

# Organogenesis and functional genomics of the endocrine pancreas

Nuria C. Bramswig · Klaus H. Kaestner

Received: 11 August 2011 / Revised: 4 December 2011 / Accepted: 29 December 2011 / Published online: 13 January 2012  
© Springer Basel AG 2012

**Abstract** Functional genomics, the analysis of the wealth of data produced by genome-wide analyses of gene expression, protein–protein, and protein–DNA interactions, has revolutionized biomedical research. Our ability to determine global gene expression profiles, transcription factor-binding sites, and histone modification maps using microarray-based technologies and next-generation sequencing applications has greatly enhanced our understanding of gene regulatory networks and the molecular wiring diagrams of cells and tissues. The organogenesis of the endocrine pancreas involves numerous signaling events within the endoderm-derived pancreatic epithelium and the surrounding mesenchyme, as well as complex transcription factor networks. Detailed understanding of the differentiation process from foregut endoderm to mature endocrine cells has enabled the rational design of in vitro differentiation protocols that coax embryonic stem cells into  $\beta$ -like cells that might enable cell replacement therapy for diabetes in the future. In this review, we summarize the research studies that have utilized genomic tools to elucidate endocrine pancreatic organogenesis.

**Keywords** Functional genomics · Gene expression · Transcription factors · Histone modifications · Endocrine pancreatic organogenesis · In vitro differentiation

## Introduction

The pancreas, a complex organ consisting of both exocrine and endocrine compartments, is critical for nutrient digestion and blood glucose homeostasis. The endocrine pancreas is organized into the islets of Langerhans, which constitute less than 2% of the mass of the pancreas, and comprise five endocrine cell types: the  $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\epsilon$ , and pancreatic polypeptide (PP) cells. The  $\alpha$  cells produce glucagon and  $\beta$  cells produce insulin, two hormones that act in an opposing fashion to maintain blood glucose homeostasis. Glucagon mobilizes glucose from peripheral tissues and therefore elevates blood glucose levels to prevent severe hypoglycemia during fasting, whereas insulin stimulates glucose storage and lowers blood glucose levels in the postprandial state.

In diabetes, there is an insufficient amount of insulin, which in type 1 diabetes is caused by autoimmune destruction of pancreatic  $\beta$  cells and absolute insulin deficiency, and in type 2 diabetes is the result of decreased peripheral insulin sensitivity and relative insulin deficiency. The resulting hyperglycemia can lead to severe complications such as stroke, heart attack, and renal failure. The incidence of diabetes, especially type 2, has increased dramatically over the past 30 years, even among adolescents [1], and it has been estimated that by the year 2030, about 439 million adults worldwide will be affected by the disease [2]. Because both  $\alpha$  and  $\beta$  cells of the endocrine pancreas are central to the control of glucose homeostasis, understanding their biology and genesis is critical for the future development of new treatment paradigms, including cell replacement therapy.

Eleven years ago, following the establishment of the so-called ‘Edmonton protocol’ for islet transplantation for severely ill type 1 diabetics, pancreatic endocrine cell

---

N. C. Bramswig · K. H. Kaestner (✉)  
Department of Genetics, Institute for Diabetes,  
Obesity and Metabolism, Perelman School of Medicine,  
University of Pennsylvania, 12-126 Translational Research  
Center, 3400 Civic Center Boulevard,  
Philadelphia, PA 19104, USA  
e-mail: kaestner@mail.med.upenn.edu

differentiation and proliferation became an intense research focus [3]. Through a novel immunosuppression regimen, cadaveric islets transplanted into a recipient liver via the portal vein could survive for years and normalize blood sugar levels, even allowing complete insulin independence. However, given the permanent shortage of organ donors and the increasing incidence of type 1 diabetes [4], the exploration of novel sources of  $\beta$  (and possibly  $\alpha$ ) cells became an urgent research endeavor. With a detailed understanding of pancreatic development, the in vitro production of functional  $\beta$  cells from human embryonic stem cells (hESC) to produce an infinite source of functional  $\beta$  cells for diabetes treatment has developed into a realistic possibility [5, 6].

The field of functional genomics focuses on the genome-wide evaluation of gene transcription, gene translation, and DNA–protein interactions [7, 8]. About 15 years ago, the sequencing of the first genomes of model organisms (such as yeast) and the development of DNA array-based technologies (such as microarrays for genome-wide expression analysis and later assays for genome-wide location of transcription factor-binding sites) revolutionized the field [9, 10]. The application of hierarchical clustering analysis made it possible to analyze the large data sets resulting from expression microarrays in a comprehensive fashion by assembling genes into groups of similar and different expression patterns [11] and elucidating tissue- and development-specific gene regulation modules. Even though these techniques have transformed biomedical research, there are certain limitations to the microarray approach because the coverage of the genome is restricted, and because this hybridization-based method has limited sensitivity [12]. The development of new sequencing technologies, including ultra high-throughput sequencing strategies such as the sequencing of mRNA (RNA-Seq) and chromatin immunoprecipitation followed by sequencing (ChIP-Seq), has overcome these restrictions, and has enabled biomedical researchers to analyze global changes in gene expression and identify transcription factor-binding sites across the entire genome with extremely high accuracy and sensitivity [12, 13].

Transcription factor networks control the maintenance of the pluripotent state in stem cells and the progression of development toward the  $\alpha$  and  $\beta$  cell fates [14, 15]. Therefore, global gene expression analysis and changes in transcription factor occupancy during the time course of development provide valuable information for in vitro differentiation protocols from hESC or other cell sources toward mature  $\alpha$  and  $\beta$  cells. Gene regulatory networks are controlled by the binding of transcription factors to promoter or enhancer regions of other transcription factors and numerous additional target genes that execute enzymatic and structural functions within the cell [16]. The ChIP-Seq

method, which will be discussed in detail below, can be applied to determine transcription factor-binding sites and to derive complex maps of histone modifications [17], thereby determining the precise epigenetic landscape of different tissues and cell types, developmental stages, and disease states. This allows for the identification of gene regulatory networks; for example, those that control development and function of the endocrine pancreas, and their epigenetic regulation [18, 19]. In-depth analysis of these networks is crucial for the understanding of physiology and pathophysiology of  $\alpha$  and  $\beta$  cells, and can provide important clues for new diabetes treatment options. This review will focus on the organogenesis of the endocrine pancreas, and summarize how functional genomics has contributed to the understanding of its development and function.

## Overview of mammalian endocrine pancreatic organogenesis

Historically, the embryonic origin of the endocrine pancreas was a source of controversy. Because pancreatic endocrine cells share many features (such as hormone secretion and electrical excitability) with neuroendocrine cells like the chromaffin cells of the adrenal medulla, scientists in the 1970s and 1980s suggested that pancreatic endocrine cells are of neural crest origin [20]. However, embryological and genetic lineage tracing experiments have established beyond doubt that the endocrine pancreas—just like its exocrine counterpart—is derived from the embryonic endoderm, which also gives rise to the gut and liver, among other organs [21].

Complex signaling networks control the proper induction of pancreatic endoderm from pre-pancreatic endoderm and the formation of the dorsal and the ventral pancreatic buds. These signaling events include the repression of sonic hedgehog (*Shh*) signaling in the pre-pancreatic endoderm through inhibitory signals such as FGF2 and activin  $\beta$  emanating from the notochord, which facilitates the formation of the dorsal pancreatic bud [22]. Furthermore, inductive signaling from the neighboring endothelium, and the mesenchyme via FGF10 for instance, is crucial for dorsal pancreatic budding and the activation of Ptf1a, a transcription factor critical for pancreas development, in the dorsal pancreatic endoderm [23, 24]. Interestingly, the emergence of the ventral pancreatic bud is controlled by different mechanisms. This was clearly demonstrated through genetic models in which only the development of the ventral pancreatic Anlage was affected. For example, *Hhex*<sup>−/−</sup> embryos lack a ventral pancreatic bud, and careful investigation of *Hhex*<sup>−/−</sup> embryos revealed that Hhex is crucial for the proliferation and correct positioning of the

ventral definitive endoderm, leading to induction of the ventral pancreatic fate [25]. Conversely, retinoic acid signaling is essential for the formation of the dorsal pancreas, as shown by mutations of the retinaldehyde dehydrogenase 2 (*Raldh2*) gene [26]. In addition, an exact interplay and switch of inhibiting and activating signaling through BMP and TGF $\beta$  between the 3–4 somite and 5–6 somite stages allows for the specification of liver and pancreas in adjacent domains [27].

Pancreatic organogenesis and the differentiation of progenitor cells to endocrine pancreatic cells take place in three stages, referred to as the primary, secondary, and tertiary transitions [28]. The primary transition occurs from embryonic day (E) 8.5 to E12.5 of mouse development, during which time the foregut endoderm thickens and buds into the surrounding mesenchyme, giving rise to the dorsal and the ventral pancreatic Anlagen [28]. The two buds contain pancreatic progenitor cells (E9), which generate a small number of early hormone-positive cells. Most of these elaborate glucagon (E9.5), but some also stain for insulin (E10.5) [21, 28, 29].

Interestingly, first-wave insulin/glucagon co-expressing cells do not seem to contribute to the mature islet, at least in the mice where this question has been studied using genetic lineage tracing [30]. However, the contribution of other first-wave hormone-positive cells to the adult islet still remains uncertain. The first wave of endocrine cell differentiation is accompanied by the beginning of pancreatic epithelial branching (E11.5) and followed by a second, more extensive, wave of endocrine cell differentiation during the secondary transition (E13–E16.5). This major wave of endocrine cell differentiation gives rise to glucagon-expressing  $\alpha$  cells, insulin-expressing  $\beta$  cells, somatostatin-expressing  $\delta$  cells, ghrelin-expressing  $\epsilon$  cells and pancreatic polypeptide expressing cells [29, 31, 32]. The proportion of endocrine cells produced varies with developmental time;  $\alpha$  cells initially predominate, and then  $\beta$  cells [33]. The second-wave endocrine cells originate from a specialized domain of multipotent progenitor cells, termed the ‘trunk’ domain, which is distinguishable molecularly around E14.5 due to the expression of different markers and transcription factors [34]. Cells in the trunk region are primarily labeled by markers of endocrine precursors, and are thought to differentiate into duct and endocrine cells, whereas cells in the tip region present labeling by exocrine markers and give rise to acinar cells [34]. The tertiary and final transition phase (E16.5 to birth) includes islet cell proliferation, migration of differentiated cells, and formation of the islets of Langerhans, the final endocrine structure [35].

In summary, the development of the pancreas is an elaborate process in which all pancreatic cell types (endocrine, acinar, and duct cells) are derived from the epithelium, itself a descendant of the endoderm germ layer

[36]. Extensive genetics research has established the indispensable role in pancreatic organogenesis of several transcription factors, such as neurogenin 3 (*Ngn3*), the forkhead box proteins A1 and A2 (*Foxa1* and *Foxa2*), regulatory factor  $\times 6$  (*Rfx6*), pancreatic and duodenal homeobox 1 (*Pdx1*), the nk2 and nk6 homeobox factors (*Nkx2.2* and *Nkx6.1*), and pancreas-specific transcription factor 1 (*Ptf1a*) [37–50]. While some of the crucial transcription factor genes are still active and have known functions in subsets of mature endocrine pancreatic cells, for example *Pdx1* [40], other essential developmental factors, such as *Ngn3*, exhibit much reduced levels of expression in differentiated pancreatic cells [51]. Thus, some transcription factors play roles in both endocrine development and maintenance of endocrine cell function, while others are only relevant during specific developmental stages.

During the last 20 years, gene targeting in the mouse (and more recently, functional genomics) have been vital for gaining in-depth knowledge about the transcriptional networks governing endocrine pancreatic development and maintenance of mature endocrine cells. Analysis of mice carrying null, point, and conditional mutations in transcription factor genes helped to establish a hierarchy of activation and function of these factors by establishing their relevance during development and in adult function. In addition, as our understanding of early endocrine cell markers deepens and cell sorting strategies are optimized, the characterization of the transcriptome of cells at early developmental stages and the comparison to mature  $\beta$  cells using DNA-microarray or RNA-Seq technology can provide important insights about the temporal importance of each transcription factor, and provide crucial knowledge for the directed differentiation of hESC (as well as other cell types) into functional  $\beta$  cells.

## Methods in functional genomics

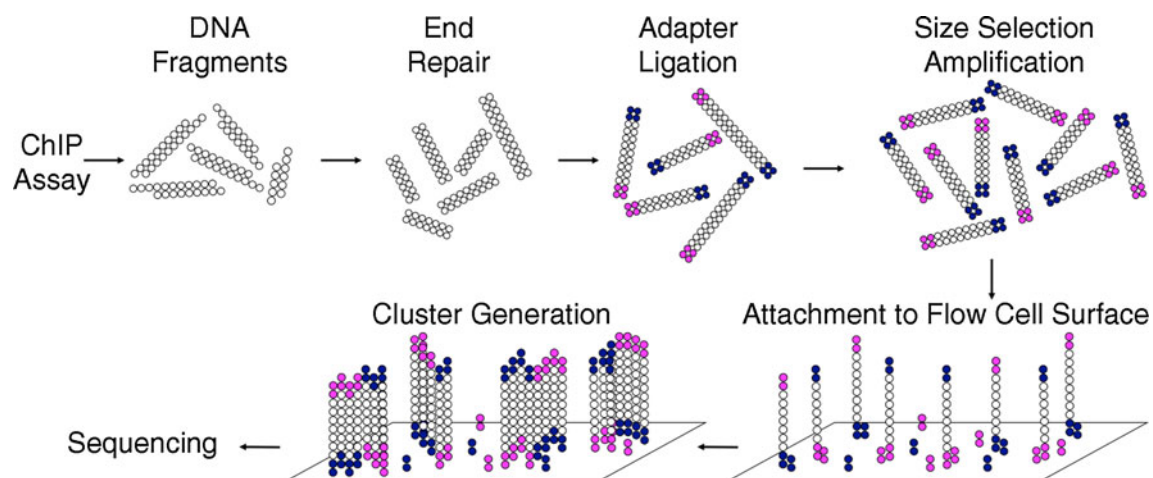
### Microarray-based technologies

The invention of microarray technology during the mid-1990s enabled for the first time massively parallel analyses of steady-state mRNA levels for thousands of genes in a process later termed ‘expression profiling.’ Briefly, short single-stranded DNA sequences are printed or synthesized on a glass slide, and one or two cDNA samples (labeled separately with fluorescent dyes) are hybridized to this ‘microarray’ slide containing many thousand DNA probes [9]. After extensive washing to remove excess of the labeled sample, the resulting fluorescence intensity at each spot on the array is determined using a high-resolution fluorescence scanner. The fluorescence intensity is then a

direct measure of the relative abundance of the corresponding mRNA in the sample. Though this method initially was used only for global mRNA expression analysis, its utility was greatly expanded through the invention of ‘ChIP-on-chip’ by Ren and colleagues in 2000 [52]. ChIP-on-chip combines chromatin immunoprecipitation and microarray analysis to make possible the simultaneous analysis of transcription factor-binding sites at thousands of loci in a given genome [52]. Labeling of an immunoprecipitated sample and a control sample (not immunoprecipitated/enriched) with two different fluorescent dyes, followed by hybridization to and scanning of this DNA-microarray allows for detection of DNA–protein interaction sites. Two limitations of the ChIP-on-chip technology, however, eventually became apparent. First, because the size of the targets on the array is relatively large and represent genomic areas containing gaps, the spatial resolution of binding site determination was limited. Second, and even more important, ChIP-on-chip is not truly genome-wide for species with complex genomes such as mice or humans, because these large genomes cannot be tiled completely even on the highest density arrays. In the last 5 years, advances in DNA sequencing technology have dramatically reduced the cost and time required for large-scale sequencing, such that at the time of this writing, a human genome can be sequenced within a week and for less than \$5,000 [53]. As a result of these methodological advances, next-generation sequencing applications, such as RNA-Seq and ChIP-Seq, have begun to replace microarray-based expression profiling and genome-wide location analysis by ChIP-on-chip [12, 13].

## Next-generation sequencing

Next-generation sequencing applications include the RNA-Seq and ChIP-Seq methods. There are multiple ultra-high-throughput sequencing technologies available today, but we will focus on one commonly used methodology, the so-called “Solexa sequencing” technology marketed by Illumina. In this section, we focus on the ChIP-Seq application, describing it in more detail and pointing out the advantages of this technique over the formerly used ChIP-on-chip assay. In ChIP-Seq, cells or tissues are first treated with formaldehyde, which crosslinks the DNA-binding proteins or histones to the DNA. Sequencing analysis of the DNA requires random shearing of the chromatin by sonication, after which the majority of the resultant DNA fragments ideally range from about 200–500 bp in size. For some applications (in particular the analysis of histone modifications), chromatin can instead be digested with micrococcal nuclease (MNase). This enzyme digests all DNA except that which is wrapped around the nucleosome core, which is made up of the histone octamer. Following obtainance of the optimal size range, the chromatin is immunoprecipitated with an antibody, for example one against a transcription factor or a specific posttranslational histone modification. After reversal of the crosslinks, a ‘library’ is prepared that allows for sequencing analysis of the DNA fragments that were initially bound by the protein of interest. Briefly, the immunoprecipitated DNA is deproteinized, modified with adapters, size-selected, and amplified (Fig. 1). Because



**Fig. 1** Schematic overview of a ChIP-seq experiment. The ends of the immunoprecipitated and deproteinized DNA fragments need to be repaired before adapter ligation. The modified DNA is size selected and amplified. The ligated adapters bind to complementary oligonucleotides on the flow cell surface after melting of the two strands, where the single bound DNA molecules are amplified and clusters are

generated. The four nucleotides, each labeled with a different fluorescent dye, are added to the DNA molecules one base at a time, and after laser excitation a separate image is captured of each base of every cluster. The saved images are further analyzed to extract the DNA sequence of each cluster



both sonication and MNase digestion produce DNA fragments with ragged ends (either 5' or 3' overhangs of unknown length and sequence) the DNA fragments require repair to produce blunt ends. This is typically achieved using T4 DNA polymerase and Klenow polymerase in the presence of deoxynucleotides. The 5' ends of the blunted ends are then phosphorylated using polynucleotide kinase before the addition of adapters of defined sequence by DNA ligation. The DNA fragments are subsequently amplified by limited PCR. Finally, the resulting library is melted to single-strand DNA. The ligated adapters allow one single DNA molecule to bind to one of billions of randomly spaced complementary 'anchor' oligonucleotides on the flow cell surface. Limited PCR then amplifies the captured molecules to millions of clusters consisting of 100s to 1,000s of identical DNA fragments (Fig. 1). Adenosine, cytosine, guanine, and thymine, each labeled with a different fluorescent dye and blocked at the 3' end to limit DNA synthesis to one base per cycle, are incorporated using DNA polymerase at the first base of each cluster. After laser excitation, a separate image for each of the four fluorophores of the first base of each cluster is captured and saved. The fluorescent signal is then quenched, and the 3' end of the added nucleotide is deblocked. Addition of a single fluorescent base, laser excitation, image capturing, signal quenching, and deblocking of the 3' end are then repeated base by base, providing the sequence of each of the DNA fragments, which are then aligned to the reference genome. Depending upon the antibody used in the ChIP experiment, this procedure can allow for detection of transcription factor-binding sites, the determination of histone modification maps, or the determination of nucleosome positions.

The ChIP-Seq method enormously increases the accuracy of our knowledge about DNA–protein interactions. Compared to the ChIP-on-chip method, the resolution is enhanced and the genome coverage is much more comprehensive, as ChIP-Seq does not depend upon sequences spotted on a microarray [17]. Another important advantage is the low amount of DNA required for a ChIP-Seq experiment. This is especially beneficial for determining exact histone modification maps and studying DNA–protein binding events in tissues at early developmental stages, in small, sorted cell populations, and in scarce samples, such as human tissue samples. Some of the first published studies using the ChIP-Seq method included the description of histone modification profiles in CD4+T cells [54] and the mapping of neuron-restrictive silencer factor (NRSF) binding sites in the human genome [55] in 2007, and this technology has now been adopted by hundreds of laboratories world-wide.

The large volume of data that results from these applications necessitates the creation of new algorithms for proper data analysis and the determination of transcription factor targets as well as histone modification profiles.

Differences in peak type, such as strong and narrow peaks resulting from classic transcription factor binding to its sequence motif, or low and broad peaks such as those for the repressive histone 3 lysine 27 trimethylation mark (H3K27me3), determine which algorithm is suitable for use. Publicly available algorithms include CisGenome [56], Model-based Analysis of ChIP-Seq (MACS) [57], PeakSeq [58], Quantitative Enrichment of Sequence Tags (QuEST) [59], GLobal Identifier of Target Regions (GLITR) [60], and Hypergeometric Optimization of Motif EnRichment (HOMER) [61, 62], discerning characteristics of which include their peak calling criteria and the use of background data [63, 64].

The possibility of studying defined cell populations at early developmental stages and comparing transcription factor targets and the epigenetic landscape to those of mature endocrine cells greatly improves our understanding of pancreatic development, and will help us to attain the goal of producing functional  $\beta$  cells in vitro. Previously unknown enhancer and promoter sites, determined by the histone 3 lysine 4 monomethylation mark (H3K4me1) and the histone 3 lysine 4 trimethylation mark (H3K4me3), can be identified across the entire genome, elucidating the regulation of crucial genes. In the cancer research field, an extensive effort to develop epigenetic therapeutics is being mounted, as abnormal epigenetic modifications can lead to dysregulation of cell differentiation and proliferation [65]. For example, histone deacetylase inhibitors are presently being tested in clinical trials and have been approved for the treatment of several cancer types [66]. Interestingly, it has been shown that histone deacetylase inhibitors can also be used to influence pancreatic development [67]. Treatment of embryonic rat pancreata with different histone deacetylase inhibitors resulted in an increase of the endocrine progenitor population and the  $\beta$  cell pool, indicating a possible use of these inhibitors in the development of new therapeutic strategies for diabetes and improving the efficacy of in vitro techniques for differentiating stem cells into  $\beta$  cells [67, 68], and highlighting the importance of epigenetic studies in diabetes research.

The RNA-Seq method is used for precise profiling of the transcriptome using high-throughput sequencing [13]. Some of the first studies using this method were published in 2008 describing the yeast transcriptome [69], the transcriptome of mouse embryonic stem cells and embryoid bodies [70], and different mouse tissues [71]. A comparison of RNA-sequencing to microarray data showed that 81% of differentially expressed genes characterized by the microarray method were also detected with the RNA-Seq method, conferring validity to this novel approach [72]. Though RNA-Seq offers many advantages over the microarray method, it likewise presents several challenges. It permits the detection of a large range of expression levels

with an enhanced sensitivity, and allows one to map exact gene boundaries, as well as to discover novel transcripts and rare or tissue-specific alternative splicing isoforms [13]. High sequencing depth and the production of large numbers of sequence reads, which lead to the desired high sensitivity and accuracy, also result in a vast amount of data that need to be analyzed, and present bioinformatical challenges [13]. The possibility of accurately determining the entire transcriptome, especially for transcripts at low expression levels, is extremely attractive for studies focusing on pancreatic organogenesis and can greatly improve our understanding about the differentiation process from pancreatic progenitor cells to mature  $\alpha$  and  $\beta$  cells, which promises to impact the development of novel strategies for diabetes treatment.

### Functional genomics: applications in endocrine pancreatic research

#### Global gene expression and transcription factor-binding profiles

The exciting possibility of determining global gene expression levels prompted many type 1 and type 2 diabetes research groups to perform microarray analyses on tissues involved in the pathogenesis of type 2 diabetes, such as white adipose tissue, liver, and pancreatic islets. One such analysis using Affymetrix GeneChip Mu6500 involved the comparison of gene expression levels in white adipose tissue of wild-type, ob/ob, and transgenic mice expressing low levels of leptin [73]. Other studies sought to further our understanding of the  $\beta$  cell transcriptome, for example by performing gene expression analysis on a pancreatic  $\beta$  cell line after the induction of growth arrest (using the mouse GDA system, Genome Systems Inc.) [74], and on primary rat  $\beta$  cells after cytokine-stimulation [75] in order to elucidate the signaling pathways and mechanisms involved in  $\beta$  cell dysfunction and to find targets for new treatment options. However, the arrays used in these early studies were small and not specifically designed for diabetes research, which presented limitations and prompted the development of the “PancChip,” a resource developed specifically for the diabetes research field [76]. This microarray contained 3,400 clones corresponding mostly to mRNAs expressed in the pancreas, but also including diabetes-related pathways and housekeeping genes, and was used for the study of pancreatic tissue at different time points during development and in adulthood [76]. Recognizing that the early expressed sequence tag (EST) sequencing efforts in the 1990s had completely ignored the endocrine pancreas, the NIDDK-sponsored Endocrine Pancreas Consortium set out to construct cDNA

libraries and sequence hundreds of thousands of transcripts from various stages of pancreatic development. This led not only to the discovery of several thousand novel genes, but also to the construction of a new microarray called “PancChip 4.0,” comprising 13,848 elements, with more than 10,000 unique genes, a development that greatly improved diabetes research resources at the time [77]. These and other array types have been extensively used to determine gene expression patterns and to elucidate the impact of transcription factors and signaling pathways on the differentiation process from endocrine progenitor cells to  $\alpha$  and  $\beta$  cells.

#### Early regulators of endocrine pancreatogenesis: *Pdx1* and *Ngn3*

Arguably the most important protein controlling pancreatic development is the homeobox transcription factor encoded by the *Pdx1* gene. Loss of *Pdx1* (termed *IPF-1* in humans) results in pancreatic agenesis in mice and humans alike, and heterozygous mutations in the human *IPF-1/PDX1* gene lead to the development of MODY4 [38–40, 78–81]. Gene expression profiling using a small custom microarray (95 genes spotted on Affymetrix GMS417 arrayer) and 60 single epithelial cells isolated from dorsal pancreatic buds at murine E10.5 showed the presence of six different gene expression patterns, the majority of which express *Pdx1*, and some of which coexpress pancreatic hormones, suggesting that even the early dorsal pancreatic bud is made up of a heterogeneous cell population [82]. Gu and colleagues analyzed the transcriptional profiles at four different stages of murine pancreas development, including unspecified endoderm at E7.5, FACS-sorted *Pdx1*-expressing cells at E10.5, *Ngn3*-expressing cells at E13.5, and mature islets of Langerhans, by microarray analysis utilizing Mu11K, Mu74Av1, and Mu74Av2 Affymetrix arrays [83]. They detected 193, 60, 71, and 217 genes with enriched expression at the respective stages, and the discovery of pathways not previously implicated in pancreatic development, such as the ‘myelin transcription factor 1’ (*Myt1*) gene which was found to be expressed in endocrine progenitors. Inhibiting the function of *Myt1* using a dominant negative *Myt1* under the *Ngn3*-promoter led to a reduction of insulin- and glucagon-expressing cells at E14.5, illustrating the utility of this “discovery research” for establishing new, testable hypotheses on gene function in pancreas development.

To decipher the pathways regulated by *Pdx1* in early pancreatic progenitor cells and to elucidate the mechanisms by which pancreatic organogenesis is prevented, global gene expression analysis was performed using two different microarray types (containing ~15,000 and ~21,000 clones from different sources) on control and *Pdx1*<sup>-/-</sup>

dorsal pancreatic buds at E10.5 [84]. The analyses revealed differential expression for 111 genes, of which 73 genes were down- and 38 genes were up-regulated. *Pax6* and *Nkx6.1* were among the down-regulated genes, and the expression of *Foxa2* and *glucagon* remained unchanged. The relatively low number of differentially regulated genes could be due to technical difficulties and the small sample size employed, which most likely limited the outcome of this experiment [84]. Adult *Pdx1*<sup>+/-</sup> mice displayed limited  $\beta$  cell mass expansion when metabolically challenged by a high fat diet [85]. This finding was attributed to enhanced ER stress-induced apoptosis, and was examined more closely by measuring global transcript levels using microarray analysis (Mouse PancChip 6.1) in MIN6 cells after acute silencing of *Pdx1*, and also by comparing *Pdx1*<sup>+/+</sup> and *Pdx1*<sup>+/-</sup> mouse islets. Additionally, ChIP-on-chip analysis (Mouse PromoterChip BCBC-5A) was used to characterize Pdx1-binding events in MIN6 cells. The expression of two genes encoding ER proteins, *Ero1lb* and *Nnat*, was down-regulated in both MIN6 cells transduced with an adenovirus expressing a short hairpin RNA targeting Pdx1 and in *Pdx1*<sup>+/-</sup> mouse islets, implying a role for Pdx1 in ER homeostasis.

In 2007, Keller and colleagues employed a promoter microarray (Mouse PromoterChip BCBC-5B), including the promoters of the miRNA genes known at the time to determine the direct genomic targets of Pdx1 in a mouse insulinoma cell line [86]. A total of 583 new *Pdx1* targets were identified, and the transcriptional regulation of the majority of a selection of genes was validated in a cell culture system expressing a dominant negative mutant of *Pdx1*. In addition, NeuroD1 targeting events were determined, and it was discovered that Pdx1 and NeuroD1 share binding sites in 440 genes, among them the gene encoding miR-375, important for glucose-stimulated insulin secretion in MIN6 cells [87] and the maintenance of proper pancreatic endocrine cell function in mice [88]. There is growing evidence that miRNAs, a group of small non-coding RNA molecules, serve as a crucial component of the gene regulatory network and play an important role in health and disease states, including those of pancreatic  $\beta$  cells [89]. The application of functional genomics will be needed to discover unknown miRNAs, predict miRNA targets and expand our knowledge about the nature of their regulation and function.

The transcription factor Ngn3 is expressed in endocrine progenitor cells and is indispensable for the development of pancreatic endocrine cells [37]. Therefore, it was important to understand the regulatory networks that depend on Ngn3. In 2006, the microarray method was used to compare the transcriptome of *Ngn3*-deficient and wild-type dorsal murine pancreatic buds at E13 and E15, capturing the secondary transition phase of embryonic

development [90]. The microarray used in this study (ArrayTAG 20k murine gene collection spotted in duplicate on CodeLink slides) allowed for the expression analysis of 12,140 genes, but unfortunately did not include some of the genes of interest, such as *insulin*, *glucagon*, *Pdx1*, and *NeuroD1*. At either or both time points, 504 genes were differentially expressed in the *Ngn3*-deficient mice. After the application of different filtering strategies, 52 genes showed significant and robust differential expression, all characterized by decreased expression in the *Ngn3*-deficient state. Among the robustly down-regulated genes were genes important for the processing of hormones, genes encoding vesicle proteins (such as synaptotagmin 7 and 13), the hormones *ghrelin* and *neuropeptide Y*, and two transcription factors ('*Iroquois related homeobox 1*' (*Irx1*) and *Myt1*). Given the complete absence of endocrine cells in the *Ngn3*-deficient state, finding these endocrine cell-expressed transcripts reduced in the mutant pancreas was not surprising. *Irx1* and *Irx2* were co-expressed in glucagon-positive, but not insulin- or somatostatin-expressing, cells at E18, suggesting the discovery of novel transcription factors specific for glucagon-expressing cells. A more recent study analyzed mRNA expression profiles from E12.5, E15.5, and E18.5 pancreata of *Ngn3*<sup>-/-</sup> and wild-type mice using the MGU74Av2 and/or the MOE430 2.0 Affymetrix microarrays, generating lists of differentially expressed genes at each time point and comparing the expression patterns to adult mouse islets and pancreatic endocrine tumor cell lines [91]. The highest number of downregulated transcripts was detected at E18.5 (645 genes) and included *MafB*, *Nkx6.1*, *Isl1*, and *Iapp*. Juhl and colleagues point out that the insulin granule zinc transporter *Znt8* (*Slc30a8*), a gene associated with type 2 diabetes risk, was decreased by 18-fold in *Ngn3*<sup>-/-</sup> mice at E18.5, reflective of the loss of endocrine cells in the mutant tissue [91].

White and colleagues chose a different approach to determine the gene expression profiles of endocrine progenitors, which are marked by Ngn3 expression, and their direct descendants [92]. The Ngn3-enhanced green fluorescent protein (EGFP) 'knock-in' mouse model was used to sort *Ngn3*<sup>+/EGFP</sup>-positive and -negative cells from E13.5, E14.5, E15.5, E16.5, and E17.5 pancreata. RNA was isolated from *Ngn3*<sup>+/EGFP</sup>-positive cells and adult mouse islets, and subjected to microarray analysis (Mouse PancChip 6.0) to define the precise mRNA expression profiles at these various developmental stages. This led to the identification of 1,029 genes, including 237 transcription factors that are temporally regulated in endocrine cells during development, and provided an extensive transcriptional profile of the endocrine cell lineage from the progenitor to the differentiated state.

### *HNF-4 $\alpha$* , the MODY1 gene

In-depth examination of the regulatory function of diabetes-related genes, as discussed above for the MODY4 gene *Pdx1*, is critical for the discovery of new aspects and important transcriptional networks in diabetes pathogenesis. In a microarray study (Affymetrix 133A and B arrays, ~44,928 genes) analyzing the global gene expression of non-diabetic and type 2 diabetic human islets, 370 genes were found to be differentially expressed, 243 up-regulated, and 137 down-regulated in the type 2 diabetic samples [93]. A marked reduction of the *Hnf-4 $\alpha$*  (*hepatic nuclear factor 4 alpha*) transcript, the gene mutated in ‘Maturity onset diabetes of the young 1’ (MODY1), was reported in type 2 diabetic islets. The expression of genes involved in the insulin-signaling pathway, such as the *insulin receptor* (*IR*) and *AKT2* (which encodes a serine/threonine-protein kinase), were significantly reduced as well, while the hormone expression levels did not differ significantly. The transcript of the *aryl hydrocarbon receptor nuclear transporter* (*ARNT*) gene showed the most significant reduction in type 2 diabetic islets, suggesting an important function of *ARNT* in  $\beta$  cell function. Mice carrying a  $\beta$  cell-specific ablation of *Arnt* showed impaired glucose tolerance and decreased glucose-stimulated insulin secretion (more pronounced in female mice) and reduced levels of *Hnf-4 $\alpha$* , further suggesting a role of *Arnt* in the regulation of this MODY gene [93]. Global gene expression profiling of mice deficient for *Hnf-4 $\alpha$*  in  $\beta$  cells in *Hnf-4 $\alpha$ <sup>LoxP/LoxP</sup>; InsCre* mice revealed up-regulation of 128 genes and down-regulation of 57 genes using the PancChip version 5.0, 13K cDNA microarray [94]. Gene ontology analysis disclosed that a large percentage of the differentially expressed genes were annotated to play a role in cell proliferation, a previously unsuspected function of *Hnf-4 $\alpha$* . While  $\beta$  cell mass was unchanged in unstressed 4-month-old *Hnf-4 $\alpha$*  mutant mice, there was a 50% decrease in  $\beta$  cell mass in 4-month-old pregnant *Hnf-4 $\alpha$* -deficient mice, and a complete loss of the pregnancy-induced increase in  $\beta$  cell replication. The expression profiling thereby formed the basis for uncovering a novel role for *Hnf-4 $\alpha$*  in the physiological  $\beta$  cell expansion in response to proliferative stimuli. Further data mining of the expression profile suggested that *Hnf-4 $\alpha$*  contributes to  $\beta$  cell proliferation at least partly through the Ras/ERK signaling pathway.

### Forkhead and homeobox transcription factors

The forkhead box A1 and A2 proteins (Foxa1 and Foxa2), two winged-helix transcription factors, are expressed and serve important roles in the development and functional maintenance of tissues derived from the foregut endoderm,

such as the pancreas and liver [95]. The loss of Foxa2 in  $\beta$  cells and its effect on gene expression patterns in islets was studied in two different mouse models, the *Foxa2<sup>LoxP/LoxP</sup>; RIP-Cre* [45] and the tamoxifen-inducible *Foxa2<sup>LoxP/LoxP</sup>; Pdx1-CreERT2* model [96] using PancChip version 4.0 and 6.1, respectively. Microarray analysis revealed the reduced expression of *Kir6.2* and *Sur1*, encoding the two subunits of the ATP-dependent K<sup>+</sup> channel, in *Foxa2* mutant islets from both studies. The latter study determined differential expression of 143 genes, among them genes that encode proteins for vesicle trafficking, granule biosynthesis, and exocytosis, highlighting the importance of Foxa2 in insulin secretion and therefore mature  $\beta$  cell function [96]. The additional simultaneous deletion of the related *Foxa1* gene in *Foxa1<sup>LoxP/LoxP</sup>; Foxa2<sup>LoxP/LoxP</sup>; Pdx1-CreERT2* mice led to a more severe phenotype with impaired glucose homeostasis with significantly altered expression levels of 566 genes in *Foxa1/a2*-mutant islets, of which 294 were down-regulated and mainly involved in metabolic processes and ion transport, and 272 were up-regulated and found to be annotated to function in neuronal differentiation (Whole Mouse Genome Oligo Microarray G4122A; Agilent Technologies) [97]. The up-regulation of neuronal genes was surprising and uncovered a repressive role of the Foxa factors in the regulation of neuronal genes in mature  $\beta$  cells. The carbohydrate response element-binding protein (ChREBP), a main regulator of carbohydrate metabolism, showed a 13-fold reduction in mRNA expression levels, which prompted Gao and colleagues to further investigate the role of this factor and its association with the Foxa factors during development. Foxa2 ChIP-Seq analysis and Foxa1 and Foxa2 ChIP-assays determined Foxa1 and Foxa2 binding regions in the first intron of the *ChREBP* gene from E13.5 onward, with the highest chromatin occupancy at E18.5. This discovery of a direct regulation of the *ChREBP* gene by Foxa1 and Foxa2 suggested a novel role of the two Foxa factors in the regulation of metabolic processes associated with ChREBP [97]. In 2008, an additional study on the regulatory functions of Foxa1 and Foxa2 with respect to *Pdx1* was performed using ChIP and ChIP-Seq analyses on fetal pancreas and adult islets [98]. Deletion of *Foxa1* and *Foxa2* starting at E9.0 in *Foxa1<sup>LoxP/LoxP</sup>; Foxa2<sup>LoxP/LoxP</sup>; Pdx1-CreEarly* mice resulted in dramatic pancreatic hypoplasia, severe hyperglycemia, and early postnatal lethality. Loss of *Pdx1* expression in the *Foxa1/a2* mutant mice suggested direct or indirect regulation of *Pdx1* by the Foxa factors. ChIP and ChIP-Seq analyses demonstrated binding of both Foxa1 and Foxa2 at the *Pdx1* locus. Interestingly, these experiments showed that the location of the Foxa1 and Foxa2 binding differed according to developmental stage, with preferred binding at the Area IV of the *Pdx1* enhancer during fetal life, and increased binding at the Area I-II-III



enhancer in adult islets. These findings proved that Foxa1 and Foxa2 regulate *Pdx1* and highlighted the importance of the two winged-helix transcription factors in pancreatic development.

The homeobox transcription factor Nkx2.2 is important for endocrine cell differentiation, particularly terminal  $\beta$  cell differentiation, as *Nkx2.2*-null mutant mice completely lack  $\beta$  cells but display small numbers of  $\alpha$  and PP cells [47]. Microarray analysis (using PancChip 6 containing 13,059 cDNAs) of control and *Nkx2.2*-null mutant pancreata at E12.5 and E13.5 was employed to determine differential gene expression patterns and elucidate the mechanism of endocrine cell specification arrest that occurs during development [99]. Differential expression of 65 genes was observed at E12.5 (49 down-regulated and 16 up-regulated genes) and an additional 15 genes were down-regulated at E13.5. Among the down-regulated genes were nine transcription factors, including ones that encode for early endocrine transcription factors such as Ngn3 and MafB. In addition, it was shown that deletion of Nkx2.2 leads to differential expression of exocrine genes, such as the upregulation of *Elastase 1* and the downregulation of *Spink3*, the functional repercussions of which are yet to be elucidated [99].

#### MafA, MafB, and NeuroD1

MafA and MafB belong to the basic leucine-basic transcription factor family and are dynamically regulated during pancreatic development [100]. Global gene expression analysis using Mouse PancChip 6.1 was performed on E18.5 murine pancreata from wild-type, *MafB*<sup>-/-</sup>, pancreas-wide *MafA* (*MafA*<sup>ΔPanc</sup>) and *MafA*<sup>ΔPanc</sup>; *MafB*<sup>-/-</sup> mutants, and revealed differing expression patterns. The authors concluded that MafB regulates genes important for mature  $\beta$  cell function, including insulin secretion and glucose sensing through *Nnat* and *Slc2a2*, but not for early  $\beta$  cell development, since transcription factors important for early endocrine differentiation, such as Ngn3, were not differentially expressed [100].

NeuroD1-deficient mice are severely diabetic at birth and die perinatally, with NeuroD1-deficient pancreata displaying a reduction of endocrine cells, especially  $\beta$  cells [101]. The importance of NeuroD1 for the development and function of  $\beta$  cells was studied utilizing two gene ablation models deleting *NeuroD1* at the beginning of  $\beta$  cell formation, and in adult  $\beta$  cells after tamoxifen-induced activation of Cre recombinase [102]. Deletion of *NeuroD* in  $\beta$  cells resulted in impaired glucose tolerance despite maintained expression of Insulin 2, suggesting that a defect in insulin secretion was responsible for the observed phenotype. Microarray analysis (Mouse PancChip 6.0) revealed the differential gene expression of 68 genes in  $\beta$

cell-specific NeuroD-deficient adult islets, among them genes encoding glycolytic enzymes, resulting in a gene expression program reminiscent of neonatal, immature  $\beta$  cells, thereby suggesting a role of NeuroD in terminal  $\beta$  cell maturation [102].

#### Expression patterns at early developmental stages

To gain insight into the transcriptional landscape and the biology of human endodermal cells in vitro, Wang and colleagues recently analyzed hESC and in vitro derived purified Sox17<sup>+</sup> cells at stage 1 (early definitive endoderm) and stage 2 (primitive gut tube endoderm) utilizing microarray (Affymetrix Mouse Human Genome Array U133 Plus 2.0) [103]. They determined that 13,209 genes are differentially expressed between these three stages, which clustered into six groups of similar expression patterns, and they discovered new surface antigens that provide a novel sorting strategy for the endoderm at the primitive gut tube stage [103]. Of course, given the dramatic alterations in cellular phenotype produced in this in vitro differentiation system, it is not surprising that nearly half the genome was affected.

#### The transcriptome of the mature islet

Analysis of gene expression patterns in mature islets promises to yield important insights into their functional state, and establishes a valuable resource for the diabetes research community. Massive parallel signature sequencing (MPSS) on two human pancreatic islet samples identified 6,941 expressed genes, of which 3,552 were shared between both samples [104]. Apart from the pancreatic hormones *INS*, *GCG*, and *SST*, the most abundant transcripts included the regenerating islet-derived 3 alpha gene (*REG3A*), *REG1B*, *REG1A*, and bone morphogenetic protein 5 (*BMP5*). Gene ontology analysis showed that the top 200 transcripts were enriched for categories such as protein synthesis and metabolism, reflecting the function of islets as protein hormone secreting cells.

Martens and colleagues examined the transcriptional landscape of human  $\beta$ -cell-enriched samples, mouse islets, and rat FACS-purified  $\beta$ -cells by microarray analysis (Affymetrix HG133A and RG230A) and provided a set of 332 conserved  $\beta$  cell marker genes [105]. The conserved  $\beta$ -cell marker genes were distributed into three clusters, with 15% of the genes being  $\beta$ -cell specific, 30% shared with immune and gut mucosal cells, and 15% with neuronal tissue. These gene sets were functionally related to hormone processing, protein synthesis, and vesicle transport, respectively. Motif analysis for transcription factor-binding sites on all identified conserved genes showed an overrepresentation of established  $\beta$ -cell transcription factors,

but also muscle-related regulators, a finding that remains unexplained. In addition, the effect of fasting was assessed by comparison of the transcriptome of  $\beta$ -cells isolated from fed and 24-h-fasted rats, which showed that especially pathways involved in protein folding and their ER/Golgi processing were affected.

In 2011, the first detailed transcriptional analysis of sorted human pancreatic  $\alpha$ ,  $\beta$ , acinar, small and large duct cells was reported (using Agilent  $4 \times 44$  Whole Human Genome Array) [106]. Genes expressed 20-fold higher in  $\beta$  cells compared to  $\alpha$  cells included *INS*, *IAPP*, *estrogen receptor 1 (ESR1)*, and the solute carrier protein *SLC17A6*, whereas *GCG*, *ARX*, *IRX2*, and the glutamate receptor *GRIN3A* were expressed at least 20-fold higher in  $\alpha$  cells. Interestingly, *MAFB*, expressed in the adult mouse specifically in  $\alpha$  cells, but not  $\beta$  cells, is active in both mature human  $\alpha$  and  $\beta$  cells, revealing important differences of transcription factor expression patterns among species. Analysis of signaling events between different cell types identified 121 ligand-receptor interactions among pancreatic cell types, and suggested that Ephrin (EFN) signaling is implicated in pancreatic function.

### Epigenetics

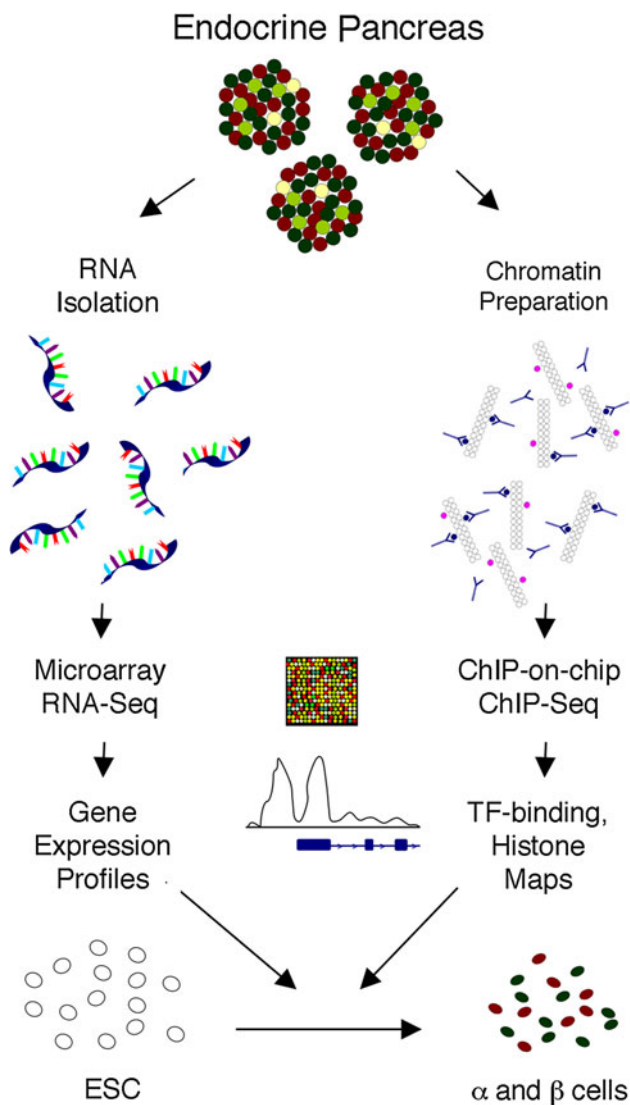
The study of epigenetics, including DNA methylation and histone modifications, has become increasingly important to many research fields [107]. The discovery that chromatin structure regulates gene transcription along with DNA replication and repair, coupled with the fact that the information stored in the chromatin structure is inheritable, prompted many research fields to study the epigenetic landscape of several tissues in different disease states and developmental stages [108]. The presence and function of different groups of histone modifications, such as acetylation, phosphorylation, and methylation, have now been characterized. Many studies have focused on the activating histone 3 lysine 4 trimethylation (H3K4me3) and the repressing H3K27me3 mark due to their regulatory role in transcription. The colocalization of the H3K4me3 and the H3K27me3 mark in the gene promoter region is referred to as a bivalent mark, and has been predominantly observed in undifferentiated cell types, such as embryonic stem cells [109, 110]. Upon differentiation, the bivalent mark is resolved at many loci via removal of the H3K27me3 mark, which results in activation of the corresponding genes.

Despite elevated interest in chromatin research and the genome-wide analysis of histone modifications, only a few related studies of the endocrine pancreas have been published to date, likely due in part to the difficulty of obtaining pure islet tissue and sorted endocrine cells. Genome-wide analysis of open chromatin regions in human pancreatic islets was performed utilizing the formaldehyde-

assisted isolation of regulatory elements with high-throughput sequencing (FAIRE-Seq) approach [111]. Approximately 80,000 regions were identified and after comparison to open chromatin sites of non-islet human cell-lines, about 3,300 were implicated as islet-specific, one of which contained a TCF7L2 variant, which is linked to type 2 diabetes [111]. Bhandare and colleagues performed ChIP-Seq analysis on human pancreatic islets and provided the first genome-wide histone modification map for the H3K4me1, H3K4me2, H3K4me3, and H3K27me3 marks in human pancreatic islets [18]. Comparison with H3K4me3 levels in CD4<sup>+</sup> T cells [54] revealed that a group of promoters exhibited tissue-specific enrichment for histone modification marks. Interestingly, a set of developmental regulators showed bivalent marks, previously thought to be characteristic of pluripotent and not differentiated cells. Furthermore, H3K4me3 enrichment levels did not always correlate with gene expression levels, but were dependent upon the specific promoter structure. While promoters with CpG islands exhibited the expected loose correlation of H3K4me3 enrichment with mRNA expression, promoters without CpG islands, such as those for *Insulin* and *Glucagon*, were not marked strongly by H3K4me3 even though they are very active in islet cells. Future studies on sorted endocrine cell populations will undoubtedly increase the resolution and power of these epigenetic analyses.

Genome-wide analysis of H3K4me1, H3K4me3, and H3K79me2 marked histones and CTCF-binding sites was performed in human pancreatic islets identifying ~18,000 putative promoters, approximately 30% of which were not previously annotated and some of which were islet-specific [112]. This study confirmed earlier findings indicating that promoters of highly transcribed genes, such as the *Insulin* and *Glucagon*, were not strongly marked by the standard activating histone modifications, suggesting alternative activating mechanisms [112].

Gene expression and histone modification (H3K4me3 and H3K27me3) profiles of different mouse tissue types (including neural, adipose, and liver) as well as pancreatic  $\beta$  and acinar cells, were determined by microarray (Affymetrix Mouse Genome 430 2.0 Array) and ChIP-Seq analyses [19]. Interestingly, the global gene expression and active H3K4me3 profiles of pancreatic  $\beta$  cells clustered with the profiles of neural tissues, and not with the acinar cell profiles. However, the repressive H3K27me3 profile of  $\beta$  cells was most similar to acinar cells, pancreatic progenitor cells, and liver tissue, which is reflective of their common embryonic origin. In concordance with one of the epigenetic studies on human pancreatic islets [18], bivalent marks were also observed in genes of differentiated tissues although they were more common in pluripotent cells [19]. These epigenetic studies provide genome-wide histone



**Fig. 2** Functional genomics in endocrine pancreatic research. The analysis of global gene expression profiles, genome-wide transcription factor-binding and histone modification maps has begun to enhance the knowledge about the organogenesis of the pancreas and provide important clues for the improvement of in vitro differentiation protocols from hESC to functional  $\alpha$  and  $\beta$  cells

modification profiles of human and mouse pancreatic tissue. They present interesting findings on the comparison of these profiles with gene expression levels and with the epigenetic landscape of other tissues, and highlight that it is crucial to investigate a combination of histone modifications and other transcription factor and coregulator binding events in order to fully elucidate the molecular mechanisms directing cell differentiation and other vital cell functions.

#### Summary and future perspectives

In summary, functional genomics technologies have complemented gene loss-of-function and gain-of-function

experiments throughout the last decade and have greatly improved our understanding of the transcriptional networks governing  $\alpha$  and  $\beta$  cell differentiation and pancreatic organogenesis and have helped to refine the in vitro differentiation protocols that are being deployed to coax hESC toward the  $\beta$  cell fate (Fig. 2). Development of the more sensitive RNASeq method and the ability to sequence RNA obtained from only a few hundred cells, and in the future possibly from single cells, will further increase this understanding and allow for global gene expression analysis even at the earliest developmental stages. In addition, these technical improvements will enable researchers to address the critical question of whether  $\alpha$  and  $\beta$  cells are homogeneous or composed of distinct subpopulations. The combination of precise gene expression profiles and detailed histone modification maps will provide insight into the molecular biology of distinct endocrine cell populations at early and late developmental stages and enhance our understanding of endocrine pancreatic organogenesis.

**Acknowledgements** We thank Amber Riblett for careful copy-editing of the manuscript. Related work in the Kaestner lab has been supported through NIH grants DK088383, DK055342, and DK089529.

#### References

1. D'Adamo E, Caprio S (2011) Type 2 diabetes in youth: epidemiology and pathophysiology. *Diabetes Care* 34(Suppl 2):S161–S165. doi:[10.2337/dc11-s212](https://doi.org/10.2337/dc11-s212)
2. Shaw JE, Sicree RA, Zimmet PZ (2010) Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract* 87(1):4–14. doi:[10.1016/j.diabres.2009.10.007](https://doi.org/10.1016/j.diabres.2009.10.007)
3. Shapiro AM, Lakey JR, Ryan EA, Korbutt GS, Toth E, Warnock GL, Kneteman NM, Rajotte RV (2000) Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *N Engl J Med* 343(4):230–238. doi:[10.1056/NEJM200007273430401](https://doi.org/10.1056/NEJM200007273430401)
4. Vehik K, Dabelea D (2011) The changing epidemiology of type 1 diabetes: why is it going through the roof? *Diabetes Metab Res Rev* 27(1):3–13. doi:[10.1002/dmrr.1141](https://doi.org/10.1002/dmrr.1141)
5. D'Amour KA, Bang AG, Eliazar S, Kelly OG, Agulnick AD, Smart NG, Moorman MA, Kroon E, Carpenter MK, Baetge EE (2006) Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat Biotechnol* 24(11):1392–1401. doi:[10.1038/nbt1259](https://doi.org/10.1038/nbt1259)
6. Kroon E, Martinson LA, Kadoya K, Bang AG, Kelly OG, Eliazar S, Young H, Richardson M, Smart NG, Cunningham J, Agulnick AD, D'Amour KA, Carpenter MK, Baetge EE (2008) Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat Biotechnol* 26(4):443–452. doi:[10.1038/nbt1393](https://doi.org/10.1038/nbt1393)
7. McKusick VA (1997) Genomics: structural and functional studies of genomes. *Genomics* 45(2):244–249
8. Rastan S, Beeley LJ (1997) Functional genomics: going forwards from the databases. *Curr Opin Genet Dev* 7(6):777–783
9. Brown PO, Botstein D (1999) Exploring the new world of the genome with DNA microarrays. *Nat Genet* 21(1 Suppl):33–37. doi:[10.1038/4462](https://doi.org/10.1038/4462)

10. Bulyk ML (2006) DNA microarray technologies for measuring protein-DNA interactions. *Curr Opin Biotechnol* 17(4):422–430. doi:[10.1016/j.copbio.2006.06.015](https://doi.org/10.1016/j.copbio.2006.06.015)
11. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95(25):14863–14868
12. Mardis ER (2007) ChIP-seq: welcome to the new frontier. *Nat Methods* 4(8):613–614. doi:[10.1038/nmeth0807-613](https://doi.org/10.1038/nmeth0807-613)
13. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1):57–63. doi:[10.1038/nrg2484](https://doi.org/10.1038/nrg2484)
14. Oliver-Krasinski JM, Stoffers DA (2008) On the origin of the beta cell. *Genes Dev* 22(15):1998–2021. doi:[10.1101/gad.1670808](https://doi.org/10.1101/gad.1670808)
15. Pesce M, Scholer HR (2001) Oct-4: gatekeeper in the beginnings of mammalian development. *Stem Cells* 19(4):271–278. doi:[10.1634/stemcells.19-4-271](https://doi.org/10.1634/stemcells.19-4-271)
16. Lee TI, Young RA (2000) Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* 34:77–137. doi:[10.1146/annurev.genet.34.1.77](https://doi.org/10.1146/annurev.genet.34.1.77)
17. Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 10(10):669–680. doi:[10.1038/nrg2641](https://doi.org/10.1038/nrg2641)
18. Bhandare R, Schug J, Le Lay J, Fox A, Smirnova O, Liu C, Naji A, Kaestner KH (2010) Genome-wide analysis of histone modifications in human pancreatic islets. *Genome Res* 20(4):428–433. doi:[10.1101/gr.102038.109](https://doi.org/10.1101/gr.102038.109)
19. van Arensbergen J, Garcia-Hurtado J, Moran I, Maestro MA, Xu X, Van de Castele M, Skoudy AL, Palassini M, Heimberg H, Ferrer J (2010) Derepression of Polycomb targets during pancreatic organogenesis allows insulin-producing beta-cells to adopt a neural gene activity program. *Genome Res* 20(6):722–732. doi:[10.1101/gr.101709.109](https://doi.org/10.1101/gr.101709.109)
20. Le Douarin NM (1988) On the origin of pancreatic endocrine cells. *Cell* 53(2):169–171
21. Slack JM (1995) Developmental biology of the pancreas. *Development* 121(6):1569–1580
22. Hebrok M, Kim SK, Melton DA (1998) Notochord repression of endodermal Sonic hedgehog permits pancreas development. *Genes Dev* 12(11):1705–1713
23. Yoshitomi H, Zaret KS (2004) Endothelial cell interactions initiate dorsal pancreas development by selectively inducing the transcription factor Ptf1a. *Development* 131(4):807–817. doi:[10.1242/dev.00960](https://doi.org/10.1242/dev.00960)
24. Jacquemin P, Yoshitomi H, Kashima Y, Rousseau GG, Lemaigre FP, Zaret KS (2006) An endothelial-mesenchymal relay pathway regulates early phases of pancreas development. *Dev Biol* 290(1):189–199. doi:[10.1016/j.ydbio.2005.11.023](https://doi.org/10.1016/j.ydbio.2005.11.023)
25. Bort R, Martinez-Barbera JP, Beddington RS, Zaret KS (2004) Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. *Development* 131(4):797–806. doi:[10.1242/dev.00965](https://doi.org/10.1242/dev.00965)
26. Martin M, Gallego-Llamas J, Ribes V, Keding M, Niederreither K, Chambon P, Dolle P, Gradwohl G (2005) Dorsal pancreas agenesis in retinoic acid-deficient Raldh2 mutant mice. *Dev Biol* 284(2):399–411. doi:[10.1016/j.ydbio.2005.05.035](https://doi.org/10.1016/j.ydbio.2005.05.035)
27. Wandzioch E, Zaret KS (2009) Dynamic signaling network for the specification of embryonic pancreas and liver progenitors. *Science* 324(5935):1707–1710. doi:[10.1126/science.1174497](https://doi.org/10.1126/science.1174497)
28. Pictet RL, Clark WR, Williams RH, Rutter WJ (1972) An ultrastructural analysis of the developing embryonic pancreas. *Dev Biol* 29(4):436–467
29. Herrera PL, Huarte J, Sanvito F, Meda P, Orci L, Vassalli JD (1991) Embryogenesis of the murine endocrine pancreas; early expression of pancreatic polypeptide gene. *Development* 113(4):1257–1265
30. Herrera PL (2000) Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. *Development* 127(11):2317–2322
31. Prado CL, Pugh-Bernard AE, Elghazi L, Sosa-Pineda B, Sussel L (2004) Ghrelin cells replace insulin-producing beta cells in two mouse models of pancreas development. *Proc Natl Acad Sci USA* 101(9):2924–2929. doi:[10.1073/pnas.0308604100](https://doi.org/10.1073/pnas.0308604100)
32. Pictet R, Rutter W (1972) Development of embryonic endocrine pancreas. In: Steiner DF, Feinkel N (eds) *Handbook of physiology*. Washington, DC
33. Johansson KA, Dursun U, Jordan N, Gu G, Beermann F, Gradwohl G, Grapin-Botton A (2007) Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Dev Cell* 12(3):457–465. doi:[10.1016/j.devcel.2007.02.010](https://doi.org/10.1016/j.devcel.2007.02.010)
34. Zhou Q, Law AC, Rajagopal J, Anderson WJ, Gray PA, Melton DA (2007) A multipotent progenitor domain guides pancreatic organogenesis. *Dev Cell* 13(1):103–114. doi:[10.1016/j.devcel.2007.06.001](https://doi.org/10.1016/j.devcel.2007.06.001)
35. Miller K, Kim A, Kilimnik G, Jo J, Moka U, Periwal V, Hara M (2009) Islet formation during the neonatal development in mice. *PLoS One* 4(11):e7739. doi:[10.1371/journal.pone.0007739](https://doi.org/10.1371/journal.pone.0007739)
36. Gu G, Dubauskaite J, Melton DA (2002) Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* 129(10):2447–2457
37. Gradwohl G, Dierich A, LeMeur M, Guillemot F (2000) Neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc Natl Acad Sci USA* 97(4):1607–1611
38. Ahlgren U, Jonsson J, Edlund H (1996) The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development* 122(5):1409–1416
39. Offield MF, Jetton TL, Labosky PA, Ray M, Stein RW, Magnuson MA, Hogan BL, Wright CV (1996) PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* 122(3):983–995
40. Stoffers DA, Heller RS, Miller CP, Habener JF (1999) Developmental expression of the homeodomain protein IDX-1 in mice transgenic for an IDX-1 promoter/lacZ transcriptional reporter. *Endocrinology* 140(11):5374–5381
41. Krapp A, Knofler M, Ledermann B, Burki K, Berney C, Zoerkler N, Hagenbuchle O, Wellauer PK (1998) The bHLH protein PTF1-p48 is essential for the formation of the exocrine and the correct spatial organization of the endocrine pancreas. *Genes Dev* 12(23):3752–3763
42. Smith SB, Qu HQ, Taleb N, Kishimoto NY, Scheel DW, Lu Y, Patch AM, Grabs R, Wang J, Lynn FC, Miyatsuka T, Mitchell J, Seerke R, Desir J, Eijnden SV, Abramowicz M, Kacet N, Weill J, Renard ME, Gentile M, Hansen I, Dewar K, Hattersley AT, Wang R, Wilson ME, Johnson JD, Polychronakos C, German MS (2010) Rfx6 directs islet formation and insulin production in mice and humans. *Nature* 463(7282):775–780. doi:[10.1038/nature08748](https://doi.org/10.1038/nature08748)
43. Kaestner KH, Katz J, Liu Y, Drucker DJ, Schutz G (1999) Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression in vivo. *Genes Dev* 13(4):495–504
44. Shih DQ, Navas MA, Kuwajima S, Duncan SA, Stoffel M (1999) Impaired glucose homeostasis and neonatal mortality in hepatocyte nuclear factor 3alpha-deficient mice. *Proc Natl Acad Sci USA* 96(18):10152–10157
45. Lantz KA, Vatamaniuk MZ, Brestelli JE, Friedman JR, Matschinsky FM, Kaestner KH (2004) Foxa2 regulates multiple



- pathways of insulin secretion. *J Clin Invest* 114(4):512–520. doi:[10.1172/JCI21149](https://doi.org/10.1172/JCI21149)
46. Sund NJ, Vatamaniuk MZ, Casey M, Ang SL, Magnuson MA, Stoffers DA, Matschinsky FM, Kaestner KH (2001) Tissue-specific deletion of *Foxa2* in pancreatic beta cells results in hyperinsulinemic hypoglycemia. *Genes Dev* 15(13):1706–1715. doi:[10.1101/gad.901601](https://doi.org/10.1101/gad.901601)
  47. Sussel L, Kalamaras J, Hartigan-O'Connor DJ, Meneses JJ, Pedersen RA, Rubenstein JL, German MS (1998) Mice lacking the homeodomain transcription factor *Nkx2.2* have diabetes due to arrested differentiation of pancreatic beta cells. *Development* 125(12):2213–2221
  48. Henseleit KD, Nelson SB, Kuhlbrodt K, Hennings JC, Ericson J, Sander M (2005) *NKX6* transcription factor activity is required for alpha- and beta-cell development in the pancreas. *Development* 132(13):3139–3149. doi:[10.1242/dev.01875](https://doi.org/10.1242/dev.01875)
  49. Sander M, Sussel L, Connors J, Scheel D, Kalamaras J, Dela Cruz F, Schwitzgebel V, Hayes-Jordan A, German M (2000) Homeobox gene *Nkx6.1* lies downstream of *Nkx2.2* in the major pathway of beta-cell formation in the pancreas. *Development* 127(24):5533–5540
  50. Soyer J, Flasse L, Raffelsberger W, Beucher A, Orvain C, Peers B, Ravassard P, Vermot J, Voz ML, Mellitzer G, Gradwohl G (2010) *Rfx6* is an *Ngn3*-dependent winged helix transcription factor required for pancreatic islet cell development. *Development* 137(2):203–212. doi:[10.1242/dev.041673](https://doi.org/10.1242/dev.041673)
  51. Wang S, Jensen JN, Seymour PA, Hsu W, Dor Y, Sander M, Magnuson MA, Serup P, Gu G (2009) Sustained *Neurog3* expression in hormone-expressing islet cells is required for endocrine maturation and function. *Proc Natl Acad Sci USA* 106(24):9715–9720. doi:[10.1073/pnas.0904247106](https://doi.org/10.1073/pnas.0904247106)
  52. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA (2000) Genome-wide location and function of DNA binding proteins. *Science* 290(5500):2306–2309. doi:[10.1126/science.290.5500.2306](https://doi.org/10.1126/science.290.5500.2306)
  53. Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev Genet* 11(1):31–46. doi:[10.1038/nrg2626](https://doi.org/10.1038/nrg2626)
  54. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129(4):823–837. doi:[10.1016/j.cell.2007.05.009](https://doi.org/10.1016/j.cell.2007.05.009)
  55. Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316(5830):1497–1502. doi:[10.1126/science.1141319](https://doi.org/10.1126/science.1141319)
  56. Ji H, Jiang H, Ma W, Johnson DS, Myers RM, Wong WH (2008) An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat Biotechnol* 26(11):1293–1300. doi:[10.1038/nbt.1505](https://doi.org/10.1038/nbt.1505)
  57. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9):R137. doi:[10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137)
  58. Rozowsky J, Euskirchen G, Auerbach RK, Zhang ZD, Gibson T, Bjornson R, Carriero N, Snyder M, Gerstein MB (2009) Peak-Seq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat Biotechnol* 27(1):66–75. doi:[10.1038/nbt.1518](https://doi.org/10.1038/nbt.1518)
  59. Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglu S, Myers RM, Sidow A (2008) Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat Methods* 5(9):829–834. doi:[10.1038/nmeth.1246](https://doi.org/10.1038/nmeth.1246)
  60. Tuteja G, White P, Schug J, Kaestner KH (2009) Extracting transcription factor targets from ChIP-Seq data. *Nucleic Acids Res* 37(17):e113. doi:[10.1093/nar/gkp536](https://doi.org/10.1093/nar/gkp536)
  61. Lin YC, Jhunjunwala S, Benner C, Heinz S, Welinder E, Mansson R, Sigvardsson M, Hagman J, Espinoza CA, Dutkowski J, Ideker T, Glass CK, Murre C (2010) A global network of transcription factors, involving *E2A*, *EBF1* and *Foxo1*, that orchestrates B cell fate. *Nat Immunol* 11(7):635–643. doi:[10.1038/ni.1891](https://doi.org/10.1038/ni.1891)
  62. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38(4):576–589. doi:[10.1016/j.molcel.2010.05.004](https://doi.org/10.1016/j.molcel.2010.05.004)
  63. Pepke S, Wold B, Mortazavi A (2009) Computation for ChIP-seq and RNA-seq studies. *Nat Methods* 6(11 Suppl):S22–S32. doi:[10.1038/nmeth.1371](https://doi.org/10.1038/nmeth.1371)
  64. Rye MB, Saetrom P, Drablos F (2011) A manually curated ChIP-seq benchmark demonstrates room for improvement in current peak-finder programs. *Nucleic Acids Res* 39(4):e25. doi:[10.1093/nar/gkq1187](https://doi.org/10.1093/nar/gkq1187)
  65. Ellis L, Atadja PW, Johnstone RW (2009) Epigenetics in cancer: targeting chromatin modifications. *Mol Cancer Ther* 8(6):1409–1420. doi:[10.1158/1535-7163.MCT-08-0860](https://doi.org/10.1158/1535-7163.MCT-08-0860)
  66. Tsai HC, Baylin SB (2011) Cancer epigenetics: linking basic biology to clinical medicine. *Cell Res* 21(3):502–517. doi:[10.1038/cr.2011.24](https://doi.org/10.1038/cr.2011.24)
  67. Haumaitre C, Lenoir O, Scharfmann R (2008) Histone deacetylase inhibitors modify pancreatic cell fate determination and amplify endocrine progenitors. *Mol Cell Biol* 28(20):6373–6383. doi:[10.1128/MCB.00413-08](https://doi.org/10.1128/MCB.00413-08)
  68. Haumaitre C, Lenoir O, Scharfmann R (2009) Directing cell differentiation with small-molecule histone deacetylase inhibitors: the example of promoting pancreatic endocrine cells. *Cell Cycle* 8(4):536–544
  69. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320(5881):1344–1349. doi:[10.1126/science.1158441](https://doi.org/10.1126/science.1158441)
  70. Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods* 5(7):613–619. doi:[10.1038/nmeth.1223](https://doi.org/10.1038/nmeth.1223)
  71. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621–628. doi:[10.1038/nmeth.1226](https://doi.org/10.1038/nmeth.1226)
  72. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 18(9):1509–1517. doi:[10.1101/gr.079558.108](https://doi.org/10.1101/gr.079558.108)
  73. Soukas A, Cohen P, Socci ND, Friedman JM (2000) Leptin-specific patterns of gene expression in white adipose tissue. *Genes Dev* 14(8):963–980
  74. Zimmer Y, Milo-Landesman D, Svetlanov A, Efrat S (1999) Genes induced by growth arrest in a pancreatic beta cell line: identification by analysis of cDNA arrays. *FEBS Lett* 457(1):65–70
  75. Cardozo AK, Kruhoffer M, Leeman R, Orntoft T, Eizirik DL (2001) Identification of novel cytokine-induced genes in pancreatic beta-cells by high-density oligonucleotide arrays. *Diabetes* 50(5):909–920
  76. Scarse LM, Brestelli JE, McWeeney SK, Lee CS, Mazzarelli J, Pinney DF, Pizarro A, Stoeckert CJ Jr, Clifton SW, Permutt MA, Brown J, Melton DA, Kaestner KH (2002) Functional genomics of the endocrine pancreas: the pancreas clone set and *PancChip*, new resources for diabetes research. *Diabetes* 51(7):1997–2004
  77. Kaestner KH, Lee CS, Scarse LM, Brestelli JE, Arsenlis A, Le PP, Lantz KA, Crabtree J, Pizarro A, Mazzarelli J, Pinney D,

- Fischer S, Manduchi E, Stoeckert CJ Jr, Gradwohl G, Clifton SW, Brown JR, Inoue H, Cras-Meneur C, Permutt MA (2003) Transcriptional program of the endocrine pancreas in mice and humans. *Diabetes* 52(7):1604–1610
78. Jonsson J, Ahlgren U, Edlund T, Edlund H (1995) IPF1, a homeodomain protein with a dual function in pancreas development. *Int J Dev Biol* 39(5):789–798
79. Ahlgren U, Jonsson J, Jonsson L, Simu K, Edlund H (1998) beta-cell-specific inactivation of the mouse *Ip1/Pdx1* gene results in loss of the beta-cell phenotype and maturity onset diabetes. *Genes Dev* 12(12):1763–1768
80. Stoffers DA, Ferrer J, Clarke WL, Habener JF (1997) Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. *Nat Genet* 17(2):138–139. doi:[10.1038/ng1097-138](https://doi.org/10.1038/ng1097-138)
81. Stoffers DA, Zinkin NT, Stanojevic V, Clarke WL, Habener JF (1997) Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nat Genet* 15(1):106–110. doi:[10.1038/ng0197-106](https://doi.org/10.1038/ng0197-106)
82. Chiang MK, Melton DA (2003) Single-cell transcript analysis of pancreas development. *Dev Cell* 4(3):383–393
83. Gu G, Wells JM, Dombkowski D, Pfeffer F, Aronow B, Melton DA (2004) Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* 131(1):165–179. doi:[10.1242/dev.00921](https://doi.org/10.1242/dev.00921)
84. Svensson P, Williams C, Lundeberg J, Ryden P, Bergqvist I, Edlund H (2007) Gene array identification of *Ip1/Pdx1*-regulated genes in pancreatic progenitor cells. *BMC Dev Biol* 7:129. doi:[10.1186/1471-213X-7-129](https://doi.org/10.1186/1471-213X-7-129)
85. Sachdeva MM, Claiborn KC, Khoo C, Yang J, Groff DN, Mirmira RG, Stoffers DA (2009) *Pdx1* (MODY4) regulates pancreatic beta cell susceptibility to ER stress. *Proc Natl Acad Sci USA* 106(45):19090–19095. doi:[10.1073/pnas.0904849106](https://doi.org/10.1073/pnas.0904849106)
86. Keller DM, McWeeney S, Arsenlis A, Drouin J, Wright CV, Wang H, Wollheim CB, White P, Kaestner KH, Goodman RH (2007) Characterization of pancreatic transcription factor *Pdx-1* binding sites using promoter microarray and serial analysis of chromatin occupancy. *J Biol Chem* 282(44):32084–32092. doi:[10.1074/jbc.M700899200](https://doi.org/10.1074/jbc.M700899200)
87. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, MacDonald PE, Pfeffer S, Tuschl T, Rajewsky N, Rorsman P, Stoffel M (2004) A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* 432(7014):226–230. doi:[10.1038/nature03076](https://doi.org/10.1038/nature03076)
88. Poy MN, Hausser J, Trajkovski M, Braun M, Collins S, Rorsman P, Zavolan M, Stoffel M (2009) miR-375 maintains normal pancreatic alpha- and beta-cell mass. *Proc Natl Acad Sci USA* 106(14):5813–5818. doi:[10.1073/pnas.0810550106](https://doi.org/10.1073/pnas.0810550106)
89. Melkman-Zehavi T, Oren R, Kred-Russo S, Shapira T, Mandelbaum AD, Rivkin N, Nir T, Lennox KA, Behlke MA, Dor Y, Hornstein E (2011) miRNAs control insulin content in pancreatic beta-cells via downregulation of transcriptional repressors. *EMBO J* 30(5):835–845. doi:[10.1038/emboj.2010.361](https://doi.org/10.1038/emboj.2010.361)
90. Petri A, Ahnfelt-Ronne J, Frederiksen KS, Edwards DG, Madssen D, Serup P, Fleckner J, Heller RS (2006) The effect of neurogenin3 deficiency on pancreatic gene expression in embryonic mice. *J Mol Endocrinol* 37(2):301–316. doi:[10.1677/jme.1.02096](https://doi.org/10.1677/jme.1.02096)
91. Juhl K, Sarkar SA, Wong R, Jensen J, Hutton JC (2008) Mouse pancreatic endocrine cell transcriptome defined in the embryonic *Ngn3*-null mouse. *Diabetes* 57(10):2755–2761. doi:[10.2337/db07-1126](https://doi.org/10.2337/db07-1126)
92. White P, May CL, Lamounier RN, Brestelli JE, Kaestner KH (2008) Defining pancreatic endocrine precursors and their descendants. *Diabetes* 57(3):654–668. doi:[10.2337/db07-1362](https://doi.org/10.2337/db07-1362)
93. Gunton JE, Kulkarni RN, Yim S, Okada T, Hawthorne WJ, Tseng YH, Roberson RS, Ricordi C, O'Connell PJ, Gonzalez FJ, Kahn CR (2005) Loss of ARNT/HIF1beta mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes. *Cell* 122(3):337–349. doi:[10.1016/j.cell.2005.05.027](https://doi.org/10.1016/j.cell.2005.05.027)
94. Gupta RK, Gao N, Gorski RK, White P, Hardy OT, Rafiq K, Brestelli JE, Chen G, Stoeckert CJ Jr, Kaestner KH (2007) Expansion of adult beta-cell mass in response to increased metabolic demand is dependent on HNF-4alpha. *Genes Dev* 21(7):756–769. doi:[10.1101/gad.1535507](https://doi.org/10.1101/gad.1535507)
95. Friedman JR, Kaestner KH (2006) The Foxa family of transcription factors in development and metabolism. *Cell Mol Life Sci* 63(19–20):2317–2328. doi:[10.1007/s00018-006-6095-6](https://doi.org/10.1007/s00018-006-6095-6)
96. Gao N, White P, Doliba N, Golson ML, Matschinsky FM, Kaestner KH (2007) Foxa2 controls vesicle docking and insulin secretion in mature Beta cells. *Cell Metab* 6(4):267–279. doi:[10.1016/j.cmet.2007.08.015](https://doi.org/10.1016/j.cmet.2007.08.015)
97. Gao N, Le Lay J, Qin W, Doliba N, Schug J, Fox AJ, Smirnova O, Matschinsky FM, Kaestner KH (2010) Foxa1 and Foxa2 maintain the metabolic and secretory features of the mature beta-cell. *Mol Endocrinol* 24(8):1594–1604. doi:[10.1210/me.2009-0513](https://doi.org/10.1210/me.2009-0513)
98. Gao N, LeLay J, Vatamaniuk MZ, Rieck S, Friedman JR, Kaestner KH (2008) Dynamic regulation of *Pdx1* enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes Dev* 22(24):3435–3448. doi:[10.1101/gad.1752608](https://doi.org/10.1101/gad.1752608)
99. Anderson KR, White P, Kaestner KH, Sussel L (2009) Identification of known and novel pancreas genes expressed downstream of *Nkx2.2* during development. *BMC Dev Biol* 9:65. doi:[10.1186/1471-213X-9-65](https://doi.org/10.1186/1471-213X-9-65)
100. Artner I, Hang Y, Mazur M, Yamamoto T, Guo M, Lindner J, Magnuson MA, Stein R (2010) MafA and MafB regulate genes critical to beta-cells in a unique temporal manner. *Diabetes* 59(10):2530–2539. doi:[10.2337/db10-0190](https://doi.org/10.2337/db10-0190)
101. Naya FJ, Huang HP, Qiu Y, Mutoh H, DeMayo FJ, Leiter AB, Tsai MJ (1997) Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in *BETA2/neuroD*-deficient mice. *Genes Dev* 11(18):2323–2334
102. Gu C, Stein GH, Pan N, Goebbels S, Hornberg H, Nave KA, Herrera P, White P, Kaestner KH, Sussel L, Lee JE (2010) Pancreatic beta cells require *NeuroD* to achieve and maintain functional maturity. *Cell Metab* 11(4):298–310. doi:[10.1016/j.cmet.2010.03.006](https://doi.org/10.1016/j.cmet.2010.03.006)
103. Wang P, Rodriguez RT, Wang J, Ghodasara A, Kim SK (2011) Targeting *SOX17* in human embryonic stem cells creates unique strategies for isolating and analyzing developing endoderm. *Cell Stem Cell* 8(3):335–346. doi:[10.1016/j.stem.2011.01.017](https://doi.org/10.1016/j.stem.2011.01.017)
104. Kutlu B, Burdick D, Baxter D, Rasschaert J, Flamez D, Eizirik DL, Welsh N, Goodman N, Hood L (2009) Detailed transcriptome atlas of the pancreatic beta cell. *BMC Med Genomics* 2:3. doi:[10.1186/1755-8794-2-3](https://doi.org/10.1186/1755-8794-2-3)
105. Martens GA, Jiang L, Hellemans KH, Stange G, Heimberg H, Nielsen FC, Sand O, Van Helden J, Gorus FK, Pipeleers DG (2011) Clusters of conserved beta cell marker genes for assessment of beta cell phenotype. *PLoS One* 6(9):e24134. doi:[10.1371/journal.pone.0024134](https://doi.org/10.1371/journal.pone.0024134)
106. Dorrell C, Schug J, Lin CF, Canaday PS, Fox AJ, Smirnova O, Bonnahr R, Streeter PR, Stoeckert CJ Jr, Kaestner KH, Grompe M (2011) Transcriptomes of the major human pancreatic cell types. *Diabetologia* 54(11):2832–2844. doi:[10.1007/s00125-011-2283-5](https://doi.org/10.1007/s00125-011-2283-5)
107. Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. *Cell* 128(4):669–681. doi:[10.1016/j.cell.2007.01.033](https://doi.org/10.1016/j.cell.2007.01.033)
108. Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* 128(4):707–719. doi:[10.1016/j.cell.2007.01.015](https://doi.org/10.1016/j.cell.2007.01.015)
109. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R,

- Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2):315–326. doi:[10.1016/j.cell.2006.02.041](https://doi.org/10.1016/j.cell.2006.02.041)
110. Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 1(3):299–312. doi:[10.1016/j.stem.2007.08.003](https://doi.org/10.1016/j.stem.2007.08.003)
111. Gaulton KJ, Nammo T, Pasquali L, Simon JM, Giresi PG, Fogarty MP, Panhuis TM, Mieczkowski P, Secchi A, Bosco D, Berney T, Montanya E, Mohlke KL, Lieb JD, Ferrer J (2010) A map of open chromatin in human pancreatic islets. *Nat Genet* 42(3):255–259. doi:[10.1038/ng.530](https://doi.org/10.1038/ng.530)
112. Stitzel ML, Sethupathy P, Pearson DS, Chines PS, Song L, Erdos MR, Welch R, Parker SC, Boyle AP, Scott LJ, Margulies EH, Boehnke M, Furey TS, Crawford GE, Collins FS (2010) Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. *Cell Metab* 12(5):443–455. doi:[10.1016/j.cmet.2010.09.012](https://doi.org/10.1016/j.cmet.2010.09.012)