *Cell stem cell* **5**, 332-342, doi:10.1016/j.stem.2009.07.011 (2009).
28 Pang, Z. *et al.* Induction of human neuronal cells by defined transcription factors. *Nature* **476**, 220-223, doi:10.1038/nature10202 (2011).
29 Antonica, F. *et al.* Generation of functional thyroid from embryonic stem cells. *Nature* **491**, 66-71, doi:10.1038/nature11525 (2012).
30 CDC. *CDC - Chronic Disease - Diabetes - At A Glance*, <http://www.cdc.gov/chronicdisease/resources/publications/AAG/ddt.htm > (2011).
31 Tisch, R. & McDevitt, H. Insulin-dependent diabetes mellitus. *Cell* **85**, 291-297, doi:10.1016/s0092-8674(00)81106-x (1996).
32 Khardori, R. & Griffing, G. Type 2 Diabetes Mellitus. *Medscape./eMedicine* http://emedicine. *medscape. com/article/117853-overview* (2011).
33 Bertuzzi, F., Verzaro, R., Provenzano, V. & Ricordi, C. Brittle type 1 diabetes mellitus. *Current medicinal chemistry* **14**, 1739-1744 (2007).
34 Shapiro, A. *et al.* Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *The New England journal of medicine* **343**, 230-238, doi:10.1056/nejm200007273430401 (2000).
35 Ryan, E. *et al.* Clinical outcomes and insulin secretion after islet transplantation with the Edmonton protocol. *Diabetes* **50**, 710-719, doi:10.2337/diabetes.50.4.710 (2001).
36 Gibly, R. *et al.* Advancing islet transplantation: from engraftment to the immune response. *Diabetologia* **54**, 2494-2505, doi:10.1007/s00125-011-2243-0 (2011).
37 Furth, M. & Atala, A. Stem cell sources to treat diabetes. *Journal of cellular biochemistry* **106**, 507-511, doi:10.1002/jcb.22000 (2009).
38 Assady, S. *et al.* Insulin production by human embryonic stem cells. *Diabetes* **50**, 1691-1697, doi:10.2337/diabetes.50.8.1691 (2001).
39 Lumelsky, N. *et al.* Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. *Science (New York, N.Y.)* **292**, 1389-1394, doi:10.1126/science.1058866 (2001).
40 Rajagopal, J., Anderson, W., Kume, S., Martinez, O. & Melton, D. Insulin staining of ES cell progeny from insulin uptake. *Science (New York, N.Y.)* **299**, 363, doi:10.1126/science.1077838 (2003).
41 Conlon, F. *et al.* A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development (Cambridge, England)* **120**, 1919-1928 (1994).
42 Kawaguchi, Y. *et al.* The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nature genetics* **32**, 128-134, doi:10.1038/ng959 (2002).

43 Gu, G., Dubauskaite, J. & Melton, D. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development (Cambridge, England)* **129**, 2447-2457 (2002).

44 Zhou, Q. *et al.* A multipotent progenitor domain guides pancreatic organogenesis. *Developmental cell* **13**, 103-114, doi:10.1016/j.devcel.2007.06.001 (2007).

45 Wessells, N. Differentiation of epidermis and epidermal derivatives. *The New England journal of medicine* **277**, 21-33, doi:10.1056/nejm196707062770107 (1967).

46 Gradwohl, G., Dierich, A., LeMeur, M. & Guillemot, F. neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1607-1611 (2000).

47 Schwitzgebel, V. Programming of the pancreas. *Molecular and cellular endocrinology* **185**, 99-108 (2001).

48 Jenny, M. *et al.* Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *The EMBO journal* **21**, 6338-6347 (2002).

49 Tremblay, K., Hoodless, P., Bikoff, E. & Robertson, E. Formation of the definitive endoderm in mouse is a Smad2-dependent process. *Development (Cambridge, England)* **127**, 3079-3090 (2000).

50 Chen, Y. *et al.* Retinoic acid signaling is essential for pancreas development and promotes endocrine at the expense of exocrine cell differentiation in Xenopus. *Developmental biology* **271**, 144-160, doi:10.1016/j.ydbio.2004.03.030 (2004).

51 Dessimoz, J., Opoka, R., Kordich, J., Grapin-Botton, A. & Wells, J. FGF signaling is necessary for establishing gut tube domains along the anterior-posterior axis *in vivo*. *Mechanisms of development* **123**, 42-55, doi:10.1016/j.mod.2005.10.001 (2006).

52 Apelqvist, A. *et al.* Notch signalling controls pancreatic cell differentiation. *Nature* **400**, 877-881, doi:10.1038/23716 (1999).

53 Hart, A., Papadopoulou, S. & Edlund, H. Fgf10 maintains notch activation, stimulates proliferation, and blocks differentiation of pancreatic epithelial cells. *Developmental dynamics : an official publication of the American Association of Anatomists* **228**, 185-193, doi:10.1002/dvdy.10368 (2003).

54 Oliver-Krasinski, J. & Stoffers, D. On the origin of the beta cell. *Genes & development* **22**, 1998-2021, doi:10.1101/gad.1670808 (2008).

55 Desgraz, R. & Herrera, P. Pancreatic neurogenin 3-expressing cells are unipotent islet precursors. *Development (Cambridge, England)* **136**, 3567-3574, doi:10.1242/dev.039214 (2009).

56 Murry, C. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661-680, doi:10.1016/j.cell.2008.02.008 (2008).

57 *File:Blastocyst English.svg - Wikipedia, the free encyclopedia*, <http://en.wikipedia.org/wiki/File:Blastocyst_English.svg> (2012).

58 *Wild Type Pancreatic islet of Langerhans*, <http://www.betacell.org/content/articlepanelview/article_id/13/panel_id/1> (2005).

59 D'Amour, K. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nature biotechnology* **24**, 1392-1401, doi:10.1038/nbt1259 (2006).

60 Nostro, M. *et al.* Stage-specific signaling through TGFβ family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development (Cambridge, England)* **138**, 861-871, doi:10.1242/dev.055236 (2011).

61 Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nature biotechnology* **26**, 443-452, doi:10.1038/nbt1393 (2008).

62 Kelly, O. *et al.* Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nature biotechnology* **29**, 750-756, doi:10.1038/nbt.1931 (2011).

63 Rezania, A. *et al.* Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. *Diabetes* **61**, 2016-2029, doi:10.2337/db11-1711 (2012).

64 Basford, C. *et al.* The functional and molecular characterisation of human embryonic stem cell-derived insulin-positive cells compared with adult pancreatic beta cells. *Diabetologia* **55**, 358-371, doi:10.1007/s00125-011-2335-x (2012).

65 Chen, S. *et al.* A small molecule that directs differentiation of human ESCs into the pancreatic lineage. *Nature chemical biology* **5**, 258-265, doi:10.1038/nchembio.154 (2009).

66 Borowiak, M. *et al.* Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell stem cell* **4**, 348-358, doi:10.1016/j.stem.2009.01.014 (2009).

67 Jiang, J. *et al.* Generation of insulin-producing islet-like clusters from human embryonic stem cells. *Stem cells (Dayton, Ohio)* **25**, 1940-1953, doi:10.1634/stemcells.2006-0761 (2007).

68 Jiang, W. *et al. In vitro* derivation of functional insulin-producing cells from human embryonic stem cells. *Cell research* **17**, 333-344, doi:10.1038/cr.2007.28 (2007).

69 Nostro, M. C. *et al.* Stage-specific signaling through TGF family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development* **138**, doi:10.1242/dev.065904 (2011).

70 Tateishi, K. *et al.* Generation of insulin-secreting islet-like clusters from human skin fibroblasts. *The Journal of biological chemistry* **283**, 31601-31607, doi:10.1074/jbc.M806597200 (2008).

71 Zhang, D., Jiang, W., Shi, Y. & Deng, H. Generation of pancreatic islet cells from human embryonic stem cells. *Science in China. Series C, Life sciences / Chinese Academy of Sciences* **52**, 615-621, doi:10.1007/s11427-009-0095-3 (2009).

72 Cheng, X. *et al.* Self-renewing endodermal progenitor lines generated from human pluripotent stem cells. *Cell stem cell* **10**, 371-384, doi:10.1016/j.stem.2012.02.024 (2012).

73 Shapiro, A. *et al.* International trial of the Edmonton protocol for islet transplantation. *The New England journal of medicine* **355**, 1318-1330, doi:10.1056/NEJMoa061267 (2006).

74 Gangemi, A. *et al.* Islet transplantation for brittle type 1 diabetes: the UIC protocol. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **8**, 1250-1261, doi:10.1111/j.1600-6143.2008.02234.x (2008).

75 Matsumoto, S. *et al.* Insulin independence after living-donor distal pancreatectomy and islet allotransplantation. *Lancet* **365**, 1642-1644, doi:10.1016/s0140-6736(05)66383-0 (2005).

76 *Human fetus Vector Clip art*, <http://dir.coolclips.com/Healthcare/Anatomical_Diagrams/Fetus/Human_fetus_vc049069.html> (2007).

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

78 Riedel, M. *et al.* Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia* **55**, 372-381, doi:10.1007/s00125-011-2344-9 (2012).

79 Rezania, A. *et al.* Production of functional glucagon-secreting α-cells from human embryonic stem cells. *Diabetes* **60**, 239-247, doi:10.2337/db10-0573 (2011).

80 Johansson, K. *et al.* Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Developmental cell* **12**, 457-465, doi:10.1016/j.devcel.2007.02.010 (2007).

81 Artner, I. *et al.* MafA and MafB regulate genes critical to beta-cells in a unique temporal manner. *Diabetes* **59**, 2530-2539, doi:10.2337/db10-0190 (2010).

82 Dohrmann, C., Gruss, P. & Lemaire, L. Pax genes and the differentiation of hormone-producing endocrine cells in the pancreas. *Mechanisms of development* **92**, 47-54, doi:10.1016/s0925-4773(99)00324-x (2000).

83 Bock, C. *et al.* Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-452, doi:10.1016/j.cell.2010.12.032http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1038/nbt.1783 (2011).

84 Elliott, A., Elliott, K. & Kammesheidt, A. High resolution array-CGH characterization of human stem cells using a stem cell focused microarray. *Molecular biotechnology* **46**, 234-242, doi:10.1007/s12033-010-9294-1 (2010).

85 Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature cell biology* **13**, 541-549, doi:10.1038/ncb2239 (2011).

86 Martinez-Fernandez, A., Nelson, T. & Terzic, A. Nuclear reprogramming strategy modulates differentiation potential of induced pluripotent stem cells. *Journal of cardiovascular translational research* **4**, 131-137, doi:10.1007/s12265-010-9250-2 (2011).

87 Hu, Q., Friedrich, A., Johnson, L. & Clegg, D. Memory in induced pluripotent stem cells: reprogrammed human retinal-pigmented epithelial cells show tendency for spontaneous redifferentiation. *Stem cells (Dayton, Ohio)* **28**, 1981-1991, doi:10.1002/stem.531 (2010).

88 Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell stem cell* **10**, 595-609, doi:10.1016/j.stem.2012.02.014 (2012).

89 Kim, K. *et al.* Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature biotechnology* **29**, 1117-1119, doi:10.1038/nbt.2052 (2011).

Chapter 2

PKC inhibitors direct human ESC-derived pancreatic progenitor differentiation toward pancreatic endocrine cells

2.1. Abstract

Directed differentiation of human embryonic stem cells (hESCs) has the potential to produce human cell types that can be used for disease modeling and cell transplantation. A key step in the differentiation from hESCs to glucose-responding β cells is the specification from pancreatic progenitors to endocrine cells. Two independent high-content chemical screens identified PKC inhibitors as inducers of endocrine cells from pancreatic progenitors. PKC inhibition at the stage of endocrine differentiation generated up to ten-fold more insulin expressing endocrine cells. Conversely PKC agonists blocked differentiation to pancreatic endocrine cells. PKC agonists were previously shown to increase differentiation into pancreatic progenitors, a developmental step immediately preceding endocrine differentiation. Together these results suggest a dynamic role of PKC at different stages during pancreatic *in vitro* differentiation. Transplantation of PKCβ inhibitor-treated pancreatic progenitors, containing higher proportions of endocrine progenitors and endocrine cells, results in mature β cells that show glucose-stimulated human c-peptide production *in vivo*.

2.2. Introduction

Human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs) represent a potentially unlimited starting material for the generation of functional pancreatic β cells. Essential in this pursuit is an efficient method for the differentiation of hESCs/iPSCs down the pancreatic lineage to produce endocrine cells. By mimicking known signals used during embryonic pancreatic development *in vivo*, efficient stepwise protocols have been developed to differentiate hESCs first into definitive endoderm and then into pancreatic progenitors (Figure 2-1a)[1-11]. However, the signals needed to produce endocrine progenitors from pancreatic progenitors, as well as insulin-expressing β cells from the endocrine progenitors, remain poorly defined.

Lineage tracing studies in mice have shown that production of hormone-expressing endocrine cells from pancreatic progenitors (marked by expression of Pancreatic and Duodenal Homeobox 1, PDX1) requires an intermediate stage, termed endocrine progenitor, in which cells express the key transcription factor Neurogenin 3 (NGN3)[12]. Following transient NGN3 expression, committed pancreatic endocrine progenitors give rise to pancreatic endocrine cells. Five types of pancreatic endocrine cells exist: α cells secrete the hormone glucagon, β cells secrete insulin, δ cells produce somatostatin, PP cells secrete pancreatic polypeptide and ε cells produce ghrelin. Although in adult pancreata the expression of each hormone is restricted to only one endocrine cell type, polyhormone expressing cells are frequently present during human fetal development[13,14]. hESC directed differentiation protocols produce both monohormonal and polyhormonal cells.

Several reports have shown that *in vitro* hESC-derived INS$^+$ cells do not exhibit functional glucose stimulated insulin secretion and lack expression of several mature β cell markers. In other words, these INS$^+$ cells are not functionally mature β cells. Moreover, transplantation of purified or enriched hESC-derived INS$^+$ cells generates primarily functional α cells, not β cells[11,15]. As a result of these observations, it has been proposed that generating immature hESC-derived INS$^+$ cells will not be beneficial for the ultimate goal of creating functional human β cells. However, the interpretation of purified endocrine or INS$^+$ cell transplantation experiments is confounded by the fact that *in vitro* derived INS$^+$ cells may require other surrounding cell types to successfully mature *in vivo*. It is therefore conceivable that hESC-derived INS$^+$ cells mature *in vivo* following transplantation and contribute beneficially to human c-peptide secretion in the mature graft. In the absence of genetic reporters it is however not possible to formally test this hypothesis. Therefore we hypothesized that increasing the percentage of hESC-derived endocrine progenitors and endocrine cells before transplantation will have a beneficial, not detrimental, effect on *in vivo* maturation to functional β cells.

Several recent studies demonstrate that blocking TGFβ signaling and BMP signaling improves the differentiation of pancreatic progenitors into endocrine cells [1,16,17]. However, the overall efficiency of creating insulin-producing cells remains very low. Moreover, protocols and hESC-lines vary dramatically in efficiencies with which INS$^+$ cells are produced. Thus, additional work needs to be done to dissect the signaling pathways

controlling the differentiation of pancreatic progenitors to endocrine cells. Based on our previous success using chemical screening to identify small molecules that direct hESC differentiation toward definitive endoderm[18] and pancreatic progenitors[19], we utilized the same approach to identify small molecules that could facilitate differentiation of pancreatic progenitors towards endocrine cells.

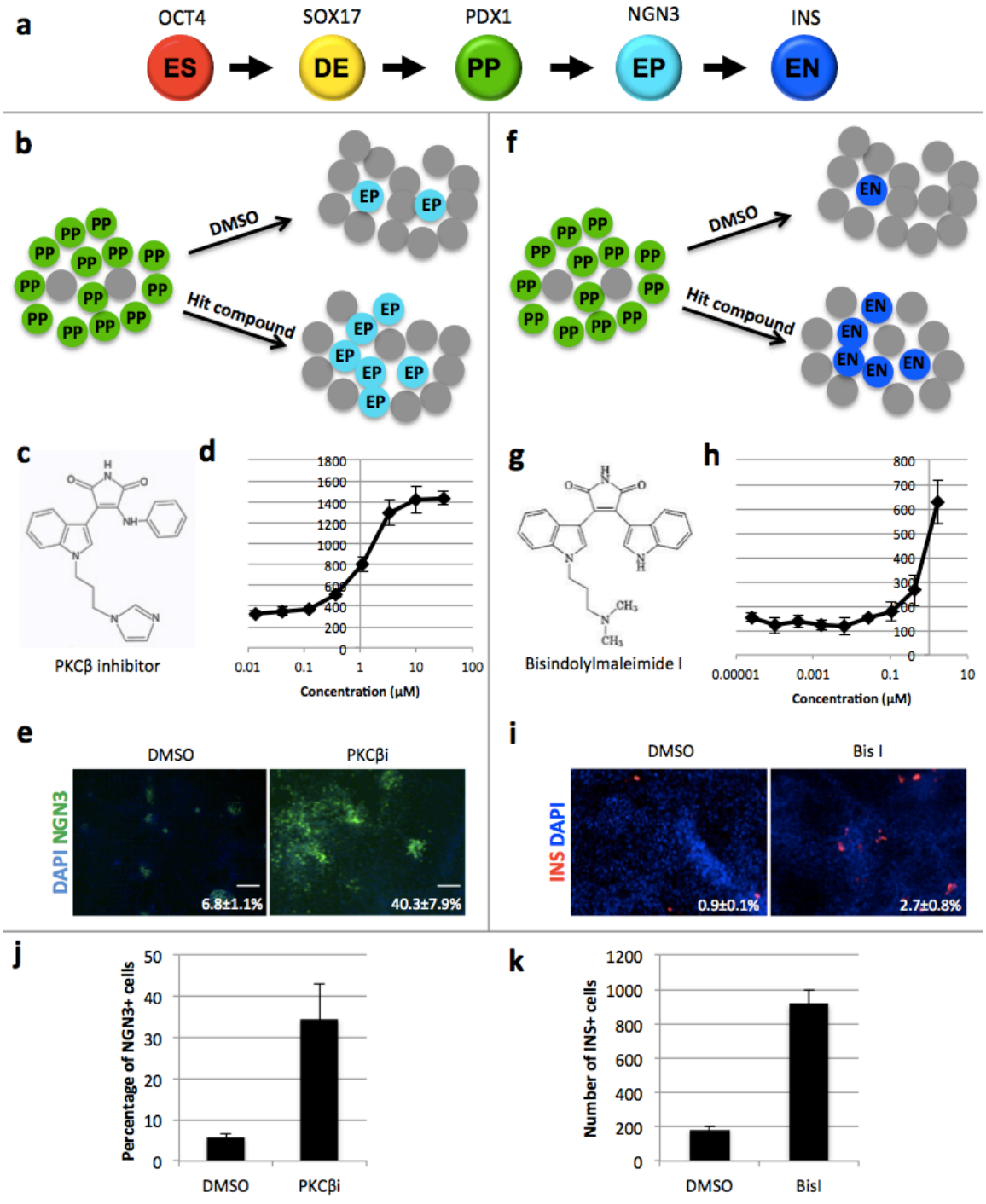
We designed two complementary screens to identify compounds that could promote endocrine differentiation. In the first instance, we screened for compounds that could increase the percentage and the total number of $NGN3^+$ endocrine progenitors produced from $PDX1^+$ hESC-derived pancreatic progenitors (Figure 2-1b). Since NGN3 expression is transient during embryonic development, it was not clear that a screen set up to detect this necessary intermediate and transient stage (NGN3 expression) could be successful. Therefore we concurrently performed a screen for compounds that could promote the appearance of insulin-expressing (INS^+) endocrine cells from $PDX1^+$ progenitors after prolonged culture. These two independent screens differed in the choice of hESC lines, directed differentiation protocols, chemical libraries and primary screening assays. In outcome it was gratifying that both screens identified PKC antagonists as inducers of the endocrine pancreatic lineage. Together with our previous work implicating PKC agonists in the induction of pancreatic progenitors, we suggest a dynamic role for PKC during pancreatic development and propose improvements to current directed differentiation protocols that result in the production of up to ten-fold more endocrine

cells, and significantly higher levels of human c-peptide production, following *in vivo* maturation.

Figure 2-1. High-content screens to increase the number of NGN3-expressing and insulin-expressing cells.

(a) Stepwise differentiation from hESCs to pancreatic endocrine cells. DE, definitive endoderm; PP, pancreatic progenitor; EP, endocrine progenitor; EN, endocrine cells. Scheme of high-content screen beginning with a population of cells that have differentiated into pancreatic progenitors. **(b,f)** Scheme of high-content screen, which begins with a population of cells that have differentiated into pancreatic progenitors **(c,g)** Chemical structures of PKCβi hit and Bisindolylmaleimide I compound. **(d,h)** Efficacy curve of the hit compounds. Bisindolylmaleimide I treatment at 6.75 μM and 27μM caused considerable auto-fluorescence and toxicity. Those data points were omitted from the graph. **(e)** PKCβi effect on HUES 8-pancreaticprogenitor population. After 14 d of differentiation, the HUES 8-pancreatic progenitor populations were treated with 10 μM PKCβi for four days and then stained with NGN 3 antibody (green). NGN3, neurogenin 3. Scale bar is 100 μm. **(i)** BisI effect on H1-pancreatic progenitor population. Representative images after 6 days of chemical treatment, insulin staining. Percentage of insulin-positive cells with s.d. **(j)** Effect of PKCβi on H1-derived pancreatic progenitor populations. After 14 d of differentiation, the H1-derived pancreatic progenitor populations were treated with 10 μM PKCβi for 6 days and then stained with NGN3 antibody. The percentage of $NGN3^{+}$ cells was analyzed with the Cellomics high content screening system . **(k)** HUES8 pancreatic progenitors were treated with DMSO or BisI for 6 days in the presence of Noggin and Alk5 inhibitors. Insulin cell number with was counted over the same area in different wells. Error bars indicate s.d. n=4.

Figure 2-1 Continued



2.3. Results

2.3.1. Screen to induce NGN3$^+$ cells

For the first screen, to detect compounds that promote appearance of NGN3$^+$ endocrine progenitors, we differentiated HUES8 hESCs using a modified version of previously published protocols[19] (see Methods, Protocol 1) to produce a population containing 71.2±6.5% PDX1$^+$ pancreatic progenitors. Compounds from a library containing 2000 chemicals, including signaling pathway regulators, kinase inhibitors, natural products, and FDA approved drugs (detailed library information is described in the Methods) were individually tested at 10 μM and 1 μM final concentrations, corresponding to 0.1% and 0.01% DMSO respectively. After six days culture, cells were stained with an antibody against NGN3 and analyzed with a Cellomics imaging reader. In DMSO control conditions, 6.8±1.1% cells stained positively for NGN3. Two primary hits increased the number of NGN3 expressing cells more than three-fold (Figure 2-2). Among these, PKCβ inhibitor (PKCβi) was selected for follow up studies due to the high efficacy and low toxicity (Figure 2-1c). PKCβi treatment increased the total number of NGN3$^+$ cells in a dose dependent manner (EC_{50}= 7.8 μM, Figure 2-1d), increasing the percentage of NGN3$^+$ cells to a maximum of 40.3±7.9%, nearly 7-fold higher than DMSO-treated controls

(Figure 2-1e). Quantitative RT-PCR analysis confirmed that the PKCβi-treated cells showed higher expression of NGN3 mRNA compared to DMSO controls (Figure 2-3).

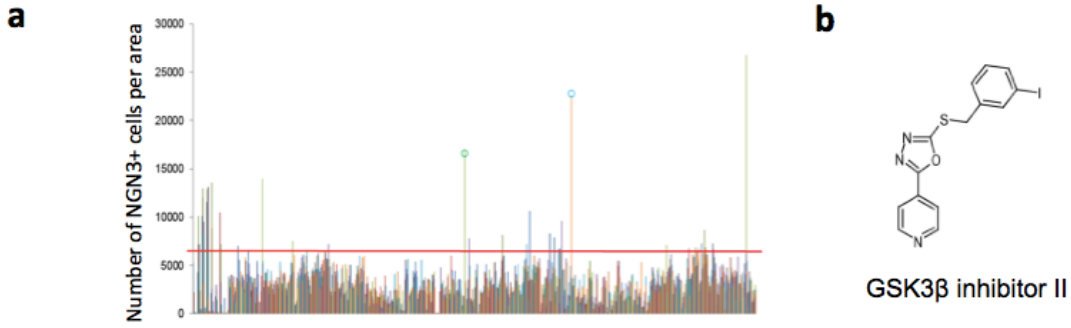


Figure 2-2. Screen1: Data analysis of the primary screen.

(a) Data of primary screen. Each line represents one compound at one concentration. 2,000 compounds were tested at two concentrations: 10 μM and 1 μM. The x-axis is the 2,000 compounds with two concentrations of each. The y-axis is the number of cells positively stained by the NGN3 antibody. Primary hits (above the red line) were designated as compounds that induced NGN3 in more than 6300 $NGN3^+$ cells/well, which is 3 times higher than the average. Subsequent tests confirmed six compounds that increase both the number and percentage of $NGN3^+$ cells. The compounds were labeled with different colors. The other dots above the red line are the compounds that only increase the percentage not the number of $NGN3^+$ cells because of compound toxicity. **(b)** Chemical Structures of other hit compounds.

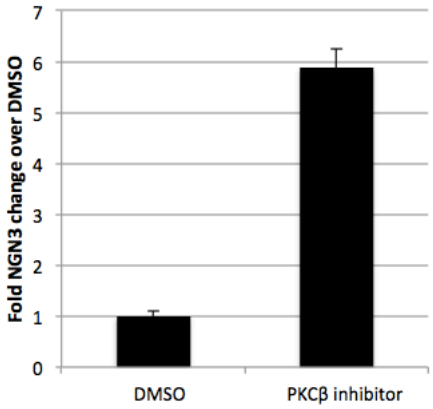


Figure 2-3. qRT-PCR analysis of the expression of NGN3 in the chemically- treated populations.

mRNA of HUES 8-derived pancreatic progenitor cells treated with DMSO was used as a control to normalize data.

2.3.2. Screen to induce INS$^+$ cells

In parallel, we carried out a second screen to increase the total number of insulin-expressing cells produced from pancreatic progenitors (Figure 2-1f). In order to discover novel pathways to increase the total number of insulin producing cells, both control and experimental conditions were treated with compounds that had been previously shown to facilitate the induction of the endocrine lineage, namely Noggin (a BMP inhibitor) and Alk5 inhibitor (a TGFβ inhibitor)[1,17]. The hESC line H1 was differentiated (see Methods Protocol 2) to stage 3 day 4 (S3D4) to produce a population containing approximately 80% PDX1$^+$ pancreatic progenitors, at which time a collection of 418 kinase inhibitors, signaling pathway regulators, and natural products were individually tested at 10 μM final concentrations, corresponding to 0.1% DMSO. After 6 days of compound treatment, cells

were fixed, stained using an insulin antibody, and analyzed with an imaging reader. This screen identified 6 compounds that increase the number of insulin expressing cells >3-fold over the DMSO controls (Figure 2-4). Notably, in light of the results of our first screen, 4 of the 6 hit compounds were classified as PKC inhibitors.

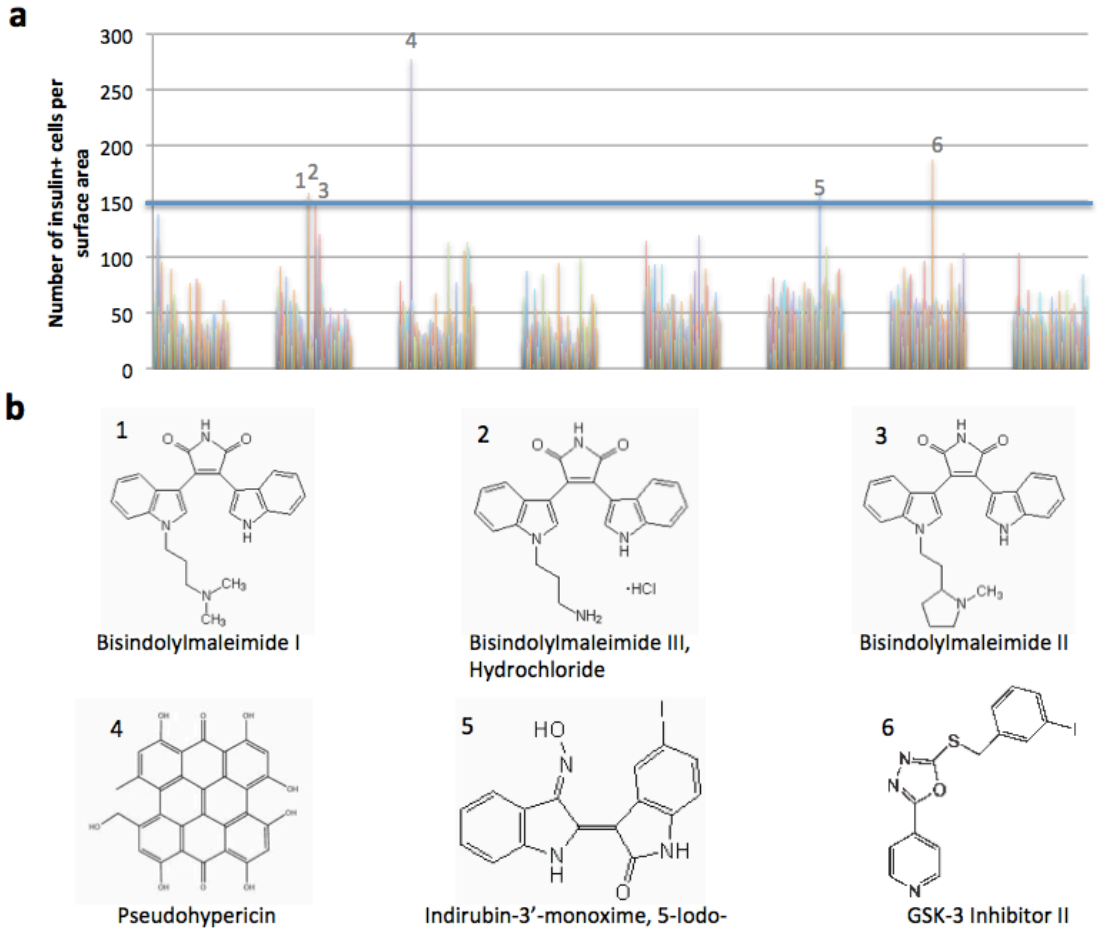


Figure 2-4. Screen 2: Data analysis of the primary screen.

(a) Data of primary screen. Each line represents one compound at 10 μM. 418 compounds and 366 DMSO controls were tested. The y-axis is the number of cells per surface area positively stained by the INS antibody. Primary hits (above the red line, numbered) were designated as compounds that induced INS in more than 148 INS^+ cells/defined area, which is 3 times higher than the average. **(b)** Chemical structures of hit compounds

Since 3 hit compounds were in the bisindolylmaleimide family, we selected Bisindolylmaleimide I (BisI) for subsequent studies (Figure 2-1g) BisI increased the total number of insulin-expressing cells in a dose dependent manner although significant auto-fluorescence and toxicity hampered dose-curve measurements at 6.75 μM and 27 μM. (Figure 2-1h). A lower dose of 1 μM BisI induced up to 2.7±0.8% insulin-expressing cells, compared with 0.9±0.1% for DMSO-treated controls (Figure 2-1i). Consistent with this finding, in a separate experiment, BisI treatment moderately increased mRNA expression of multiple hormones expressed downstream of NGN3, including insulin (2.6±0.5-fold), glucagon (2.2±0.4-fold), and somatostatin (1.25±0.1 fold) after a three day treatment (Figure 2-5).

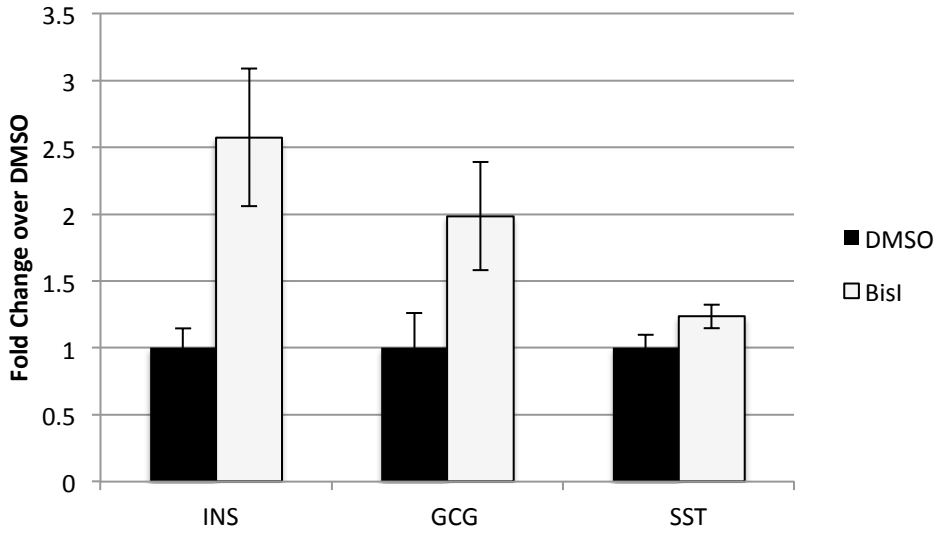


Figure 2-5. BisI induces expression of multiple endocrine markers

H1-derived pancreatic progenitors were treated for 3 days with DMSO or 3 μM BisI. qRT-PCR for INS, GCG and SST performed and shown as fold change over DMSO control.

2.3.3. Robust endocrine induction using independent hESC lines

To confirm that the effects of individual PKC inhibitors were not cell line specific, we tested them on the differentiation of both HUES8 and H1 cell lines. H1 pancreatic progenitors treated with PKCβi had a higher percentage of NGN3-expressing cells (34±8%) than DMSO treated controls (5.7±1%) (Figure 2-1j). Similarly, BisI increased the numbers of insulin-expressing cells in HUES8 up to 7.5-fold in the presence of Noggin and Alk5 inhibitor (Figure 2-1k). Thus, our parallel screening approaches identified a broad utility for PKC inhibition in the induction of the endocrine pancreatic lineage from hESC-derived pancreatic progenitors.

2.3.4. Characterization of BisI induced insulin-expressing cells

Previous work showed that insulin-producing cells generated by *in vitro* differentiation of hESCs are not fully functional β cells [1,3,19]. We therefore sought to determine whether insulin-expressing cells produced through the inhibition of PKC more closely resembled adult β cells. We observed that insulin-expressing cells differentiated in the presence of BisI displayed many features common to other insulin-expressing cells produced using *in vitro* differentiation[1,3,4,15,17,20,21], in that they are often polyhormonal, and lack expression of mature β cell markers Nkx6-1 and Ucn3 (Figure 2-6**a**). We therefore conclude that the insulin-expressing cells produced *in vitro* following PKC inhibition

resemble those which have been previously reported and are not true β cells. Consistent with literature, hESC-derived INS^+ do not exhibit functional glucose-stimulated insulin secretion (data not shown).

2.3.5. In vitro potential of PKCβi induced $NGN3^+$ cells

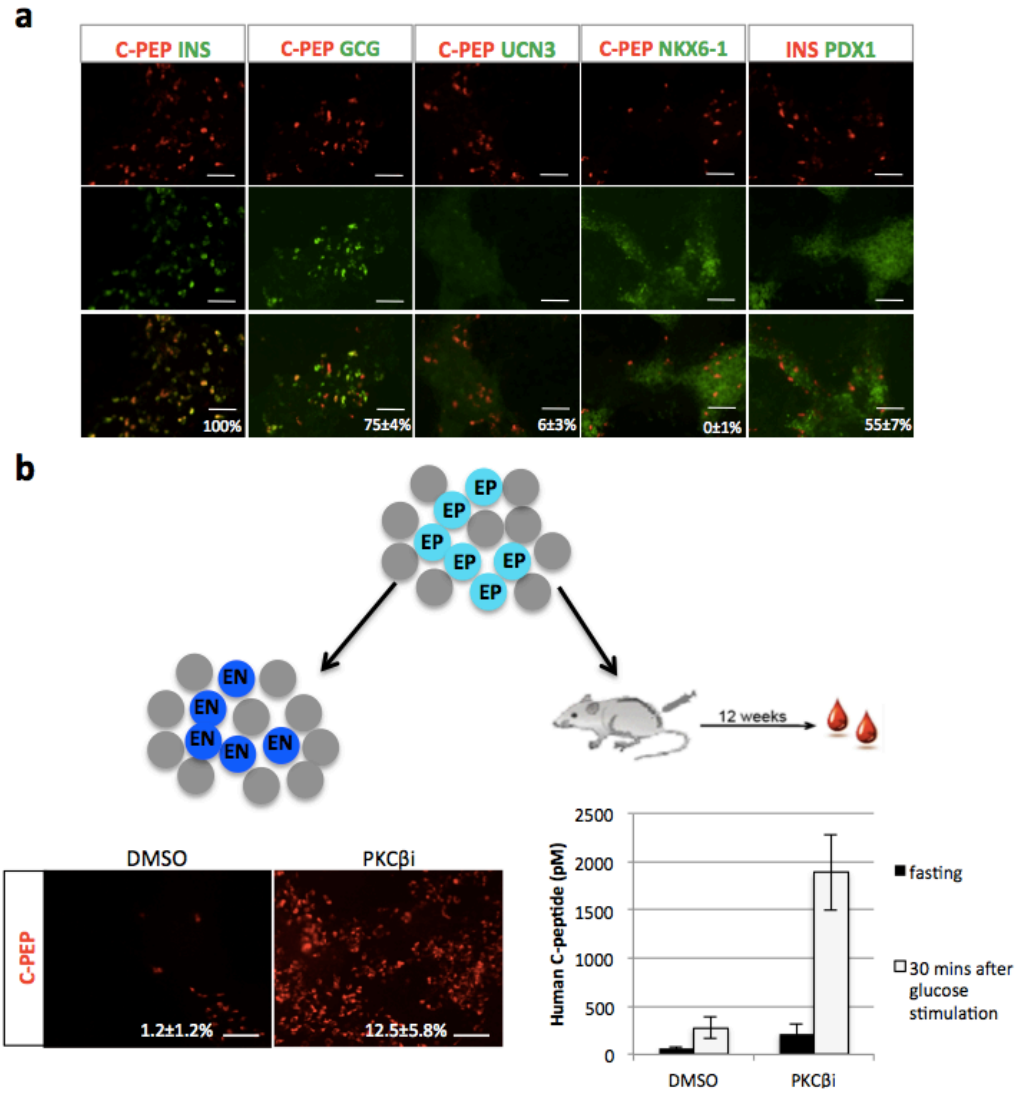
We next investigated the developmental potential of cell populations enriched for $NGN3^+$ cells that were induced by PKCβ inhibition. Endocrine progenitors generated after a 6-day treatment with PKCβi were further differentiated *in vitro* for an additional 6 days in spontaneous differentiation medium (DMEM+B27 medium without additional chemical or growth factors). More c-peptide (a byproduct of insulin biosynthesis) -expressing cells were detected in the derivatives of PKCβi treated cells (12.5±5.8%) than the derivatives of control cells (DMSO-treated, 1.2±1.2%; Figure 2-6b).

Figure 2-6. Characterization of cells derived by PKC inhibitor treatment

(a) H1 hESC-derived pancreatic progenitors were treated for 3 days with 3μM BisI. Cells were stained with insulin, c-peptide, glucagon, UCN3, PDX1 and NKX6-1 antibodies. All insulin-expressing cells express c-peptide. Few insulin-expressing cells express NKX6-1 or UCN3, markers of mature human β cells. Many insulin-expressing cells co-express another endocrine marker – glucagon. Scale bars are 100 μm. **(b)** The PKCβi-treated populations can further differentiate into endocrine cells *in vitro* and *in vivo*. Scheme beginning with cells that have been treated with hit chemicals, a population containing many NGN3-expressing cells. The HUES 8-derived pancreatic progenitor cells treated with DMSO were used as negative controls. Starting populations were cultured in DMEM+B27 medium for 6 days and stained with c-peptide antibody. C-PEP: c-peptide. The PKCβi -treated population differentiates into glucose-responding cells after transplantation under the kidney capsule of SCID-Beige mice. The PKCβi -treated populations were collected and implanted into the left kidney of SCID-Beige mice. The DMSO-treated population was used as a negative control. 12 weeks later, the mouse sera collected at fasting condition and at 30 mins after glucose stimulation were analyzed using ELISA to measure human c-

peptide expression. Error bars indicate s.d. Scale bars are 100 μm.

Figure 2-6 Continued



2.3.6. *In vivo maturation of PKCβi treated pancreatic progenitors*

Next, the developmental competence of pancreatic progenitors was assessed using an *in vivo* transplantation assay[18,19]. Control or PKCβi-treated cells were transplanted under

the kidney capsule of SCID-beige mice. Following a 12-week incubation period, the cells were assayed for insulin secretion by a glucose-stimulated c-peptide assay. Human c-peptide levels in the serum were significantly higher (8.71-fold, $p<0.01$, N=3) upon glucose injection than during fasting, indicating that PKCβi treated endocrine progenitor stage cells are capable of giving rise to mature β-like cells *in vivo* (Figure 2-6b). Moreover, glucose stimulated human c-peptide levels were significantly higher (6.82-fold, $p<0.01$, N=3) in animals implanted with PKCβi-treated than DMSO-treated pancreatic progenitors. The average human c-peptide levels in animals transplanted with PKCβi-treated cells 30 minutes after glucose injection were 1892±392 pM (N=3, S.D.)

2.3.7. PKC agonists block endocrine induction

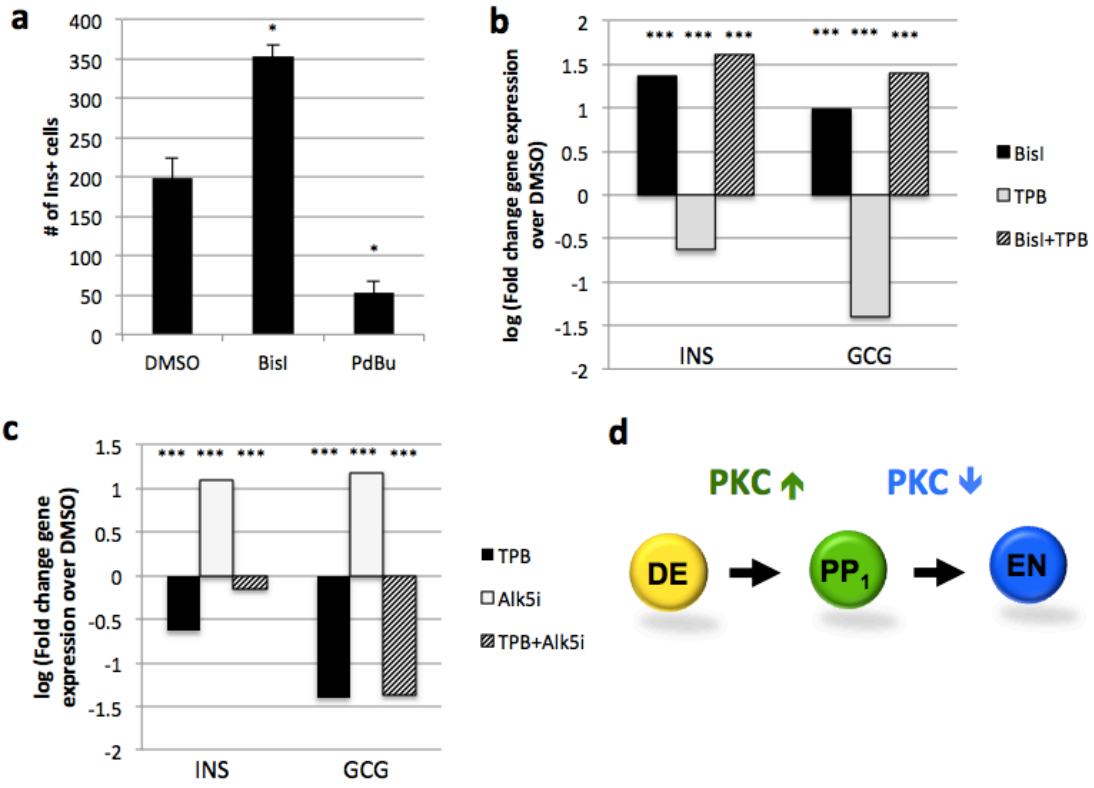
Since PKC antagonists induce endocrine progenitors, the converse treatment, namely PKC activation, should block endocrine induction. Indeed, PKC activation with phorbol 12,13-dibutyrate (PdBu) caused a dramatic (3.7-fold) decline in endocrine differentiation as measured by the number of insulin$^+$ cells after a three-day treatment of H1-derived pancreatic progenitors (see Methods Protocol 2) (Figure 2-7a). To confirm this result a different PKC agonist, (2S,5S)-(E,E)-8-(5-(4-(trifluoromethyl)phenyl)-2,4-pentadienoylamino)benzolactam (TPB) was shown by qRT-PCR to decrease the expression of endocrine hormones insulin and glucagon. The effect of TPB was reversed in the presence of BisI indicating that both TPB and BisI act specifically through PKC activation or inhibition (Figure 2-7b). Recently Rezania et al. using a different directed differentiation

protocol showed that strong PKC activation decreased NGN3 and NEUROD1 expression, consistent with these findings[10]. Several groups have reported that TGFβ inhibition is sufficient to increase the numbers of pancreatic endocrine cells[1,16,17]. We observed that PKC agonists were sufficient to block endocrine induction in the presence of TGFβ inhibitors, suggesting that PKC inactivation is required for endocrine induction by TGFβ inhibition (Figure 2-7c).

Figure 2-7. PKC agonists block the formation of insulin-expressing cells.

(a) hESC-derived pancreatic progenitors cells were treated with DMSO, 1 μM Bisindolylmaleimide I and 1 μM PDBu for 6 days. Wells were stained for insulin and cell number counted in representative portions of the well by the Cellomics high content screening system. Error bars represent s.d. Two-tailed T-test $p<0.05$ **(b)** Fold change of insulin (INS) and glucagon (GCG) transcripts by qRT-PCR of cells after 3 day treatment with 3 μM BisI, 5 μM TPB (PKC agonist) and BisI+ TPB. Endocrine induction is blocked by PKC agonist TPB and reversed by BisI. Data displayed as $\log_2$ of the fold change over DMSO. $P<0.001$. The decrease in INS and GCG expression due to TPB treatment is reversed by BisI. **(c)** Fold change of insulin (INS) and glucagon (GCG) transcripts by qRT-PCR of cells after 3 day treatment with TPB, Alk5 inhibitor or a combination of both. PKC agonist TPB is sufficient to block the effect of Alk5 inhibitors on endocrine induction. Data displayed as $\log_2$ of the fold change over DMSO. $P<0.001$. **(d)** Summary of the model where PKC agonists induce the differentiation towards pancreatic progenitors while PKC antagonists induce the differentiation towards endocrine cells.

Figure 2-7. Continued



2.4. Discussion

Current directed hESC differentiation protocols generate pancreatic progenitors, marked by expression of PDX1, with very high efficiency (70-99%). However, only a fraction of the pancreatic progenitors differentiate *in vitro* into endocrine cells. Recently, BMP and TGFβ pathway inhibitors were shown to significantly increase the differentiation of pancreatic progenitors into endocrine cells[1,16,17]. Using these inhibitors, up to 25% of insulin-expressing endocrine cells could be generated from pancreatic progenitors.

However, the efficiency of generating endocrine cells varied dramatically depending on the choice of cell line and directed differentiation protocol[1,22].

We hypothesized that additional pathways could play a role in the differentiation of pancreatic progenitors into endocrine cells. To uncover these pathways we carried out two chemical screens using different hESC-lines and directed differentiation protocols. To cover a broad chemical space, different small molecule libraries were used for each screen. Both screens identified PKC antagonists as inducers of endocrine differentiation. The first screen identified a specific PKCβ inhibitor that induced in a dose-dependent manner $NGN3^+$ endocrine progenitors and subsequently INS^+ cells. The second screen identified 6 hit compounds that increased the number of INS^+ cells over 3 fold. Strikingly, 4 of the 6 compounds were PKC inhibitors. Since 3 PKC inhibitors were in the bisindolylmaleimide family, Bisindolylmaleimide I (BisI) was chose for subsequent studies. Significant variation in the differentiation of different hES-lines has been observed, with protocols frequently optimized for the differentiation of one specific cell line. We therefore tested and confirmed that PKCβi and BisI had a reproducible effect on two different hESC lines.

To better understand the effect of BisI on pancreatic progenitors, we characterized the INS^+ cells produced after BisI treatment. Treated INS^+ cells lacked expression of mature β cells markers NKX6-1 and UCN3 and lacked mature function assessed by glucose stimulated insulin secretion (data not shown) suggesting that BisI increases the number of INS^+ cells and does not affect their immature phenotype. PKCβi treatment increases the numbers of $NGN3^+$ endocrine progenitors and subsequently INS^+ cells suggesting

endocrine induction instead of replication as the mechanism behind the increase in INS^+ cells from pancreatic progenitors.

Since the loss of function for PKC using two distinct inhibitors lead to the *in vitro* increase in INS^+ cells, we decided to test whether the converse, PKC activation, blocks endocrine induction. Indeed, evaluated by qRT-PCR and INS^+ cell numbers, two distinct PKC agonists blocked endocrine induction. As expected based on their molecular targets, the combined application of the PKC inhibitor BisI and PKC agonist TPB increased endocrine markers suggesting that modulation of PKC activity is specifically responsible for both the induction and inhibition of differentiation into endocrine cells. Surprisingly, the effect of TGFβ receptor (Alk5) inhibitor, a broadly used inducer of the endocrine lineage, was abolished and reversed by the PKC agonist TPB suggesting that PKC inhibition may be necessary for endocrine induction by TGFβ inhibitors.

Transplantation of pancreatic progenitors under the kidney capsule of immune compromised mice is a well-established assay for spontaneous *in vivo* differentiation and maturation of pancreatic progenitors. PKCβi-treated pancreatic progenitors transplanted into SCID-Bg mice secreted significantly higher levels of c-peptide 30 minutes after glucose injection as compared to fasting. This confirms successful *in vivo* maturation. Moreover, transplanted PKCβi-treated pancreatic progenitors had significantly higher c-peptide levels than transplanted DMSO-treated progenitors. This is likely due to a higher proportion of $NGN3^+$ progenitors and INS^+ cells at the time of transplant. This explanation is consisted with the observation that ALK5i-treated pancreatic progenitors give rise to higher human c-

peptide serum levels 8, 12 and 16 weeks post-transplant [10]. It is also possible, although less likely, that PKCβi treatment additionally affects the speed of maturation or the survival of the transplanted pancreatic progenitors. 12 weeks post-transplant of PKCβi-treated pancreatic progenitors, glucose stimulated human c-peptide levels were 1892±392 pM, significantly higher than our previously reported transplants using the same transplantation protocol and HUES8 cell line (human c-peptide levels <250pM) [19]. Streptozotocin (STZ) treatment is a common methods to ablate mouse β cells and induce diabetes. Three studies have to date shown rescue of STZ-treated mice with hESC-derived pancreatic progenitor grafts[4,10,23]. Compared to these studies, we report similarly high levels of human c-peptide 12wk post-transplant of PKCβi-treated and not DMSO-treated pancreatic progenitors.

Our work suggests that PKC inhibitors increase the numbers of committed endocrine progenitors and immature INS^+ cells before transplantation and result in increased amounts of c-peptide secretion from grafts following *in vivo* maturation. These findings would suggests that generating more INS^+ cells *in vitro* is a beneficial approach to generate functional β cells grafts and achieve high levels of human c-peptide production. Lineage tracing experiments are required to determine whether the *in vivo* matured β cells arise directly from the $NGN3^+$ or INS^+ cells present at the time of transplantation.

In sum, these experiments show the application of high-content chemical screening for investigating pathways capable of directing the differentiation of hESCs toward the pancreatic lineage. We performed two high-content chemical screens and both identified

PKC antagonists that enhanced differentiation of hESC-derived pancreatic progenitors toward the endocrine lineage. Conversely, PKC agonists blocked differentiation towards insulin-producing cells even in the presence of TGFβ inhibitors. Previously we reported that (-)-Indolactam V, a PKC activator, promotes the generation of pancreatic progenitors[19]. Together with our current findings, we suggest that PKC plays a dynamic role in human pancreatic development: PKC signaling promotes the generation of pancreatic progenitors, while PKC inhibition in necessary for the subsequent differentiation towards pancreatic endocrine cells (Figure 2-7d).

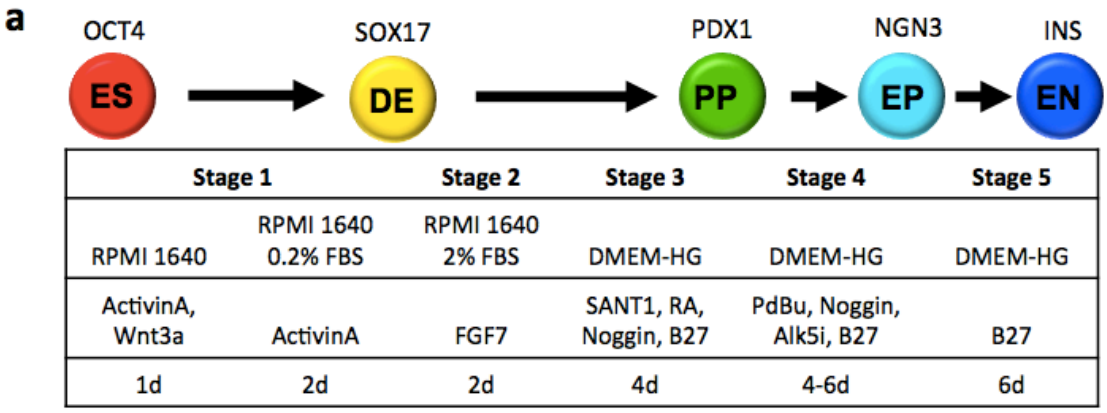
Unbiased chemical screens are a useful approach to identify pathways and reagents to control and improve hESC differentiation. Treatment with a specific PKCβ inhibitor increased the numbers of endocrine progenitors and upon transplantation generated grafts that produce high levels of human c-peptide, typically sufficient to maintain normoglycemia and rescue STZ induced diabetes. We anticipate the chemicals identified from these screens will help to establish a more efficient strategy to differentiate hESCs/iPSCs into functional β cells. Remaining challenges include finding small molecules that can further improve the generation of insulin-expressing cells, and importantly, generate functionally mature β cells *in vitro*.

2.5. Materials and Methods

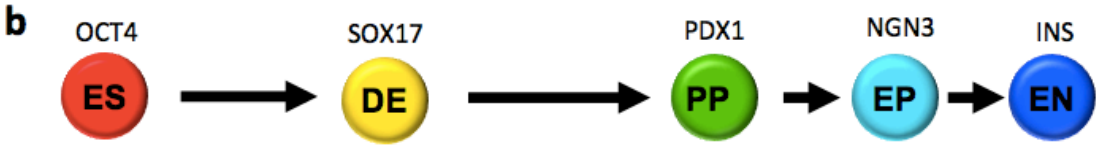
2.5.1. hESC culture and differentiation. Protocol 1.

HUES 8 and H1 cells are cultured on irradiated MEF feeder cells in KnockOut DMEM (Invitrogen) supplemented with 10% (v/v) KnockOut serum replacement (Invitrogen), 0.5% (v/v) human plasma fraction (Talecris Biotech), 2 mM L-glutamine (Invitrogen), 1.1 mM 2-mercaptoethanol (Invitrogen), 1 mM nonessential amino acids (Invitrogen), 1×penicillin/streptomycin (PS, Invitrogen) and 10 ng/ml bovine FGF (Invitrogen). Cells are passaged at the ratio of 1:6 every 5 d by using 0.05% trypsin (Invitrogen). To generate the HUES-pancreatic progenitor population, HUES cells were cultured on MEF feeder cells until they are 80–90% confluent, then treated with 25 ng/ml Wnt3a (R&D systems) + 100 ng/ml activin A (R&D systems) in RPMI (Invitrogen) supplemented with 1×L-glutamine and 1×PS for 1 d, and then 100 ng/ml activin A in RPMI supplemented with 1×L-glutamine, 1×PS and 0.2% (v/v) fetal bovine serum (FBS, Invitrogen) (Figure 2-8a). The medium was changed 2 d later to 50 ng/ml/ FGF7 (R&D systems) in RPMI supplemented with 1×L-glutamine, 1×PS and 2% (v/v) FBS, and maintained for an additional 2 d. Cells were then transferred to 100 ng/ml Noggin (R&D systems) + 0.25 μM SANT-1 (Sigma) + 2 μM retinoic acid (Sigma) in DMEM supplemented with 1×L-glutamine, 1×PS and 1×B27 (Invitrogen) and cultured for an 4 d, followed by treatment with 100 ng/ml Noggin (R&D systems) + 1 μM ALK5 inhibitor (Enzo) + 100 nM PDBu (Sigma) in DMEM supplemented with 1×L-glutamine, 1×PS and 1×B27 (Invitrogen) and cultured for an additional 4 d. All HUES-pancreatic progenitor cells were

derived using the same protocol as described above. The percentage and number of PDX1^{+} cells and NGN3^{+}cells were analyzed with the Cellomics high content screening system (Thermo Scientific). PKCβi was purchased from Calbiochem.



Stage 1		Stage 2	Stage 3	Stage 4	Stage 5
RPMI 1640	RPMI 1640 0.2% FBS	RPMI 1640 2% FBS	DMEM-HG	DMEM-HG	DMEM-HG
ActivinA, Wnt3a	ActivinA	FGF7	SANT1, RA, Noggin, B27	PdBu, Noggin, Alk5i, B27	B27
1d	2d	2d	4d	4-6d	6d



Stage 1		Stage 2	Stage 3	Stage 4	Stage 5
RPMI 1640 2% BSA	RPMI 1640 2% BSA	RPMI 1640 2% BSA	DMEM-HG 0.1% BSA	DMEM-HG 0.1% BSA	DMEM-HG 0.1% BSA
ActivinA, Wnt3a bFGF	ActivinA bFGF	FGF7	FGF7, ActivinA, SANT1, RA, Noggin, ITS-X	SANT1, PdBu, Noggin, ITS-X	Noggin, Alk5, ITS-X
1d	2d	2d	4d	3d	3d

Figure 2-8. Schematic of directed differentiation protocols

Stepwise differentiation from hESCs to pancreatic endocrine cells. DE, definitive endoderm; PP, pancreatic progenitor; EP, endocrine progenitor; EN, endocrine cells. Table contains reagents used during each stage of directed differentiation.

2.5.2. High-content screen 1

The quality of all compounds was assured by the vender as greater than 90% pure. The library includes compounds from Sigma LOPAC

(http://www.sigmaaldrich.com/chemistry/drug-discovery/validation-libraries/

lopac1280-navigator.html/), and activators and inhibitors of signaling pathways. Data were normalized as fold change compared with DMSO control. Hit compounds were defined as those higher than threefold compared to the DMSO control. Individual samples of hit compounds were picked from the original library and confirmed with the same method as in the primary screen for three times. Three compounds were confirmed for further assay.

2.5.3. Screen 1: Generation of efficacy curve.

The HUES 8- pancreatic progenitor population was generated and plated onto 96-well plates with the same method as the high-content screen. After overnight incubation, the hit compounds were added at final concentrations of 30 μM, 10 μM, 3.3 μM, 1.1 μM, 370 nM, 123 nM, 41 nM, and 14 nM. Six days later, the cells were stained with NGN3 antibody and the plates were analyzed with the Cellomics high content screening system (Thermo Scientific).

2.5.4. Screen1: Immunostaining.

Cells were fixed with 10% (v/v) formalin solution (Sigma) for 20 min at room temperature (22–24 °C). Immunostaining was carried out with standard protocols. The following primary antibodies were used: goat anti-PDX1 (1:500, R&D system, AF2419); guinea pig anti-insulin (1:1,000, Dako, A0564); rabbit anti-glucagon (1:200, Cell Signaling, 2760); rabbit anti-c-peptide (1:500, Linco, 4020-01); and sheep anti-NGN3 (1:100, RnD systems, AF3444). Alexa-488–, Alexa-555– and Alexa-647–conjugated donkey antibodies against mouse, rabbit, guinea pig, goat, and sheep (Invitrogen), were used at 1:500 dilution.

2.5.5. hESC culture and differentiation. Protocol 2.

Human ESCs were routinely cultured on hESC-certified matrigel (BD Biosciences) in mouse embryonic fibroblast conditioned hES media (MEF-CM). MEF-CM media was produced by conditioning hESC media for 24 days on a confluent layer of mouse embryonic fibroblasts and subsequently adding 20ng/ml bFGF (Invitrogen). hES media was composed of DMEM/F12 (GIBCO) media supplemented with 20% KnockOut Serum Replacement (GIBCO), 2 mM L-glutamine (L-Glu, GIBCO), 1.1 mM 2-mercaptoethanol (GIBCO), 1 mM nonessential amino acids (GIBCO), 1x penicillin/strepto-mycin(P/S,GIBCO).Cells were passaged at the ratio of 1:6–1:20 every 4-7 days using TrypLE Express (Invitrogen).

To initiate differentiation the cells were dissociated using TrypLE Express to single cells and seeded at 150,000 cell/cm2 onto 1:30 dilution of growth factor reduced matrigel (BD Biosciences) in DMEM/F12 in MEF-CM media with 10uM Y27632 (StemGent). Two days following seeding the differentiation was started.

Cells were exposed to RPMI 1640 (Invitrogen) supplemented with 2% reagent grade BSA (Proliant) and 20ng/ml human Wnt3a (R&D Systems) + 100ng/ml rhActivinA (R&D Systems) + 8ng/ml bFGF(Invitrogen) for the first day (stage 1.1) (Figure 2-8b). During day 2 and 3 the day1 media was used with the exception of Wnt3a (stage1.2) . During days 4-5 cells were treated with RPMI + 2% BSA + 50ng/ml FGF7 (Peprotech) (stage 2). For days 6-9 cells were treated with DMEM-HG (Invitrogen) + 0.1% Albumax BSA (Invitrogen) + 1:200 ITS-X (Invitrogen) + 50ng/ml FGF7 (Peprotech) + 2μM RA (Sigma) + 0.25μM SANT-1 (Sigma) + 20ng/ml rhActivinA (R&D Systems) + 100ng/ml rhNoggin (R&D Systems) (stage 3). During days 10-12 cells were treated with DMEM-HG (Invitrogen) + 0.1% Albumax BSA (Invitrogen) + 1:200 ITS-X (Invitrogen) + 100ng/ml rhNoggin (R&D Systems) + 0.25μM SANT-1 (Sigma) + 100nM PDBu (EMD Biosciences) (stage 4). During days 13-15 cells were treated with DMEM-HG (Invitrogen) + 0.1% Albumax BSA (Invitrogen) + 1:200 ITS-X (Invitrogen) + 100ng/ml rhNoggin (R&D Systems) + 1μM Alk5 inhibitor (Axxora) (stage 5). PKC inhibitor Bisindolylmaleimide I (VWR) was added during stage 5.

2.5.6. High-content screen 2

The compound libraries used for this study: 400 compounds, including bioactive molecules, natural products, and 400 compounds that are known modulators of development or signaling pathways. For the chemical screen the day 8 cells were dispersed into single cells, using TrypLE, and replated at 150,000 cells/well of a 96 well plate in the presence of stage 3 media and 10μM Y27632. Compounds were added to the wells on day 10 in stage 4 media, then again on days 12 and 14 in stage 5 media. Media was changed every other day suing the treatment. Cells were fixed using 4% paraformaldehyde (PFA, Sigma) on day 15 and stained using mouse anti-glucagon and rabbit anti-c-peptide antibodies.

2.5.7. Screen 2: Immunostaining.

Following 4% PFA fixation, cells were 3x5 min washed in PBS and blocked with 10% donkey serum (Jackson Immunoresearch) in PBS/0.3% Triton X.

Primary antibodies were incubated overnight at 4C. Secondary antibodies were incubated for 1 hr at room temperature. The following primary antibodies and dilutions were used: rabbit anti-c-peptide (1:1000, BCBC), rat anti-c-peptide (1:500, DSHB, GN-ID4), guinea pig anti-insulin (1:1000, DAKO, A0564), guinea pig anti-glucagon (1:500, DAKO), mouse anti-glucagon (1:500, Sigma Aldrich, G2654), goat anti-PDX1 (R&D

Systems AF2419), mouse anti-NKX6.1 (DSHB, F55A12), rabbit anti-UCN3 (1:500, Phoenix Pharmaceuticals, H-019-28).

2.5.8. Kidney capsule implantation and tissue preparation.

All animal experiments were performed following an approved protocol of Harvard University under assurance # A3593-1 (protocol 93-15). The cells were washed with DPBS 3 times and collected for transplantation. The cells were then lifted with a cell scraper, collected by centrifugation and resuspended in 50 μl PBS. About 40 μl ($3\text{-}4\times10^6$ cells) of cell clumps were implanted into the left kidney of avertin-anesthetized SCID-Beige mice. 12 weeks later, the mouse sera were collected after overnight fasting condition or 30 mins after 3g/kg D-Glucose treatment by IP injection. The c-peptide levels in mouse sera were measured using the human c-peptide ELISA kit (Alpco Diagnostics or Millipore).

2.6. References

1 Nostro, M. *et al.* Stage-specific signaling through TGFβ family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development (Cambridge, England)* **138**, 861-871, doi:10.1242/dev.055236 (2011).

2 D'Amour, K. *et al.* Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature biotechnology* **23**, 1534-1541, doi:10.1038/nbt1163 (2005).

3 D'Amour, K. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nature biotechnology* **24**, 1392-1401, doi:10.1038/nbt1259 (2006).

4 Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nature biotechnology* **26**, 443-452, doi:10.1038/nbt1393 (2008).

5 Wells, J. & Melton, D. Vertebrate endoderm development. *Annual review of cell and developmental biology* **15**, 393-803, doi:10.1146/annurev.cellbio.15.1.393 (1999).

6 Jiang, J. *et al.* Generation of insulin-producing islet-like clusters from human embryonic stem cells. *Stem Cells* **25**, 1940-1953, doi:10.1634/stemcells.2006-0761 (2007).

7 Jiang, W. *et al. In vitro* derivation of functional insulin-producing cells from human embryonic stem cells. *Cell research* **17**, 333-344, doi:10.1038/cr.2007.28 (2007).

8 Tateishi, K. *et al.* Generation of insulin-secreting islet-like clusters from human skin fibroblasts. *J Biol Chem* **283**, 31601-31607, doi:10.1074/jbc.M806597200 (2008).

9 Zhang, D. *et al.* Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells. *Cell research* **19**, 429-438, doi:10.1038/cr.2009.28 (2009).

10 Rezania, A. *et al.* Maturation of Human Embryonic Stem Cell-Derived Pancreatic Progenitors into Functional Islets Capable of Treating Pre-existing Diabetes in Mice. *Diabetes*, doi:10.2337/db11-1711 (2012).

11 Basford, C. *et al.* The functional and molecular characterisation of human embryonic stem cell-derived insulin-positive cells compared with adult pancreatic beta cells. *Diabetologia* **55**, 358-371, doi:10.1007/s00125-011-2335-x (2012).

12 Gu, G., Dubauskaite, J. & Melton, D. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development (Cambridge, England)* **129**, 2447-2457 (2002).

13 Lukinius, A., Ericsson, J., Grimelius, L. & Korsgren, O. Ultrastructural studies of the ontogeny of fetal human and porcine endocrine pancreas, with special reference to colocalization of the four major islet hormones. *Developmental biology* **153**, 376-385, doi:10.1016/0012-1606(92)90122-w (1992).

14 Riedel, M. *et al.* Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia* **55**, 372-381, doi:10.1007/s00125-011-2344-9 (2012).
15 Kelly, O. *et al.* Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nature biotechnology* **29**, 750-756, doi:10.1038/nbt.1931 (2011).
16 Kunisada, Y., Tsubooka-Yamazoe, N., Shoji, M. & Hosoya, M. Small molecules induce efficient differentiation into insulin-producing cells from human induced pluripotent stem cells. *Stem cell research* **8**, 274-284, doi:10.1016/j.scr.2011.10.002 (2012).
17 Rezania, A. *et al.* Production of functional glucagon-secreting α-cells from human embryonic stem cells. *Diabetes* **60**, 239-247, doi:10.2337/db10-0573 (2011).
18 Borowiak, M. *et al.* Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell stem cell* **4**, 348-358, doi:10.1016/j.stem.2009.01.014 (2009).
19 Chen, S. *et al.* A small molecule that directs differentiation of human ESCs into the pancreatic lineage. *Nature chemical biology* **5**, 258-265, doi:10.1038/nchembio.154 (2009).
20 Blum, B. *et al.* Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nature biotechnology* **30**, 261-264, doi:10.1038/nbt.2141 (2012).
21 Maehr, R. *et al.* Generation of pluripotent stem cells from patients with type 1 diabetes. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15768-15773, doi:10.1073/pnas.0906894106 (2009).
22 Micallef, S. *et al.* INS(GFP/w) human embryonic stem cells facilitate isolation of *in vitro* derived insulin-producing cells. *Diabetologia* **55**, 694-706, doi:10.1007/s00125-011-2379-y (2012).
23 Schulz, T. *et al.* A scalable system for production of functional pancreatic progenitors from human embryonic stem cells. *PloS one* **7**, doi:10.1371/journal.pone.0037004 (2012).

Chapter 3

Functional β cells maturation is marked by an increase in the glucose threshold for insulin secretion and by expression of urocortin3

3.1. Abstract

Insulin expressing cells that have been differentiated from human pluripotent stem cells *in vitro* lack the glucose responsiveness characteristic of mature β cells. β cell maturation in mice was studied to find genetic markers that enable screens for factors that induce bona fide β cells *in vitro*. We find that functional β cell maturation is marked by an increase in the glucose threshold for insulin secretion and by expression of the gene urocortin 3.

3.2. Body

The directed differentiation of human pluripotent stem cells (HPSCs) has the potential to produce β cells for transplantation into diabetics. However, the available protocols for *in vitro* differentiation produce only "β-like" cells. These "β-like" cells do not perform the accurate glucose-stimulated-insulin-secretion (GSIS) found in mature β cells unless they are transplanted into mice and allowed to further differentiate for many weeks[1]. During normal development, insulin-expressing β cells appear around embryonic day 13.5 in mice or week 8-9 post-conception in humans[2,3], but regulated GSIS has been observed only days after birth. The signals and mechanisms governing β cell maturation, either during postnatal development or after transplantation, are unknown.

We aim to define functional β cell maturation based on GSIS parameters, and to identify markers of functionally mature β cells that could be used to make functional HPSC-derived β cells in culture.

Traditionally, GSIS is measured by the fold change in insulin secretion between low (2.8-5mM) and high (>10mM) glucose concentrations[4]. In this assay, neonatal β cells display a high basal insulin secretion at low glucose concentrations, and stimulation with a high concentration of glucose results in a small fold increase in insulin secretion. These data could be explained if neonatal β cells have uncontrolled insulin "leakiness" at low

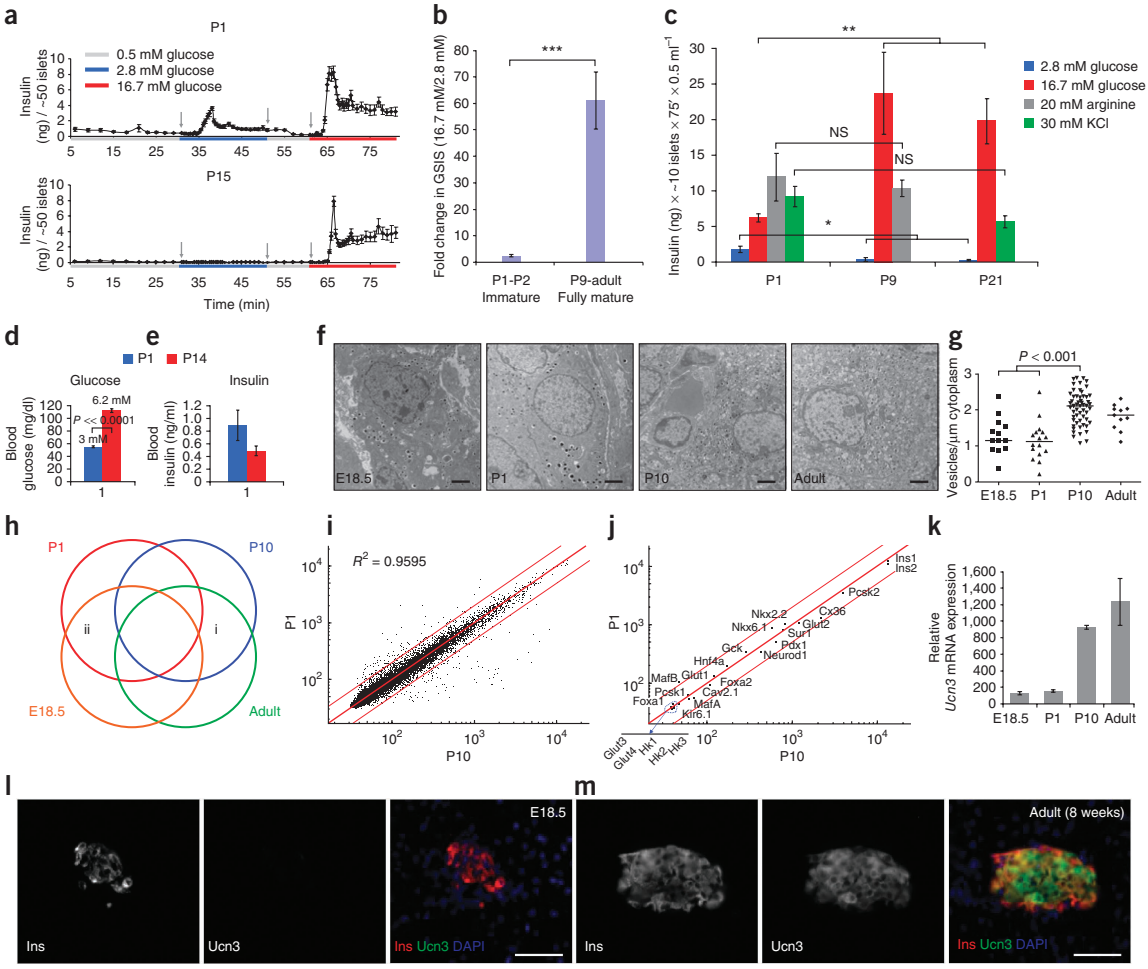
glucose concentrations, or alternatively, if they have a lower glucose concentration threshold at which they secrete insulin. To distinguish between these two hypotheses, we performed dynamic GSIS on neonatal (P1) and older (P15) mouse islets using a very low baseline glucose level of 0.5mM. The data show that neonatal P1 islets execute a full GSIS response (both first and second phases of insulin secretion) at low (2.8mM) glucose concentrations, whereas P15 islets show no response (no insulin secretion) at this concentration (Figure 3-1a). These results show that immature β cells are not "leaky", but rather have a reduced threshold for GSIS, secreting insulin in response to a lower glucose concentration than mature β cells.

To determine when β cells acquire a mature GSIS capacity, we tested mouse islets isolated from P1 to adult for their response to low (2.8mM) and high (16.7mM) glucose concentrations. Islets from neonatal mice, ages P1 and P2, secreted 2.6±0.5-fold more insulin in high glucose than in low glucose, whereas islets from P9 to adult secreted, on average, 60.9±10.7-fold more insulin in high glucose than in low glucose (Figure 3-1b). Thus the dramatic change in GSIS response between low and high glucose that characterizes β cell maturation occurs between P2 and P9. Islets of mice younger than P2 display an “immature” response, whereas islets from mice older than P9 respond as "mature" β cells. Between P3 to P8, a mixed (intermediate) GSIS phenotype was observed.

Figure 3-1. β cell maturation is defined by a decrease in GSIS sensitivity to low glucose levels and by the expression of Ucn3

(a) Three independent sets of 50 islets each, from P1 or P15 mice, were sequentially perfused with basal (0.5mM, gray), low (2.8mM, blue) or high (16.7mM, red) glucose in a dynamic GSIS assay. Arrows indicate the time points at which solutions were changed. P1 islets display complete first and second phases of GSIS in response to low glucose, whereas P15 islets do not secrete insulin at this glucose concentration. **(b)** Triplicates of 10 islets from P1 to adult were assayed for GSIS using low glucose (2.8mM) and high glucose (16.7mM). Two age groups can be distinguished according to their stimulation index (fold change in GSIS). ***, $P<5x10^{-5}$. **(c)** Three independent sets of ten islets each from P1, P9 or P21 were assayed for GSIS using low glucose (2.8mM, blue), high glucose (16.7mM, red), 20mM arginine (gray) or 30mM KCl (green). The difference in the amount of insulin secreted between mature and immature islets is specific to glucose. *, $P<0.05$; **, $P<0.001$; NS, not significant). **(d)** Blood glucose and **(e)** insulin levels in immature (P1, blue) and mature (P14, red) mouse pups. Insulin levels in the immature pups are higher than in the mature ones, although their blood glucose levels are lower. **(f)** Electron micrograph of insulin vesicles in β cells at various ages. Scale bars = 2μm. **(g)** Quantification of the number of insulin vesicles vs. β cell area of the data shown in F. **(h)** A scheme representing the microarray approach. Genes differentially expressed in both mature age groups compared to both immature age groups (I and II) are chosen as candidate markers. **(i)** Representative scattered plot from the microarray. Note high similarity (R^2) in gene expression between the mature (P10) and immature (P1) samples. **(j)** The expression levels of most β cell markers are unchanged during GSIS maturation. Scatter plots of global gene expression from microarrays on FACS-sorted immature (P1) and mature (P10) β cells. Red lines mark a 2-fold difference in expression and, with the exception of MafB, gene expression is not significantly different between these stages. **(k)** The expression of Ucn3 mRNA at various ages as detected in the microarray. **(l)** Immunostaining of Ucn3 (green) and insulin (red) on pancreata from E18.5 and adult mice. Nuclei are stained with DAPI (blue). Scale bars = 50μm **(m)** Ucn3 is undetectable in E18.5 embryo. Ucn3 is detected at high levels and co-localizes with adult β cells.

Figure 3-1 Continued



The differences in insulin secretion between mature and immature β cells is specific for glucose. The amount of insulin secreted by P1 and P9 islets in response to 30mM KCl was 11.9±3.5ng and 10.3±1.1ng, respectively. The amount of insulin secreted from P1 and P21 islets in response to 20mM arginine was 9.17±1.4ng and 5.66±0.9ng, respectively. These differences are not statistically significant (Figure 3-1c). In contrast,

islets from P1 mice secreted only 6.2±0.6ng insulin, during 75min in 0.5ml high (17.7mM) glucose, while the same number of islets from P9 or P21 secreted 23.7±5.7ng and 19.7±3.2ng insulin, respectively ($P<0.001$). At low glucose levels, the opposite trend was observed: P1 islets secrete 1.8±0.5ng insulin at 2.8mM glucose, whereas P9 and P21 islets secrete only 0.4±0.3ng and 0.3±0.1ng insulin, respectively ($P<0.05$) (Figure 3-1c).

We examined the physiological consequences *in vivo* of the differences observed *in vitro* between mature and immature β cells' response to glucose. In agreement with previous reports[4], P1 pups had significantly lower blood glucose levels than P14 pups. The average blood glucose concentration at P1 is 3mM whereas blood glucose at P14 averages 6.2mM ($P<2.5x10^{-24}$) (Figure 3-1d). Notably, the average blood glucose level in P1 pups is higher than the glucose concentration that causes insulin secretion *in vitro* in P1 islets. If the *in vitro* observation that immature β cells secrete insulin at low glucose levels (Figure 3-1a) holds true *in vivo*, one should see higher insulin in the blood of neonates. Consistent with this prediction, the P1 pups had nearly two-fold higher levels of insulin in their blood than P14 animals (Figure 3-1e), although we note that there is a high variability of blood insulin in non-fasted animals. We also examined insulin granules in β cells at each stage using electron microscopy (Figure 3-1f-g). P1 β cells contained approximately 2-fold fewer insulin granules compared to P10 β cells, suggesting a mechanistic difference in insulin secretion.

To find molecular markers whose expression pattern correlates with β cell maturation, β cells expressing Pdx1-EGFP from P1 or P10 animals were sorted by FACS and their global gene expression patterns were compared using transcriptional arrays. The Pdx1-EGFP strain was used instead of the insulin-EGFP strain due to the slightly diabetic phenotype observed in the latter animals. We also analyzed β cells from E18.5 embryos and adult mice (Figure 3-1h) to further reduce the number of genes whose transcriptional differences are not related to GSIS maturation. Remarkably, the gene expression profiles of functionally mature and immature β cells tested this way is very similar (Figure 3-1i).

Various molecular mechanisms have been proposed to explain the ineffective GSIS observed in fetal and neonatal β cells as compared to adult β cells. These include insensitivity of the ATP-regulated K^+ channel[5,6], reduced expression of glucose transporters[6], low activity of glucokinase[7] or low levels of the β cell selective gap junction protein Connexin36[8]. Recently, genetic ablation of the transcription factors NeuroD1 in adult mouse β cells[9] or the combined deletion of Foxa1 and Foxa2[10] resulted in β cells with an immature-like GSIS phenotype. We thus first assessed known β cell genes whose expression levels could explain the functional difference between mature and immature β cells. We examined expression levels of β cell transcription factors (Pdx1, Nkx2.2, Nkx6.1, NeuroD1, Foxa1, Foxa2, MafA, MafB and Hnf4a), key proteins involved in glucose sensing and insulin secretion (Glucokinase, Glut2, Cav6.1, Kir6.1, Sur1, Pcsk1 and Pcsk2), the β cell-selective gap junction gene Connexin36 and the genes for Insulin1 and Insulin2. We

also looked at tissue specific glucose transporters (Glut 1, 3 and 4) and hexokinases (Hexokinase 1, 2 and 3) (Figure 3-1j). The RNA expression levels of most of these genes does not change significantly between immature and mature cells (or change expression by less than two-fold making them unsuitable for on/off detection of mature β cells). One exception is the transcription factor MafB which is expressed at 2.5-fold higher levels in immature β cells, consistent with previous reports[11].

We next examined all genes for which expression changes by more than two-fold between immature and mature cells. We excluded genes for which a significant change in expression also occurred between E18.5 and P1 or P10 and adult, thereby focusing on genes that change expression specifically within the time window of β cell maturation (groups I and II in Figure 3-1h). We found 71 genes (81 probes) that were up-regulated and 66 genes (72 probes) that were down-regulated during β cell maturation. Of the former group, 36 genes were acinar-related genes which is best explained by the rapid expansion of exocrine tissue at this stage, thereby increasing the probability of a small acinar cell contamination during FACS sorting and resulting in the misleading indication that acinar genes are up-regulated. We chose 16 genes (Table 3-1) for which β cell expression had previously been reported and analyzed their protein expression levels using western blotting and immunohistochemistry. From all these analyses, one strong candidate emerged: the gene Urocortin 3 (Ucn3).

Table 3-1. List of genes differentially expressed between immature (E18.5 and P1) and mature (P10 and adult) β cells.

Gene symbol	E18.5	P1	P10	Adult	Fold Change
Klk6	57	74	564	829	12.3
Angptl7	68	105	383	1043	8.8
Ucn3	127	154	923	1234	7.2
Syt4	80	62	252	734	6.9
Rgs9	133	107	361	456	3.1
Dlk1	2028	2051	126	39	-23.3
Npy	3464	2001	235	77	-17.8
Ghrl	1199	643	77	40	-17.1
Nnat	1677	1286	257	69	-8.8
Slc38a5	7037	5761	1181	421	-8.0
Lgi2	823	833	131	166	-5.8
Nxf	525	510	89	94	-5.7
Egr4	514	434	130	46	-5.3
Cryba2	4385	4145	1153	624	-4.8
Arc	321	288	90	46	-4.6
Atf3	198	115	47	43	-4.4

Numbers represent normalized expression levels in the microarray.

The levels of Ucn3 mRNA increase more than 7-fold between immature and mature β cells, and nearly 10-fold between E18.5 and adult (Figure 3-1k and Table 3-1). Immunofluorescence staining showed that Ucn3 is highly expressed in all adult β cells, but is undetectable in islets from E18.5 embryos (Figure 3-1l-m). As with insulin, the signal intensity of Ucn3 protein varies from cell to cell in the adult islet. This variation does not correlate with the variation in insulin intensity as cells that show high staining intensity for insulin show both high and low staining intensities for Ucn3, and vice versa. No co-

localization of Ucn3 with glucagon, somatostatin or pancreatic polypeptide was observed (Figure 3-2).

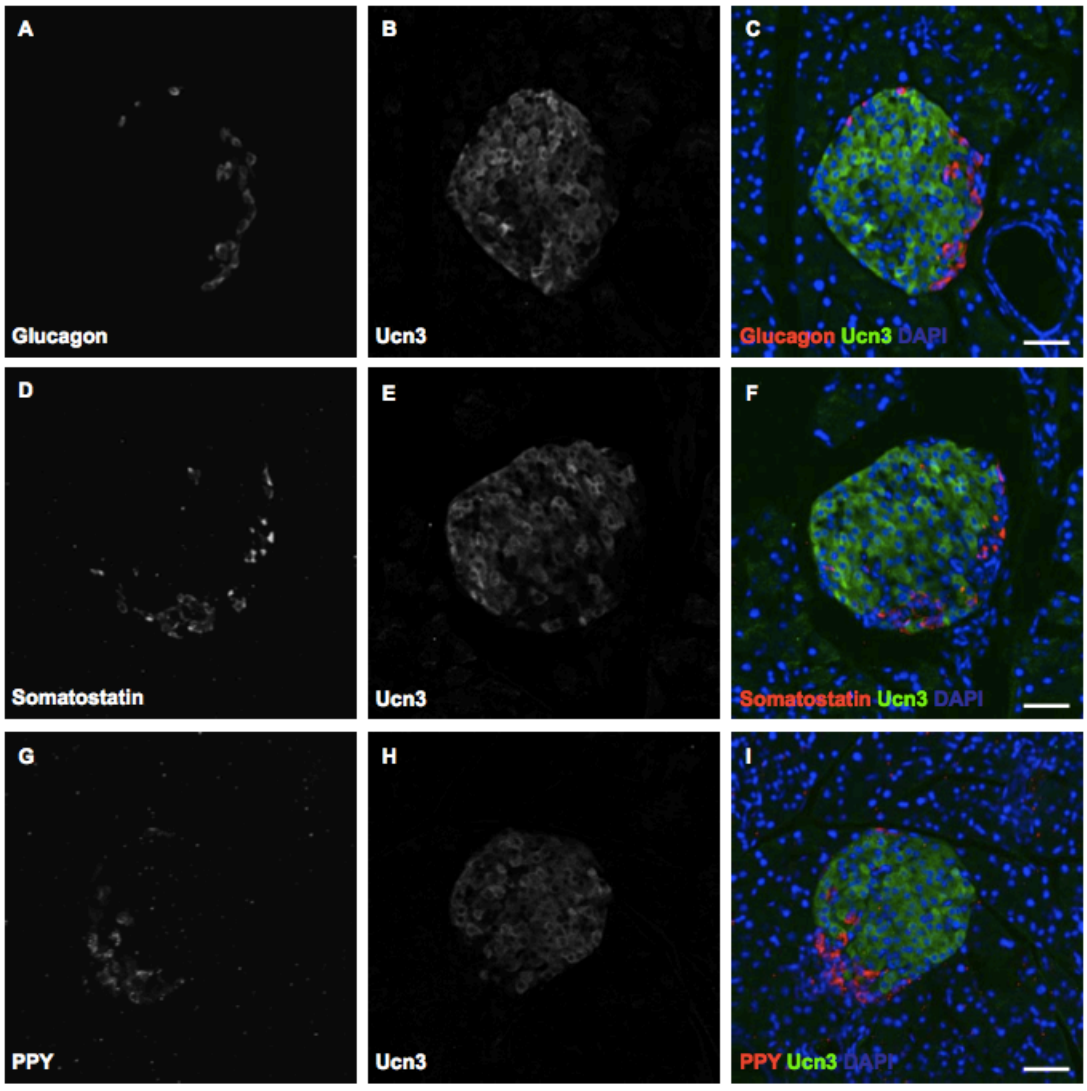


Figure 3-2. Ucn3 expression in mouse islets is restricted to β cells.

(A-C) Confocal images showing immunostaining of Ucn3 (green) and glucagon (red) on adult mouse pancreatic sections. **(D-F)** Ucn3 (green) and somatostatin (red) **(G-I)** Ucn3 (green) and pancreatic polypeptide (PPY, red). Nuclei are stained with DAPI (blue). Scale bars = 50μm. No co-localization of Ucn3 is seen with any of the islet hormones (other than insulin – see Fig 2 and 3).

Ucn3 is a secreted protein expressed in regions of the brain and in the pancreas, and was reported to be exclusively expressed by β cells but not other endocrine cells in the islet[12]. Li *et al.*[12,13] found that secretion of Ucn3 from β cells is induced by high glucose in adult mice, and that the gene has a positive effect on GSIS at high glucose concentrations. We tried to mature fetal β cells *in vitro* by culturing them in the presence of recombinant Ucn3 protein for several days, but did not observe any effect of the recombinant protein on the GSIS profile of the cells, suggesting that Ucn3 by itself can not induce β cell maturation (data not shown). It remains to be determined whether Ucn3 knockout mice have β cell maturation defects.

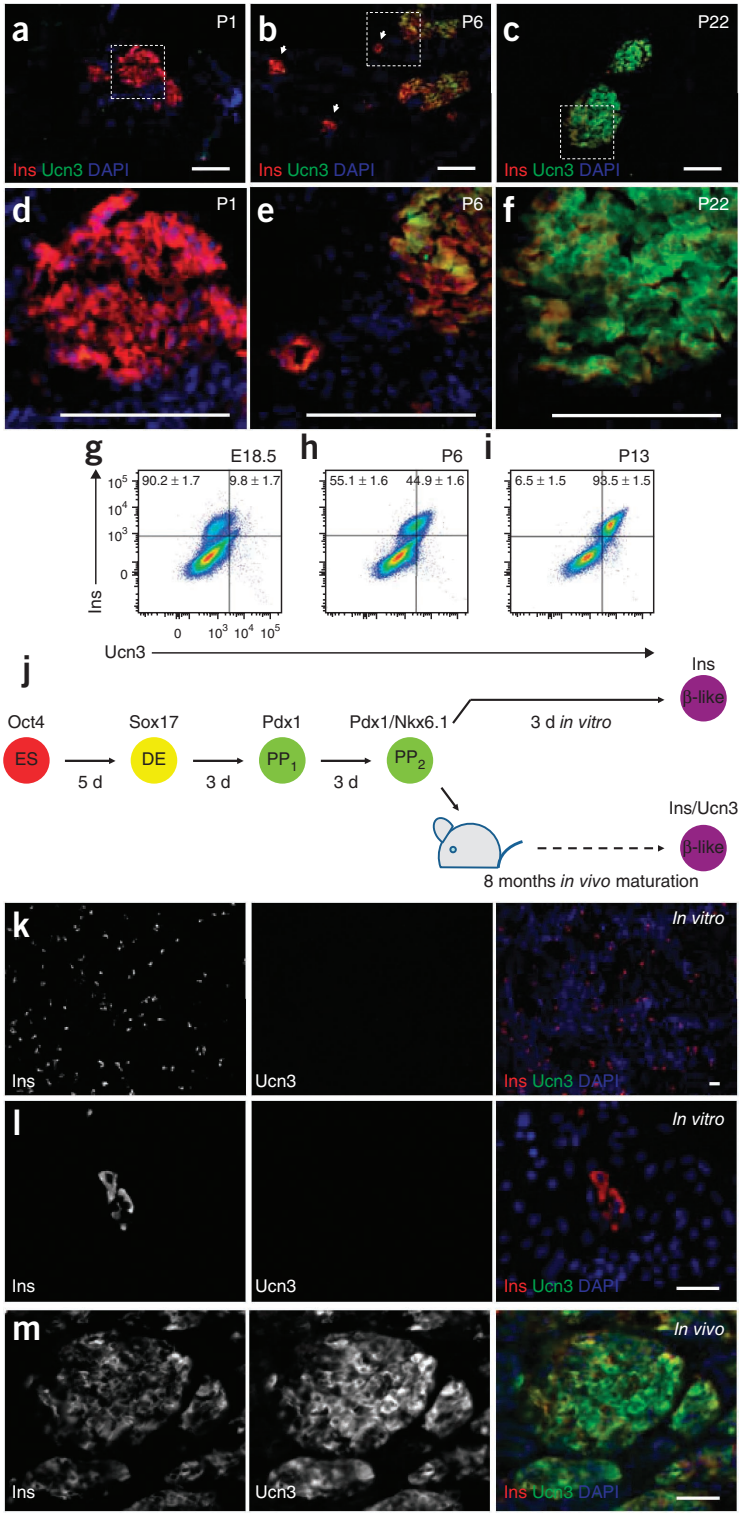
We next examined the patterns of Ucn3 expression at additional time points during the period of β cell maturation. Ucn3 protein was not detected in any islets of P1 pups (Figure 3-3a and d). At P6, Ucn3 expression is found primarily in large islets, not in small β cell aggregates (Figure 3-3b and e). By P22, Ucn3 protein is strongly detected in all β cells (Figure 3-3c and f). Intra-cellular FACS analysis with antibodies against Ucn3 and insulin was performed to quantify the percentage and levels of Ucn3 expression in β cells. At E18.5, 90.2±1.7% of β cells express insulin alone while 9.8±1.7% also stain weakly for Ucn3 (Figure 3-3g). The low expression level of Ucn3 in the small population detected by FACS at this age is probably too low to be detected by conventional immunofluorescence on tissue sections. At P6, 55.1±1.6% of β cells are either negative for Ucn3 or express low

levels of the protein, while 44.9±1.6% of β cells express high levels of both Ucn3 and insulin (Figure 3-3h). By P13, just at the end of the maturation window, 93.5±1.5% of the β cells express high levels of Ucn3 and only 6.5±1.5% express insulin alone (Figure 3-3i). The increase in Ucn3 in β cells during maturation is gradual, as can be seen by the shift in the mean Ucn3 signal intensity (Figure 3-4a). The signal intensity of insulin is unchanged, indicating that expression of insulin protein remains constant throughout this time period (Figure 3-4b). This mixed pattern of Ucn3 expression may explain why a "marginally mature" phenotype was observed between P2 and P9.

Figure 3-3. Ucn3 expression gradually increases during the course of mouse β cell maturation *in vivo* and is expressed in hESC-derived β-like cells after differentiation and maturation *in vivo*, but not after differentiation *in vitro*.

(a-c) Immunostaining of Ucn3 (green) and insulin (red) on pancreata from P1, P6 and P22 mice. **(d-f)** Enlargement insets shown in A-C, respectively. Nuclei are stained with DAPI (blue). Scale bars = 50μm. **(a, d)** Ucn3 in not detected at P1 even in large islets. **(b, e)** At P6, some large islets express Ucn3, but small aggregates do not express the peptide (arrows). **(d, f)** At P22, Ucn3 is highly expressed in all islets. **(g-i)** Intra-cellular FACS analysis of insulin and Ucn3 at E18.5, P6 and P13. Numbers in upper quadrants represent the percentage of insulin only (left) or insulin and Ucn3 co-expressing cells (right) of all insulin-expressing cells (two upper quadrants), calculated as average±sem of three independent biological repeats (three separate litters) for each age group. **(j)** An outline of the experimental approach on hESCs differentiation. hESCs (ES, red) marked by Oct4 were differentiated *in vitro* into definitive endoderm (DE, yellow) marked by Sox17 and subsequently to pancreatic progenitors (PP, green), marked by the expression of Pdx1 and NKX6.1. The cells were transplanted into SCID-beige mice to complete maturation *in vivo*. **(k, l)** Immunostaining for Ucn3 (green) and insulin (red) on the *in vitro* differentiated cells shown at two magnifications (k, low magnification; L, high magnification). *In vivo* differentiated (transplanted) cells are shown in **(k)**. Nuclei are stained with DAPI (blue). Scale bars = 50μm. Ucn3 is expressed in the *in vivo* matured cells, but not in *in vitro* differentiated insulin-positive β-like cells.

Figure 3-3 Continued



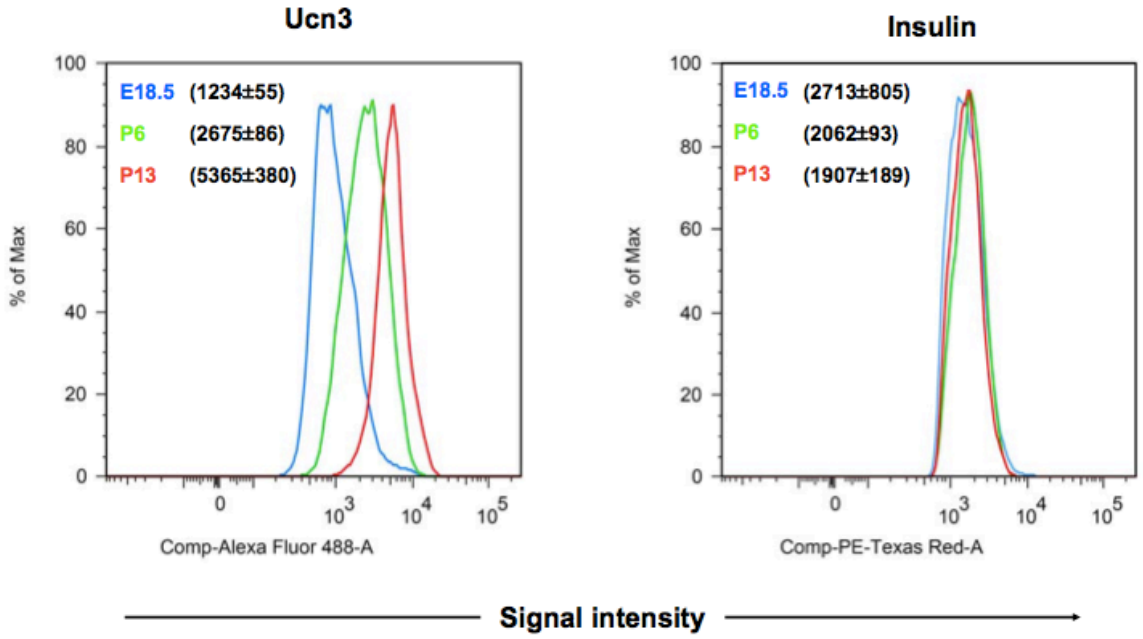


Figure 3-4. Ucn3 expression levels increase gradually in all β cells during maturation, whereas insulin content stays constant.

Intra-cellular FACS analysis of insulin and Ucn3 in E18.5 (blue), P6 (green) and P13 (red). Histograms present the signal intensity of Ucn3 (A) and insulin (B) plotted against the percentage of all insulin expressing cells. Numbers in brackets show mean intensity ± sem of three independent biological repeats (three separate litters) for each age group.

Finally, we wished to determine if Ucn3 could serve as a marker for functionally mature β cells derived from human pluripotent stem cells (HPSCs). Immunoassaying with antibodies against Ucn3 on pancreatic sections obtained from an adult human donor revealed that the gene is expressed by all insulin-positive β cells, and is excluded from glucagon-expressing α cells. A small fraction of somatostatin- and PPT-expressing cells also express Ucn3 (Figure 3-5).

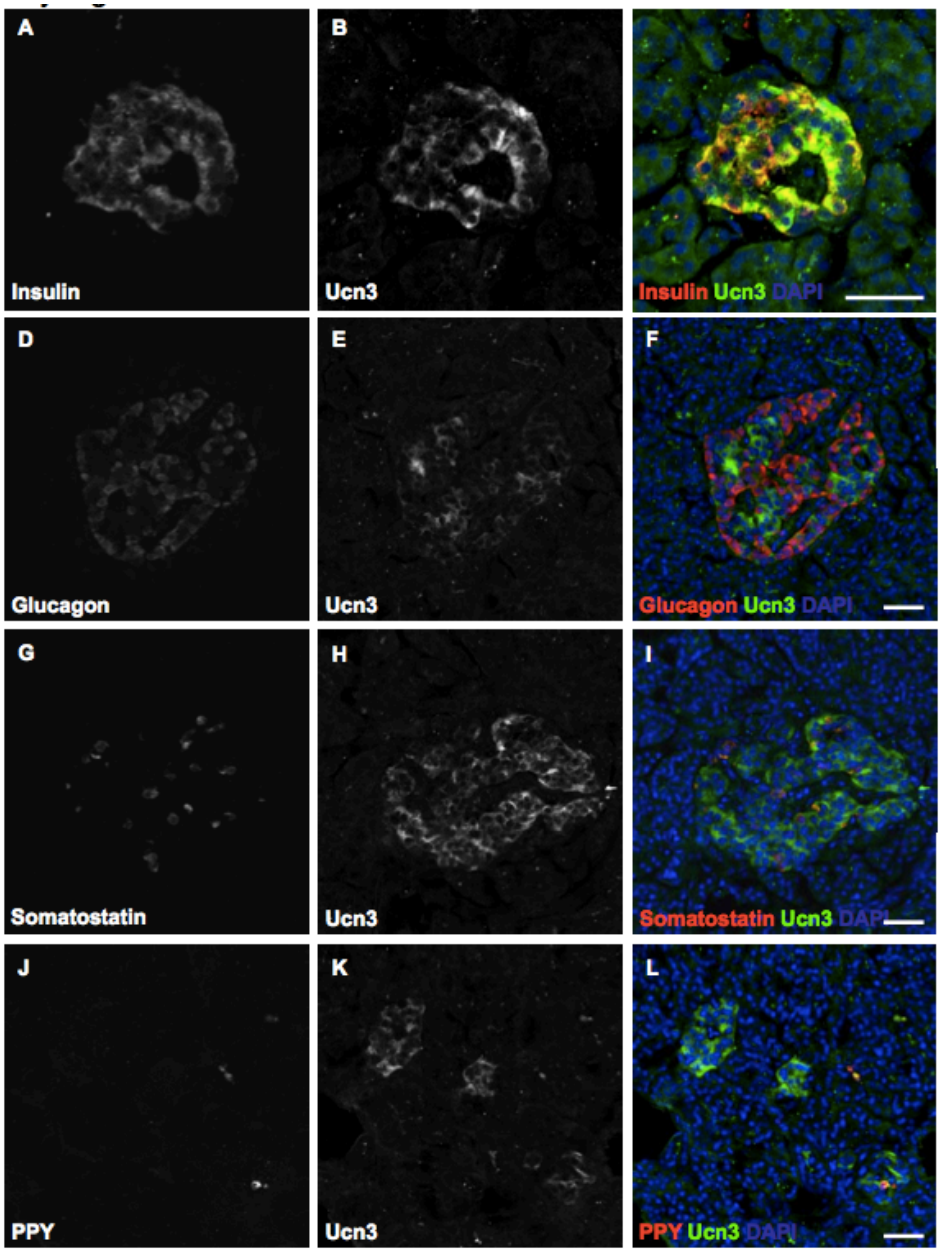


Figure 3-5. Ucn3 expression in human pancreas

(A-C) Confocal images showing immunostaining of Ucn3 (green) and insulin (red) on pancreatic sections from an adult human. **(D-F)** Ucn3 (green) and glucagon (red)**. (G-I)** Ucn3 (green) and somatostatin (red) **(J-L)** Ucn3 (green) and pancreatic polypeptide (PPY, red). Nuclei are stained with DAPI (blue). Scale bars = 50μm.

To see whether Ucn3 is induced during the maturation of human ESC-derived β cells following transplantation, human embryonic stem cells were differentiated using a 4-step protocol to Pdx1 and NKX6.1-positive pancreatic progenitors[14]. These cells were then differentiated *in vitro* for 3 more days to become insulin-positive β-like cells (see Figure 3-3j and material and methods for details). Separately, stage 4 clusters of Pdx^+ $Nkx6.1^+$ pancreatic progenitors, containing a few insulin-positive β-like cells, were transplanted to the kidney capsule of SCID-beige mice where they differentiate further and mature *in vivo* (Figure 3-3j). A glucose tolerance test, performed on transplanted animals, showed fasting human C-peptide levels above background 12 weeks after transplantation (Figure 3-6). Despite high variation in fasting human C-peptide between the transplanted mice, all but one animal (6/7) showed an increase in blood human C-peptide between 1.7-fold to 7.6-fold (average 2.8±0.9-fold), demonstrating that the transplanted human embryonic stem cell- (hESC)-derived cells matured to glucose-responsive β cells. Immunostaining showed that while the *in vitro* differentiated β-like cells express insulin, they are negative for Ucn3 staining (Figure 3-3k-l). Conversely, the *in vivo* matured cells stained positive for both insulin and Ucn3 proteins (Figure 3-3m). This expression of human Ucn3 in the transplant is exclusive to the β cells; the Ucn3 protein is not detected in any glucagon-, somatostatin- or PPY-expressing cells (Figure 3-7).

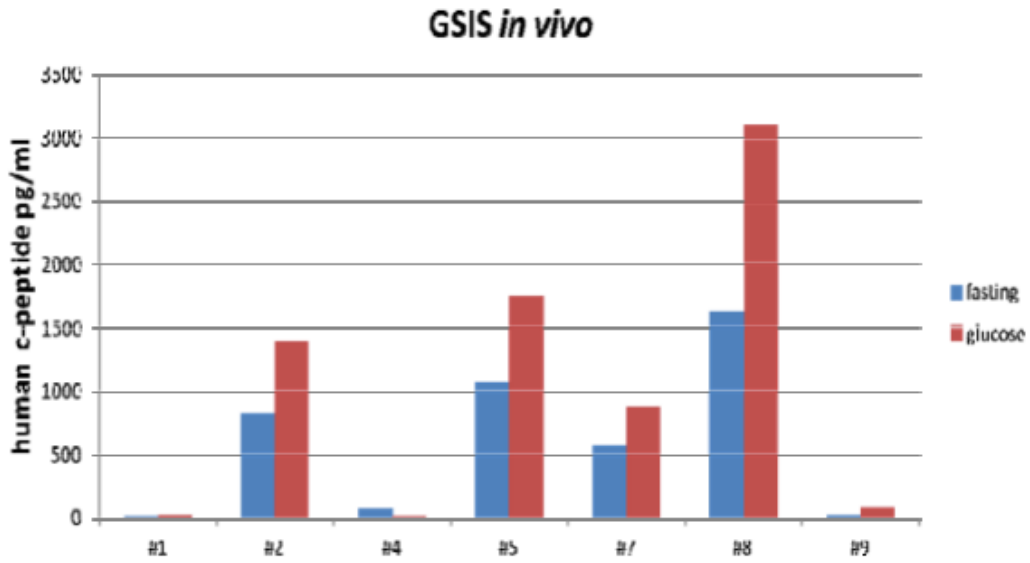


Figure 3-6. hESC-derived β cells secrete human C-peptide in response to glucose challenge

Mice transplanted with 5 million hESC-derived pancreatic progenitors were fasted 12 weeks after transplantation over night and injected with 2mg/kg glucose. The levels of human C-peptide before (fasting, blue) and one hour after glucose administration (glucose, red) are shown. Despite variable basal levels of human C-peptide, all animals except mouse #4, showed glucose-stimulated secretion of human C-peptide.

In summary, we propose an operational definition for mature β cells based on changing glucose thresholds for GSIS response during development. We also identify a molecular marker, Ucn3, that distinguishes mature from immature β cells. Notably, we find that Ucn3 is induced in hESC-derived β cells following maturation *in vivo*. High-throughput screening can now utilize the difference in GSIS and the expression of Ucn3 as benchmarks in studies aimed at finding conditions to induce functional β cell maturation *in vitro*.

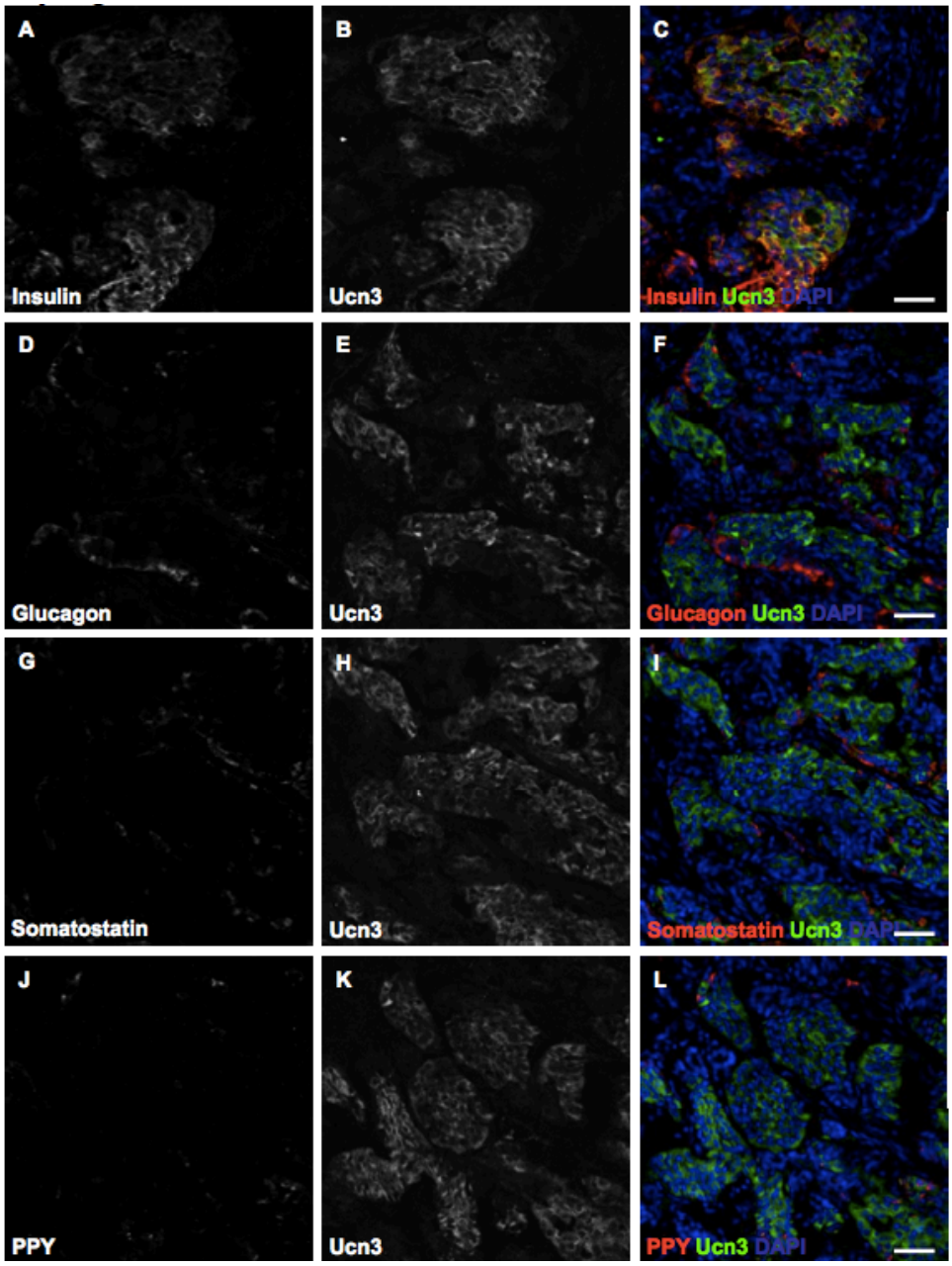


Figure 3-7. Ucn3 expression in hESC-derived β cells after maturation *in vivo*

(A-C) Confocal images showing immunostaining of Ucn3 (green) and insulin (red) on hESC-derived graft 8 months post transplantation. **(D-F)** Ucn3 (green) and glucagon (red)**. (G-I)** Ucn3 (green) and somatostatin (red) **(J-L)** Ucn3 (green) and pancreatic polypeptide (PPY, red). Nuclei are stained with DAPI (blue). Scale bars = 50μm.

3.3. Materials and methods

3.3.1. Animal experiments and islet isolation

All animal experiments were performed in compliance with the Harvard University International Animal Care and Use Committee (IACUC) guidelines. Mouse strains used in this study were ICR, Swiss-Webster wild type and Pdx1–GFP transgenic and SCID- beige mice. Blood glucose levels were measured using OneTouch Ultra2 glucometer (LifeScan). Blood insulin levels were measured with an Ultrasensitive Insulin ELISA kit (Alpco). For glucose tolerance test, animals were fasted over-night and blood was taken form tail tips before, and 1 hour after, injection of 2gr/kg body weight glucose. Human C- peptide levels were measured using Human C-peptide ELISA kit (Alpco). For islet isolation, adult pancreata were perfused through the common bile duct with 0.8mM Collagenase P (Roche) and fetal and neonatal pancreata were dissected wholly without perfusion. Pancreata were digested with 0.8mM Collagenase P (Roche) and purified by centrifugation in Histopaque gradient (Sigma).

3.3.2. Glucose stimulated insulin secretion (GSIS) assays

Isolated islets were recovered over night in islet media (DMEM containing 1gr/L glucose, 10% v/v FBS, 0.1% v/v Penicillin/Streptomycin). Islets were picked manually under a fluorescent dissecting microscope according to their GFP fluorescence. Care was taken to pick islets of approximately the same size from all ages. For dynamic GSIS, approximately 50 islets were hand picked and assayed on a fully automated Perifusion

System (BioRep). Chambers were sequentially perfused with 0.5mM, 2.8mM or 16.7mM glucose in KRB buffer (128mM NaCl, 5mM KCl, 2.7mM CaCl2, 1.2mM MgCl2, 1mM Na2HPO4, 1.2mM KH2PO4, 5mM NaHCO3 10 HEPES, 0.1% BSA) at a flow rate of 0.1ml/min. Fractions were collected and kept at -80°C until analysis. For static GSIS assays, approximately 10 islets were hand picked, incubated for 2 hours in KRB buffer at 37°C, 5%CO2, and then incubated for 75min with 2.8mM or 16.7mM glucose in the same conditions. Insulin concentrations in the supernatant were determined using Ultrasensitive Insulin ELISA kit (Alpco). Analysis of the results was done using Matlab software.

3.3.3. Electron microscopy

Samples were fixed with 4% paraformaldehyde and 0.1% glutaraldehyde for 2 h at room temperature and further refixed with a mixture of 1% osmiumtetroxide (OsO4) plus 1.5% potassium ferrocyanide (KFeCN6) for 2 h, washed in water and stained in 1% aqueous uranyl acetate for 1 h followed by dehydration in grades of alcohol (50%, 70%, 95%, 2 × 100%) and propyleneoxide (1 h). Samples were then infiltrated in propyleneoxide:Epon 1:1 overnight and embedded in TAAB Epon (Marivac Canada Inc.). Ultrathin sections (about 60–80 nm) were cut on a Reichert Ultracut-S microtome,

picked up on to copper grids, stained with 0.2% lead citrate and examined in a Tecnai G2 Spirit BioTWIN transmission electron microscope. Images were taken with an AMT CCD camera. The number of insulin vesicles and cell area were determined using ImageJ software.

3.3.4. Microarray analysis

Islets were isolated as above from heterozygous Pdx1–GFP (crossed with ICR) animals and further dissociated into single cells with 0. 25% Trypsin-EDTA (Invitrogen). GFP^+ cells were isolated using FACSaria (BD Bioscience). Total RNA was extracted with RNeasy RNA extraction kit (Qiagen). Biotinylated cRNA was prepared from ≥ 100 ng of isolated RNA using Illumina TotalPrep RNA Amplification Kit (Ambion) and hybridized to the Illumina mouse genome Bead Chips (MouseRef8). All samples were prepared as four biological replicates. Data were acquired with Illumina Beadstation 500 and were evaluated using BeadStudio Data Analysis Software (Illumina).

3.3.5. Immunohistochemistry and FACS analyses

For immunohistochemistry, pancreata were fixed by immersion in 4% paraformaldehyde overnight at 4 °C. Samples were washed with PBS, incubated in 30% sucrose solution overnight and embedded with optimal cutting temperature compound (Tissue-Tek). 10μm sections were blocked with 10% donkey serum (Jackson Immunoresearch) in PBS/0.1% Triton X and incubated with primary antibodies overnight at 4°C. Secondary antibodies were incubated for 1 hr at room temperature. The following primary antibodies and dilutions were used: rabbit anti-mouse or anti-human Ucn3 (1:600-1:800, both from Phoenix Pharmaceuticals), rabbit anti-human Ucn3 (1:600, a gift from Dr. Wylie Vale, Salk Institute), Guinea Pig anti-insulin (1:800, DAKO), Guinea Pig anti-glucagon (1:200, Linco), Goat anti-Somatostatin (1:200, Santa Cruz) and Goat anti-PPY

(1:200, Novus). Secondary antibodies were: Alexa Fluor 488 donkey anti-rabbit (1:400, Invitrogen), Alexa Fluor 647 donkey anti- goat (1:400, Invitrogen) and DyLight 649 donkey anti-guinea pig (1:400, Jackson Immunoresearch). Nuclei were visualized with DAPI. It is important to note that, in our hands, anti-human immunostaining was successful only on human tissues fixed over-night with 4% paraformaldehyde directly after surgery. Efforts to use the abovementioned anti-human Ucn3 antibodies on flash- frozen-unfixed cryosections, acetone-fixed cryosections or formalin-fixed-paraffin- embedded samples resulted in either close-to-background or non-specific staining. Images were taken using an Olympus IX51 Microscope or Zeiss LSC 700 confocal microscope. For intra-cellular FACS analysis, islets were isolated as above and further dissociated into single cells with 0. 25% Trypsin-EDTA (Invitrogen). The cells were then fixed with Cytofix/Cytoprem solution (BD Biosciences) at 4°C for 30min, washed once with Perm Wash Buffer (BD Biosciences), and stained with Guinea Pig anti-insulin (1:800, DAKO) and rabbit anti-Ucn3 (1:600, Phoenix Pharmaceuticals) for 1 hour at room temperature. The cell were then washed once with Perm Wash Buffer (BD Biosciences), incubated with TexasRed donkey anti-guinea pig (1:400, Jackson Immunoresearch) and Alexa Fluor 488 donkey anti-rabbit (1:400, Invitrogen) for 45 min at room temperature, washed with PBS, filtered through a nylon mash, and analyzed LSR-II FACS machine (BD Biosciences). Analysis of the results was done using FlowJo software.

3.3.6. hESC Culture and Differentiation

Human ESCs (WA1) were cultured on Matrigel (BD Biosciences) in mouse embryonic fibroblast conditioned media (MEF-CM). MEF-CM media was produced by conditioning media for 24 days on a confluent layer of mouse embryonic fibroblasts and subsequently adding 20 ng/ml bFGF (Invitrogen). The media was composed of DMEM/F12 (GIBCO) media supplemented with 20% KnockOut Serum Replacement (GIBCO), 2mM L- glutamine (L-Glu, GIBCO), 1.1mM 2-mercaptoethanol (GIBCO), 1mM nonessential amino acids (GIBCO), 1x penicillin/streptomycin (GIBCO). Cells were passaged at the ratio of 1:6–1:20 every 4-7 days using TrypLE Express (Invitrogen). To initiate differentiation, the cells were cultured as previously described1 onto 1:30 dilution of growth factor reduced matrigel (BD Biosciences) in MEF-CM. Two to three days following seeding the differentiation was initiated as follows: cells were exposed to RPMI 1640 (Invitrogen) supplemented with 0.2% fetal bovine serum (FBS) (Hyclone, Utah), 100 ng/mL activin-A (AA; Pepro-tech; Rocky Hill, NJ), and 20 ng/mL of Wnt3A (R&D Systems) for day one only. For days 2-3, cells were cultured in RPMI with 0.5% FBS and 100 ng/mL AA (stage 1). During days 4-5 cells were treated with DMEM-F12 medium containing 2% FBS and 50ng/ml FGF7 (Peprotech) (stage 2). For days 6-9 cells were treated with DMEM-HG (Invitrogen), 1% (v/v) B27 (Invitrogen), 2uM RA (Sigma), 0.25uM SANT-1 (Sigma), and 100ng/ml rhNoggin (R&D Systems) (stage 3). During days 10-13 cells were treated with DMEM-HG (Invitrogen) + 1% (v/v) B27 (Invitrogen), 100ng/ml rhNoggin (R&D Systems), 50nM TPB (PKC activator, EMD Biosciences), and 1 μM ALK5 inhibitor II (Axxora, San Diego, CA) (stage 4). On day 14, cells were treated

with 5 mg/mL Dispase for 5 min, followed by gentle pipetting to mix and break the cell clumps into small clusters (< 100 micron). The cell clusters were cultured for one day in a 125 ml Spinner Flask (Corning) at 80-100 rpm overnight with DMEM-HG supplemented with 1 μM ALK5 inhibitor II, 100 ng/mL of Noggin and 1% B27.

For transplantation into mice, 10 million cells in clusters were transplanted under the kidney capsule of SCID-Bg mice (Jackson Laboratory). 8 months following transplant the graft was surgically extracted from under the mouse kidney capsule, fixed in 4% paraformaldehyde (PFA, Sigma), equilibrated in 30% sucrose, embedded in O.C.T., cryopreserved and sectioned.

3.4. References

1 Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nat Biotechnol* **26**, 443-452 (2008).

2 Pan, F. C. & Wright, C. Pancreas organogenesis: from bud to plexus to gland. *Dev Dyn* **240**, 530-565.

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

4 Rozzo, A., Meneghel-Rozzo, T., Delakorda, S. L., Yang, S. B. & Rupnik, M. Exocytosis of insulin: *in vivo* maturation of mouse endocrine pancreas. *Ann N Y Acad Sci* **1152**, 53-62 (2009).

5 Rorsman, P. *et al.* Failure of glucose to elicit a normal secretory response in fetal pancreatic beta cells results from glucose insensitivity of the ATP-regulated K+ channels. *Proc Natl Acad Sci U S A* **86**, 4505-4509 (1989).

6 Richardson, C. C. *et al.* Low levels of glucose transporters and K+ATP channels in human pancreatic beta cells early in development. *Diabetologia* **50**, 1000-1005 (2007).

7 Taniguchi, S., Tanigawa, K. & Miwa, I. Immaturity of glucose-induced insulin secretion in fetal rat islets is due to low glucokinase activity. *Horm Metab Res* **32**, 97-102 (2000).

8 Carvalho, C. P. *et al.* Beta cell coupling and connexin expression change during the functional maturation of rat pancreatic islets. *Diabetologia* **53**, 1428-1437.

9 Gu, C. *et al.* Pancreatic beta cells require NeuroD to achieve and maintain functional maturity. *Cell Metab* **11**, 298-310 (2010).

10 Gao, N. *et al.* Foxa1 and Foxa2 maintain the metabolic and secretory features of the mature beta-cell. *Mol Endocrinol* **24**, 1594-1604 (2010).

11 Artner, I. *et al.* MafA and MafB regulate genes critical to beta-cells in a unique temporal manner. *Diabetes* **59**, 2530-2539 (2010).

12 Li, C. *et al.* Urocortin III is expressed in pancreatic beta-cells and stimulates insulin and glucagon secretion. *Endocrinology* **144**, 3216-3224 (2003).

13 Li, C., Chen, P., Vaughan, J., Lee, K. F. & Vale, W. Urocortin 3 regulates glucose-stimulated insulin secretion and energy homeostasis. *Proc Natl Acad Sci U S A* **104**, 4206-4211 (2007).

14 Rezania, A. *et al.* Production of functional glucagon-secreting alpha-cells from human embryonic stem cells. *Diabetes* **60**, 239-247.

Chapter 4

Differentiated human embryonic stem cells resemble fetal, not adult β cells

4.1. Abstract

Human pluripotent stem cells (hPSCs) have the potential to generate unlimited amounts of any human cell type. Direct comparisons between differentiated cell types produced *in vitro*, and their counterparts produced during human development *in vivo*, is an essential step towards validating hPSC-derived cells. In order to compare insulin-expressing (INS^+) cells produced during human development and by directed differentiation of hPSCs, we developed a Method for Analyzing RNA following Intracellular Sorting (MARIS). Genome-wide transcriptional analysis of sorted INS^+ cells derived from three independent hPSC-lines suggest that despite variation in efficiency, different lines produce highly similar INS^+ cells. Furthermore, *in vitro* derived INS^+ cells resemble human fetal β cells, which are distinct from adult β cells. This study presents the first purification and genome-wide profiling of a human fetal and adult cell type for which robust cell surface markers do not exist, and a direct comparison of normal human cells with cells produced from hPSCs in culture. Our data reveal differences in gene expression that may account for the functional differences between hPSC-derived INS^+ cells and true β cells. MARIS may be applicable for transcriptional profiling of any cell type that can be sorted based on intracellular immunofluorescent antigen staining.

4.2. Introduction

Human pluripotent stem cells (hPSCs), including human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs), are characterized by their capacity for unlimited self-renewal and their ability to differentiate into any human cell type [1-4]. Stepwise differentiation protocols, designed to mimic sequential developmental signals, attempt to generate specific cell types from hPSC-lines for use in transplantation therapy and disease modeling [5-7].

Significant variation in differentiation efficiency has been observed between different hPSC-lines, with some lines more readily differentiating into a particular cell type than others [8-11]. This may be attributable in part to variations in gene expression and DNA methylation patterns that impact the efficiency of differentiation[12]. Long-term *in vitro* culture may result in the accumulation of genetic mutations, translocations and chromosomal abnormalities, which may also affect the differentiation propensity [13]. Additionally, hiPSCs retain epigenetic memory of their previous mature cells state, undergo slow erosion of X-chromosome inactivation and often contain partially silenced viral integrations as byproducts of the reprogramming process [14-18]. Due to these differences in differentiation propensity, directed differentiation protocols often require laborious optimization for specific hPSC-lines [19].

Large variation during directed differentiation poses the question of the degree to which cells of a desired type produced using different protocols and different cell lines resemble each other at the transcriptome level. Significant variation in gene expression

between supposedly identical cell-types generated from different hPSC-lines would bring into question the robustness of directed differentiation and its applications in disease modeling and cell therapy. To date, however, experiments addressing this question have been nearly impossible due to the scarcity of reporter cell lines that facilitate the purification and analysis of desired cell populations.

In addition to questions regarding the similarity of differentiation products produced from different cell lines, the extent to which any differentiated cell produced *in vitro* resembles its counterpart produced during normal human development remains unknown. Directed differentiation protocols are often generated using mouse embryonic development as a guide, due to the relative paucity of information about human development [20]. However, it is well established that significant differences in transcriptional regulation exist between these two species [21,22]. Understanding of transcriptional changes during human development would greatly benefit directed differentiation efforts to produce mature human cells *in vitro*. However, in the absence of established cell surface markers, it near impossible to isolate and purify most human fetal and adult cell types

Efforts to produce functional pancreatic β cells from hPSCs for use in disease modeling and cell transplantation therapy have been significantly hampered by variability between cell lines and an incomplete knowledge of the relevant human developmental biology. Current directed differentiation protocols generate INS^+ cells (hPSC-INS^+) that lack expression of several key β cell genes and fail to secrete insulin in response to glucose [20,23-33]. It is currently unknown whether these cells are a culture artifact or an *in vitro*

equivalent of human fetal β cells. In the absence of tools to sort and analyze INS$^+$ cells from directed differentiation cultures and human pancreata at various developmental time-points, these two hypotheses cannot be resolved.

To address the robustness of hPSC-differentiation and its relation to human development, we sought to develop the tools to isolate high-quality RNA from purified INS$^+$ cells, derived from different genetically unmodified hPSC-lines as well as from human fetal and adult pancreata. Previously, RNA of sufficient quality for FISH, nuclease protection assays and microarray analysis has been obtained following fixation, intracellular immunofluorescent staining, and FACS [34-36]. However, it was unclear whether these relatively harsh manipulations would produce biased results due to crosslinking and partial degradation of RNA. A method to sort cells based on intracellular immunofluorescence and preserve the RNA for whole transcriptome analysis would be of considerable benefit to many biological disciplines.

Here we present a Method for Analyzing RNA following Intracellular Sorting (MARIS). MARIS can obtain RNA of sufficient quality for transcriptome profiling methods such as microarray analysis and RNA-Seq following cellular fixation, intracellular immunofluorescent staining and FACS. Using MARIS, we analyzed the global gene expression profile of sorted INS$^+$ cells from three different hPSC-lines to determine the degree of similarity between INS$^+$ cells derived from different hESC and hiPSC-lines. We also compared hPSC-INS$^+$ cells to human adult and fetal β cells to determine the degree to which directed differentiation generates cell types present during normal human

development. We anticipate MARIS will be of particular utility to the stem cell field for the purification and analysis of hPSC-derived cell populations and their *in vivo* counterparts. Broadly, MARIS can be used for the purification and transcriptional characterization of any cell type based on immunofluorescent detection of a marker gene in the absence of reporter lines and/or sortable cell surface markers.

4.3. Results

4.3.1. RNA isolation from fixed, stained and sorted cells

We developed a protocol for the extraction of high quality RNA from cells sorted based on intracellular immunofluorescence (Figure 4-1a, Materials and Methods). Several assays were used to compare the quality of the RNA isolated using our protocol and RNA isolated from live (fresh, unfixed) cells. hESCs differentiated to the final stage of our pancreatic differentiation protocol (modified from [23]) were used as starting material (Stage 6,Figure 4-2). RNA was extracted from both live cells and cells following fixation, permeabilization, primary/secondary antibody staining and FACS (processed cells). RNA isolated from these two populations achieved RIN scores of 8.1 (live) and 8.0 (processed, Figure 4-1b). The RNA quality was highly reproducible across independent preparations and different cell types with average RIN score 8.2±0.8 (Figure 4-3).

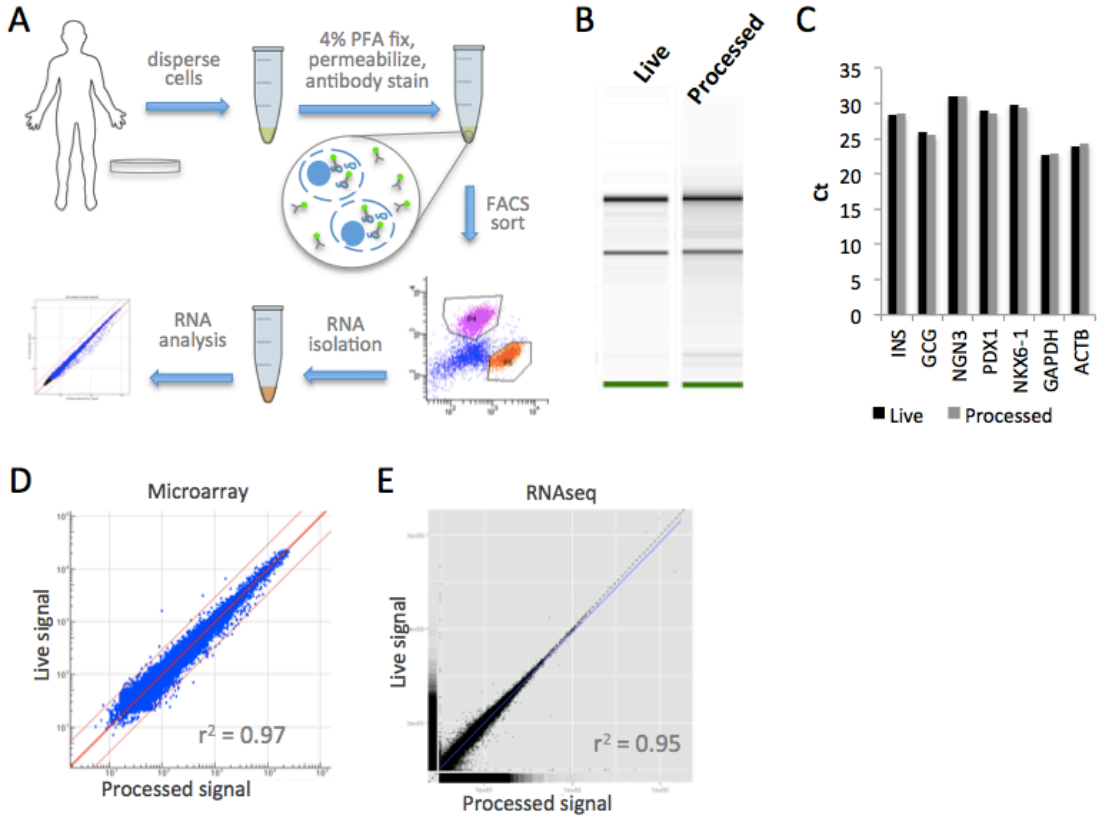


Figure 4-1. High quality RNA isolation and profiling from fixed and stained cells

(a) Outline of the developed protocol. *In vivo* or *in vitro*-derived cells are dispersed, fixed in 4% PFA, permeabilized, stained using standard immunofluorescent antibodies and FACS sorted. Total RNA is isolated using a modified RNA extraction protocol (see methods) and can be analyzed by qRT-PCR, microarrays or RNA-Seq. **(b,c,d,e)** RNA was isolated and analyzed from hESC-derived Stage 6 cells before fixation (live) or following fixation, staining and sorting (processed). **(b)** Bioanalyzer gel plot suggests no degradation of total RNA; RIN value 8.1 for live, 8.0 for processed sample. **(c)** qRT-PCR on live and processed Stage 6 cells for pancreatic and housekeeping genes. **(d)** Logarithmic scatter plot of Illumina microarray data between live and processed samples, r^2=0.97 for all detected probes (detection $p<0.05$). Red lines represent 2-fold change. **(e)** Samples prepared and paired-end sequenced using TruSeq chemistry on a HiSeq 2000 (Illumina). Shown are 100 million reads, logarithmic plot. r^2=0.95.

Stage 1		Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
MCDB-131 0.1% BSA 10mM Glucose	MCDB-131 0.1% BSA 10mM Glucose	MCDB-131 0.1% BSA 10mM Glucose	MCDB-131 2% BSA 8mM Glucose	MCDB-131 2% BSA 8mM Glucose	MCDB-131 2% BSA 8mM Glucose	MCDB-131 2% BSA 8mM Glucose
ActivinA, CHIR99021	ActivinA	FGF7, Vitamin C	FGF7, ActivinA, SANT1, RA, Noggin, ITS-X, Vitamin C	SANT1, PdBu, Noggin, ITS-X, Vitamin C	Noggin, Alk5, ITS-X, Vitamin C	ITS-X, Vitamin C
1d	3d	2d	4d	3d	3d	3d+

Figure 4-2. Directed differentiation protocol

Stepwise differentiation from hESCs to pancreatic endocrine cells. DE, definitive endoderm; PP, pancreatic progenitor; EP, endocrine progenitor; EN, endocrine cells. Table contains reagents used during each stage of directed differentiation.

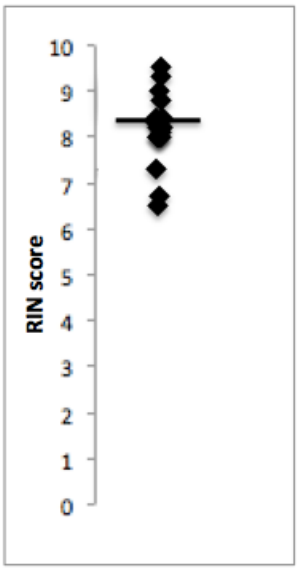


Figure 4-3. RIN scores from multiple experiments

RNA quality from 18 experiments following fixation, staining and FACS. Error bars represent standard deviation.

Having confirmed the integrity of RNA isolated using MARIS, we next assessed whether the protocol changed the representation of individual transcripts. We first performed qRT-PCR analysis of RNA extracted from live and processed Stage 6 cells for several housekeeping genes, as well as genes specific to the pancreatic lineage (Figure 4-1c). There was no systemic difference in cycle threshold values between live and processed cells. Some genes showed slightly lower cycle thresholds while others slightly higher as a result of processing. The largest detected cycle threshold difference was 0.45 cycles corresponding to a 37% change in gene expression.

To evaluate the impact of MARIS on transcript levels at the whole-genome level, RNA from live and processed cells was analyzed using the Illumina microarray platform. Across all detected genes, expression between live and processed samples was nearly identical (r^2=0.97) (Figure 4-1d). Finally, analysis by RNA-Seq showed very similar gene expression between live and processed cells (r^2=0.95, Figure 4-1e). These results confirm our RT-PCR findings and demonstrate that MARIS does not alter the relative abundance of transcripts across the genome.

Because RNA-Seq allows for the examination of RNA representation across whole transcripts, analysis of RNA-Seq results can be used to determine whether the starting RNA had undergone significant degradation prior to analysis. Specifically, degradation of RNA samples results in increased detection of transcripts at the 3' end relative to the 5' end (3'-bias)[37]. To determine whether MARIS resulted in 3'-bias, we plotted the relative distribution of detected transcripts among all annotated genes, grouped by transcript

length (Figure 4-4). In transcripts longer than 2.5knt (% of genome annotations), the processed sample 3' bias was greater than the live sample by roughly 2% (AUC differential). Together, these analyses confirm that MARIS produces high-quality RNA and has little effect on the representation of transcripts as analyzed at the level of the individual gene, or through two methods of genome-wide analysis, making it amenable for accurate gene expression profiling of purified cell populations.

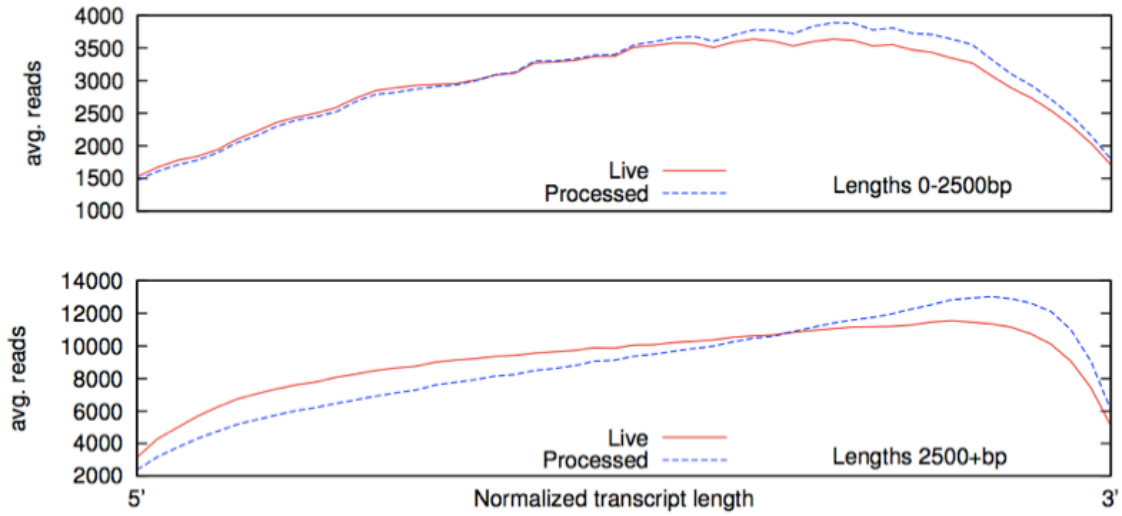


Figure 4-4. Relative RNA-Seq coverage of all annotated transcripts shows 3' bias in longer length genes

Per-bp RNA-Seq read coverage was normalized by total mean log counts and summed over all exonic Ensembl regions. Transcripts were separated according to exonic length and a historgram of average per-transcript coverage is plotted, adjusted for length.

4.3.2. Global transcriptional profile of INS$^+$ cells from several hESC and iPSC lines

Having established MARIS as a robust and reliable way to prepare RNA from fixed, stained cells, we next used it to determine the degree to which hPSC-derived INS$^+$ cells

produced from different cell lines resemble each other at the transcriptome level. We differentiated hESC lines H1 and HUES 8, and hiPSC line iPS-17b to Stage 6, at which point a small percentage of cells expressed insulin and other pancreatic hormones (Figure 4-2a) [2,38]. Stage 6 cells were stained for insulin, glucagon, and somatostatin, and all insulin-positive cells (even those co-expressing other hormones) were sorted for RNA isolation and analysis (Figure 4-5a). INS^+ cells comprised approximately 1% of all Stage 6 cells. A large proportion of INS^+ cells also co-expressed the pancreatic hormones glucagon and somatostatin, consistent with previous reports [20,23]. qRT-PCR for insulin, glucagon and somatostatin indicated significant enrichment of all three endocrine hormones in the sorted populations, confirming successful purification of INS^+ cells (Figure 4-5b).

Next, we analyzed RNA isolated from the hPSC-INS^+ cell populations using the Illumina microarray platform. RNA isolated from live undifferentiated pluripotent cells from each line was included as a control. Hierarchical clustering across all genes identified three distinct, statistically confirmed groups of samples (Figure 4-5c). All INS^+ cells clustered together, suggesting that there were fewer differences between INS^+ cells derived from different cell lines than differences between unsorted Stage 6 cells and sorted INS^+ cells within each cell line. Moreover, expression profiles between INS^+ cells derived from different hPSC lines were as highly correlated as those between hPSC-lines at the pluripotent stage (Figure 4-5d). Together this data suggest that INS^+ cells derived from different hPSC-lines are highly similar.

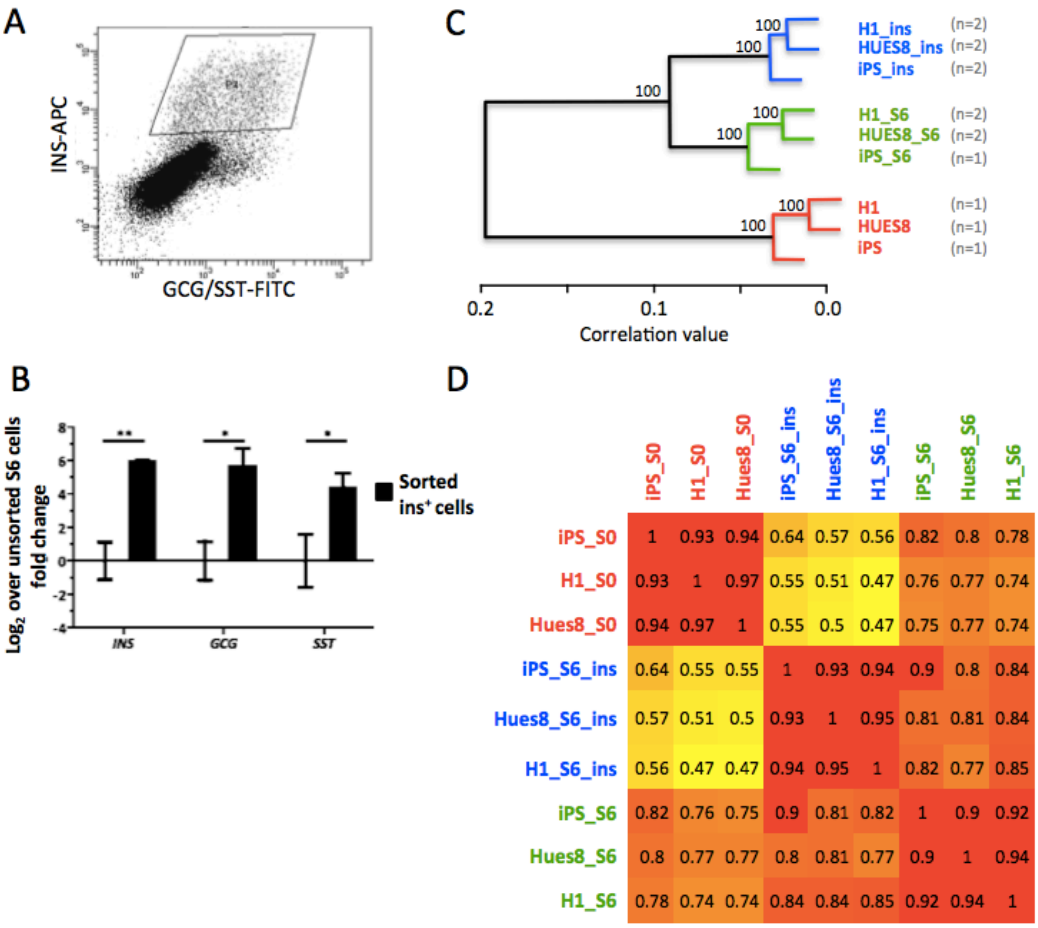


Figure 4-5. RNA profiling of sorted hESC-derived insulin expressing cells

(a) FACS plot of Stage 6 H1-derived cells sorted for insulin-APC. **(b)** qRT-PCR of unsorted and insulin-sorted Stage 6 hESC-derived cells for pancreatic hormone genes INS (insulin), GCG (glucagon), SST (somatostatin) suggests significant enrichment of mRNA specific for pancreatic hormones in the insulin-APC sorted population (* $p<0.05$, ** $p<0.01$). **(c)** Three human pluripotent stem cell lines, HUES8, H1 and iPS-17b were differentiated to stage 6 and sorted for INS^+ cells. RNA was isolated from undifferentiated cells, stage 6 cells and sorted INS^+ cells for all three cell-lines. Global gene expression for each sample was analyzed using the Illumina microarray platform. Numbers indicate biological replicates. Hierarchical clustering identified three major groups of samples. Lengths in the dendrogram represent correlation value. Approximately Unbiased (AU) p-values are displayed. INS^+ cells from different cell lines form a statistically significant cluster. **(d)** R^2 values based on microarray data across all genes is shown. The average r^2 value between stage 0 cells, 0.947, is similar to the average r^2 value between sorted $insulin^+$ cells, 0.94.

To further analyze hPSC-INS$^+$ cells we focused on 152 genes known for their role in pancreatic development, endocrine hormone secretion and glucose metabolism [20,28,39] (Table 4-1). Hierarchical clustering and correlation values based on these 152 genes confirmed a high degree of similarity between hESC-derived and hiPSC-derived INS$^+$ cells (Figure 4-6).

Table 4-1. Gene list, 152 pancreatic lineage gene

Secretory	Transcription factors			Glucose processing, insulin secretion		Membrane channels		Other
CHGA	ALX3	ISL1	ONECUT1	ACOX2	RIMS4	ABCC8	KCNJ11	CDH1
CHGB	ARX	LHX6	ONECUT2	ACSS1	SLC2A14	CACNA1A	KCNJ2	CDH2
GAST	CMIP	LMX1A	ONECUT3	ACSS3	SLC30A8	CACNA1B	KCNJ4	COL1A1
GCG	EN1	LMX1B	PAX2	ALDOA	SNAP25	CACNA1C	KCNJ6	CTNNB1
GHRL	ESR1	MAF	PAX4	ATP5G3	STX1A	CACNA1E	KCNK1	FFAR1
GIP	FOXA1	MAFA	PAX6	ATP6V0E1	STXBP1	HCN1	KCNK12	RAP1A
IAPP	FOXA2	MAFB	PBX1	COX4I1	SYP	HCN2	KCNK17	CDH1
INS	FOXA3	MLXIPL	PDX1	GPI	SYT4	HCN3	KCNK3	CDH2
NTS	FOXO1	MNX1	POU3F4	NNT	SYT7	HCN4	KCNMA1	DLL3
PPY	GATA4	MYT1	POU5F1	PCSK1	SYT8	KCNA3	KCNMB2	DPP4
PYY	GATA6	NANOG	PROX1	PCSK2	VAMP2	KCNA5	KCNN1	FGF4
SCT	GLIS3	NEUROD1	RFX2	PDIA6	VAMP3	KCNA6	KCNN3	GCGR
SST	HES1	NEUROG3	RFX3	PDK3	G6PC2	KCNB1	SCN1A	MMP2
	HNF1A	NKX2-2	RFX6	PGAP1	GCK	KCNC2	SCN2A	RAP1A
	HNF1B	NKX6-1	SNAI1	PRKCSH	SLC2A1	KCNC4	SCN3A	UCN3
	HNF4A	NKX6-2	SNAI2	RIMS1	SLC2A2	KCNG4	SCN4A	WNT2
	INSM1	NKX6-3	SOX4	RIMS2	SLC2A3	KCNH2	SCN7A	WNT3
	IRX2	NOTCH1	SOX9	RIMS3		KCNH6		WNT4

152 pancreatic lineage genes were selected based on published literature for their relevance in pancreatic development or β cell function.

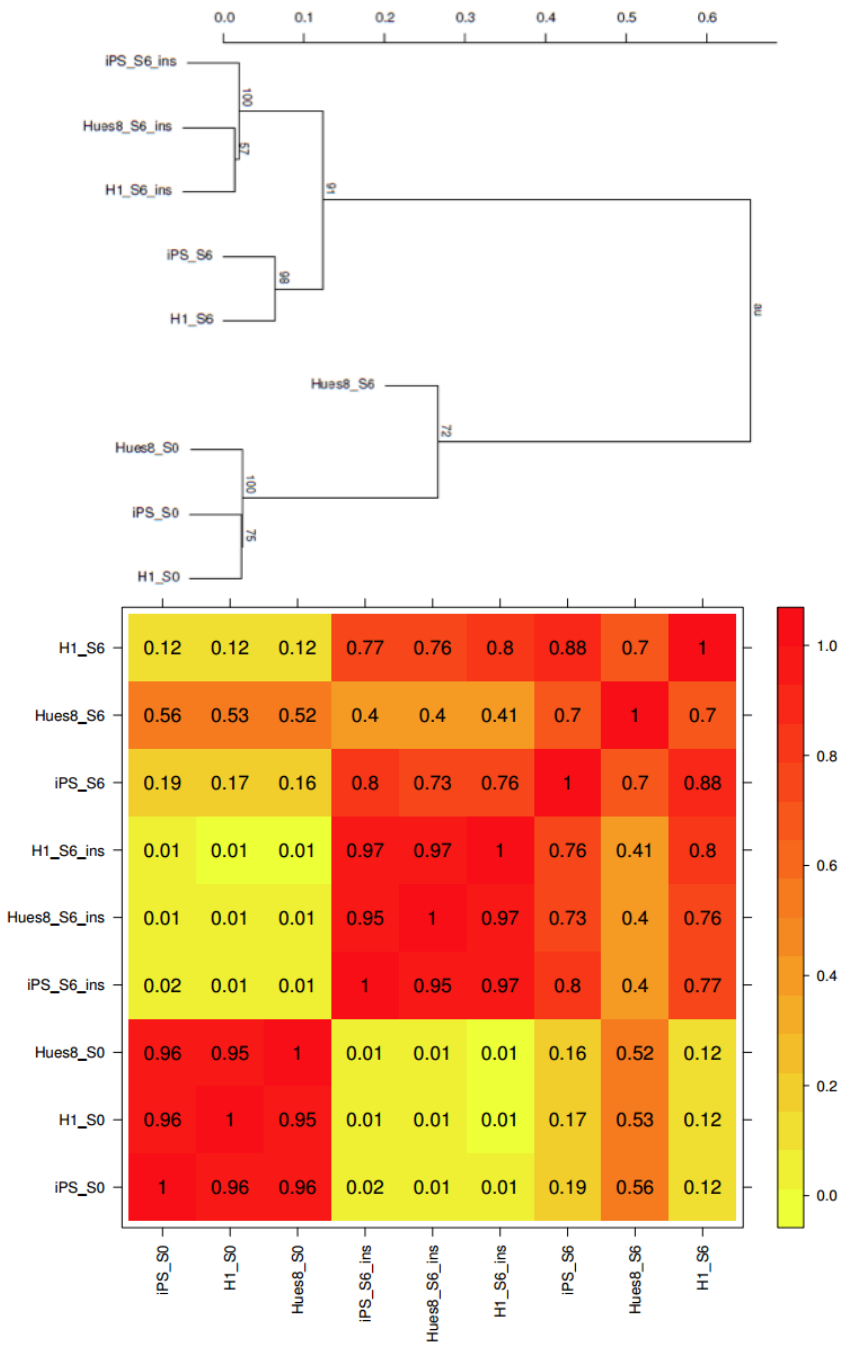


Figure 4-6. Correlation and clustering of cells derived by directed differentiation

R^2 values and hierarchical clustering based on microarray data across 152 pancreatic lineage genes is shown. Lengths in the dendrogram represent correlation value. Approximately Unbiased (AU) p-values are displayed. INS^+ cells from different cell lines form a statistically significant cluster.

4.3.3. Robust molecular signature of sorted hPSC-derived INS$^+$ cells

Recently Micallef et al. reported the generation of an insulin-GFP knock-in hESC-reporter line allowing for the profiling of sorted INS$^+$ cells [29]. Using the insulin-GFP reporter line, Basford et al. performed microarray analysis of sorted INS$^+$ cells using a different directed differentiation protocol and hESC-line than presented here [28]. We wondered whether our analysis of fixed and sorted INS$^+$ cells would generate a similar gene expression signature. Basford et al. analyzed 28 genes that were differentially expressed between insulin-positive and insulin-negative cells. In our analysis, 27 of the 28 identified genes had the same pattern of enrichment in each of the three hPSC-lines (Figure 4-7)[28]. This data further strengthens our conclusion that INS$^+$ cells derived from different hPSC-lines, using different protocols, display highly similar molecular signatures.

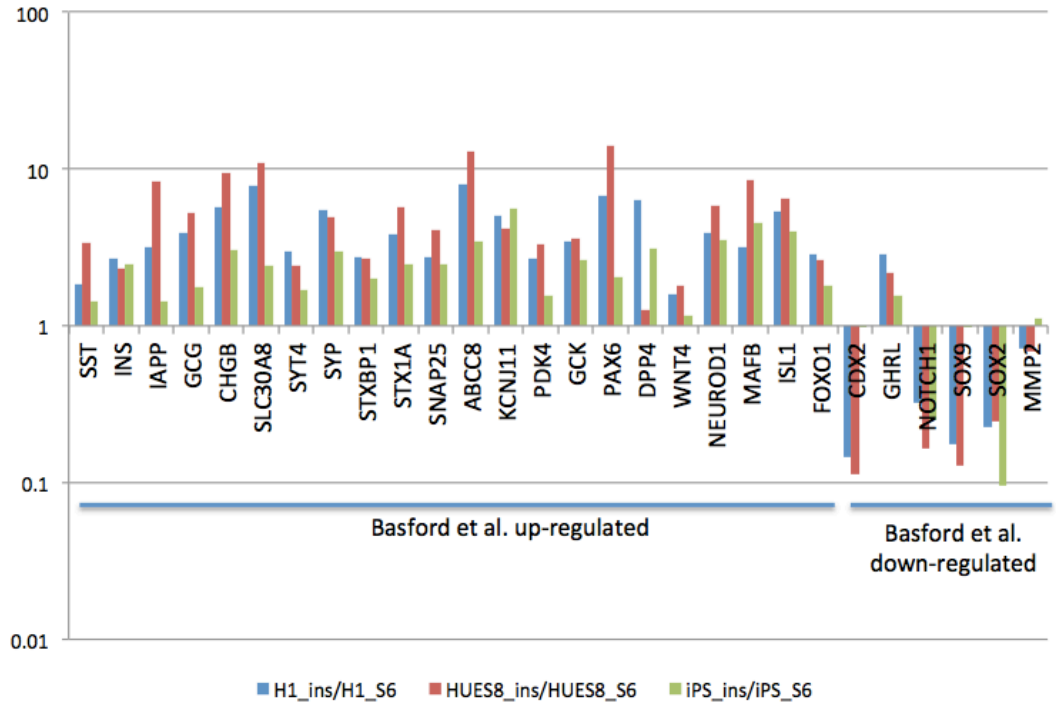


Figure 4-7. Comparable gene expression

Basford et al. identified 28 genes that were differentially expressed between hESC-derived INS^+ cells and ins- cells. Here shown is fold change in gene expression between hPSC-INS^+ and unsorted stage 6 cells for all 28 genes. 27 of the 28 genes had the same pattern of gene expression. GHRL (starred) was down-regulated in Basford et al. while up-regulated in all cells lines in our study.

4.3.4. Human β cell maturation differs from mouse β cell maturation

Study of human fetal development has been hampered by the absence of reporters and cell surface markers to permit sorting of specific cell types, as well as the scarcity of material for study. Researchers have thus relied on studies in model organisms, mainly the mouse, as the basis for understanding human development and optimizing directed differentiation. Recently, our laboratory identified gene expression signatures that

distinguish fetal β cells from adult β cells in the mouse [40]. In order to determine the extent to which these gene expression signatures might be conserved between mouse and human, we performed the first purification and transcriptome-wide molecular characterization of human fetal and adult β cells during human β cell maturation.

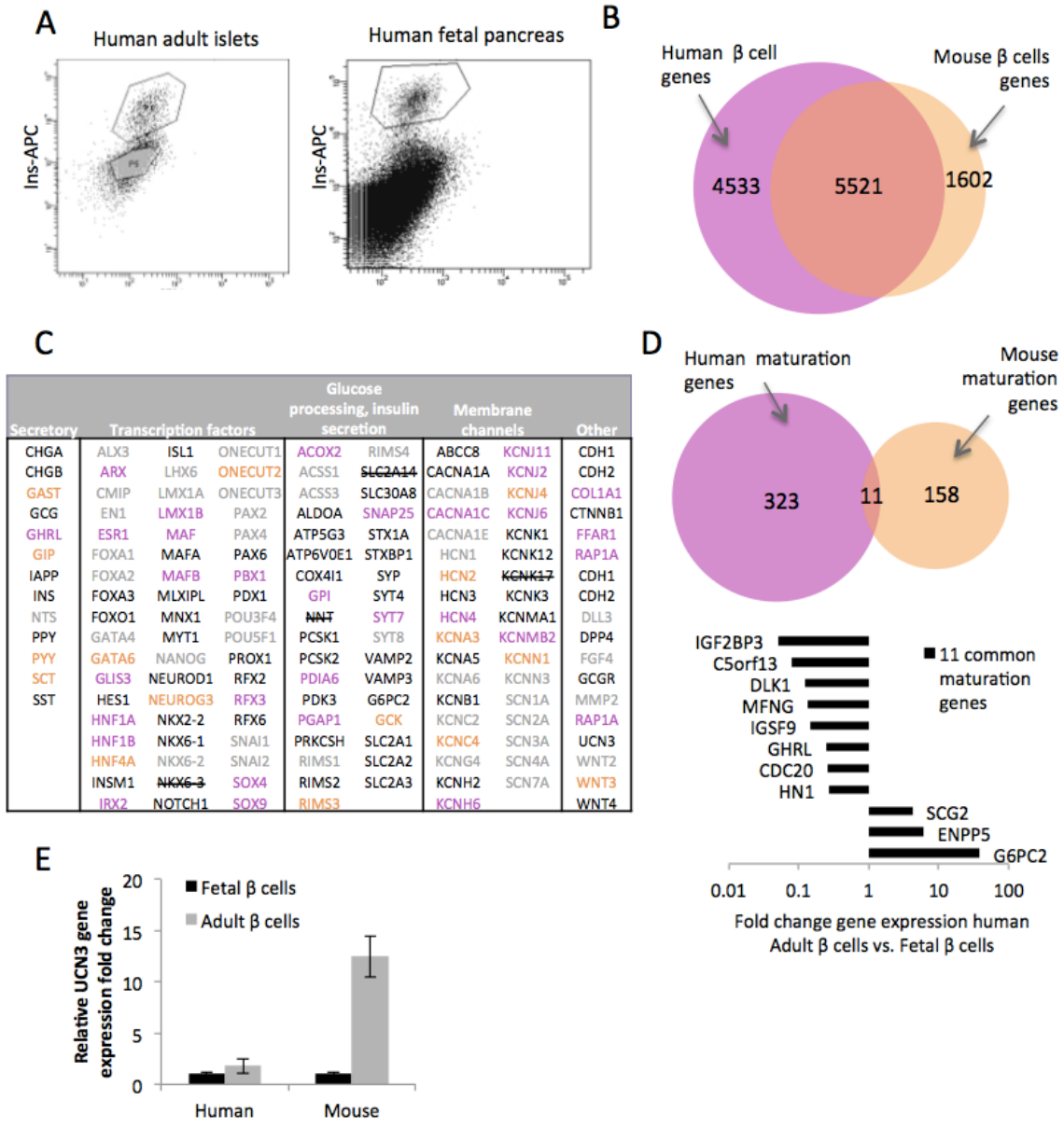
Human pancreata of 15-16 weeks gestational age were used, since β cells at this stage were previously found to be immature and glucose non-responsive [41,42]. Adult human cadaveric islets and fetal pancreata were dispersed, stained for insulin and FACS sorted. RNA was isolated and analyzed by Illumina microarrays (Figure 4-8a).

We first compared the list of genes that were expressed in adult mouse and human β cells (Illumina gene detection $p<0.05$). Analysis was limited to genes with known homology between the two species using the NCBI Homologene tool[43]. 5521 genes were expressed in β cells from both species while 4533 and 1602 genes were uniquely expressed in respectively human and mouse β cells (Figure 4-8b). We next focused the comparison on the previously described set of 152 pancreatic lineage genes. Among them, 62 genes were detected in β cells from both mouse and human, 40 genes from neither, 30 were detected only in human β cells and 16 only in mouse β cells (Figure 4-8c). Our whole-genome expression analysis indicates that although human and mouse β cells express many genes in common, there are also significant differences between them. Understanding of those differences may be critical for the translation of the research on model organisms into human therapy.

Figure 4-8. Comparison of human and mouse β cells

(a) FACS plots of human adult islets and human fetal pancreata sorted for INS^+ cells (APC^+). **(b)** Detected homologous genes ($p<0.05$) in human β cells (n=3) and mouse β cells (n=3). 5521 genes are present in both species. 4533 and 1602 genes are unique to human β cells and mouse β cells respectively. **(c)** The list of 152 genes is colored for genes expressed in both mouse and human β cells (black); genes not expressed in either mouse or human β cells (gray); genes expressed only in human β cells (magenta); genes expressed only in mouse β cells (orange). **(d)** 334 genes are differentially expressed (fold change >3, $p<0.05$) between human fetal (n=2) and adult β cells (n=3). 169 genes are differentially expressed (fold change >3, $p<0.05$) between mouse E18.5 (n=3) and mouse adult β cells (n=3) based on data from[40]. 11 genes were identified as differentially expressed during both mouse and human β cell maturation. A list of 11 common maturation genes ordered by fold change in gene expression between human adult and fetal β cells. **(e)** Relative expression of UCN3 in mouse and human fetal and adult β cells. Expression normalized to fetal levels in each species.

Figure 4-8 Continued



Secretory	Transcription factors			Glucose processing, insulin secretion		Membrane channels		Other
CHGA	ALX3	ISL1	ONECUT1	ACOX2	RIMS4	ABCC8	KCNJ11	CDH1
CHGB	ARX	LHX6	ONECUT2	ACSS1	~~SLC2A14~~	CACNA1A	KCNJ2	CDH2
GAST	CMIP	LMX1A	ONECUT3	ACSS3	SLC30A8	CACNA1B	KCNJ4	COL1A1
GCG	EN1	LMX1B	PAX2	ALDOA	SNAP25	CACNA1C	KCNJ6	CTNNB1
GHRL	ESR1	MAF	PAX4	ATP5G3	STX1A	CACNA1E	KCNK1	FFAR1
GIP	FOXA1	MAFA	PAX6	ATP6V0E1	STXBP1	HCN1	KCNK12	RAP1A
IAPP	FOXA2	MAFB	PBX1	COX4I1	SYP	HCN2	~~KCNK17~~	CDH1
INS	FOXA3	MLXIPL	PDX1	GPI	SYT4	HCN3	KCNK3	CDH2
NTS	FOXO1	MNX1	POU3F4	~~NNT~~	SYT7	HCN4	KCNMA1	DLL3
PPY	GATA4	MYT1	POU5F1	PCSK1	SYT8	KCNA3	KCNMB2	DPP4
PYY	GATA6	NANOG	PROX1	PCSK2	VAMP2	KCNA5	KCNN1	FGF4
SCT	GLIS3	NEUROD1	RFX2	PDIA6	VAMP3	KCNA6	KCNN3	GCGR
SST	HES1	NEUROG3	RFX3	PDK3	G6PC2	KCNB1	SCN1A	MMP2
	HNF1A	NKX2-2	RFX6	PGAP1	GCK	KCNC2	SCN2A	RAP1A
	HNF1B	NKX6-1	SNAI1	PRKCSH	SLC2A1	KCNC4	SCN3A	UCN3
	HNF4A	NKX6-2	SNAI2	RIMS1	SLC2A2	KCNG4	SCN4A	WNT2
	INSM1	~~NKX6-3~~	SOX4	RIMS2	SLC2A3	KCNH2	SCN7A	WNT3
	IRX2	NOTCH1	SOX9	RIMS3		KCNH6		WNT4

Analysis of human β cells from fetal and adult samples revealed that the functional changes that occur between gestational week 16 and adulthood are characterized by changes in the expression of 334 genes ($p<0.05$, fold change >3). Surprisingly, only 11 of

the 334 differentially expressed genes were also differentially expressed between mouse E18.5 and adult β cells, indicating a significant divergence between human and mouse β cell maturation between these developmental time points (Figure 4-8d). *Ucn3*, which we identified previously as a marker of mouse β cell maturation[40] was expressed 1.82-fold higher in sorted human adult β cells over fetal β cells, indicating that while this gene is up-regulated during human β cell maturation, the magnitude of the change is less than what is observed in mouse (Figure 4-8e).

Together, this data provides the first transcriptome-wide molecular characterization of human β cells and gene expression changes during human maturation. Since the observed differences between mouse and human maturation may reflect a disparity between analyzed stages, analysis of β cells at multiple time-points during human and mouse development is necessary to further elucidate potential inter-species differences.

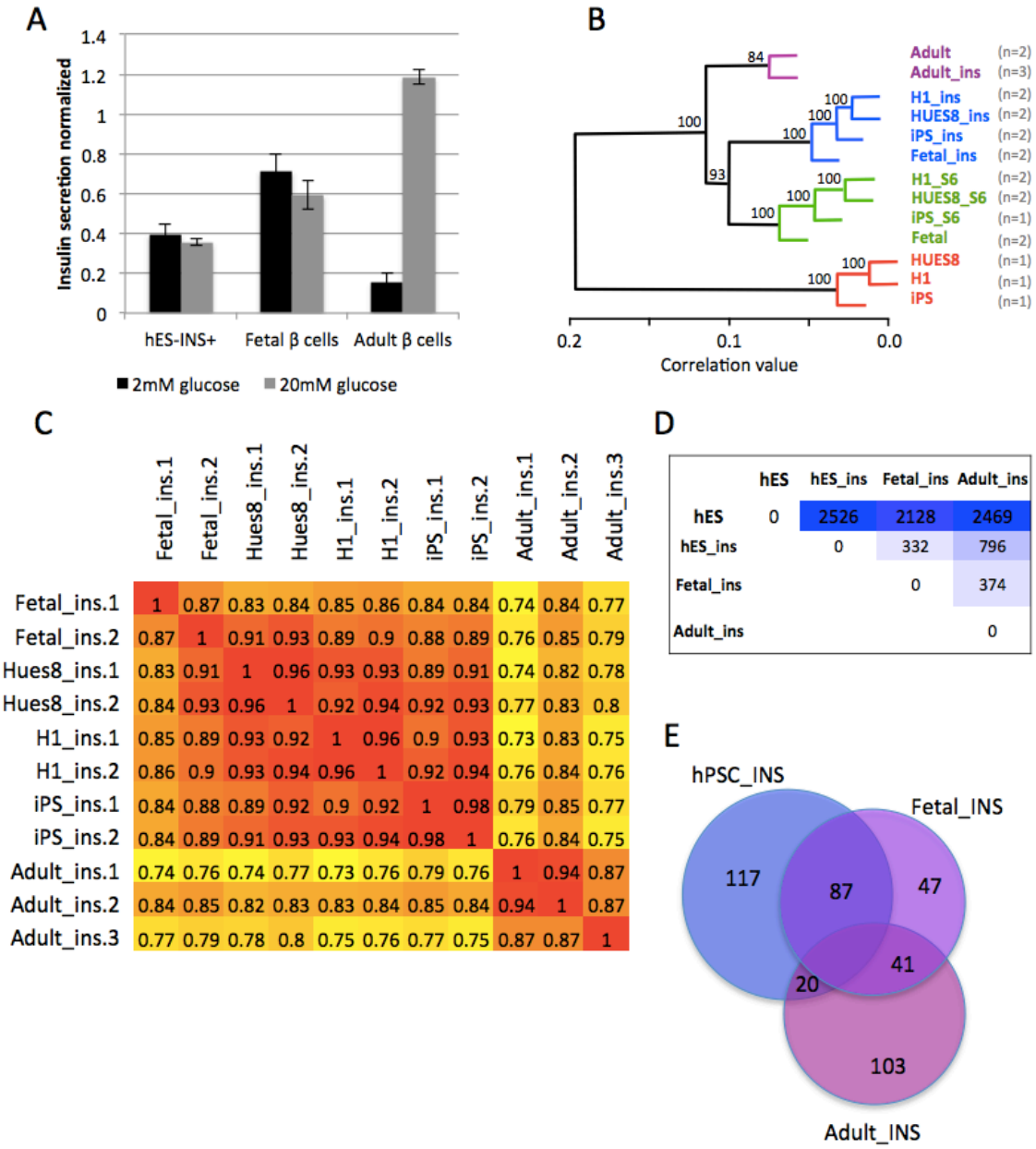
4.3.5. hPSC-derived INS$^+$ cells resemble human fetal β cells

Directed differentiation from hPSCs attempts to recreate human development *in vitro*. The extent to which this has been achieved remains unknown, because direct comparisons with human fetal cells have been either difficult or impossible. We first functionally tested dispersed hPSC-INS$^+$ cells, fetal and adult β cells for glucose stimulated insulin secretion (GSIS). In contrast to adult β cells, hPSC-INS$^+$ cells and human fetal β cells have elevated basal glucose secretion, and do not display a robust GSIS response (Figure 4-9a).

Figure 4-9. hESC-derived insulin expressing cells resemble human fetal β cells

(a) Glucose stimulated insulin secretion of dispersed cells. In contrast to adult β cells, fetal β cells and hESC-INS$^+$ cells both appear functionally immature as indicated by increased basal glucose secretion and lack of glucose stimulation. **(b)** Hierarchical clustering based on global gene expression across all 47325 microarray probes indicated that hESC-INS$^+$ cells cluster closely with human fetal and not adult β cells. Numbers in parentheses indicate biological replicates. Lengths in the dendrogram represent correlation distances. **(c)** R^2 values based on microarray data across all genes are shown. Each row and column represents one sample. R^2 values between biological replicates of adult β cell samples (Adult_ins) are on average 0.89±0.04. R^2 values between sorted hPSC-derived insulin$^+$ stage 6 cells and sorted fetal β cells are 0.87±0.03. The biological variation between adult β cells is not statistically smaller then the variation between fetal β cells and hPSC-derived insulin$^+$ stage 6 cells ($p=0.34$). This indicates that a high degree of similarity between hPSC-derived insulin$^+$ cells and human fetal β cells. **(d)** Numbers of genes differentially expressed ($p<0.05$, fold change >3) between groups of samples. **(e)** Venn diagram of differentially expressed genes ($p<0.05$, fold change >3). Numbers that belong only to one group represent genes that are differentially higher expressed in that sample group over each of the other two groups. Numbers in the overlap between two groups represent genes that are differentially higher expressed in those two groups over the third.

Figure 4-9 Continued



To molecularly compare hPSC-INS^{+} cells with human fetal and adult β cells, we performed whole genome expression analysis. Unsorted Stage 6 cells and undifferentiated

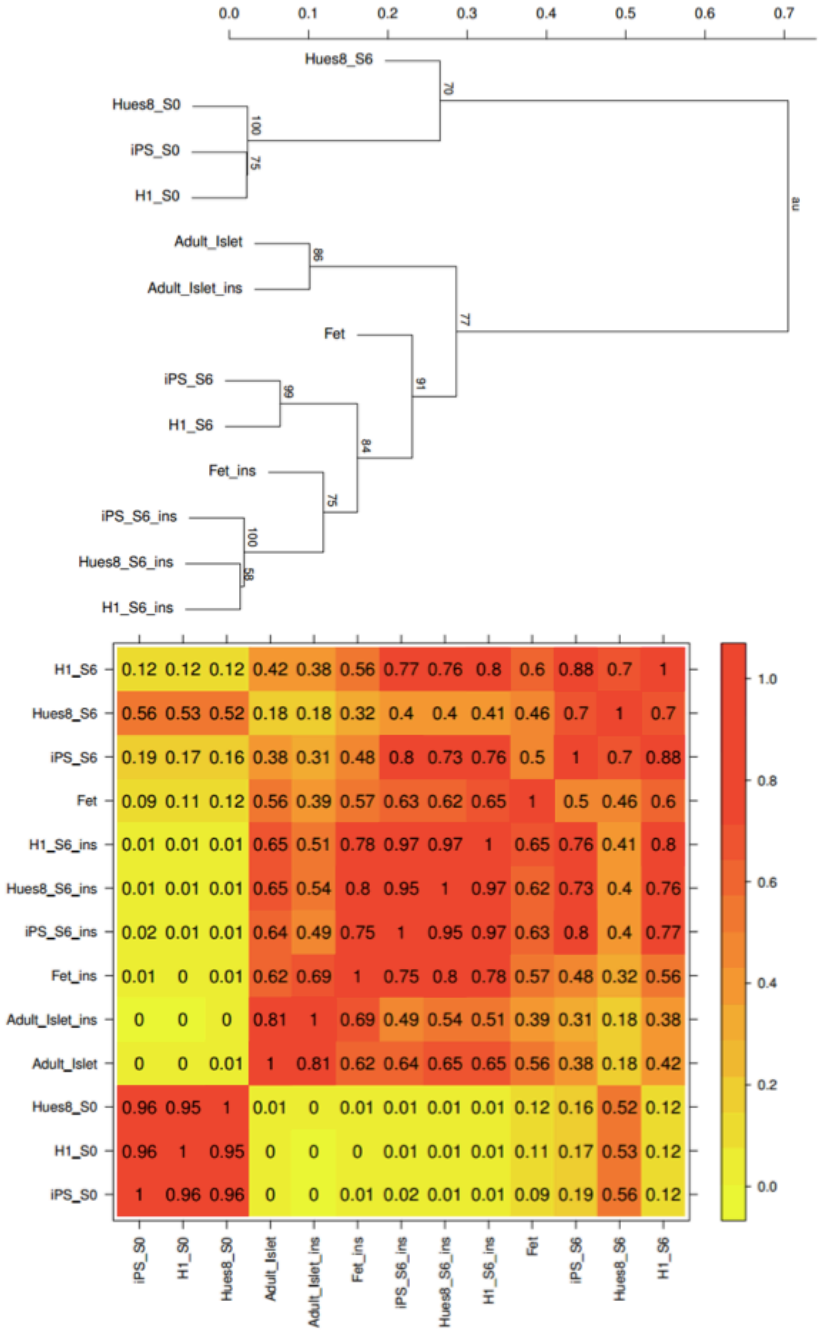
pluripotent stem cells were included as control groups. Hierarchical clustering across all genes identified four distinct groups of samples (Figure 4-9b). hPSC-derived INS^+ cells clustered more closely with human β cells than with hPSCs. Importantly, fetal β cells clustered together with hPSC-derived INS^+ cells and not human adult β cells. The correlation between three biological replicates of adult β cells (r^2 = 0.89±0.04) was not significantly different from the correlation between fetal β cells and hPSC-INS^+ cells (r^2=0.87±0.03) (two-tailed test, p=0.34), indicating a high degree of similarity between the latter two cell types (Figure 4-9c). Hierarchical clustering and correlation values based on 152 endocrine lineage genes confirmed this result (Figure 4-10).

Differential gene expression confirmed a much higher similarity of hPSC-INS^+ to human β cells than to hPSCs (Figure 4-9d). Additionally, fewer genes were differentially expressed between hPSC-INS^+ and fetal β cells (332) than between hESC-INS^+ cells and adult β cells (796). Together this data shows that INS^+ cells derived from different hPSC-lines are similar to fetal β cells and not adult β cells.

Figure 4-10. Correlation and clustering of hPSC-INS^+ cells, fetal and adult β cells

R^2 values and hierarchical clustering based on microarray data across 152 pancreatic lineage genes is shown. Lengths in the dendrogram represent correlation value. Approximately Unbiased (AU) p-values are displayed. INS^+ cells from different cell lines form a statistically significant cluster. R^2 values and hierarchical clustering based on microarray data across 152 pancreatic lineage genes is shown.

Figure 4-10 Continued



4.3.6. Human fetal β cells as an intermediate phenotype between hES-derived INS$^+$ cells and human adult β cells

To uncover cell-type specific differences in gene expression, we identified genes that had differentially higher or lower expression (>3 fold, $p<0.05$) in one cell type compared to the other two (Figure 4-9e). We observe a large number of uniquely overexpressed genes in hPSC-INS$^+$ cells (117) and adult β cells (103). The gene expression of fetal β cells largely overlapped with either hPSC-INS$^+$ cells or adult β cells. Only 47 genes had differentially high expression and 20 genes differentially low expression in fetal β cells over hPSC-INS$^+$ cells and adult β cells. We therefore propose that week 16 human fetal β cells may represent an intermediate phenotype between hPSC-INS$^+$ cells and adult β cells. Thus, our study raises the possibility that hPSC-INS$^+$ cells may more closely resemble cells at an earlier stage of human fetal development. Future analysis of human fetal β cells from younger gestational ages will be needed to resolve this possibility. Our study does not address whether hPSC-INS$^+$ cells are true progenitors of human fetal or adult β cells, or that they can give rise to these cell types upon further differentiation. The advent of more robust lineage tracing tools for use in hPSCs would facilitate the direct testing of this hypothesis.

4.3.7. Transcriptional differences between hES-derived INS$^+$ cells and human adult β cells

The goal of hPSC directed differentiation is to generate functional human β cells from hPSCs. We therefore compared gene expression by microarray and RNA-Seq between hESC-INS$^+$ cells and adult human β cells. We first focused our analysis on 152 pancreatic lineage genes. 28 of the 152 genes were significantly differentially expressed between adult β cells and hESC-INS$^+$ cells based on microarray analysis (>3 fold, $p<0.05$). RNA-Seq analysis of two samples of human adult INS$^+$ cells and two samples of HUES8-INS$^+$cells confirmed 27/28 genes as differentially expressed >3-fold (Figure 4-11a). Among these, we focused on transcription factors known to be crucial for β cell identity and function. Using immunofluorescent staining, we confirmed the lack of expression of NKX6-1 and relatively heterogeneous expression of PDX1 and MAFA in hESC-INS$^+$ cells (Figure 4-11b). PDX1, MAFA and NKX6-1 are robustly expressed in normal human β cells and *in vivo* matured hPSC-derived INS$^+$ cells [44].

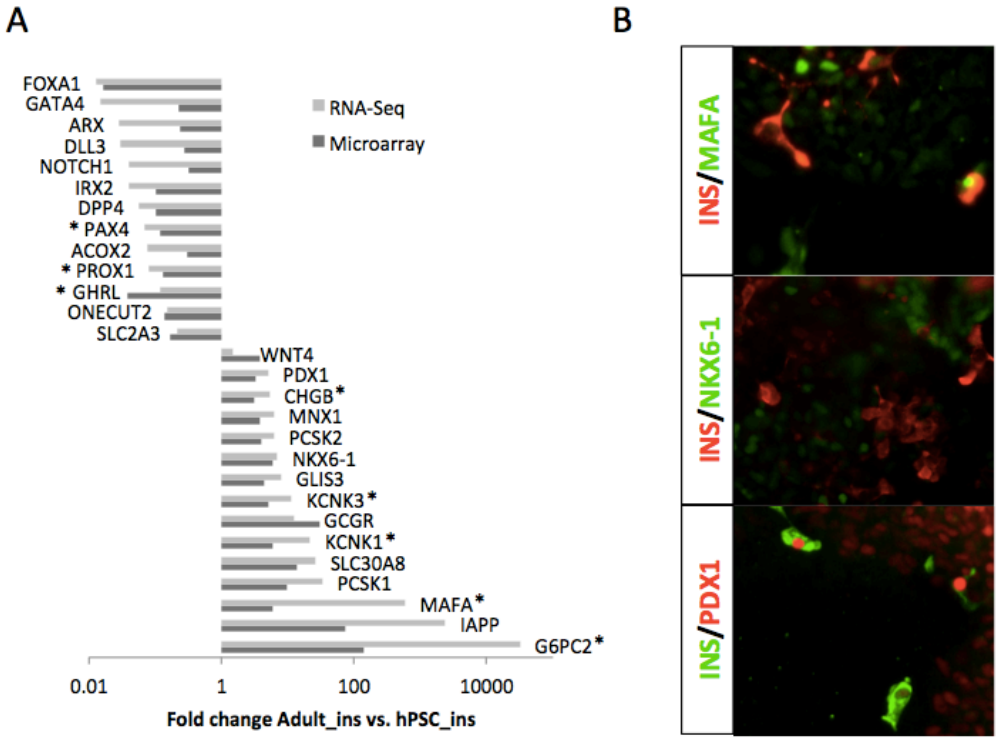


Figure 4-11. Differentially expressed genes between adult β cells and hPSC-INS$^+$ cells

(a) 28 genes from the list of 152 pancreatic lineage genes were differentially expressed between adult β cells and hPSC-derived INS$^+$ cells (microarray data, fold change >3) (dark gray bars). 27 of the 28 genes were confirmed by RNA-Seq between HUES8-INS$^+$ cells and adult β cells (fold change > 3) (light gray bars). Genes are ranked by fold change of expression in adult β cells over expression in hPSC-INS$^+$ cells. Genes marked by asterix are also differentially expressed between fetal β cells and adult β cells **(b)** Immunofluorescence. hPSC-INS$^+$ cells with PDX1, NKX6-1 and MAFA.

To understand the functional significance of incorrect expression of these 28 genes, we surveyed the β cell literature for known overexpression or knockdown phenotypes. A striking number of differentially expressed transcription factors had known roles in endocrine subtype specification. PDX1, NKX6-1 and MNX1 are necessary for the

determination of β cell fate and have over 3-fold lower expression in hPSC-INS$^+$ than human adult β cells[45-47]. Conversely FOXA1, ARX and IRX2, determinants of α cell fate, were miss-expressed 4-40-fold higher in hPSC-INS$^+$ cells than adult β cells (Figure 4-11a) [48-50].

Additionally, several identified genes had known effects on β cell physiology and glucose stimulated insulin secretion (GSIS). The overexpression of MAFA is sufficient to induce mature GSIS in immature P2 rat β cells[51]. MAFA had the largest differential gene expression by RNA-Seq among transcription factors (587-fold lower in hPSC-INS$^+$ cells than adult β cells). Conversely, prolonged expression of PAX4 in mouse adult β cells was shown to blunt GSIS [52]. PAX4 had 14.7-fold higher expression in hPSC-INS$^+$ cells than adult β cells. Several other notable genes may be responsible for the lack of functional GSIS in hPSC-INS$^+$ cells: presence of ghrelin (GHRL) suppresses GSIS [53]; CHRB knockout animals have reduced GSIS and elevated basal insulin secretion[52]; PROX1 is associated with insulin secretion abnormalities [54]; lack of tandem pore domain potassium channels KCNK1 and KCNK3 may elevate resting membrane potential cause hyperactivity and higher basal insulin secretion.

Since both hPSC-INS$^+$ and fetal β cells lack mature GSIS, we wondered whether the differentially expressed genes associated with β cell function were also miss-expressed in human fetal β cells. Indeed, MAFA, PAX4, GHRL, PROX1, CHRB, KCNK1 and KCNK3 were also differentially expressed ($p<0.05$) between fetal β cells and adult β cells (Figure

4-11a). The modulation of these genes along with others with the same pattern of expression may be critical for the acquisition of mature glucose response.

Interestingly, transcription factors with known roles in endocrine subtype specification, namely PDX1, NKX6-1, MNX1, FOXA1, ARX and IRX2, were miss-expressed in hPSC-INS$^+$ cells but not fetal β cells. Together, the analysis of gene expression patterns suggests two possibly distinct challenges to produce functional human β cells from hPSC-INS$^+$ cells: 1) β cell lineage commitment and 2) functional maturation. Our observations support the hypothesis that hPSC-INS$^+$ cells resemble earlier stage human fetal cells potentially not committed to the β cell lineage.

Finally, we expanded our analysis to the whole genome. Of the 796 genes that were differentially expressed ($p<0.05$, >3 fold) between hPSC-INS$^+$ cells and human β cells, 639 were confirmed by RNA-Seq (>3 fold). We present a list of all differentially expressed transcription factors as they are of particular interest for their role in modulating cell fates (Figure 4-12). Genes identified here can be used as markers for differentiation into human β cells or may be manipulated to direct the conversion of hPSC-INS$^+$ cells into human functional β cells.

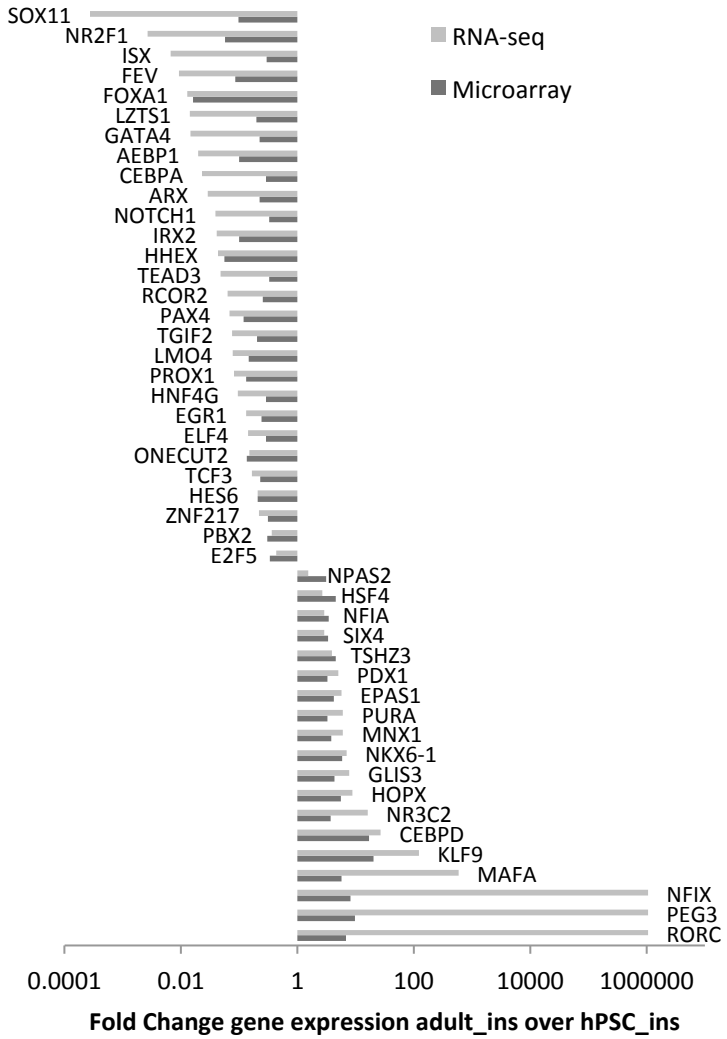


Figure 4-12. Differential expression of transcription factors between human β cells and hPSC-derived insulin expressing cells

A list of 47 transcription factors that were differentially expressed based on microarray data between human adult β cells and hPSC-derived INS$^+$ cells (fold change >3, $p<0.05$) and confirmed by RNA-Seq between HUES8-INS$^+$ cells and adult β cells (fold change > 3).

4.4. Discussion

This study presents a method for transcriptional profiling of cells following fixation, immunoflorescent staining and FACS. The developed method (MARIS) produces high quality RNA that can be used for qRT-PCR, microarray and RNA-Seq analysis without detectable loss of fidelity. MARIS permits, for the first time, comparison of genetically unmodified cells produced by differentiating hPSCs *in vitro*, and the direct comparison of these cells to corresponding cell types isolated from human fetal and adult tissues. We have successfully used this approach to analyze the transcriptome of hPSC-derived pancreatic cells produced *in vitro*, and compare the observed gene expression patterns with those of human fetal and adult β cells.

We first determined the degree of similarity between INS^+ cells derived from multiple hPSC-lines. It has been long known that different hPSC-lines vary in their propensity to differentiate into a desired cell type, but the degree to which the generated cells vary across different cell lines remained unexplored. Significant gene expression differences between different hPSC-lines would question the robustness of the results thus far generated from hPSCs. Here we present the analysis of INS^+ cells from two hESC-lines and one hiPSC-line. Our data suggests a high degree of similarity between INS^+ cells derived from all three cell-lines. The degree of correlation between INS^+ cells derived after more then 20 days of directed differentiation resembled the degree of correlation between different pluripotent stem cells. Furthermore, correlation between INS^+ cells was similar to the degree of correlation between two human β cell samples obtained from two different

donors. The degree of similarity we observe between pancreatic cells derived from different hPSC lines presents an important proof-of-principle observation for hPSC directed differentiation. MARIS opens the door for similar comparisons of other cell types.

The methodology that we have developed allows, for the first time, direct analysis of purified human fetal and adult cell types that cannot be sorted by cell surface markers, and comparison of these cells with cells from mice or other model organisms. In the absence of tools to study human development, model organisms have informed our understanding of human development and biology, although the degree to which human development resembles the development of other organisms has generally not been rigorously investigated. Using MARIS we have been able to compare human fetal and adult β cells, and identify genes that are differentially expressed during human β cell maturation. We discovered similarities and large differences in gene expression between mouse and human β cells. Moreover, that gene expression changes during human β cell maturation did not resemble changes in mouse β cell gene expression during the late prenatal and early postnatal period. A possible explanation might be that we are comparing two different stages of β cell maturation. Alternatively, there may be intrinsic developmental differences between mouse and human. Further study of multiple stages during human and mouse development using RNA-Seq in addition to microarrays would help elucidate the answer to this question.

It has been widely thought that hPSC-directed differentiation produces cells that have a fetal, or immature, phenotypes, rather than adult cells[20]. These assessments are

generally made based on the expression of a handful of fetal-specific genes, or the absence of some number of adult markers. Because the number of genes analyzed in these earlier studies is quite small, the degree to which cells derived *in vitro* actually resemble true human fetal cells has remained unknown. To answer this question, we carried out a genome wide expression comparison of hPSC-INS$^+$ cells from three different pluripotent stem cells lines with human fetal and adult β cells. There are many differences in gene expression between hPSC-INS$^+$ cells, human fetal β cells and human adult β cells, suggesting that no two cell-types are fully equivalent. However, hierarchical clustering based on genome-wide gene expression showed close clustering of hPSC-INS$^+$ cells with human fetal β cells and not human adult β cells. hPSC-INS$^+$ cells and fetal β cells were no more different from each other then the variance between biological replicates of human adult β cells. This result was confirmed by correlation analysis based on 152 pancreatic lineage genes. We conclude that hPSC-derived INS$^+$ cells highly resemble fetal β cells and not adult β cells.

Through the technology developed here, we have been able to directly assess the gene expression differences that remain between the INS$^+$ cells produced *in vitro*, and bona fide adult human β cells. By focusing on genes known for their role in pancreatic development and β cell function, we identified by microarray and RNA-Seq 27 genes that are differentially expressed between these two cell types, and may have a functional role in converting the cells produced *in vitro* into the desired fully mature phenotype. Several transcription factors involved in endocrine subtype specification were miss-expressed in

hPSC-INS$^+$ cells but not fetal β cells. hPSC-INS$^+$ cells and fetal β cells lack mature glucose stimulated insulin secretion. Aberrant expression of seven genes in hPSC-INS$^+$ and fetal β cells correlated with this phenotype. Four of the seven genes, MAFA, PROX1, PAX4 and GHRL had been associated with defects in insulin secretion. Two genes, KCNK1 and KCNK3 are potassium channels whose properties suggest a role in maintaining resting membrane potential. The last gene G6PC2 is involved in glucose metabolism.

Our data suggests at least two classes of genes whose regulation may be critical in generating functional β cells. The first class of genes regulates the lineage commitment of hPSC-INS$^+$ cells towards β cells and away from other pancreatic endocrine cell types. The second class of genes, miss-expressed in both human fetal β cells and hPSC-INS$^+$ cells may be responsible for functional β cells maturation.

At the end we expand our analysis across the genome and present a list of 47 differentially expressed transcription factors that may be of particular interest to the β cell community. Overexpression and knockdown studies are needed to determine which of these genes are necessary or sufficient to direct the differentiation of hPSC-INS$^+$ cells to functional mature β cells.

In conclusion we developed a Method to Analyze RNA following Intracellular Sorting (MARIS) and carried out the first transcriptome wide analysis of sorted human fetal and adult β cells. Our analysis of human and mouse β cells and genes marking human and mouse β cell maturation indicates significant inter-species differences and points to a need to further study human development. Transcriptome analysis of sorted INS$^+$ cells

derived from two hESC-lines and one hiPSC-line shows for the first time that different cell lines produce highly similar INS^+ cells. We demonstrate that non-functional hPSC-derived INS^+ cells resemble human fetal β cells, not adult β cells. Finally we generate a list of candidate genes whose regulation may be critical for the generation of a) lineage committed and b) functional β cells from hPSCs.

4.5. Materials and Methods

4.5.1. Directed Differentiation

hPSCs were routinely cultured on hESC-certified matrigel (BD Biosciences) in mTeSR medium (Stemcell Technologies). Cells were passaged at the ratio of 1:6–1:20 every 4-7 days using TrypLE Express (Invitrogen). Two different basal medias were used during differentiation. Basal media 1 (BM-1) contained MCDB-131 (Invitrogen) supplemented to 10mM glucose, 1x GlutaMAX (Gibco, Life Technologies), 2.35g/l $NaHCO_3$ and 0.1% reagent grade BSA (Proliant). Basal media 2 (BM-2) contained MCDB-131 (Invitrogen) supplemented to 8mM glucose, 1x GlutaMAX (Gibco, Life Technologies), 2.93g/l $NaHCO_3$, 2% reagent grade BSA (Proliant), 1:200 ITS-X (Invitrogen) and 44mg/l Vitamin C.

To initiate differentiation the cells were dissociated using TrypLE Express to single cells and seeded at 150,000 cell/cm2 onto 1:30 dilution of growth factor reduced matrigel

(BD Biosciences) in DMEM/F12 in mTeSR media with 10uM Y27632 (StemGent). Two days following seeding the differentiation was started.

Day 1 (stage 1.1) cells were exposed to BM-1 supplemented with and 3uM CHIR-99021 (Stemgent) + 100ng/ml rhActivinA (R&D Systems). Day 2-3 (stage 1.2): BM-1 + 100ng/ml rhActivinA (R&D Systems). Day 4-5 (stage 2): BM-1 + 50ng/ml FGF7 (Peprotech) (stage 2). Day 6-9 (stage 3) BM-2 + 50ng/ml FGF7 (Peprotech) + 2μM RA (Sigma) + 0.25μM SANT-1 (Sigma) + 20ng/ml rhActivinA (R&D Systems) + 100ng/ml rhNoggin (R&D Systems). Day 10-12 (stage 4) BM-2 + 100ng/ml rhNoggin (R&D Systems) + 0.25μM SANT-1 (Sigma) + 100nM PDBu (EMD Biosciences) (stage 4). Day 13-15 (stage 5) BM-2 + 100ng/ml rhNoggin (R&D Systems) + 1μM Alk5 inhibitor (Axxora). Day 15 onwards cells were kept in BM-2 media awaiting analysis.

4.5.2. Staining and FACS

hPSC-derived cells and human islets were dispersed to a single cell suspension using TrypLE Express (Invitrogen). Human fetal pancreata were mechanically dispersed in the presence of 1mg/ml Dispase (Roche) and 1mM Collagenase P (Roche). All cells were passed through a 40um filter and washed with PBS at least twice. Cells were fixed with 4% PFA (Electron Microscopy Sciences) in molecular grade PBS (Ambion) supplemented with 1:100 RNasin Plus RNase Inhibitor (Promega, N2615) for 30' at 4°C. All the subsequent steps were carried out at 4°C. Cells were pelleted by centrifugation at 3000g for 3' 4°C and washed in Wash Buffer: PBS containing 0.2% BSA (Gemini Bio-Products), 0.1% saponin

(Sigma-Aldrich), 1:100 RNasin Plus RNase Inihibitor. Primary antibody staining was carried out while 3D rocking for 30' at 4°C in Staining buffer containing PBS with 1% BSA, 0.1% saponin and 1:20 RNasin Plus RNase Inhibitor. Cells were washed twice in Wash Buffer followed by secondary antibody staining in Staining buffer. Following secondary antibody staining cells were washed twice in Wash buffer and resuspended in Sort buffer containing PBS, 0.5% BSA, and 1:20 RNasin Plus RNase Inhibitor.

Cells were sorted on the FACSAria (BD Biosciences) using FACSDiva software. Gates were set with reference to negative controls. The sorting speed was adjusted to ensure sorting efficiency above 90%. Cells were collected in tubes that were coated with a small amount of Sort buffer.

4.5.3. RNA isolation

After sorting, cells were pelleted by centrifugation at 3000g for 5' at 4°C. The supernatant was discarded. Total RNA was isolated from the pellet using the RecoverAll Total Nucleic Acid Isolation kit (Ambion), starting at the protease digestion stage of manufacturer-recommended protocol. The following modification to the isolation procedure was made: instead of incubating cells in digestion buffer for 15 minutes at 50°C and 15 minutes at 80°C, we carried out the incubation for 3 hours at 50°C. Cell lysates were frozen at -20°C overnight before continuing the RNA isolation by the manufacturer's instructions.

After isolation, RNA quantity was evaluated using a NanoDrop 2000 (NanoDrop Technologies). RNA quality was assessed by microcapillary electrophoresis on the Bioanalyzer 2100 (Agilent) using RNA Nano 6000 or RNA Pico 6000 chips, depending on the RNA concentration. Agilent 2100 Expert software was used to visualize the electropherograms and calculate the RNA Integrity Number (RIN), a standardized categorization of total RNA quality on a scale of 1 (worst) to 10 (best)[55].

4.5.4. Quantitative RT-PCR

Complementary DNA (cDNA) was made from 4 ng unamplified total RNA with random hexamer priming using the High Capacity cDNA Reverse Transcription with RNase Inhibitor kit (Applied Biosystems). One-fourth of the volume of cDNA was used for each TaqMan quantitative reverse transcription polymerase chain reaction (qRT-PCR). Reactions contained transcript-specific TaqMan probes (Applied Biosystems) and Fast Universal PCR Master Mix with no AmpErase UNG (Applied Biosystems). The following probes were used: ACTB (Hs99999903_m1), INS (Hs00355773_m1), GCG (Hs00174967_m1) and SST (Hs00356144_m1). Reactions were run on an Applied Biosystems 7900HT Fast Real-Time PCR System with default settings. Detection thresholds were automatically computed by SDS 2.3 software (Applied Biosystems). Threshold data were analyzed in DataAssist 3.0 (Applied Biosystems) using the Comparative Ct (ΔΔCt) relative quantitation method, using β-actin as the endogenous control.

4.5.5. Global gene expression analysis - microarray

Using the Illumina TotalPrep RNA Amplification kit (Ambion), double-stranded cDNA was generated following reverse transcription from 100 ng of total RNA. *In vitro* transcription overnight with biotin-labeled nucleotides created amplified mRNA (cRNA), which was concentrated by vacuum centrifugation at 30°C. 750 ng cRNA per sample was then hybridized to Human HT-12 Expression BeadChips (Illumina) using the Whole-Genome Expression Direct Hybridization kit (Illumina). Finally, chips were scanned on the Illumina Beadstation 500. The chip annotation manifest was version 4, revision 1. For differential expression analysis and the generation of gene lists for functional annotation and pathway analysis, microarray data were processed in GenomeStudio (Illumina, V2011.1). Raw data were adjusted by background subtraction and rank-invariant normalization. Before calculating fold change, an offset of 20 was added to all probe set means to eliminate negative signals. The p- values for differences between mean signals were calculated in GenomeStudio by t-test and corrected for multiple hypotheses testing by the Benjamini-Hochberg method in combination with the Illumina custom false discovery rate (FDR) model.

4.5.6. Global gene expression analysis – RNA-Seq

Isolated RNA was obtained from 2 biological replicates of HUES8-derived INS^+ cells and human adult β cell, as well as one replicate of live and processed stage 6 cells each. Samples were poly-A purified and converted to cDNA libraries using the Illumina TruSeq protocol, and prepared into Illumina libraries using the Beckman Coulter Genomics SPRI-works system using custom adapters. 6nt 3' barcodes were added during PCR enrichment and the resulting fragments were evaluated using Agilent BioAnalyzer 2100. Samples were multiplexed 2-per-lane for sequencing using the Illumina HiSeq 2000 platform with paired-end read lengths of 80nt, resulting in 68M to 112M paired reads per sample, and an average biological fragment length of 168-179nt. Reads were aligned to the human genome (GRCh37/hg19) using TopHat (version 2.0.4) [56,57] guided by Ensembl gene annotations. RNA-Seq enrichment in annotated Ensembl transcripts was determined by Cuffdiff [56,57] (version 2.0.2) which performed a maximum likelihood estimate of transcript abundance measured in fragments per kilobase of exon per million fragments mapped (FPKM). Statistically significant transcript differential expression was calculated by Cuffdiff using the default negative binomial model, along with associated p-values.

4.5.7. RNA-Seq transcript integrity analysis

To identify any RNA fragment length bias introduced by the MARIS protocol, TopHat aligned RNA-Seq read counts from before and after were analyzed at Ensembl annotated exon locations (GRCh37/hg19 version 68)[58]. Per-bp RNA-Seq read coverage was normalized by total mean log counts and summed over all exonic regions. Transcripts were

binned by exonic length and average per-transcript coverage was adjusted for length (Figure 4-4). In transcripts longer than 2.5knt (% of genome annotations), the processed sample 3' bias was greater than the live sample by roughly 2% (AUC differential).

4.5.8. Microarray expression clustering

Gene level microarray expression values were generated by GenomeStudio, using rank invariant normalization with background subtraction, and analyzed using the R package pvclust [59]. Per-sample and per-condition (averaging gene levels) hierarchical clustering was performed based on Pearson correlation and dendogram visualizations were drawn. Pvclust's multiscale bootstrap resampling was used (with 500k iterations) to estimate "approximately unbiased" (AU) p-values indicating the significance of each subcluster choice in the hierarchy given the underlying data.

4.5.9. Glucose stimulated insulin secretion

Approximately 10^5 dispersed hPSC-derived S6 cells or dispersed fetal cells were plated per well of a 96-well plate and allowed to attach overnight. Alternatively, approximately 5000 dispersed islet cells we plated amidst 1×10^5 hESCs (for cell-cell contact and attachment). Cells were then washed two times in PBS. Cells were incubated for 1 hour in Kreb's buffer with no glucose (128 mM NaCl, 5 mM KCl, 2.7 mM CaCl2, 1.2 mM MgCl2, 1 mM Na2HPO4, 1.2 mM KH2PO4, 5 mM NaHCO3, 10 mM HEPES, 0.1%

BSA) at 37°C, 5% CO2, and then incubated for 60 min in Kreb's buffer with either a) 2mM (low) glucose, b) 20 mM (high) glucose, or c) 2mM glucose with 30 mM KCl. Supernatant fractions after each exposure to glucose were collected and stored at -80°C until analysis. Insulin concentrations were measured using the Mouse Ultrasensitive Insulin ELISA kit (Alpco, 80-INSMSU-E01), which cross-reacts >100% with human insulin. Concentrations were calculated from cubic spline interpolation of a standard curve and normalized to input cell number.

4.6. References

1 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-933, doi:10.1016/j.cell.2007.11.019 (2007).

2 Thomson, J. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)* **282**, 1145-1147 (1998).

3 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.)* **318**, 1917-1937, doi:10.1126/science.1151526 (2007).

4 Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141-146, doi:10.1038/nature06534 (2008).

5 Colman, A. & Dreesen, O. Pluripotent stem cells and disease modeling. *Cell stem cell* **5**, 244-251, doi:10.1016/j.stem.2009.08.010 (2009).

6 Daley, G. Stem cells: roadmap to the clinic. *The Journal of clinical investigation* **120**, 8-18, doi:10.1172/jci41801 (2010).

7 Rubin, L. Stem cells and drug discovery: the beginning of a new era? *Cell* **132**, 549-601, doi:10.1016/j.cell.2008.02.010 (2008).

8 Chin, M. *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell stem cell* **5**, 111-134, doi:10.1016/j.stem.2009.06.008 (2009).

9 Doi, A. *et al.* Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature genetics* **41**, 1350-2703 (2009).

10 Hu, B.-Y. *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4335-4375, doi:10.1073/pnas.0910012107 (2010).

11 Osafune, K. *et al.* Marked differences in differentiation propensity among human embryonic stem cell lines. *Nature biotechnology* **26**, 313-318, doi:10.1038/nbt1383 (2008).

12 Bock, C. *et al.* Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-491, doi:10.1016/j.cell.2010.12.032 (2011).

13 Elliott, A., Elliott, K. & Kammesheidt, A. High resolution array-CGH characterization of human stem cells using a stem cell focused microarray. *Molecular biotechnology* **46**, 234-242, doi:10.1007/s12033-010-9294-1 (2010).

14 Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature cell biology* **13**, 541-549, doi:10.1038/ncb2239 (2011).

15 Martinez-Fernandez, A., Nelson, T. & Terzic, A. Nuclear reprogramming strategy modulates differentiation potential of induced pluripotent stem cells. *Journal of cardiovascular translational research* **4**, 131-137, doi:10.1007/s12265-010-9250-2 (2011).

16 Hu, Q., Friedrich, A., Johnson, L. & Clegg, D. Memory in induced pluripotent stem cells: reprogrammed human retinal-pigmented epithelial cells show tendency for spontaneous redifferentiation. *Stem cells (Dayton, Ohio)* **28**, 1981-1991, doi:10.1002/stem.531 (2010).

17 Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell stem cell* **10**, 595-609, doi:10.1016/j.stem.2012.02.014 (2012).

18 Kim, K. *et al.* Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature biotechnology* **29**, 1117-1119, doi:10.1038/nbt.2052 (2011).

19 Nostro, M. *et al.* Stage-specific signaling through TGFβ family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development (Cambridge, England)* **138**, 861-932, doi:10.1242/dev.055236 (2011).

20 D'Amour, K. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nature biotechnology* **24**, 1392-1401, doi:10.1038/nbt1259 (2006).

21 Ginis, I. *et al.* Differences between human and mouse embryonic stem cells. *Developmental biology* **269**, 360-440, doi:10.1016/j.ydbio.2003.12.034 (2004).

22 Odom, D. *et al.* Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nature genetics* **39**, 730-732, doi:10.1038/ng2047 (2007).

23 Rezania, A. *et al.* Maturation of Human Embryonic Stem Cell-Derived Pancreatic Progenitors into Functional Islets Capable of Treating Pre-existing Diabetes in Mice. *Diabetes*, doi:10.2337/db11-1711 (2012).

24 Maehr, R. *et al.* Generation of pluripotent stem cells from patients with type 1 diabetes. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15768-15773, doi:10.1073/pnas.0906894106 (2009).

25 Chen, S. *et al.* A small molecule that directs differentiation of human ESCs into the pancreatic lineage. *Nat Chem Biol* **5**, 258-265, doi:10.1038/nchembio.154 (2009).

26 Kelly, O. *et al.* Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nature biotechnology* **29**, 750-756, doi:10.1038/nbt.1931 (2011).

27 Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nature biotechnology* **26**, 443-495, doi:10.1038/nbt1393 (2008).

28 Basford, C. *et al.* The functional and molecular characterisation of human embryonic stem cell-derived insulin-positive cells compared with adult pancreatic beta cells. *Diabetologia* **55**, 358-429, doi:10.1007/s00125-011-2335-x (2012).

29 Micallef, S. *et al.* INS(GFP/w) human embryonic stem cells facilitate isolation of *in vitro* derived insulin-producing cells. *Diabetologia* **55**, 694-1400, doi:10.1007/s00125-011-2379-y (2012).

30 Nostro, M. C. *et al.* Stage-specific signaling through TGF family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development* **138**, doi:10.1242/dev.065904 (2011).

31 Jiang, J. *et al.* Generation of insulin-producing islet-like clusters from human embryonic stem cells. *Stem Cells* **25**, 1940-1953, doi:10.1634/stemcells.2006-0761 (2007).

32 Tateishi, K. *et al.* Generation of insulin-secreting islet-like clusters from human skin fibroblasts. *J Biol Chem* **283**, 31601-31607, doi:10.1074/jbc.M806597200 (2008).

33 Zhang, D. *et al.* Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells. *Cell research* **19**, 429-438, doi:10.1038/cr.2009.28 (2009).

34 Dobson, J., Da'as, S., McBride, E. & Berman, J. Fluorescence-activated cell sorting (FACS) of whole mount in situ hybridization (WISH) labelled haematopoietic cell populations in the zebrafish. *British journal of haematology* **144**, 732-737, doi:10.1111/j.1365-2141.2008.07527.x (2009).

35 Pan, Y., Ouyang, Z., Wong, W. & Baker, J. A new FACS approach isolates hESC derived endoderm using transcription factors. *PloS one* **6**, doi:10.1371/journal.pone.0017536 (2011).

36 Pechhold, S. *et al.* Transcriptional analysis of intracytoplasmically stained, FACS-purified cells by high-throughput, quantitative nuclease protection. *Nature biotechnology* **27**, 1038-1080, doi:10.1038/nbt.1579 (2009).

37 Auer, H. *et al.* Chipping away at the chip bias: RNA degradation in microarray analysis. *Nature genetics* **35**, 292-293, doi:10.1038/ng1203-292 (2003).

38 Cowan, C. *et al.* Derivation of embryonic stem-cell lines from human blastocysts. *The New England journal of medicine* **350**, 1353-1356, doi:10.1056/NEJMsr040330 (2004).

39 Bonal, C. & Herrera, P. Genes controlling pancreas ontogeny. *The International journal of developmental biology* **52**, 823-858, doi:10.1387/ijdb.072444cb (2008).

40 Blum, B. *et al.* Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nature biotechnology* **30**, 261-264, doi:10.1038/nbt.2141 (2012).

41 Hayek, A. & Beattie, G. Experimental transplantation of human fetal and adult pancreatic islets. *The Journal of clinical endocrinology and metabolism* **82**, 2471-2476 (1997).

42 Beattie, G., Butler, C. & Hayek, A. Morphology and function of cultured human fetal pancreatic cells transplanted into athymic mice: a longitudinal study. *Cell transplantation* **3**, 421-426 (1994).

43 Coordinators, N. R. Database resources of the National Center for Biotechnology Information. *Nucleic acids research*, doi:10.1093/nar/gks1189 (2012).

44 Rezania, A. *et al.* Maturation of Human Embryonic Stem Cell-Derived Pancreatic Progenitors into Functional Islets Capable of Treating Pre-existing Diabetes in Mice. *Diabetes*, doi:10.2337/db11-1711 (2012).

45 Dalgin, G. *et al.* Zebrafish mnx1 controls cell fate choice in the developing endocrine pancreas. *Development (Cambridge, England)* **138**, 4597-4608, doi:10.1242/dev.067736 (2011).

46 Sander, M. *et al.* Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas. *Development (Cambridge, England)* **127**, 5533-5540 (2000).

47 Yang, Y.-P., Thorel, F., Boyer, D., Herrera, P. & Wright, C. Context-specific α- to-β-cell reprogramming by forced Pdx1 expression. *Genes & development* **25**, 1680-1685, doi:10.1101/gad.16875711 (2011).

48 Shih, D., Navas, M., Kuwajima, S., Duncan, S. & Stoffel, M. Impaired glucose homeostasis and neonatal mortality in hepatocyte nuclear factor 3alpha-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 10152-10157, doi:10.1073/pnas.96.18.10152 (1999).

49 Petri, A. *et al.* The effect of neurogenin3 deficiency on pancreatic gene expression in embryonic mice. *Journal of molecular endocrinology* **37**, 301-316, doi:10.1677/jme.1.02096 (2006).

50 Collombat, P. *et al.* Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes & development* **17**, 2591-2603, doi:10.1101/gad.269003 (2003).

51 Aguayo-Mazzucato, C. *et al.* Mafa expression enhances glucose-responsive insulin secretion in neonatal rat beta cells. *Diabetologia* **54**, 583-593, doi:10.1007/s00125-010-2026-z (2011).

52 Obermüller, S. *et al.* Defective secretion of islet hormones in chromogranin-B deficient mice. *PloS one* **5**, doi:10.1371/journal.pone.0008936 (2010).

53 Gelling, R. *et al.* Lower blood glucose, hyperglucagonemia, and pancreatic alpha cell hyperplasia in glucagon receptor knockout mice. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1438-1443, doi:10.1073/pnas.0237106100 (2003).

54 Ingelsson, E. *et al.* Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. *Diabetes* **59**, 1266-1275, doi:10.2337/db09-1568 (2010).

55 Schroeder, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology* **7**, 3, doi:10.1186/1471-2199-7-3 (2006).

56 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562-578, doi:10.1038/nprot.2012.016 (2012).

57 Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**, 511-515, doi:10.1038/nbt.1621 (2010).

58 Flicek, P. *et al.* Ensembl 2012. *Nucleic acids research* **40**, 90, doi:10.1093/nar/gkr991 (2012).

59 Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics (Oxford, England)* **22**, 1540-1542, doi:10.1093/bioinformatics/btl117 (2006).

Chapter 5

Discussion

The work presented in this dissertation began in August 2008, one decade after the derivation of the first human embryonic stem cell line [1] and less than a year after the derivation of human induced pluripotent stem cells [2-4]. Differentiation of hESCs provided the first opportunity to study human development. The unique *in vitro* setting of differentiation enabled the use of genetic modification and high throughput screening applications for the dissection cell states and lineage commitment decisions. The ultimate goal was the generation of unlimited quantities of any human cell type. Our lab and many others, academia and industry alike, set on a goal to create functional human β cells from hPSCs *in vitro* for disease modeling and cell transplantation therapy into diabetic patients. I joined Dr. Melton's lab two years after the publication of the first directed differentiation

protocol to produce INS$^+$ cells [5] and only two months following the first rescue of diabetic mice using hESC-derived pancreatic progenitors [6]. Despite these landmark achievements, no one had been able to generate functional human β cells from hPSCs *in vitro*.

Early on, technical reproducibility of directed differentiation protocols became one of the greatest challenges. Consequently, many research groups including ours invested several years optimizing the technical intricacies of directed differentiation. As a result, many reports have improved the efficiency of generation of hPSC-INS$^+$ cells [7-12]. One group independently confirmed rescue of diabetic mice by transplanting hESC-derived pancreatic progenitors [11]. However, despite many independent attempts no one has still, to our knowledge, generated *in vitro* hPSC-derived functional human β cells capable of autonomously regulating serum glucose.

Several challenges remain on the path to achieve that goal. 1) Are we on the right track: It remains unclear whether immature and non-functional hPSC-INS$^+$ cells can further differentiate into mature β cells following transplantation. 2) How to make more β cells: More efficient generation of human β cells is needed to achieve better efficacy following transplantation into patients. 3) The ability of directed differentiation to robustly recapitulate human development has been called into question due to large variations in directed differentiation across cell lines and the inability to generate functional β cells. 4) How to generate functional β cells from hPSCs: A genome-wide comparison between hPSC-INS$^+$ and functional β cells may provide a roadmap for the generation of functional hPSC-derived β cells *in vitro*. Work presented in this dissertation aims to address (directly

or indirectly) all the above-mentioned questions through a series of studies conducted over the past 4 years. Each of the challenges is discussed below. Although we have not been able to generate functional human β cells, we present a series of findings that advance our understanding of directed differentiation and benefit future efforts to generate functional human cell types from human pluripotent stem cells.

5.1. Competence of *in vitro* hPSC-INS$^+$ cells to generate functional β cells

hPSC-derived pancreatic progenitors differentiate into functional human β cells 16 weeks after transplantation under the mouse kidney capsule or fat pad [6,11]. *In vivo* transplantation and maturation is the only known method to generate functional human β cells from hPSCs. *In vitro* hPSC-INS$^+$ cells on the other hand, appear to be non-functional fetal-like polyhormonal cells [5]. Transplantation of sorted *in vitro* hPSC-INS$^+$ or sorted *in vitro* hPSC-derived endocrine cells generates grafts containing only human α cells [10,13]. Consequently, it is commonly believed that hPSC-INS$^+$ do not have the competence to differentiate into functional β cells and that improving the efficiency of hPSC-INS$^+$ cells will only contribute to the generation of α cells and be detrimental to the *in vivo* differentiation into functional β cells.

Our work challenges this view and provides evidence to the contrary. We find that PKC inhibitors increase the numbers of endocrine progenitors and hPSC-INS$^+$ cells before transplantation and result in increased amounts of c-peptide secretion from grafts following *in vivo* maturation. These findings suggest that improving the efficiency of hPSC-

INS$^+$ cells is beneficial to the *in vivo* differentiation into functional β cells. This interpretation is supported by data published in two independent studies. In the first, Rezania et al. transplant pancreatic progenitors treated with Alk5 inhibitor and untreated controls. Progenitors treated with Alk5 inhibitor contained 4-fold more endocrine cells (>40% of the transplanted cells) than the control, untreated cells. Grafts derived from Alk5 treated progenitors released over 5 fold higher levels of c-peptide than grafts derived from control, untreated cells 12 and 16 weeks following transplantation [11]. A second study transplanted CD142 enriched pancreatic progenitors and non-enriched controls containing higher proportions of endocrine cells and endocrine progenitors. 10 weeks following transplantation, the endocrine-depleted transplants contained 3-fold lower levels of circulating human c-peptide [13]. Together, multiple studies suggest a positive correlation between the amount of endocrine cells present at transplant and levels of human c-peptide 10-16 weeks following transplantation. Since INS$^+$ cells present 12 weeks after transplant appear functional and express adult β cell markers, we propose that hPSC-INS$^+$ cells differentiate into functional β cells when transplanted together with pancreatic progenitors and without dissociation.

Several alternative explanations include: 1) that hPSC-INS$^+$ cells facilitate the differentiation of pancreatic progenitors into functional β cells; 2) that hPSC-INS$^+$ cells facilitate the survival or vascularization of the graft following transplantation; 3) that higher proportion of hPSC-INS$^+$ cells prevents the differentiation of transplanted cells into non-endocrine lineages; 4) that chemical treatment with Alk5 inhibitor, PKC inhibitor or the

process of sorting cells effect the outcome of the graft independent of the presence of immature hPSC-INS$^+$ cells.

Lineage tracing experiments are required to determine whether the *in vivo* matured β cells arise directly from the NGN3$^+$ or INS$^+$ cells present at the time of transplantation.

5.2. Generation of more endocrine cells

Directed differentiation protocols are customized to a particular hPSC-line to generate the highest efficiencies of differentiation. Several recently published protocols generate INS$^+$ cells with efficiencies over 10% and as high as 25% [7,9,10]. However, these protocols cannot achieve similar efficiencies using different hPSC-lines. We have therefore hypothesized that additional pathways may be involved in the differentiation of pancreatic progenitors to pancreatic endocrine cells. Two independent high content chemical screens uncovered PKC inhibitors as inducers of the endocrine lineage in two hESC-lines. The number of endocrine cells increased up to 3 fold. Conversely, we found that PKC agonists block the formation of endocrine cells.

Previously our lab identified PKC agonists as inducers of pancreatic progenitors [12]. Together we suggest a dynamic role of the PKC pathway in pancreatic differentiation. The role of PKC signaling in pancreatic development remains unknown.

In the context of other published work, our study suggests a complex interplay between pathways involved in the process of differentiating pancreatic progenitors to endocrine cells. To date, TGFβ inhibition, BMP inhibition, Notch inhibition and now

PKC inhibition have been shown to increase the differentiation toward pancreatic endocrine cells [6,14-16]. However, it is hard to imagine mechanisms by which an embryo would provide several specific inhibitory signals to induce the formation of endocrine cells. Our discovery of the dynamic role for PKC and the observation that PKC agonists block endocrine induction even in the presence of TGFβ inhibitors, which were previously described as sufficient to induce the endocrine lineage, lead us to consider an alternative hypothesis. Pancreatic progenitors could be maintained and self-renewed in the developing pancreatic epithelium by several growth factor signals, some resulting in PKC activation. The absence of these signals may lead to cell cycle arrest and provide permissive conditions for spontaneous differentiation towards endocrine cells. Modulating cell cycle and blocking multiple growth factors signaling pathways may therefore result is more efficient differentiation towards pancreatic endocrine cells.

Interestingly, TGFβ inhibitors and PKC inhibitors added to pancreatic progenitors promote the formation of dense ridge-like and island-like cell structures. These denser cell structures appear to be the source of new endocrine progenitors, indicating a potential requirement for epithelial arrangements of pancreatic progenitors in endocrine induction. Indeed, during mouse development $NGN3^{+}$ cells emerge from epithelial cords, and migrate into the surrounding mesenchyme. Directed differentiation efforts could further explore the requirements of epithelium formation prior to endocrine induction to improve the efficiency of generation of endocrine cells.

The specification of different endocrine cell subtypes remains a mystery. Recent work suggests that a single mouse NGN3$^+$ endocrine progenitor can only give rise to one type of endocrine cell [17]. Pathways that control endocrine subtype specification remain unknown. Our high-content chemical screen used both insulin and glucagon immunofluorescence as a primary assay to additionally uncover pathways which modulate endocrine subtype specification. While PKC antagonists increased the numbers of both insulin and glucagon-expressing cells, PKA activator forskolin increased the numbers of insulin$^+$ relative to glucagon$^+$ cells (data not shown). Since current directed differentiation protocols generate several fold more endocrine cells than insulin$^+$ cells, understanding endocrine subtype specification could greatly improve the efficiency of generation of insulin-expressing cells.

In addition to providing a permissive or inductive environment for pancreatic endocrine differentiation, it is important to consider that not all PDX1$^+$ pancreatic progenitors may be competent to generate endocrine cells. Despite the generation of >90% PDX1$^+$ pancreatic progenitors, only a fraction become endocrine or INS$^+$ cells. Indeed, data not presented here suggests that viral overexpression of NGN3 in hESC-derived PDX1$^+$ pancreatic progenitors is not sufficient to commit all NGN3-expressing pancreatic progenitors to endocrine cells. Multiple genetic markers may be required to identify pancreatic progenitors competent to differentiate into INS$^+$ cells.

Since sorted hPSC-INS$^+$ cells differentiate into glucagon$^+$ cells *in vivo*, one may conclude that generating INS$^+$ cells *in vitro* is detrimental for the *in vivo* differentiation into

functional β cells [10]. Our work identified that PKC inhibitors increase the number of endocrine cells *in vitro* and following *in vivo* maturation generate significantly higher levels of human c-peptide, possibly sufficient to rescue diabetic mice. We suggest that *in vitro* endocrine induction has a positive effect on the development on functional β cells and therefore anticipate further studies aimed at producing *in vitro* hPSC-INS$^+$ cells with higher efficiencies.

5.3. Is directed differentiation robust and does it generate cell types present during human development?

Reproducibility of results across hPSC-lines is of critical importance for hPSC research. Recently, much attention had been given to the description and explanation of differences between various hESC and hiPSC-lines [18]. These differences result in large variations in propensity of differentiation between lines. Beyond the practical implications, the observations raise great concerns about the reproducibility of findings across hPSC-lines. Specifically, the degree to which cells of the same cell type produced from different hESC and hiPSC-lines resemble each other, has not been known. To address this question we developed a Method for Analyzing RNA following Intracellular Sorting (MARIS). Using MARIS, we show a very high similarity (R^2 = 0.95) in global gene expression profile of sorted INS$^+$ cells from two hESC-lines and one hiPSC-line. This result provides the first confirmation of the robustness of directed differentiation protocols across multiple cells lines.

The goal of directed differentiation is to recapitulate human development *in vitro*. To date, however, experiments addressing the degree to which this has been achieved have been nearly impossible due to the scarcity of reporter cell lines and sorting methods that facilitate the purification and analysis of desired cell populations. Using MARIS, we compared hPSC-INS$^+$ cells to human adult and fetal β cells and determined a high degree of similarity between cells derived from directed differentiation and cells present during normal human development. To the degree analyzed here, our data indicate that hPSC-INS$^+$ cells resemble human fetal β cells and not human adult β cells.

These observations provide experimental support for the notion of epigenetically favorable cell-states during development and differentiation [19,20]. Despite the differences and stochastic variation in the differentiation of various pluripotent stem cells, differentiated cells appear to fall into stable states similar to those present during development.

5.4. Generation of functional β cells from hPSCs *in vitro*

Generating functional hPSC-derived β cells *in vitro* would constitute a major breakthrough in stem cells biology and diabetes treatment. There are two possible models for the generation of mature β cells. In the first, immature hPSC-INS$^+$ cells can differentiate into mature β cells. In the second, hPSC-derived pancreatic progenitors require further differentiation into progenitors competent to give rise to mature β cells. Irrespective of the differentiation model, understanding the differences between immature

and mature β cells can elucidate the biological pathways and key genetic factors which underlie the process of maturation. In order to understand maturation, we carried out three studies: 1) The comparison of mouse fetal and early postnatal (E18.5/P1) β cells with pre-weaning and adult (P10, P21) β cells; 2) The comparison of immature human week 16 fetal β cells to human adult β cells and 3) the comparison of hPSC-INS^{+} cells to adult human β cells.

5.4.1. Mouse maturation

The term maturation is applied to developmental processes involved in the generation of adult β cells. Consequently and confusingly, many studies have been published on mouse β cell maturation. In general they describe chronological changes to insulin-expressing cells at different time points between their first appearance at E9.5 and adult stages. Together, these changes can collectively be called maturation. However, it is important to recognize that maturation, much like earlier development, involves multiple distinct stages.

The formation of insulin-expressing cells, like all pancreatic endocrine cells, requires the expressing of NGN3 and its downstream effectors [21,22]. The first insulin-expressing cells appear in the pancreatic epithelium at E9.5. At this stage most insulin-expressing cells are polyhormonal, also expressing glucagon [23]. Lineage tracing suggests these polyhormonal cell may not differentiate into adult endocrine cells [24]. At E13.5, following the fusion of the ventral and dorsal pancreatic bud, a large wave of single

hormonal insulin-expressing cells appears. This is referred to as the secondary transition. A subset of INS^+ cells begin to express MafA, a transcription factor necessary for normal β cell function. [25,26]. The remainder of embryonic development is characterized by vascularization, β cell replication and aggregation of endocrine cells into islet-like structures [27]. Shortly after birth, pups are required to autonomously regulate blood glucose levels for the first time. This period is characterized by a decrease in MafB and Pax4 expression [25,28], increase in sympathetic innervation and islet capsule formation [29].

Our work functionally defined the differences between early postnatal (P1) and mid/late postnatal (P9/P21) β cells. This stage of maturation is characterized by an increase in glucose threshold for insulin secretion and an increase in the amount of insulin secreted in response to high glucose. UCN3 was discovered as a molecular marker distinguishing β cells at E18.5 and P1 with mid weaning and adult β cells.

5.4.2. Human maturation

The degree to which human β cell maturation resembles mouse β cell maturation remains unknown. Human fetal β cells between weeks 9-16 of gestation appear functionally immature [30,31]. Over 30% of human fetal insulin-expressing cells between weeks 9-16 also express glucagon and lack expression of MAFA, PDX1 and NKX6-1 suggesting a pre-secondary transition stage [32]. However, cells are organized into islet like structures suggesting of later stages of development. Adult human β cells also differ from adult mouse

β cells. They express MAFB and compared to mice secrete higher levels of insulin in baseline glucose and lower levels on insulin upon glucose stimulation [33].

To further understand differences between mouse and human fetal and adult β cells, we used MARIS. Our data comparing global gene expression changes during mouse and human development suggests a large difference between the two species. There were approximately as many genes that are expressed in both mouse and human β cells, as they are genes that are expressed in only one species and not the other. Further, gene expression changes between E18.5 and adult mouse β cells do not resemble the gene expression changes between fetal week 16 and adult human β cells. A possible explanation might be that we are comparing two different stages of β cell maturation. Alternatively, we may have uncovered intrinsic developmental differences between mice and humans. Understanding the differences between mouse and human biology is critical to translational medical research. Study of additional stages during human and mouse development using MARIS is necessary to elucidate the answer to this important question.

5.4.3. Differences between hPSC-INS$^+$ cells and adult β cells

Using MARIS we have sorted and transcriptionally profiled hPSC-INS$^+$ cells from multiple cell lines. Microarray analysis, confirmed by RNA-Seq revealed genes differentially expressed between hPSC-INS$^+$ cells and adult β cells. Gene expression patterns indicate that week 16 human fetal β cells may present an intermediate step between hPSC-INS$^+$ cells and adult β cells.

Overexpression and knockdown studies of differentially expressed transcription factors may indicate which transcription factors are necessary or sufficient to generate hPSC-INS$^+$ cells more closely resembling functional β cells.

5.4.4. In vivo matured hPSC-derived functional β cells

Transplantation of hPSC-derived pancreatic progenitors into immunodeficient mice generates functional β cells [6,11]. The *in vivo* graft may presents a permissive environment or provide instructive signals for the generation of functional β cells. Immunofluorescent staining of grafts isolated at different stages following transplantation suggests that most markers of mature β cells are absent 1 month after transplantation and only appear 3 months following transplantation. These include the prevalence of single hormonal INS$^+$ cells and their co-expression with PDX1, NKX6-1, MAFA and PC1/3. 6 months following transplantation, the graft secretes sufficient levels of insulin to autonomously maintain normal blood glucose levels in diabetic mice.

It is however unclear whether the functional hPSC-derived INS$^+$ cells resemble adult human β cells, or perhaps a later stage of human fetal or early postnatal development. MARIS can be used to address this question. Additionally, it would be informative to monitor transcriptional changes in INS$^+$ cells at different stages during *in vivo* maturation. Together with transcriptional profiles of human fetal and adult β cells the combined dataset may generate a complete roadmap for the *in vitro* maturation of hPSC-derived INS$^+$ cells.

5.5. Applications of MARIS

Most cell types cannot be isolated due to the absence of specific cell surface markers and dyes. Instead they can be identified and isolated based on the expression of intracellular markers. Intracellular staining requires permeabilization, chemical fixation and use of reagents that degrade RNA, hindering downstream analysis. MARIS was developed to combine intracellular immunofluorescent staining and FACS with the isolation of almost intact RNA.

Although in essence MARIS is a very simple method, it has the potential to revolutionize our understanding of biology by allowing us to probe gene expression in defined previously inaccessible cell types. We anticipate MARIS eventually becoming a common lab technique for any molecular biology lab.

Presented are a few potential applications. In the context of human biology, MARIS may allow us to expand the ENCODE project to provide detailed reference transcriptional data of all human cells types. In the context of disease, MARIS may be used to isolate and analyze specific cell types affected by disease or protected from disease (pancreatic β cells surviving the autoimmune attack). MARIS may provide a new understanding for the evolution of cell types by associating genomic differences with transcriptional differences in specific cell types across species. Single cell transcriptome analysis is uncovering previously unappreciated heterogeneity between cells. Optimizations of MARIS to single cell analysis would provide the ultimate tool to probe gene expression

of any defined single cell. The list of potential applications is vast. We therefore speculate that MARIS may become a routine research technique across biological disciplines.

5.6. Future directions

5.6.1. Cell transplantation therapy

The main goal of directing the differentiation of hPSCs to β cells is to provide a new source of β cells for transplantation therapy. We anticipate two major areas of active research in the years to come: 1) Generation of β-like cells capable of regulating blood glucose immediately following transplantation and 2) Long-term protection of the graft from immune attack and protection of the host from rare and unexpected complications with the graft (e.g. tumor formation).

Currently, functional β cells can be generated 3 months following transplantation into immuno-compromised mice. As researchers attempt to produce functional, adult-like β cells or islets *in vitro*, it will be critical to test their ability to maintain blood glucose levels following transplantation. We therefore anticipate that time-to-insulin-independence following transplantation will be become a critical measure of success. The gold standard, human adult islets are capable of normalizing blood glucose levels within days following transplantation into diabetic mice. Creating functional, adult-like β cells *in vitro* may require further understanding of β cell metabolism, electrophysiology and the role of islet structure. The task is rendered more difficult by the absence of cell culture conditions for

the maintenance of functional β cells *in vitro*. Researchers may therefore increasingly rely on transplantation as a functional readout for generating β cells. Additionally, the quantity of insulin secretion will be critical to success, as the transplanted β cells need to secrete sufficient insulin to maintain normoglycemia.

Many challenges still remain concerning the immune rejection of adult islets following the Edmonton protocol [34]. Similar challenges are likely to affect the transplantation of hPSC-derived cells. Additionally, there is considerable risk to tumor formation following transplantation of hPSC-derived cells. To address safety concerns, cells may be encapsulated in a nutrient permeable device, which would prevent the migration of engrafted cells outside of the capsule. However, such encapsulation may hinder vascularization and cell survival following transplantation. Moving from bench to bedside will require addressing these and several other challenges, including ethical conversations surrounding the use of human embryonic stem cells.

Despite the challenges, there is a great and growing need to generate a sustainable source of human β cells for cell transplantation therapy. This challenge can only be solved through highly interdisciplinary collaborative projects across academia and the commercial sector. During my dissertation work, I was very privileged to be a part of one such highly collaborative and goal driven effort to create functional human β cells from hPSCs.

5.6.2. Disease modeling

Isolation of patient specific hiPSC lines allows for generation of type I diabetic, type II diabetic or mature onset diabetes of the young (MODY) β cells. Disease related phenotypes may be studied *in vitro* or inside a humanized mouse where multiple cell types from the same patient can be introduced and allowed to interact in the same host. Additionally, *in vitro* differentiation of hPSCs is amenable to high throughput screening and powerful genetic engineering tools. Gene editing approaches and hiPSCs allow us to test the role of specific disease loci found in genome-wide association studies (GWAS). Transplantation of patient-specific hiPSC-derived β cells along with other relevant hiPSC-derived tissues (thymus and bone marrow) into immuno-compromised mice may generate "humanized mice" and allow the study of complex human disease such as type I diabetes [35]. Together, these approaches may elucidate the cellular mechanisms behind different forms of diabetes and help identify ways to prevent or reverse the course of disease.

5.6.3. Study human development

Our work and work of others show significant differences between genes expressed in mouse and human β cells. In order to understand human development and disease, we therefore cannot solely rely on model organisms and must expand the study of human cells. MARIS allows us to analyze gene expression patterns of previously inaccessible cells in developing fetuses and adults.

hPSCs have opened a window into the study of human development. Genetic reporter lines allow the isolation, transplantation and lineage tracing experiments addressing, among other questions, the developmental competence of human progenitors. Gene knockout studies may uncover novel genes implicated in human development and physiology. Grafts containing hPSC-derived cells can be used as an *in vivo* model for the study of human physiology, homeostasis, drug response, replication, regeneration, organ formation, disease progression etc.

Experiences with hPSC directed differentiation lead some researchers to question whether the *in vitro* system is sufficiently robust and whether its recapitulates human development. Our work provides evidence to suggest that hPSC-directed differentiation is both robust and representative of human development. We anticipate pluripotent stem cell differentiation becoming a powerful tool in the study of human biology in the years to come.

5.7. Conclusion

In closing, this dissertation addressed several key challenges concerning the generation of functional β cells from hPSCs. First, using high-content chemical screening we uncovered a dynamic role of PKC during pancreatic differentiation. PKC inhibitors increased the number of INS^+ cells *in vitro* by up to 10-fold and generated 8-fold higher levels of serum c-peptide 12 weeks following transplantation. Next, we defined the process of early postnatal mouse β cell maturation and discovered UCN3 as a molecular marker of

mouse maturation. We developed a Method to Analyze RNA following Intracellular Sorting (MARIS) and carried out the first transcriptome wide analysis of sorted human fetal and adult β cells. The small overlap in genes marking mouse and human β cell maturation indicates significant inter-species differences and the need to further study human development. Genome-wide transcriptional analysis of sorted INS^+ cells derived from two hESC-lines and one hiPSC-line shows for the first time that different cell lines produce highly similar INS^+ cells. Further, we demonstrate that non-functional hPSC-derived INS^+ cells resemble human fetal β cells, which are distinct from adult β cells. Our gene expression analysis of *in vitro* derived INS^+ cells, human fetal and adult β cells may provide a roadmap for the conversion of hPSC-derived INS^+ cells into functional β cells.

5.8. References

1 Thomson, J. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)* **282**, 1145-1147 (1998).

2 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872, doi:10.1016/j.cell.2007.11.019 (2007).

3 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.)* **318**, 1917-1920, doi:10.1126/science.1151526 (2007).

4 Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141-146, doi:10.1038/nature06534 (2008).

5 D'Amour, K. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nature biotechnology* **24**, 1392-1401, doi:10.1038/nbt1259 (2006).

6 Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nature biotechnology* **26**, 443-452, doi:10.1038/nbt1393 (2008).

7 Zhang, D., Jiang, W., Shi, Y. & Deng, H. Generation of pancreatic islet cells from human embryonic stem cells. *Science in China. Series C, Life sciences / Chinese Academy of Sciences* **52**, 615-621, doi:10.1007/s11427-009-0095-3 (2009).
8 Tateishi, K. *et al.* Generation of insulin-secreting islet-like clusters from human skin fibroblasts. *The Journal of biological chemistry* **283**, 31601-31607, doi:10.1074/jbc.M806597200 (2008).
9 Jiang, W. *et al. In vitro* derivation of functional insulin-producing cells from human embryonic stem cells. *Cell research* **17**, 333-344, doi:10.1038/cr.2007.28 (2007).
10 Nostro, M. *et al.* Stage-specific signaling through TGFβ family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development (Cambridge, England)* **138**, 861-871, doi:10.1242/dev.055236 (2011).
11 Rezania, A. *et al.* Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. *Diabetes* **61**, 2016-2029, doi:10.2337/db11-1711 (2012).
12 Chen, S. *et al.* A small molecule that directs differentiation of human ESCs into the pancreatic lineage. *Nature chemical biology* **5**, 258-265, doi:10.1038/nchembio.154 (2009).
13 Kelly, O. *et al.* Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nature biotechnology* **29**, 750-756, doi:10.1038/nbt.1931 (2011).
14 Jensen, J. *et al.* Control of endodermal endocrine development by Hes-1. *Nature genetics* **24**, 36-44, doi:10.1038/71657 (2000).
15 Tulachan, S. *et al.* TGF-beta isoform signaling regulates secondary transition and mesenchymal-induced endocrine development in the embryonic mouse pancreas. *Developmental biology* **305**, 508-521, doi:10.1016/j.ydbio.2007.02.033 (2007).
16 Rezania, A. *et al.* Production of functional glucagon-secreting α-cells from human embryonic stem cells. *Diabetes* **60**, 239-247, doi:10.2337/db10-0573 (2011).
17 Desgraz, R. & Herrera, P. Pancreatic neurogenin 3-expressing cells are unipotent islet precursors. *Development (Cambridge, England)* **136**, 3567-3574, doi:10.1242/dev.039214 (2009).
18 Bock, C. *et al.* Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-452, doi:10.1016/j.cell.2010.12.032http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1038/nbt.1783 (2011).
19 Kauffman, S. Homeostasis and differentiation in random genetic control networks. *Nature* **224**, 177-178, doi:10.1038/224177a0 (1969).
20 Waddington, C. H. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology : With an Appendix by H. Kacser*. (G. Allen & Unwin, 1957).
21 Gu, G., Dubauskaite, J. & Melton, D. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development (Cambridge, England)* **129**, 2447-2457 (2002).

22 Gradwohl, G., Dierich, A., LeMeur, M. & Guillemot, F. neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1607-1611 (2000).

23 Teitelman, G., Alpert, S., Polak, J., Martinez, A. & Hanahan, D. Precursor cells of mouse endocrine pancreas coexpress insulin, glucagon and the neuronal proteins tyrosine hydroxylase and neuropeptide Y, but not pancreatic polypeptide. *Development (Cambridge, England)* **118**, 1031-1039 (1993).

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

25 Artner, I. *et al.* MafA and MafB regulate genes critical to beta-cells in a unique temporal manner. *Diabetes* **59**, 2530-2539, doi:10.2337/db10-0190 (2010).

26 Zhang, C. *et al.* MafA is a key regulator of glucose-stimulated insulin secretion. *Molecular and cellular biology* **25**, 4969-4976, doi:10.1128/mcb.25.12.4969-4976.2005 (2005).

27 Brissova, M. *et al.* Pancreatic islet production of vascular endothelial growth factor-a is essential for islet vascularization, revascularization, and function. *Diabetes* **55**, 2974-2985, doi:10.2337/db06-0690 (2006).

28 Dohrmann, C., Gruss, P. & Lemaire, L. Pax genes and the differentiation of hormone-producing endocrine cells in the pancreas. *Mechanisms of development* **92**, 47-54, doi:10.1016/s0925-4773(99)00324-x (2000).

29 Cabrera-Vásquez, S., Navarro-Tableros, V., Sánchez-Soto, C., Gutiérrez-Ospina, G. & Hiriart, M. Remodelling sympathetic innervation in rat pancreatic islets ontogeny. *BMC developmental biology* **9**, 34, doi:10.1186/1471-213x-9-34 (2009).

30 Beattie, G., Butler, C. & Hayek, A. Morphology and function of cultured human fetal pancreatic cells transplanted into athymic mice: a longitudinal study. *Cell transplantation* **3**, 421-425 (1994).

31 Hayek, A. & Beattie, G. Experimental transplantation of human fetal and adult pancreatic islets. *The Journal of clinical endocrinology and metabolism* **82**, 2471-2475 (1997).

32 Riedel, M. *et al.* Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia* **55**, 372-381, doi:10.1007/s00125-011-2344-9 (2012).

33 Dai, C. *et al.* Islet-enriched gene expression and glucose-induced insulin secretion in human and mouse islets. *Diabetologia* **55**, 707-718, doi:10.1007/s00125-011-2369-0 (2012).

34 Gibly, R. *et al.* Advancing islet transplantation: from engraftment to the immune response. *Diabetologia* **54**, 2494-2505, doi:10.1007/s00125-011-2243-0 (2011).

35 Shultz, L., Brehm, M., Garcia-Martinez, J. & Greiner, D. Humanized mice for immune system investigation: progress, promise and challenges. *Nature reviews. Immunology* **12**, 786-798, doi:10.1038/nri3311 (2012).