TECHNISCHE UNIVERSITÄT MÜNCHEN

FAKULTÄT FÜR MEDIZIN

Modeling monogenic diabetes using induced pluripotent stem cells

Dissertation
zur Erlangung des Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

Xianming Wang

München 2019

TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät Für Medizin

Modeling monogenic diabetes
using induced pluripotent stem cells

Xianming Wang

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende: Prof. Dr. Gabriele Multhoff

Prüfer der Dissertation:

1. Prof. Dr. Heiko Lickert

2. Prof. Dr. Wolfgang Wurst

Die Dissertation wurde am 18.07.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 12.05.2020 angenommen.

Contents

Abstract

Pancreas/duodenum homeobox protein 1 (PDX1) is a transcription factor (TF) essential for pancreas development and adult β-cell function. Loss-of-function mutations of PDX1 affecting both alleles result in pancreatic agenesis, while heterozygous mutations are linked to maturity-onset diabetes of the young 4 (MODY4) or type 2 diabetes (T2D). However, it remains unclear how PDX1 coordinates human pancreas development at the molecular level. Moreover, it is not well understood how some very common missense mutations such as the P33T and C18R mutations in the transactivation domain of PDX1 predispose to diabetes mellitus. In order to address these questions, we here generated induced pluripotent stem cell (iPSC) lines from patients' fibroblasts carrying heterozygous $PDX1^{P33T/+}$, $PDX1^{C18R/+}$ mutations. To circumvent the effects related to the genetic background of the patient-derived line and study the impact of different PDX1 doses, we also engineered isogenic cell lines carrying homozygous $PDX1^{P33T/P33T}$, $PDX1^{C18R/C18R}$ mutations and a heterozygous $PDX1$ loss-of-function mutation ($PDX1^{+/-}$) using CRISPR/Cas9 gene editing. We showed that those iPSC lines were pluripotent and could be differentiated into pancreatic progenitors (PPs) and β-like cells using an established *in vitro* protocol. In order to find PDX1 target genes, we performed chromatin immunoprecipitation sequencing. Remarkably, we identified several novel PDX1 target genes in human including regulatory factor 3 (*RFX3*) and delta like canonical Notch ligand 1 (*DLL1*). Moreover, we found that both heterozygous and homozygous P33T and C18R mutations result in impaired β-cell differentiation and function. Mechanistically, in the $PDX1^{+/-}$ and $PDX1^{P33T/P33T}$ mutant lines PDX1 failed to activate its target genes including TFs motor neuron and pancreas homeobox 1 (*MNX1*) and *PDX1* as well as insulin resistance gene carboxylesterase 1 (*CES1*) already at the PP stage and reduced the differentiation efficiency of PPs. Finally, both $PDX1^{P33T/+}$ and $PDX1^{P33T/P33T}$ mutations resulted in downregulation of the expression of PDX1-bound genes at the PP stage, including the insulin synthesis and secretion genes maternally expressed 3 (*MEG3*) and neuronatin (*NNAT*).

In summary, our data provide a clear map of stage-specific target genes of PDX1 in PPs. Furthermore, we show evidence that the regions carrying the T2D-associated SNPs are active in PPs and three of these SNPs are bound by PDX1. Finally, our data reveal mechanistic details of how common coding mutations in *PDX1* affect human pancreatic endocrine lineage formation and β-cell function.

1 Introduction

1.1 Stem cells

Stem cells (SCs) have the ability to self-renew and differentiate into other cell lineages. Based on the level of potency, SCs are further classified as totipotent, pluripotent, multipotent and unipotent cells (Figure 1.1) (De Los Angeles et al., 2015). Totipotent stem cells can give rise to all kinds of cells of an organism, which includes both placental and extraembryonic cells of the trophoblast. Pluripotent stem cells can differentiate into any cell type of the embryo, but not into extraembryonic lineages. Mouse embryonic stem cells (mESCs) and human embryonic stem cells (hESCs) are derived from the inner cell mass of the blastocyst and are classified as pluripotent stem cells. mESC and hESC lines were isolated and established in 1981 and 1998, respectively (Martin, 1981; Thomson, 1998). ESC lines have been successfully used to generate insulin-producing β cells, different types of neurons, cardiac myocytes, liver hepatocytes among other cell types.
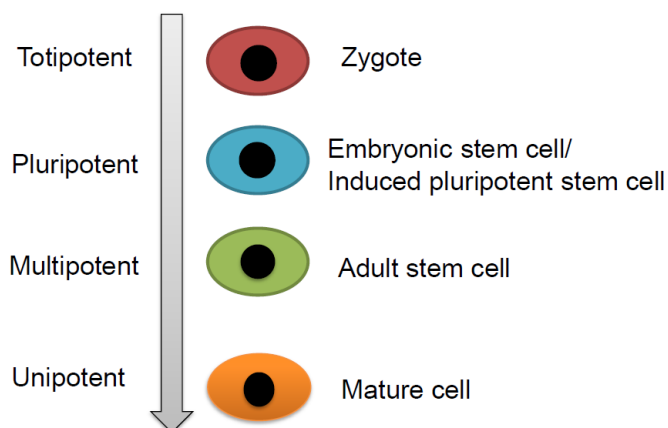
Figure 1.1 Potency levels of stem cells

1.2 Signaling pathways and pluripotent TFs

Understanding how embryonic stem cells maintain their pluripotency at the molecular level and which signaling pathways maintain the pluripotent identity is crucial knowledge to keep their pluripotency in culture maintained. For mouse ES cells, the two key factors, leukemia inhibitory factor (LIF) and bone morphogenetic protein 4 (BMP4), play an important role in maintaining pluripotency (Ying et al., 2003). LIF activates the canonical JAK/STAT3 pathway

and BMP4 inhibits extracellular receptor kinase (ERK) and p38 mitogen-activated protein kinase (MAPK) pathways and thereby maintains pluripotency (Matsuda et al., 1999; Ying et al., 2003). However, hESCs need different signaling pathways to maintain their pluripotency. hESCs do not require LIF, but instead, they rely on fibroblast growth factor 2 (FGF2) and activin A (James, 2005; Vallier, 2005). Key pluripotency factors including Oct4, Sox2, and Nanog play a pivotal role in maintaining pluripotency in both mESCs and hESCs. Those core TFs orchestrate a pluripotency transcriptional network and cooperate with each other to activate their target genes to sustain pluripotency and repress differentiation genes (Boyer et al., 2005). Besides these core TFs, several other TFs, lncRNAs, and miRNAs fine-tune and contribute to regulating pluripotency (Leonardo et al., 2012; Rosa and Ballarino, 2016; Zhang et al., 2008).

1.3 History of cellular reprogramming

In the 1950s, Briggs and King established the somatic cell nuclear transfer technique (SCNT) for reprogramming of a somatic cell to acquire pluripotency (Briggs and King, 1952). SCNT is also called cloning, which means that the nucleus of a somatic cell is transplanted into an enucleated oocyte for the creation of a new individual, indicating that the oocyte cytoplasm contains reprogramming factors to reverse a differentiated nucleus back to a totipotent status. The first organism used in SCNT was tadpoles of Rana pipiens, followed by a frog Xenopus laevis and finally, SCNT gave rise to Dolly the sheep in 1996 (Briggs and King, 1952; GURDON, 1962; Wilmut et al., 1997). The cloning of Dolly and other mammals from adult cells demonstrated that silenced genes in terminally differentiated cells can be fully re-activated and the chromatin code connected with cell identity can be erased and modified to make terminal differentiated cells become totipotent again by nuclear reprogramming. Due to the low efficiency of successful cloning and the ethical concerns regarding harvesting human unfertilized eggs, this technology is not suitable for use in humans. Another method for reprogramming a somatic cell to pluripotency is by fusing it with an ESC, as first reported for hESCs in 2005 (Cowan et al., 2005). However, these pluripotent cells generated by fusion also have an abnormal complement of chromosomes.

1.4 Human induced pluripotent stem cells (hiPSCs)

Somatic cell nuclear transfer and cell fusion with pluripotent stem cells approaches suggest that cell fate can be changed by exposing cell nuclei to reprogramming factors. Working on

this concept, Takahashi and Yamanaka in 2006 found these reprogramming factors that can be used for reprogramming. Those initial TFs used for reprogramming are OCT4, SOX2, KLF4, c-MYC. At first, they used these factors to reprogram mouse somatic cells to be pluripotent stem cells, known as induced pluripotent stem cells (Takahashi and Yamanaka, 2006). One year later, the same group reported that they can reprogram human fibroblasts into pluripotent stem cells using the same combination of transcriptions factors (Figure 1.2) (Takahashi et al., 2007). The fully reprogrammed iPSCs are similar to ESCs including morphology, pluripotency gene expression pattern, DNA demethylation of the promoter regions of *OCT4* and *NANOG*, global patterns of histone methylation and the function such as the ability to differentiate into all germ layers (Takahashi and Yamanaka, 2016). However, the mechanisms of iPSCs reprogramming are still not fully understood (Takahashi and Yamanaka, 2016). Due to several issues, it is not feasible to derive patient-specific ESCs for disease modeling. Induced pluripotent stem cells have solved these issues. Because these iPSCs can be easily obtained from different somatic cells. Especially, these patient-specific iPSCs provide a valuable platform to gain the mechanistic insight into various diseases. Cells differentiated from patient-specific iPSCs can be used for disease modeling, for instance drug screening approaches. Importantly, patient-specific-mutated iPSCs can be corrected *in vitro* using gene-editing strategies. The corrected patient-specific iPSCs could then be directly differentiated into the affected somatic cell type and subsequently transplanted back into the patient for cell-therapy (Figure 1.3). Due to the importance of this discovery, Shinya Yamanaka received the Nobel Prize in Physiology and Medicine in 2012 for pioneering the iPSCs technology.
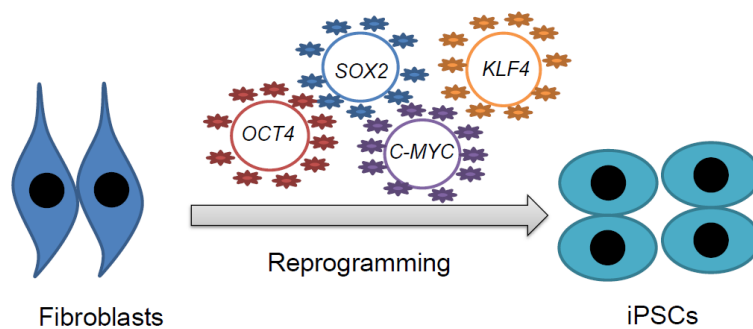
Figure 1.2 Generation of iPSCs using Yamanaka factors
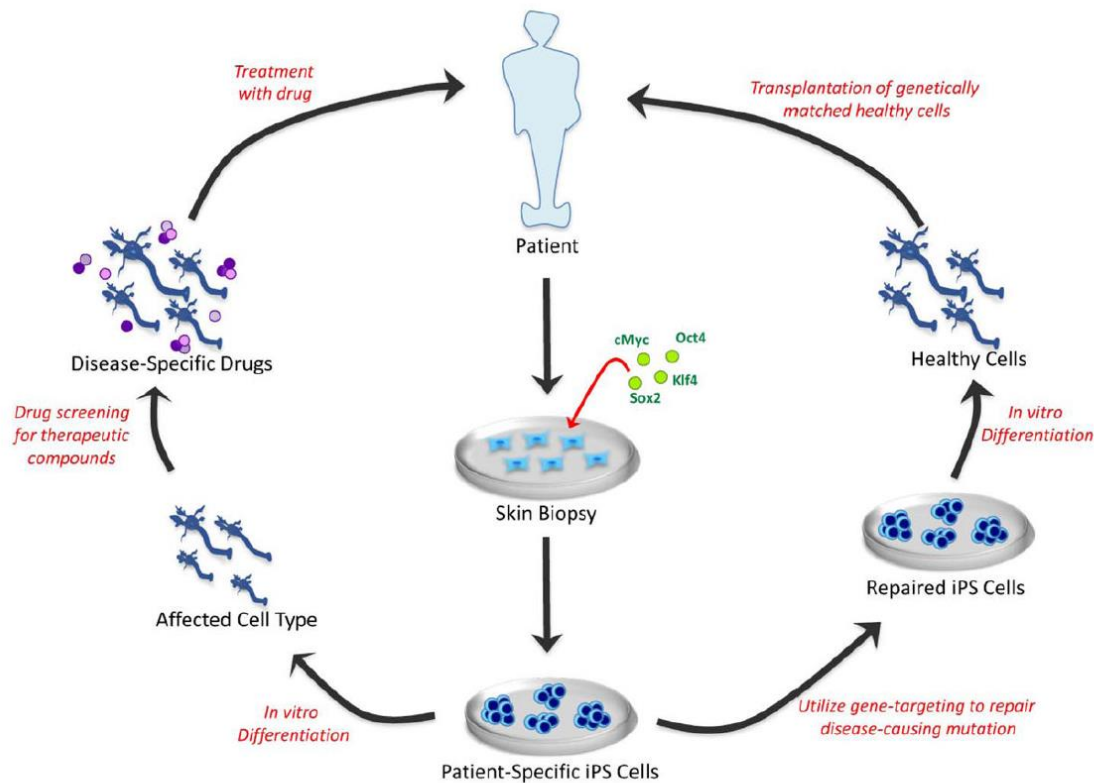
Figure 1.3 Application of hiPSCs for disease modeling, drug screening and cell-therapy (The figure is adapted, with permission from Robinton and Daley, 2012)

Several approaches have been used to express pluripotency factors in somatic cells to reprogram them into stem cells. However, the efficiency of reprogramming and the safety of iPSCs are different depending on the method used. The original method for hiPSCs reprogramming used active retroviral vectors that stably integrated into the host cell genome to introduce OCT4, SOX2, KLF4 and c-MYC expression cassettes (Takahashi et al., 2007). The retroviral transgenes are then silenced during the reprogramming process. However, the reprogramming process is often not fully completed, resulting in partially reprogrammed hiPSCs, in which the endogenous pluripotency genes are not fully activated. More recently, researchers used lentiviral vectors for reprogramming to improve the reprogramming efficiencies (Yu et al., 2007). However, lentiviral vectors also integrate into the genome and the insertion of vectors can lead to mutations in the iPSCs; therefore, this approach is also not suitable for human therapy. In order to overcome these obstacles for cell-based therapy, researchers developed integration-free reprogramming methods of hiPSCs involving non-viral reprogramming methods such as mRNA transfection, miRNA transfection, episomal plasmids and protein delivery (Anokye-Danso et al., 2011; Kim et al., 2009; Warren et al.,

2010; Yu et al., 2011). Overall, episomal reprogramming is a relatively safe strategy to generate footprint-free iPSCs. Efficiencies, however, are low compared to viral vectors. These advances in reprogramming will help us to use hiPSCs for cell therapy purposes. For instance, it has been shown that hiPSCs-derived neural progenitor cells can be used to treat spinal cord injury disease in non-human primate models(Kobayashi et al., 2012).

1.5 Gene-editing in stem cells

The isolation of human embryonic stem cells and the discovery of hiPSCs reprogramming have been a big breakthrough in stem cell biology, *in vitro* disease modeling and drug discovery (Avior et al., 2016). For disease modeling, studies normally compared differentiated cells generated from patient-derived hiPSCs with control hiPSCs derived from healthy donors. Due to the genetic diversity in humans, it is possible that the observed phenotypes result from differences in the genetic background of cells from the control and the patient. A very powerful approach to overcome this problem is to use zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALEN) and clustered regularly interspaced short palindromic repeat and CRISPR-associated protein 9 (CRISPR/Cas9) technologies that enable precise modification of the genomic sequences of endogenous human pluripotent stem cell (hPSC) to generate the isogenic controls.

A significant breakthrough in gene targeting came with the discovery of CRISPR/Cas9 system (Ran et al., 2013). The Cas endonuclease protein was originally found in bacteria as a part of an acquired immune system against foreign DNA. The CRISPR/Cas9 system has become a method of choice for gene editing in human pluripotent stem cells due to its efficiency and simplicity. This gene editing system is based on two components to achieve a DSB: the Cas9 nuclease protein and a single guide RNA (sgRNA). The first twenty nucleotides of a sgRNA guide Cas9 to a specified complementary DNA target sequence via RNA-DNA hybridization. The DNA target sequence is located upstream of an invariant protospacer adjacent motif (PAM) sequence. After correct hybridization the Cas9 nuclease creates a DSB located 3-base pair (bp) upstream of the PAM site. DSBs are being induced at the point of interest in the target genome, and custom alterations by homology-directed repair-mediated cellular repair response are initiated. This novel targeting technology has been successfully used to create knock-outs and knock-ins in hiPSCs (Hendriks et al., 2016) (Figure 1.4). Nevertheless, progress in gene editing tools now enables more rapid engineering of the genome in hPSCs using platforms such as ZFN, TALEN, and CRISPR/Cas9.
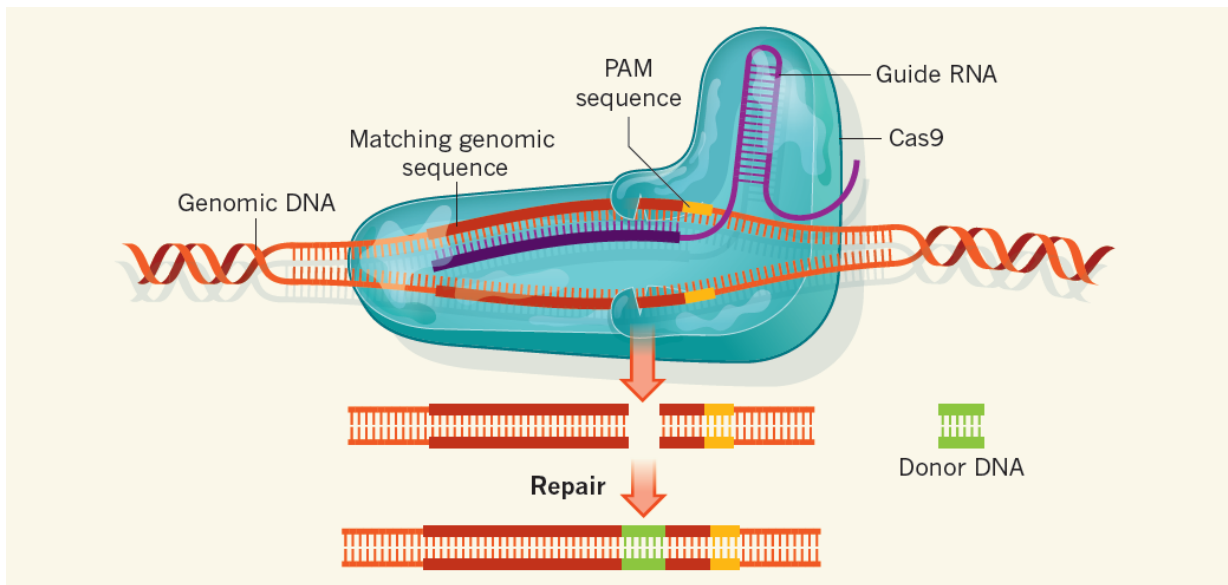
Figure 1.4 Site-specific genome editing with CRISPR/Cas9. (The figure is adapted, with permission from Charpentier and Doudna, 2013).

1.6 Diabetes

Diabetes mellitus is defined as a group of complex, heterogeneous, chronic and metabolic disorders characterized by hyperglycemia. This disease is caused by the loss and/or dysfunction of the pancreatic insulin-producing β-cells. According to the diagnostic criteria, etiology and genetics, diabetes mellitus can be classified as type 1 diabetes (T1D), type 2 diabetes (T2D), gestational diabetes (GD), and monogenic diabetes caused by genetic defects of β-cell function such as maturity onset diabetes of the young (MODY) and neonatal diabetes mellitus (NDM). Today over 451 million people worldwide suffer from different types of diabetes and this number is predicted to reach 693 million by 2045 (International Diabetes Foundation, 2017).

T1D, one of the most common chronic diseases manifesting in the childhood, is caused by insulin deficiency as a result of autoimmune destruction of the insulin-producing pancreatic β-cells (DiMeglio et al., 2018). However, the most common form of diabetes is T2D, a multifactorial disease caused by a complex interplay including genetic, epigenetic and environmental factors. T2D mostly develops in adulthood. The reduced insulin sensitivity makes β cells to increase their insulin secretion. Finally, the β cells are dysfunctional and the

remaining β cells cannot produce enough insulin for glucose maintenance (Kahn et al., 2014). GD is characterized by glucose intolerance arising during the second or third trimester of pregnancy (Chollet and Pettitt, 2006).

Monogenic diabetes is caused by mutations in single genes linked to β-cell development function (Hattersley and Patel, 2017). This type of diabetes includes MODY and NDM. MODY is inherited in an autosomal dominant inheritance pattern and develops early in life (normally before the age of 25 year), and it was first discovered by Tattersall in 1974 (Tattersall, 1974). MODY is caused by defects in developing a sufficient number of β cells or failure in β-cell function. There are at least 14 types of MODY diagnosed in 1-2% of diabetes patients (Table 1.1). These types are caused by mutations in several essential genes, such as *GCK*, *ABCC8*, *NEUROD1*, *HNF1A*, *HNF1B* and *PDX1* (Heuvel-Borsboom et al., 2016). MODY2 and MODY3 are the most common types of MODY, caused by the mutations of *GCK* and *HNF1A*, respectively (Fajans et al., 2002). MODY4 is caused by heterozygosity in the *PDX1* gene and so far only a single family of patients has been identified (Staffers et al., 1997). In terms of treatment for MODY patients, sulfonylureas and insulin have been used to treat *HNF1A*-, *HNF4A*- and *HNF1B*-MODY patients (Murphy et al., 2008; Pearson et al., 2003).

Table 1.1 MODY subtypes and their clinical features

MODY TYPE	GENE LOCUS	CLINICAL FEATURE
1	*HNF4A*	Defects in insulin secretion
2	*GCK*	Fasting hyperglycemia
3	*HNF1A*	Defects in insulin secretion, Renal glycosuria
4	*PDX1*	Defects in insulin secretion, obesity
5	*HNF1B*	Pancreatic hypoplasia
6	*NEUROD1*	Moderate to severe β-cell dysfunction
7	*KLF11*	Pancreatic atrophy
8	*CEL*	Exocrine pancreatic insufficiency
9	*PAX4*	Ketosis prone to diabetes
10	*INS*	Wide clinical spectrum
11	*BLK*	Obesity
12	*ABCC8*	Progressive insulin secretory defect
13	*KCNJ11*	Progressive insulin secretory defect
14	*APPL1*	Obesity

1.7 The transcription factor PDX1 and MODY4

Pdx1, formerly known as insulin promoter factor 1 (*Ipf1*), encodes a TF protein containing a transactivation domain and a DNA-binding homeodomain. In mice, *Pdx1* is located on chromosome 5qG3; in humans it is located on the syntenic region of chromosome 13q12.1 (Fiedorek and Kay, 1995; Inoue et al., 1996; Stoffel et al., 1995). This gene has two exons and the promoter region contains several highly conserved regulatory elements. Many endoderm and pancreas specific TFs including Hnf1β, Foxa1/2, Mafa as well as Pdx1 itself bind to these regions to initiate and maintain *Pdx1* expression (Gao et al., 2008; Gerrish K. et al., 2004). During embryonic development, Pdx1 protein starts to be detectable in the dorsal and ventral pancreatic endoderm and duodenum at E8.5 (Ahlgren et al., 1996; Guz et al., 1995). During the later stages of development and adulthood, Pdx1 is synthesized in β-and δ-cells of the pancreatic islets and in some cells of the duodenum and stomach (Ahlgren et al., 1998; Guz et al., 1995; Offield et al., 1996).

Pdx1 is a master regulator of pancreas development and genetic lineage-tracing studies showed that Pdx1 is important for differentiation of all exocrine and endocrine cells (Gu et al., 2002). Homozygous Pdx1-deficient mice fail to generate a pancreas, while heterozygous animals do develop a pancreas, but later on become diabetic (Brissova et al., 2002; Holland et al., 2005; Johnson et al., 2003; Jonsson et al., 1994). In humans, homozygous loss-of-function mutations of PDX1 also lead to pancreas agenesis, whereas certain heterozygous mutations lead to MODY4 (Staffers et al., 1997; Stoffers et al., 1997, 1998). However, not all heterozygous mutations cause monogenic diabetes, most of them being associated with T2D (Gragnoli et al., 2005; Hani et al., 1999; Macfarlane et al., 1999). So far, there are more than 150 missense coding mutations described for PDX1 (gnomad.broadinstitute.org); however, how these PDX1 mutations predispose to diabetes mellitus is not well understood.

1.8   Pancreas structure, development and function

The pancreas is of high importance for the regulation of energy consumption and metabolism and has two distinct components: the exocrine compartment and the endocrine compartment. The exocrine compartment contains acinar cells producing the nutrient-digestive enzyme juice including lipases, proteinases and amylases and ductal cells transporting enzymes into the small intestine and the duodenum to break down fats, proteins and carbohydrates for absorption. The endocrine compartment plays an important role in maintaining glucose homeostasis by producing a cocktail of hormones that are released into the blood circulation.

The endocrine compartment includes five endocrine cell types, which reside in small clusters called islets of Langerhans (Figure 1.5). These five endocrine cell types are: glucagon-producing α-cells, insulin-producing β-cells, somatostatin-producing δ-cells, pancreatic polypeptide-producing PP-cells and ghrelin-producing ε-cells (Bakhti et al., 2019; Jennings et al., 2015; Pan and Brissova, 2014; Shih et al., 2013; Zhou and Melton, 2018). Defects of these hormone-producing cells can contribute to diabetes mellitus.
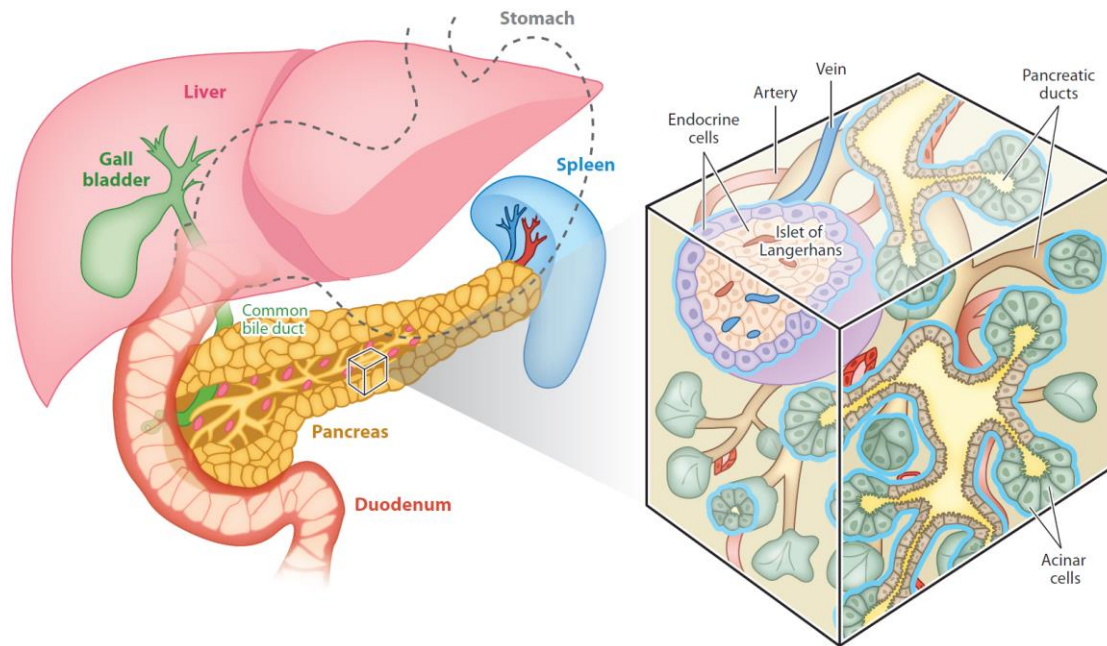
Figure 1.5 The anatomy of the pancreas. The location of the pancreas is very close to the stomach and the duodenum. The acinar cells produce the digestive enzymes that are transported to the duodenum and intestine by ductal cells. The islets of Langerhans contain five endocrine cell types (glucagon-producing α-cells, insulin-producing β-cells, somatostatin-producing δ-cells, pancreatic polypeptide-producing PP-cells and ghrelin-producing ε-cells) that secrete the hormones to regulate the blood glucose levels. (The figure is adapted, with permission from Shih et al., 2013)

1.8.1   Pancreas development, induction and specification in mice

Pancreas development is well understood in mice where pancreas formation starts with the specification of dorsal and ventral pancreatic buds from foregut endoderm around embryonic day (E) 9. Later, these two pancreatic buds elongate alongside the duodenum and stomach and eventually fuse into a single organ at E12.5, which give rise to pancreatic endocrine and exocrine cells. At the same time, the size of the pancreatic buds increases due to cell proliferation. In parallel, the pancreatic buds change their morphology and begin to form

branched tubules/tubes in a process called branching morphogenesis. This early phase of pancreatic development is known as the primary transition. This stage is characterized by the massive proliferation of pancreatic progenitors to generate a stratified epithelium and the formation of multiple microlumens and their subsequent fusion. The first glucagon-producing α-cells start to appear in the dorsal bud at this stage of development (Herrera, 2000; Kesavan et al., 2009; Pictet et al., 1972; Villasenor et al., 2010).

Several TFs control pancreas induction and formation from multipotent progenitors in the foregut endoderm. Among the earliest TFs that mark the pancreatic region are Pdx1, Ptf1a and Sox9 (Ahlgren et al., 1996; Guz et al., 1995; Krapp et al., 1998; Seymour et al., 2007). Furthermore, a number of signaling pathways are involved in the early pancreas development. For instance, Notch signaling regulates the expression of *Hes1* and *Sox9*, which plays an important role in the early pancreas specification (Kageyama et al., 2007; Shih et al., 2012). Hes1 deletion during pancreas development forces progenitors to differentiate into endocrine cells, which in turn results in pancreatic hypoplasia (Horn et al., 2012; Jensen et al., 2000). Moreover, the deletion of Sox9 during early pancreas development decreases early pancreatic growth (Seymour et al., 2007).

Pancreas development requires communication between the pancreatic epithelium and the surrounding mesenchyme. Seminal studies in the 1960s demonstrated the importance of the mesenchyme environment for early pancreatic formation (Golosow and Grobstein, 1962; RUTTER et al., 1964; Wessells and Cohen, 1967). However, we are only beginning to understand which mesenchymal signals control the specification of the early pancreas. The mesenchyme produces several signaling factors, such as FGF10, EGF and Wnt, which have been shown to be important for pancreas development (Ahnfelt-Rønne et al., 2010; Attali et al., 2007; Bhushan et al., 2001; Jonckheere et al., 2008; Kim and Hebrok, 2001; Tulachan et al., 2006). FGF10 activates the FGF signaling pathway via its receptor FGFR2 that induces proliferation and growth of pancreatic progenitor by enhancing and maintaining *Pdx1* and *Ptf1a* expression (Bhushan et al., 2001). During the early growth phase of the pancreatic buds, *Sox9*, *FGF10* and *FGFR2* build a feed-forward loop. In this loop, *FGF10* maintains *Sox9* expression and *Sox9,* in turn, controls *FGFR2* expression to allow FGF10 signaling. Disruption of this loop at any point results in loss of identity of pancreatic epithelia cells, which then activates a liver gene program (Seymour et al., 2012).

During the secondary transition, which starts around E12.5, the microlumina expand and start to fuse to form a continuous network. The endocrine, acinar and ductal cells are generated during this stage. By the end of this secondary transition, endocrine cells together with endothelial, mesenchymal and neuronal cells are clustering together to form the islets of Langerhans (Cleaver and Dor, 2012; Thorens, 2014). One of the most important steps

occurring during this stage is the segregation of multipotent progenitor cells (MPCs) into tip and trunk domains. The tip domain will further differentiate into acinar cell lineages and express *Ptf1a*, *c-Myc*, *Nr5a2* and *Cpa*, whereas the trunk domains form bi-potent endocrine or duct progenitor cells expressing *Nkx6.1*, *Sox9*, *Hnf1b, Nkx2.2* and *Pdx1* (Zhou et al., 2007). The segregation into tip or trunk domains is mainly regulated by the TFs Nkx6.1 and Ptf1a. Nkx6.1 expression induces trunk formation by repressing the tip fate, whereas the expression of Ptf1a induces tip domain formation by blocking the trunk program (Schaffer et al., 2010). Besides' this, the Notch pathway also regulates the tip-trunk patterning by regulating *Nkx6.1* and *Ptf1a* expression. Furthermore, the mesenchymal and endothelial cells also play an important role in determining the tip and trunk fate. However, how the Notch pathway and the interconnection between epithelium and endothelial cells regulate the tip-trunk patterning is still poorly understood.

1.8.2   The formation of pancreatic exocrine and endocrine cell types in mice

95% of the pancreas consists of acinar and ductal cells. Acinar cell differentiation starts from the distal tip epithelium at E13.5 by upregulating expression of Ptf1a, while downregulating the expression of TFs that maintain the MPC identity. This process is regulated by the TFs Ptf1a, Rbp-jl and Nr5a2/LRH-1 (Holmstrom et al., 2011; Masui et al., 2007, 2010). More genes were found to contribute to the acinar cell differentiation, expansion and maintenance, which includes *c-Myc* and *β-catenin* (Murtaugh, 2005; Sánchez-Arévalo Lobo et al., 2018; Wells et al., 2007). During the secondary transition, some cells in the bi-potent trunk progenitor pool start to express the TF Neurogenin3 (Ngn3), which marks the initiation of the endocrine cell differentiation (Gradwohl et al., 2000). The remaining cells of the trunk domain that do not activate Ngn3 expression will eventually form the ductal network (Beucher et al., 2012; Wang et al., 2010). Several TFs including Sox9, Hes1, Hnf1b and Glis3 play important roles in ductal cell fate determination (Cereghini et al., 2015; Delous et al., 2012; Kang et al., 2010; Shih et al., 2012). Knockout of either of these TFs in mice causes malfunction of ductal cells and leads to duct-related defects. Furthermore, the Notch signaling pathway is important for the ductal cell determination. High level of Notch signaling blocks the activation of Ngn3 and promotes ductal cell differentiation, whereas low level of Notch signaling increases the expression of Sox9 and further activates the expression of Ngn3 (Shih et al., 2012). The mechanisms regulating differential Notch signaling activity within the progenitor pool have been investigated recently (Löf-Öhlin et al., 2017; Mamidi et al., 2018).

As discussed above, the endocrine progenitors are derived from the bi-potent trunk domain. These progenitors transiently express Ngn3 (Gradwohl et al., 2000; Gu et al., 2002). Ngn3

acts upstream of several TFs, including Pax4, Arx, Rfx6, NeuroD1, Pax6, Isl1, Nkx2.2 (Petri et al., 2006). The Ngn3[+] cells give rise to five different cell types: glucagon-producing α-cells, insulin-producing β-cells, somatostatin-producing δ-cells, pancreatic polypeptide-producing PP-cells and ghrelin-producing ε-cells. Endocrine cell differentiation is regulated by several signaling pathways including Notch signaling, Wnt signaling and sphingosine-1-phosphate signaling (Kim and Hebrok, 2001).

1.8.3 Endocrine lineage allocation, differentiation and maturation in mice

It is unclear when exactly the decision to differentiate to a specific hormone-producing cell is made. We just begin to understand how specific TFs control the endocrine subtype allocation, development and maturation. Specification of glucagon-producing α-cells requires a set of TFs including Arx, Pax6, Rfx6, Foxa2 and Mafb, whereas Pax4, Pdx1, and Nkx6.1 are essential for the formation of insulin-producing β-cells (Collombat et al., 2003; Gannon et al., 2008; Henseleit, 2005; Kesavan et al., 2009). Little is known about TFs controlling the δ-cells, PP-cells and ε-cells development.

After specification, the endocrine cells acquire physiological ability in a maturation process. In α- and β-cells, this process is guided by two essential TFs, Mafa and Mafb (Artner et al., 2006, 2010; Hang and Stein, 2011). Specifically, Mafb is expressed in both cell types, however, during the maturation process, its expression becomes restricted to α-cells to maintain α-cell identity (Artner et al., 2007; Conrad et al., 2015). In contrast, maturing β-cells cease to express MafB and start expressing MafA instead, which is essential for β-cell and its identity (Nishimura et al., 2006). This switch from MafB[+] to MafA[+] in β-cells also increases *Pdx1* expression (Artner et al., 2006; Nishimura et al., 2006). Other TFs including Pdx1, Nkx6.1, NeuroD1, Nkx2.2, Pax6 and Rfx6 are also involved in this maturation process (Oliver-Krasinski and Stoffers, 2008; Pan and Wright, 2011). Furthermore, mature β-cells express markers such as Glucokinase (Gck), Slc2a2 (Glut2), Urocortin3 (Ucn3) and Flattop (Fltp) (Bader et al., 2016; Blum et al., 2012; Taniguchi et al., 2000; Thorens, 2015).

1.8.4 Human pancreas development and regeneration

In contrast to mouse pancreas development, the formation of the human embryonic pancreas remains unclear due to the limited access to human material. In general, the key events and molecular markers of pancreatic differentiation are conserved in mouse and human, although differences in timing and developmental factors have been reported.

Recently, three reviews compared major specification events and their time of appearance in mouse and human pancreas development (Jennings et al., 2015; Nair and Hebrok, 2015; Pan and Brissova, 2014). Unlike mouse pancreas development, the human pancreas only has a single wave of endocrine cell formation during development. Furthermore, the human pancreatic progenitors do not express the TF NKX2.2; however, this protein is later expressed in endocrine progenitors. Another substantial difference between human and mouse is the composition and morphology of the islet. In human, α- and β-cells are evenly distributed throughout the islet and present at a ratio of 1:1 (α: β), while in mice, α- and β-cells are located at the islet periphery and core, respectively and their ratio is approximately 1:4 (α: β). These differences in composition and morphology between mouse and human islets implicate distinct functions. For example, it has been shown that human islets do not show synchronized Ca^{2+} oscillations in response to glucose compared to the mouse islets. Moreover, due to the high number of α-β contacts in human islets it was proposed that alpha cells might regulate β-cell function, which is in line with α-β communication positively regulating glucose-stimulated insulin secretion (GSIS) (Cabrera et al., 2006; Huypens et al., 2000; Wojtusciszyn et al., 2008). The knowledge gained from human development will help us to generate fully functional β-cells from stem cells in the coming years, which could be used for cell therapy and will aid the understanding of the human pancreas development.

1.9   β-cell differentiation from pluripotent stem cells

Understanding human pancreas development is fundamental to treat endocrine pancreas-associated diseases. One promising treatment is replacing or re-generating β-cell mass. Due to limited human islets from cadaveric donors, directed differentiation of hESCs or hiPSCs into mature insulin-producing cells is a promising approach and might provide an unlimited source of material for cell therapy in the future. The first report of insulin-producing cells generated from hESCs was published in 2001 using spontaneous differentiation (Assady et al., 2001). More recently, D'Amour et al established a protocol to differentiate hESCs into cells of the pancreatic lineages using small molecules (D'Amour et al., 2006). Similar protocols were then published by other groups (Rezania et al., 2012; Xu et al., 2011). Those protocols relied on Activin A and WNT for definitive endoderm induction, and retinoic acid (RA) and inhibitors of sonic hedgehog (Shh) and bone morphogenetic protein (BMP) for inducing the pancreatic differentiation. However, a low percentage of the generated cells were insulin positive. Furthermore, these cells also expressed other hormones, such as glucagon (GCG) and somatostatin (SST), but not β-specific TFs, such as PDX1, NKX6.1. Importantly, when these *in vitro* generated pancreas progenitors were transplanted into mice,

they were fully differentiated into mature endocrine and exocrine cells. Specifically, these endocrine cells were functional and responsive to glucose and were able to regulate glucose homeostasis in diabetic animal models.

In order to generate mono-hormonal and functional β-cells *in vitro*, researchers have developed new protocols to improve differentiation. In 2014, the Melton and Kieffer groups both reported the generation of glucose-responsive β-cells from hPSCs through several stages that mimic pancreas development *in vivo* (Figure 1.6 and Figure 1.7) (Pagliuca et al., 2014; Rezania et al., 2014). Later on, the Hebrok group also developed a short and more simple protocol to generate β-cells from hPSCs *in vitro* (Figure 1.8) (Russ et al., 2015). Most of the hPSC-derived β-cells from these three labs are mono-hormonal, express the relevant TFs, including the PDX1, NKX6.1 and MAFA, and are responsive to glucose in a static assay. Furthermore, diabetes reversal in a mouse model following hPSC-derived β-cells transplantation was faster compared to the previous publications using hPSC-derived pancreas progenitors for transplantation. However, these stem cell-derived β-cells (SC-β) were less functional compared to the human islets. Specifically, the SC-β cells failed to rapidly secrete insulin in dynamic perfusion assays. In order to improve the β-cell function, the Hebrok group used a method of isolating and re-aggregating immature SC-β cells to form islet-sized enriched β-clusters to improve the maturity (Nair et al., 2019). Also recently, the Millman group modulated transforming growth factor β signaling, using an enriched serum-free media to generate SC-β cells with first- and second-phase dynamic insulin secretion similar to that of human islets (Velazco-Cruz et al., 2019). Although the SC-β cells generated by those two groups have greatly improved β-cell function, allowing us to study the β-cell physiology and functional toxicology, the expression of maturation markers such as MAFA and UCN3 is still much lower compared to that of human islets. This suggests that we still have a long way to go before making fully functional mature β cells from stem cells.
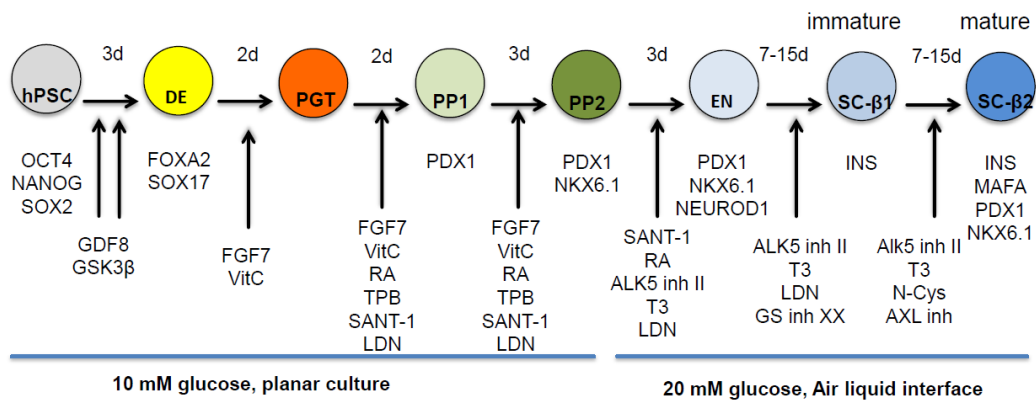
Figure 1.6 Overview of the seven-stage differentiation protocol from the Kieffer lab.
(The figure is modified from Rezania et al., 2014).
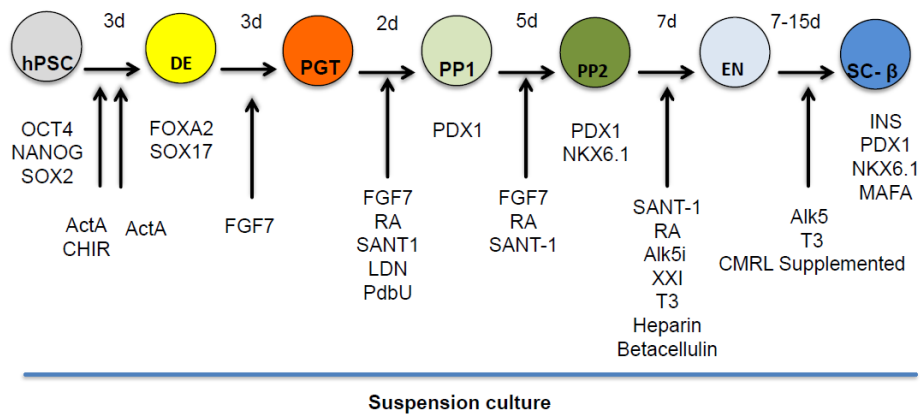

Figure 1.7 Overview of the differentiation protocol from the Melton lab. (The figure is modified from Pagliuca et al., 2014).
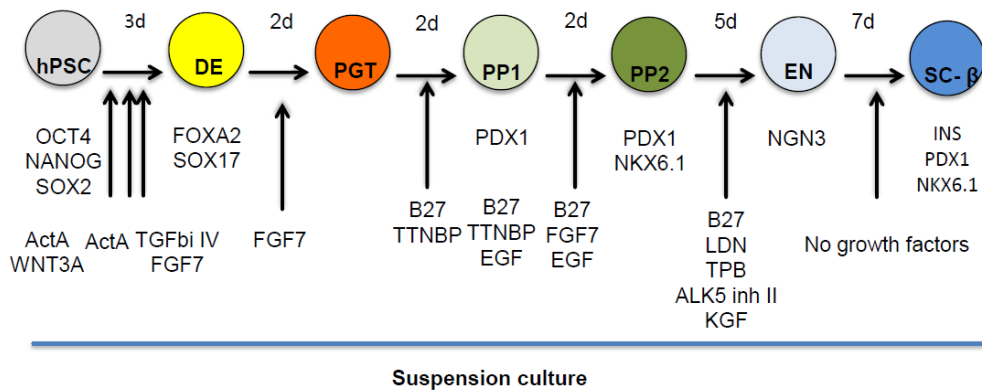
Figure 1.8 Overview of the differentiation protocol from Hebrok lab. (The figure is modified from Russ et al., 2015).

1.10   Diabetes disease modeling using human iPSCs

Patient-derived iPSCs can be used to study disease mechanisms and to develop novel therapies. Therefore, iPSCs have been generated from patients carrying mutations in a single gene that causes MODY types 1, 2, 3, 5 and 8, or Wolfram syndrome (Kondo et al., 2018). The Gadue and Huangfu groups investigated the function of GATA6 during human pancreas development using human PSCs and gene editing technologies (Shi et al., 2017; Tiyaboonchai et al., 2017). Haploinsufficient *GATA6* mutations are associated with human pancreatic agenesis. However, mice lacking *Gata6* expression do not develop obvious pancreatic defects (Carrasco et al., 2012). Both studies consistently showed that *GATA6* haploinsufficiency has only a mild effect on pancreas development but instead affects β-cell differentiation and function.

The Egli group generated a patient iPSC line with Wolfram syndrome (Shang et al., 2014), which is caused by *WFS1* mutations leading to diabetes mellitus. However, the phenotype observed from animal models for Wolfram syndrome is different from that observed in humans, necessitating the use of human model systems. The investigators showed increased endoplasmic reticulum (ER) stress level and a decrease in β-cell numbers in differentiations from Wolfram syndrome iPSCs. These cells also failed to respond to glucose. Subsequently, the Egli group reported the generation of a permanent neonatal diabetes mellitus (PNDM) patient iPSC line carrying a point mutation in the start codon of the insulin gene (Ma et al., 2018). Furthermore, gene-editing technology was used to correct the mutation in the patient iPSCs (Ma et al., 2018). It was found that pancreatic endocrine cells derived from patient iPSCs expressed β-cell makers and insulin mRNA, but did not have

insulin protein. In contrast, the corrected iPSCs could be differentiated to pancreatic endocrine cells that produced the same levels of insulin compared to those of controls (Ma et al., 2018). At the same time, the Otonkoski group also reported the establishment of a different patient-derived iPSC model of neonatal diabetes (Balboa et al., 2018). In this study, the author's generated iPSCs from patients carrying *INSULIN* gene mutations and engineered isogenic CRISPR/Cas9 mutation-corrected lines, which were differentiated into β-like cells. Insulin-mutant β-like cells showed reduced insulin secretion and increased ER-stress upon transplantation into mice. Moreover, cell size, mTORC1 signaling and respiratory chain subunits expression were reduced in insulin-mutant β-like cells. Furthermore, the Otonkoski group investigated the function of STAT3 during human pancreas development using hiPSCs (Saarimäki-Vire et al., 2017). Activating germline mutations in *STAT3* was recently identified as a cause of neonatal diabetes mellitus associated with β-cell autoimmunity. This study revealed that the K392R *STAT3* mutation could cause premature endocrine differentiation through direct induction of NEUROG3 expression. In summary, stem cell differentiations combined with gene-editing technologies enable the study of human β-cell formation and pathomechanisms, and may provide a valuable source of β-cells for therapy in the future (Zeng et al., 2016).

1.11   Aims of the thesis

Maturity Onset Diabetes of the Young (MODY) is a monogenic form of diabetes with an autosomal dominant inheritance pattern, however the pathomechanisms of MODY4 have not been understood well. Therefore, in this thesis, I used hiPSCs to model monogenic diabetes in order to understand the following questions:

Aim 1: Analysis of PDX1 target genes in human pancreatic progenitors

In human, homozygous loss-of-function mutations of PDX1 leads to pancreas agenesis, whereas heterozygous mutations can contribute to MODY4 or T2D. How PDX1 controls human pancreas development is not understood well. It is clear that the role of PDX1 is different in pancreas progenitors (PPs) and in adult β cells; however, the PDX1 target genes and binding patterns in PPs and in adult β-cells have not been well investigated. Furthermore, it has been reported that single nucleotide polymorphisms (SNPs) may contribute to T2D and are located in islet enhancers (Pasquali et al., 2014). However, whether these regions with SNPs are bound by PDX1 or these regions are active in PPs has not been investigated so far. The first aim of the project was to identify the PDX1 target genes at the PP stage and to analyze the PDX1 binding pattern in PPs and in adult β-cells. Furthermore, the additional aim of this project was to screen PDX1-bound SNPs and to check whether these regions carry SNPs or PDX1-bound SNPs are active. This will help us to understand how PDX1 coordinates early pancreas development at the molecular level, with the purpose of improving *in vitro*-directed differentiation of functional β-cells from stem cells in the future.

Aim 2: Analysis of the consequences of point mutations in the PDX1 transactivation domain during human β-cell development and function

Heterozygous amino acid exchanges in the coding region of PDX1 predispose to type 2 diabetes mellitus but the underlying mechanisms are poorly understood. The second aim of the project was to screen human subjects with a high risk of developing diabetes and identify patients carrying mutations in the *PDX1* gene. Furthermore, the additional aim of this project was to investigate how these common mutations (PDX1 C18R and P33T) affect human pancreas development and function using stem cell, CRISPR/Cas9 and *in vitro* differentiation technologies. This will help to better understand disease progression on the molecular level and might improve the diagnosis and therapy of diabetic patients.

2 Discussion

2.1 Genome-wide analysis of PDX1 target genes in human pancreatic progenitors

One promising approach for the treatment of diabetes is to restore β-cell mass using cells generated from directed differentiation of hESCs and hiPSCs (Benthuysen et al., 2016). Stem cells have the potential to provide unlimited material; however, the efficient generation of fully functional and mature β-cells still remains a major challenge (Pagliuca et al., 2014; Rezania et al., 2014; Russ et al., 2015). Due to the incomplete understanding of the molecular pathways orchestrating pancreas development, current protocols used for the generation of β-cells fail to fully mimic the *in vivo* development. Therefore, it is crucial to understand the molecular mechanisms and gene regulatory networks controlling the specification of early human pancreatic lineages, in order to generate β-cells with similar physiological functions as in mature human islets. PDX1 plays a crucial role during pancreas development, yet the temporal expression patterns of its target genes during human development are not well characterized. In the first study, we performed a genome-wide analysis of PDX1 target genes in human pancreatic progenitors (PPs). This identified novel target genes that are important in human pancreas development and might help to improve the directed differentiation of functional β-cells *in vitro*.

2.1.1 The PDX1 binding sites at pancreatic progenitors

Although the target genes of Pdx1 are well studied in mice (Ohlsson et al., 1993; Oliver-Krasinski et al., 2009; Svensson et al., 2007), it is not clear what are the PDX1 target genes at the pancreatic progenitor (PP) stage in humans. Furthermore, as PDX1 plays different roles in PPs and β-cells, it remains unclear how PDX1 target genes change during the process of pancreas and endocrine cell development. To address these questions, we generated a novel integration-free iPSC line from a healthy individual and established a highly efficient differentiation protocol towards early pancreatic progenitors. With this protocol, we differentiated our new iPSC line into early pancreatic progenitors, and performed PDX1 and H3K27ac chromatin immunoprecipitation (ChIP-seq) and mRNA profiling. Global characterization of PDX1 binding sites in iPSC-derived PPs identified in total 8088 PDX1-binding sites that map to 5644 genes. Those genes included developmental regulatory factors, *GATA4* and *HNF1B*, and TFs such as *PDX1*, *NKX2-2* and *RFX6,* which

are important for mature β-cell function. Many of the PDX1-bound genes are differentially expressed between iPSCs and PPs. Moreover, several identified PDX1 target genes have not been previously reported in humans. Among these, *RFX3* and *DLL1* play important roles in mouse pancreas development (Ait-Lounis et al., 2010; Apelqvist et al., 1999; Kim et al., 2010). *RFX3* is a member of the regulatory factor X family of transcription factors and contains a highly conserved winged helix DNA-binding domain (Reith et al., 1994). Homozygous Rfx3 knockout mice have reduced numbers of β-cells with decreased mRNA levels of *insulin*, *Slc2a2* (*Glut2*) and *Gck*, resulting in glucose intolerance and impaired glucose-stimulated insulin secretion in these animals. These findings demonstrate that *Rfx3* is important for the differentiation and function of mature β-cells in mice (Ait-Lounis et al., 2010). *DLL1* is one of the Notch ligands. Notch signaling is important for the differentiation and maintenance of pancreatic cells. *Dll1* Knockout mice have increased expression of *Ngn3*, which causes premature depletion of the endocrine progenitor pool and accelerates differentiation of pancreatic progenitors to the endocrine cells, ultimately leading to abnormal pancreatic development (Apelqvist et al., 1999; Kim et al., 2010). However, the roles of *RFX3* and *DLL1* in human pancreas development have not been investigated so far.

To our knowledge, here we report the PDX1 CHIP-seq on hiPSC-derived early PPs. Surprisingly, however, we found significant differences in PDX1 binding pattern compared to previous PDX1 datasets from hESC-derived PPs (Teo et al., 2015; Wang et al., 2015). Many theories could explain these different results. First, the efficiency of differentiation towards the PDX1$^+$ cells at PP stage is different between our system and the previous reports. For instance, Teo et al. (2015) generated around 65% PDX1$^+$ cells at PP stage compared to around 85% PDX1$^+$ cells with our differentiation protocol. Second, the basic medium and factors used for pancreatic progenitor differentiation protocol differ, Teo et al. (2015) used Activin A, BMP4 for endoderm and FGF2, Retinoic acid, Nicotinamide and DAPT for PP stage; in contrast, our protocol uses Activin A and Wnt3A for endoderm, and FGF7, SANT-1, Retinoic acid and Noggin for PP. These different protocols may produce different patterning of gut tube towards ventral/dorsal pancreas and pancreas/duodenum/stomach progenitors Moreover, one of the best protocols from the Melton group still produced gut endocrine cells in the islet cells differentiation (Veres et al., 2019), suggesting that the correct patterning of early pancreas development is not yet fully understood. Lastly, it is likely that the cells used for ChIP-seq experiments in different laboratories differ regarding their exact stage; this could explain many of the differences between our data and those from Wang et al., 2015, as their cells already expressed NKX6.1. It is well known that TF targets are dynamically activated over time (early response genes vs late response genes), thus the difference might reflect temporal changes of TF target gene activation. Collectively, the reported PDX1 binding sites

differ between labs with different protocols and different cell lines; therefore, it is important to compare all the PDX1 binding datasets and use the overlap of the PDX1 binding sites.

2.1.2 The functions of PDX1 in development and adult

During pancreas development, PDX1 starts to be expressed in the early pancreatic buds and duodenum, but its expression becomes restricted to β-and δ-cells of the pancreatic islets and the selected cells of the duodenum and stomach. This suggests that the functions of PDX1 in the early pancreatic progenitors and in the mature islets can be different. Therefore, we compared the PDX1 binding sites in iPSC-derived PPs to those in the adult human islets. Our results showed that only 15% of PDX1-bound genes are shared between PPs and adult human islets. In adult islets, specific PDX1-bound genes are related to the β-cell function such as insulin secretion, energy metabolism and β-cell maturation. These islet-specific PDX1-bound genes include the K$^+$ channel *KCNJ11* and the glucose transporter *SLC2A2,* which are essential for adult β-cell function, and TFs such as *NKX6-1* and *MAFA,* which play an important role in β-cell maturation. On the other hand, in PPs, the specific PDX1-bound genes are related to early pancreas development. These PP-specific PDX1-bound genes include TFs *GATA4*, *HHEX* and *HNF1A*. The common PDX1-bound genes include TFs *NKX2-2* and *PAX4,* which are associated with endocrine development. This data suggest that PDX1 binding is dynamic during pancreas development to meet different stage-specific functional requirements. Identifying the PDX1 target genes in PPs and in human islets will help us to understand stage specific function of PDX1 and also help us to identify genes and pathways that are dysregulated in dedifferentiated diabetic β-cells.

2.1.3 The diabetes risk of early pancreas development

Single-nucleotide polymorphisms (SNPs) are a very common type of genetic variation. Most SNPs are located in the non-coding regions, hence they often have no effect on human development or health (Shastry, 2002). However, some SNPs can increase the risk of a particular disease. Most T2D-associated SNPs occur in the non-coding regions and can increase the susceptibility to T2D. Furthermore, it has been shown that the active regulatory regions in the adult islets harbor a plethora of T2D-associated SNPs (Pasquali et al., 2014). Some T2D-associated SNPs can affect DNA-binding of a given TF or change the activity of an enhancer (Pasquali et al., 2014). Interestingly, we also found the active regulatory regions at the pancreatic progenitor stage to be enriched in T2D-associated SNPs. This suggests

that the genetic variation of these regions may interfere with the regulation of gene expression at the PP stage and further affect the developmental programs. Consequently, T2D-associated SNPs might cause decreased β-cell mass at birth and therefore the susceptibility to T2D is increased. This is in agreement with the previous findings that the β-cell mass in children is different and individuals with lower β-cell mass have higher risk of diabetes (Meier and Bonadonna, 2013; Meier et al., 2008). We identified three SNPs within stage-specific binding site of PDX1 at PP stage. Among them is the SNP rs11263763, located in the first intron of *HNF1B* and a recent study showed that rs11263763 can downregulate the mRNA level of *HNF1B* (Painter et al., 2015). In addition, there are several SNPs in the first and second intron of *HNF1B,* which are associated with T2D. In mice, Hnf1β plays an important role in the regulation of proliferation and survival in multipotent pancreatic progenitors and the loss of this gene leads to pancreatic hypoplasia. Furthermore, Hnf1β regulates the Ngn3 expression, which in turn drives the specification of endocrine progenitors (Cereghini et al., 2015). In humans, heterozygous mutations in the *HNF1B* gene cause MODY5 (Horikawa et al., 1997); however, the function of HNF1B in human pancreas development is not well understood. A recent study showed that *HNF1B*$^{S148L/+}$ causes pancreatic hypoplasia and has a compensatory up-regulation of key TFs such as FOXA2, SOX17 and PDX1 during early development (Teo et al., 2016). Another two SNPs, rs12762233 and rs7087006 within PP-specific PDX1 binding sites are very closely located to the gene *TCF7L2*. This gene, also known as *Tcf4,* is a transcription factor involved in the canonical Wnt pathway (Lickert et al., 2000). GWAS studies showed that this gene is the most significant T2D risk gene and SNPs within the *TCF7L2* locus have the highest risk for developing T2D in Europeans (McCarthy and Zeggini, 2009). In mouse, Tcf7l2 plays a crucial role in β-cell development and function, and the regulation of β-cell mass (da Silva Xavier et al., 2012). In human, *TCF7L2* regulates β-cell survival and promotes β-cell generation, and individuals carrying the polymorphisms have impaired insulin secretion (Schäfer et al., 2007; Shu et al., 2008). TFs HNF1B and TCF7L2 play an important role in the pancreas development. These three SNPs in PDX1-bound regions might affect early pancreas developmental processes that increase the susceptibility to develop T2D later in adulthood. In the future, it will be interesting to study whether these three SNPs affect human pancreas development using stem cell and gene editing technologies.

2.2   Point mutations in the PDX1 transactivation domain impair human β-cell development and function

Due to differences among species, it is important to understand the mechanism behind human endocrine cell formation, function and failure. However, due to the scarcity of human primary material, it is difficult to study the dynamics of pancreas development in humans (Bakhti et al., 2019). By using established protocols for *in vitro* differentiation of β-cells from pluripotent stem cells (Pagliuca et al., 2014; Rezania et al., 2014; Russ et al., 2015), we can mimic human pancreas development in a culture dish and establish diabetes models that can be used to identify pathomechanisms and for drug screenings. Furthermore, these technologies can be used to study the effects of mutations of MODY genes during the pancreas development, which will help to find new target genes for drug screening. For example, it has been shown that sulfonylurea can be used for treating diabetes caused by the mutation of the potassium channel (Hattersley and Ashcroft, 2005).

2.2.1 The impact of P33T and C18R on PDX1 expression

We generated human iPSCs from patients carrying common heterozygous mutations (P33T, C18R) in the *PDX1* gene and used *in vitro* differentiation system to study the impact of these mutations on human pancreas development and function. To exclude the effect from the genetic background and to study dose-dependent effects, we also generated homozygous iPSC lines of these two mutations from the control iPSC line. We found that both mutations affected human β-cell development and function. Interestingly, $PDX1^{P33T/+}$, $PDX1^{C18R/C18R}$ and $PDX1^{C18R/+}$ mutations did not affect early pancreas development. However, $PDX1^{P33T/P33T}$ impacted the early stage of pancreas development by reducing the protein level of PDX1, which in turn decreased NKX6.1 expression at the PP2 stage. It has been shown that Pdx1 can regulate its own expression in a positive auto-regulatory manner (Gerrish et al., 2001; Gerrish K. et al., 2004; Marshak et al., 2000). In the P33T mutation, replacement of proline with threonine may cause severe changes in the structure of the PDX1 transactivation domain and further affect the recruitment of transcriptional co-activators required for the induction of target gene expression, including *PDX1* itself. We also found that expression of lncRNA *PLUTO* was upregulated in the $PDX1^{P33T/P33T}$ mutant PP cells. It has been reported that *PLUTO* can regulate *PDX1* expression (Akerman et al., 2017). In Endoc-βH1 cells and human islets, knockdown of *PLUTO* leads to decreased *PDX1* mRNA expression. In our system, upregulated expression of the lncRNA *PLUTO* in the $PDX1^{P33T/P33T}$ mutant PP cells might act to compensate for the reduced PDX1 expression. Due to reduction in PDX1 levels at PP1 stage, the differentiation efficiency of $PDX1^{P33T/P33T}$ iPSCs towards the PP2 stage was significantly decreased. This is in line with the fact that the high protein levels of PDX1 are important for the induction of endocrine progenitors during endocrine-lineage specification

(Mamidi et al., 2018). Thus, the reduced levels of PDX1 in human $PDX1^{P33T/P33T}$ mutant progenitors perturb the endocrine-lineage specification. This is also in line with the phenotype we found in the $PDX1^{P33T/P33T}$ and $PDX1^{+/-}$ iPSC lines, which suggests that the P33T mutation causes a significant loss-of PDX1 function. Interestingly, PDX1 levels at PP1 and PP2 stages are not affected in $PDX1^{C18R/C18R}$ mutants. These differences might be due to the different effects on these two mutations on the PDX1 protein structure.

2.2.2 The down-regulated genes in mutant PPs

Through RNA-seq profiling, we found 21 genes that were consistently down-regulated in all isogenic mutant PPs compared to the control. These include insulin resistance-associated genes *MEG3* and *CES1* and diabetes risk genes *LARGE1* and *ANPEP*. *MEG3* is a maternally expressed and imprinted gene, which codes for a long noncoding RNA. In the mouse, *Meg3* is located in the distal region of chromosome 12, and in humans, it is located in the syntenic region of chromosome 14. It has been shown that expression of *Meg3* is important for the function of pancreatic β-cells including insulin synthesis and secretion (You et al., 2016). In humans, islets from T2D patients have lower levels of *MEG3* gene expression compared to islets from healthy subjects, which suggests *MEG3* might be important for human β-cell development and/or function (Kameswaran et al., 2014). Furthermore, *MEG3* is bound by PDX1. The next deregulated gene is *Carboxylesterase 1* (*CES1*). It is highly expressed in fat tissue, gall bladder, lung and stomach. Furthermore, *Ces1* is associated with obesity, hepatic steatosis and hyperlipidemia. *Ces1* knockout mice show abnormal insulin tolerance in muscles and adipose tissue (but not in the liver) and impaired energy homeostasis (Quiroga et al., 2012). In humans, dysregulation of *CES1* is linked to abnormal insulin resistance and an increased risk of obesity (Marrades et al., 2010). Furthermore, it has been reported that some mutations in *LARGE1* could increase the risk of diabetes by enhancing insulin resistance, and some SNPs in the *ANPEP* gene are linked to increased risk of diabetes (Grarup et al., 2018; Locke et al., 2015). Although *MEG3*, *CES1* and *LARGE1* are associated with diabetes and they regulate insulin resistance, it has not been investigated whether these genes play a role in human β-cell development and function. It will be interesting to study these genes using stem cell models in the future. Due to the homozygous P33T mutation affecting PDX1 protein levels, many genes are commonly downregulated in the $PDX1^{P33T/P33T}$ and $PDX1^{+/-}$ PPs, including the TF *MNX1* and the imprinted gene *NEURONATIN (NNAT)*. During embryonic and postnatal stages in mouse, MNX1 initiates and maintains β-cell homeostasis (Pan et al., 2015). Homozygous mutation in *MNX1* in humans can cause permanent neonatal diabetes mellitus (PNDM) (Flanagan et al.,

2014); however, how MNX1 coordinates human β-cell development needs to be investigated. *NNAT* is an imprinted gene and is strongly expressed in the brain, endocrine tissues, placenta and metabolic tissues including muscles, pancreas, duodenum and small intestine. Moreover, expression of *NNAT* can be regulated by hormones and glucose (Millership et al., 2018). SNPs in the *NNAT* gene can contribute to risk of obesity (Vrang et al., 2010). In mouse, knockout of *Nnat* in β-cells shows reduced insulin content and affects glucose-stimulated insulin secretion (Millership et al., 2018), which suggests a role in β-cell function. However, the function of NNAT in human pancreas development has not been investigated yet. Considering the fact that Nnat is important for mouse pancreas development and *NNAT* is bound by PDX1, it would be interesting to study its function in human pancreas development and β-cell function.

2.2.3 The common deregulated genes in PDX1$^{P33T/+}$ and PDX1$^{P33T/P33T}$ mutant PPs

Since the *PDX1$^{P33T/+}$* patient was diagnosed with gestational diabetes before, we further investigated differentially expressed genes in the heterozygous P33T mutant cell line. Moreover, we were interested to find common deregulated genes in *PDX1$^{P33T/+}$* and *PDX1$^{P33T/P33T}$* PPs. In total, 88 genes were deregulated in *PDX1$^{P33T/+}$* PPs compared to the control PPs, of which 21 genes were bound by PDX1. Most of the deregulated genes are associated with pancreas development and insulin secretion. The comparison between *PDX1$^{P33T/+}$* PPs and *PDX1$^{P33T/P33T}$* PPs showed 5 genes that have the same expression pattern. These five genes are *ZNF676*, *C19orf33*, *MEG3*, *NNAT* and *POSTN*. It has been shown that the *ZNF676* can regulate human telomere homeostasis (Mangino et al., 2012). *MEG3* and *NNAT* are linked to insulin secretion. Periostin, which is encoded by the *POSTN* gene, plays an important role in bone, tooth and heart development and also stimulates EMT (epithelial-mesenchymal transformation) (Yan and Shao, 2006). *Postn* knockout mice show impaired β-cell regeneration and the injection of Periostin into the adult pancreas can enhance the regeneration of β-cells and further improve the glucose tolerance (Smid et al., 2015). Therefore, it is evident that *MEG3*, *NNAT* and *POSTN* are playing an important role in pancreas and β-cell biology. Especially, MEG3 and NNAT might regulate the mature function of human β-cells and Periostin might be important for the differentiation transition from pancreas progenitors to endocrine progenitors by undergoing transient EMT; however, their exact function in human pancreas development still needs to be investigated in future studies.

2.2.4 The impact of P33T, C18R on human β-cell development and function

Both $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations impair human β-cell development and function by downregulating gene expression of important TFs, such as *NEUROD1* and *ISL1*. Furthermore, the expression levels of *UCN3* are downregulated in $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ β-cells. These data suggest that the proper PDX1 TF function is important for β-cell development and maturation, which is in line with the previous studies of Pdx1 in rodents (Bastidas-Ponce et al., 2017). Our group has shown that loss of proper TF Pdx1 activity impairs the β-cell maturation in male double knock-in reporter mice and causes diabetes and downregulation expression levels of *Ucn3* (Bastidas-Ponce et al., 2017). Furthermore, female mice develop gestational diabetes during pregnancy but are generally protected from the development of T2D (Bastidas-Ponce et al., 2017). As mentioned previously, the $PDX1^{P33T/+}$ patient had gestational diabetes and homozygous P33T mutation reduces the *UCN3* expression in iPSCs-derived β-cells. This suggests that β-cells from this patient might be very heterogeneous in terms of their functionality, which can become obvious under stress conditions. This is very similar to the above-described mouse model, which suggests an evolutionarily conserved function of PDX1 in β-cell development and function.

The β-cells from PDX1 mutant iPSC lines were less functional compared to the control as shown by static GSIS. It has been shown that reduction of Pdx1 in mice affects GSIS by reducing the Glut2 expression (Brissova et al., 2002). *INS*, *ABCC8*, *KCNJ11*, *SLC30A8* were downregulated in β-like cells carrying $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations. This finding highlights the importance of this TF in insulin synthesis and secretion in human β-cells. This is also in line with the result that even heterozygous P33T or C18R mutant PDX1 protein can decrease the binding activity to the human insulin promoter and reduce the *insulin* expression level in INS-1 and NES2Y cell lines (Gragnoli et al., 2005; Macfarlane et al., 1999).

In summary, we found common heterozygous and homozygous missense mutations in the transactivation domain of PDX1 resulting in the reduction of target gene expression in pancreatic progenitors and insulin-producing β-cells, which further impairs β-cell development and function. Furthermore, we found that the more common P33T (compared to C18R) amino-acid exchange leads to a loss of target gene activation already at the pancreatic progenitor stage in a heterozygous and homozygous background. This systematic functional analysis of coding mutations in genes associated with diabetes predicts an increased risk and predisposition for T2D.

3 Materials and methods

3.1 Human iPSC lines generation

Primary fibroblasts from patients and controls were provided by the Medical Faculty of the Eberhard Karls University, Tübingen. Fibroblasts were reprogrammed into pluripotent stem cells by using a non-integrating Episomal iPSC Reprogramming Kit (Invitrogen, Cat.no.A14703). This kit was mixed with three vectors having the oriP/EBNA-1 (Epstein-Barr unclear antigen-1) backbone that delivers six reprogramming TFs: OCT4, SOX2, NANOG, LIN28, KLF4, and L-MYC. Fibroblasts were transfected using the Amaxa 4D-Nucleofector transfection system. After transfection, Cells were plated onto geltrex-coated culture dishes and incubated in supplemented fibroblast medium. This medium was prepared with knockout DMEM/F-12 (Life technologies), 10% FBS of ESC-qualified (Life technologies), 1% MEM non-essential amino acids (Life technologies), 10 uM HA-100 (Santa Cruz) and 4 ng/ml bFGF (Life technologies). 24 hours after transfection, the medium was switched to N2B27 medium including 0.5 µM PD0325901 (Stemgent), 3 µM CHIR99021 (Stemgent), 0.5 µM A-83-01 (Stemgent), 10 µM HA-100 (Santa Cruz), 10 ng/ml hLIF (Life technologies) and 100 ng/ml bFGF. The basic N2B27 medium was prepared with DMEM/F12 with HEPES (Life technologies), 1x N2 supplement (Life technologies), 1x B27 supplement (Life technologies), 1% MEM non-essential amino acids, 1x Glutamax (Life technologies) and 1x β-Mercaptoethanol (Life technologies). On day 15 after transfection, the medium was switched to the Essential 8 medium. Around 3 weeks after transfection, undifferentiated iPS colonies were picked and expanded.

3.2 iPSC characterization

DNA from hiPSCs was extracted using in-house standard procedure. In order to exclude the transgene integration, markers for the episomal backbone were amplified by semi-quantitative PCR. Primers are as follows: *oriP* forward: TTCCACGAGGCTAGTGAACC. *oriP* reverse: TCGGGGGTGTTAGAGACAAC; *EBNA-1* forward: ATCGTCAAAGCTGCACACAG. *EBNA-1* reverse: CCCAGGAGTCCCAGTAGTCA. For karyotype analysis, the hiPSCs growing in logarithmic phase were used for experiments. Once the hiPSCs were ready and treated with colcemid for 2 hours. After that, hiPSCs were trypsinized, treated with hypotonic solution (0.075 M KCl) for 20 min and fixed with methanol: acetic acid (3:1). Metaphases of hiPSCs were spread on microscope slides. Using the standard G banding technique,

chromosomes were classified according to the International System for Human Cytogenic Nomenclature. For teratomas, 2×10^6 iPSCs were injected into the right hind leg of immunocompromised NOD/SCID mice. After 2 months, the teratomas were excised, fixed, embedded in paraffin, sectioned and stained with hematoxylin/eosin.

3.3 Cell culture

hiPSCs were cultured on Matrigel (BD Biosciences, CA, Cat #354277) in mTeSR™1 medium (STEMCELL technologies, Cat #85850). At ~70–80% confluency, cells were washed once with 1× DPBS without Mg^{2+} and Ca^{2+} (Invitrogen, Cat #14190) and incubated with TrypLE Select Enzyme (1×) (Life Technologies, Cat #12563011) for 3–5 min at 37 °C. Single cells were rinsed with mTeSR™1 medium, and spun at 1,000 rpm for 3 min. The resulting cell pellet was re-suspended in mTeSR™1 medium supplemented with 10 µM Y-27632 (Sigma-Aldrich; MO, Cat #Y0503), and the single cell suspension was seeded at ~0.75 × 10^5 cells/cm^2 on Matrigel-coated surfaces. Cultures were fed every day with mTeSR™1 medium and differentiation was initiated 72 h following seeding with ~90% starting confluency. All the cell lines were mycoplasma free.

3.4 1st pancreatic lineage differentiation protocol

3.4.1 S1: definitive endoderm

Cells were first washed once with 1× DPBS without Mg^{2+} and Ca^{2+} and then switched to MCDB 131 medium (Life Technologies, Cat #10372-019) further supplemented with 1.5 g/l sodium bicarbonate (Sigma, MO, Cat #S6297), 1× Glutamax (Life Technologies, Cat #35050-079), 10 mM final glucose (Sigma, Cat #G8769) concentration, 0.5% bovine serum albumin fraction V, fatty acid free (Sigma, Cat # 10775835001), 100 ng/ml Activin-A (R&D Systems Inc, Cat #338-AC-050/CF), and 3 µM or 5 µM of CHIR-99021 (GSK3β inhibitor, SelleckChem, Cat #S2924) for day 1. At day 2, MCDB 131 medium with 0.5% BSA, 1.5 g/l sodium bicarbonate, 1× Glutamax, 10 mM glucose, 100 ng/ml Activin-A and 0.3 µM CHIR-99021 was added to the cells. At day 3, MCDB 131 with 0.5% BSA, 1.5 g/l sodium bicarbonate, 1× Glutamax, 10 mM glucose and 100 ng/ml Activin-A was added to the cells.

3.4.2 S2: primitive gut tube

Cells were washed once with 1× DPBS without Mg^{2+} and Ca^{2+} and further cultured in MCDB 131 medium further supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 10 mM final glucose concentration, 0.5% BSA, 0.25 mM ascorbic acid (Sigma, Cat #A4544), 50 ng/ml FGF7 (R & D Systems, Cat #251-KG-010/CF) or/and 1.25 µM IWP-2 (Tocris Bioscience, Cat #3533) for 2 d.

3.4.3 S3: posterior foregut

Cells were then exposed to MCDB 131 medium supplemented with 2.5 g/l sodium bicarbonate, 1× Glutamax, 10 mM glucose concentration, 2% BSA, 0.25 mM ascorbic acid, 50 ng/ml FGF7, 0.25 µM SANT-1 (Sigma, Cat #S4572), 1 µM retinoic acid (RA; Sigma, Cat #R2625), 100 nM LDN193189 (LDN; BMP receptor inhibitor, Stemgent, CA, Cat #04-0019), 1:200 ITS-X (Life Technologies, Cat #51500056), and 200 nM TPB for 2 d (PKC activator, custom synthesis, ChemPartner).

3.4.4 S4: pancreatic endoderm

MCDB 131 medium supplemented with 2.5 g/l sodium bicarbonate, 1× Glutamax, 10 mM final glucose concentration, 2% BSA, 0.25 mM ascorbic acid, 2 ng/ml of FGF7, 0.25 µM SANT-1, 0.1 µM retinoic acid, 200 nM LDN193189, 1:200 ITS-X, and 100 nM TPB was added to the cells for 3 d.

After 3 d of culture, 10 µM Y-27632 was added to the S4 cells for 4 h before further experiments. Cells were then washed once with 1× DPBS without Mg^{2+} and Ca^{2+}. TrypLE Select Enzyme (1×) (Life Technologies, Cat #12563011) was added to the cells for 3–5 min at 37 °C. The released cells were rinsed once with basal MCDB 131 medium and centrifuged at 1,000 rpm for 3 min. Cell pellets were exposed to S5-7 media with S5 chemical supplements and put onto the membrane of transwell insert filters (6-well plate: Corning 3414) for air-liquid interface culture. We put 10 µl/spot and ~10 spots per well. Each spot was around 0.5 million cells. For each well, ~1.5 ml S5-7 media with supplements were added to the bottom of each insert.

3.4.5 S5: pancreatic endocrine precursors

S4 Cells were further exposed to MCDB 131 medium supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 20 mM glucose, 2% BSA, 0.25 µM SANT-1, 0.05 µM retinoic acid, 100 nM LDN193189, 1:200 ITS-X, 1 µM T3 (3,3′,5-Triiodo-L-thyronine sodium salt, Sigma, T6397), 10 µM ALK5 inhibitor II (Enzo Life Sciences, NY, Cat #ALX-270-445), 10 µM zinc sulfate (Sigma, Z0251) and 10 µg/ml of heparin (Sigma, H3149) for 3 d.

3.4.6 S6: immature β-like cells

MCDB 131 medium further supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 20 mM glucose, 2% BSA, 100 nM LDN193189, 1:200 ITS-X, 1 µM T3, 10 µM ALK5 inhibitor II, 10 µM zinc sulfate, 100 nM gamma secretase inhibitor XX (EMD MilliPore, MA, Cat #565789) and 10 µg/ml of heparin was added to the S5 cells for 7 d.

3.4.7 S7: mature β-like cells

MCDB 131 medium supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 20 mM final glucose concentration, 2% BSA, 1:200 ITS-X, 1 µM T3, 10 µM ALK5 inhibitor II, 10 µM zinc sulfate, 1 mM N-acetyl cysteine (N-Cys, Sigma, Cat #A9165), 10 µM Trolox (Vitamin E analogue, EMD, Cat #648471), 2 µM R428 (AXL inhibitor, SelleckChem, Cat #S2841) and 10 µg/ml of heparin was added to the S6 cells for 14 d. For all stages of differentiation, the medium was changed every day.

3.5   2nd generation pancreatic progenitor differentiation protocol

3.5.1   S1: definitive endoderm

Day 1, iPSCs were first washed once with 1× DPBS without Mg^{2+} and Ca^{2+} (Invitrogen) and then exposed to RPMI 1640 medium (Invitrogen, Cat #21875-034) further supplemented with 1.2 g/L sodium bicarbonate (Sigma), 0.2% ESC-qualified FBS (Life Technologies, Cat #16141-079), 100 ng/mL Activin-A (R&D Systems), and 20 ng/mL of Wnt3A (R&D Systems,

Cat #5036-WN-010) for day 1 only. Cells were switched to RPMI with 0.5% FBS, 1.2g/L sodium bicarbonate and 100 ng/mL Activin-A at day 2 & 3.

3.5.2 S2: primitive gut tube

Cells were washed once with 1× DPBS (without Mg^{2+} and Ca^{2+}) once and DMEM-F12 (Life Technologies, Cat #21041025) medium further supplemented with 2 g/L sodium bicarbonate, 2% FBS and 50 ng/mL of FGF7 (R&D Systems) was added to the cells for 3 days.

3.5.3 S3: posterior foregut

DMEM-HG medium (Life Technologies, Cat #41966029) supplemented with 0.25 µM SANT-1 (Sigma-Aldrich), 2 µM retinoic acid (RA; Sigma-Aldrich), 100 ng/mL of Noggin (R&D Systems, Cat #6057-NG-025), and 1% B27 (Invitrogen, Cat #17504044) was added to the cells for another 4 days.

3.6   Generation of clonal hiPSC mutant lines using gRNA or gRNA/ssDNA transfection in hiPSCs

The established protocol was used to generate all the mutant cell lines. We got the plasmid for transfection (pU6-(BbsI) sgRNA_CAG-venus-bpA, Addgene ID86985) as a gift from Dr. Ralf Kühn. This plasmid had BbsI site for single or multiple gRNAs, the Venus genes and the Cas9 gene. hiPSCs were cultured for two days and then dissociated using TrypLE select. The single hiPSCs were replated onto Matrigel-coated plates and after half an hour, hiPSCs were transfected in suspension with gRNAs or gRNA/ssDNA using Lipofectamine RNAiMAX (Thermo Fisher Scientific, Cat #13778-150) according to manufacturer's instructions. 0.75×10^6 hiPSCs were seeded for one well of 6-well plates. 2.5 ug plasmid of gRNA was used for one transfection for the $PDX1^{+/-}$ knockout cell lines. For the PDX1 point mutations, 2.5 ug plasmid with the gRNA together with 30 pmol ssDNA were added for one transfection. First, gRNAs or gRNA/ssDNA and Lipofectamine RNAiMAX were diluted separately in Opti-MEM (Thermo Fisher Scientific, Cat # 31985070). After that, the two buffers were mixed together for 5 min at room temperature and added to the hiPSCs drop by drop.

3.7 Establishment of clonal hiPSC mutant lines

48 hours after the transfection, hiPSCs were dissociated into single cells using TrypLE select. We sorted the cells positively expressing Venus by FACS. After that, the sorted Venus-positive cells were re-plated at a low density (500-1,000 cells per 10 cm dish). The left cells were used to extract genomic DNA. We performed the T7 endonuclease I assay to assess the CRISPR mutagenesis efficiency. Individual colonies were picked, mechanically disaggregated and replated in mTeSR™1 medium into individual wells of 96-well plates. The PCR was performed using Herculase II Fusion DNA Polymerase (Agilent Technologies, Cat #600679) followed by Sanger sequencing to screen mutant clones. Primers for *PDX1* are as follows: forward: *TACCTGGGCCTAGCCTCTTAGTG*, reverse: *TGAGAACCGGAAAGGAGAAAGGG.* Clonal cell lines having desired mutations were further expanded.

3.8 Immunofluorescence imaging

Cells were fixed with 4% paraformaldehyde, permeabilized and blocked. After that, cells were incubated with primary antibodies overnight at 4 °C and later on with the secondary antibodies for 1 hour at room temperature. Images were acquired on a TCS SP5 laser-scanning microscope (Leica). The following antibodies and dilutions were used for our experiments: goat anti-OCT-3/4 (1:500, Cat #sc-8628, Santa Cruz), goat anti-SOX2 (1:500, Cat #sc-17320, Santa Cruz), mouse anti-TRA-1-81 (1:50, Cat #MAB4381, Millipore), mouse anti-SSEA4 (1:500, Cat #4755, Cell Signaling), rabbit anti-FOXA2 (1:250, Cat #8186, Cell Signaling), goat anti-SOX17 (1:500, Cat #GT15094, Acris/Novus), goat anti-PDX1 (1:500, Cat #AF2419, R&D Systems), rabbit anti-NKX6.1 (1:300, Cat #NBP1-82553, Acris/Novus), guinea pig anti-C-Peptide (1:100, Cat #ab30477, Abcam), mouse anti-Glucagon (1:500, Cat #G2654, Sigma), rat anti-Somatostatin (1:300, Cat #MA5-16987, Invitrogen).

3.9 Flow cytometry

Cells were dissociated to have single-cell solution using 1× TrypLE Select Enzyme. Cells were washed, fixed and permeabilized. After that, the cells were incubated with rabbit anti-FOXA2 (1:200, #8186, Cell Signaling), goat anti-SOX17 antibody (1:200, Cat #GT15094, Acris/Novus), goat anti-PDX1 antibody (1:100, Cat #AF2419, R&D Systems), rabbit anti-NKX6.1 antibody (1:200, Cat #NBP1-82553, Acris/Novus), guinea pig anti-C-Peptide

antibody (1:100, Cat #ab30477, Abcam), mouse anti-Glucagon (1:500, Cat #G2654, Sigma), for 30 min at room temperature and then stained with appropriate AlexaFluor-555 and -647 secondary antibody for 30 min at room temperature. We performed Flow cytometry using FACS-Aria III (BD Bioscience). FACS data were analyzed using the software of FlowJo. The PDX1 positive population was first gated and then median fluorescence intensity (MFI) for PDX1 was calculated using BD FACS software.

3.10   RNA isolation and qPCR

Total RNA from cells was extracted using the miRNeasy mini kit (Qiagen). The high-capacity RNA-to-cDNA kit (Applied Biosystems) was used to perform the cDNA synthesis. ViiA7 (Applied Biosystems) and TaqMan Fast Advanced Master Mix (Applied Biosystems) were used for qPCR. The housekeeping genes 18S ribosomal RNA (*RNA18S*) and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) were used as references genes. Taqman probes (Applied Biosystems): *INS*, Hs02741908_m1; *GCG*, Hs01031536_m1; *SST*, Hs00356144_m1; *PDX1*, Hs00236830_m1; *GAPDH*, Hs02758991_g1; *18S*, Hs99999901_s1; *NKX6.1*, Hs01055914_m1; *MAFA*, Hs01651425_s1; *NEUROD1*, Hs01922995_s1; *UCN3*, Hs00846499_s1; *PAX6*, Hs01088114_m1; *ABCC8*, Hs01093752_m1; *KCNJ11*, Hs00265026_s1; *SLC30A8*, Hs00545183_m1.

SYBR qPCR was performed under standard conditions using ViiA7 (Applied Biosystems) and SYBR™ Green PCR Master Mix (Applied Biosystems). Samples were normalized to the housekeeping gene *ACTB.* Primers were as follows:

HHEX forward: ACGGTGAACGACTACACGC.

HHEX reverse: CTTCTCCAGCTCGATGGTCT;

MEIS1 forward: GGGCATGGATGGAGTAGGC.

MEIS1 reverse: GGGTACTGATGCGAGTGCAG;

ONECUT1 forward: GAACATGGGAAGGATAGAGGCA.

ONECUT1 reverse: GTAGAGTTCGACGCTGGACAT;

RFX6 forward: AAGCAGCGGATCAATACCTGT.

RFX6 reverse: ACCGTGGTAAGCAAACTCCTT;

ACTB forward: CCCAGAGCAAGAGAGG.

ACTB reverse: GTCCAGACGCAGGATG.

3.11   Affymetrix microarray

The miRNeasy Mini kit (Qiagen, Cat# 217004) was used to extract the total RNA for microarray analysis. The quality of RNA integrity was checked using Agilent 2100 Bioanalyzer (Agilent RNA 6000 Pico Kit) and cDNA was amplified with the Ovation PicoSL WTA System V2 (Nugen, 3312) together with the Encore Biotin Module (Nugen, USA). All subsequent computational analysis was analyzed in R using Bioconductor packages. Using the oligo package (version 1.38.0), expression data was RMA normalized and probe sets were defined using the package hugene20sttranscriptcluster.db (version 8.5.0). Differential expression analyses were analyzed on prefiltered data, containing the 30562 probe sets having highest expression values. P-values were adjusted for multiple testing by Benjamini-Hochberg correction using the limma package (version 3.30.7). If the adjusted p-value (FDR) was below a threshold of 0.1 and the fold-change was greater than or equal to 2, a gene was considered differentially expressed. Functional enrichments were performed using HOMER. The term affiliations of functional annotations were based on literature research and HOMER.

3.12   RNA-seq and analysis

At the PP1 stage, RNA samples from *PDX1$^{+/-}$*, *PDX1$^{P33T/P33T}$*, *PDX1$^{C18R/C18R}$* and XM001 iPSC lines were collected using the above-mentioned RNA isolation method. All of samples for RNA-seq were produced by 1st differentiation protocol. Using miRNeasy Mini kit (Qiagen, #217004), total RNA from *PDX1$^{+/-}$*, *PDX1$^{P33T/P33T}$*, *PDX1$^{C18R/C18R}$* and XM001 lines was extracted and RNA integrity was evaluated using Agilent 2100 Bioanalyzer (Agilent RNA 6000 Pico Kit). Libraries were processed using the TruSeq Stranded mRNA Library Prep (Illumina). The RNA-seq analysis was described in (Wang et al., 2019).

3.13   Static glucose-stimulation insulin secretion

Static glucose-stimulated insulin secretion (GSIS) of the generated β-like cells was done according to previous protocols (Shi et al., 2017; Zhu et al., 2016). Briefly, 5-8 aggregates were hand-picked and washed three times with KRBH buffer (129 mM NaCl, 4.8 mM KCl, 2.5

mM CaCl2, 1.2 mM MgSO4, 1 mM Na2HPO4, 1.2 mM KH2PO4, 5 mM NaHCO3, 10 mM HEPES and 0.1% BSA in deionized water and sterile filtered) and then cultured in KRBH buffer at 37 °C for 1 h. After that, aggregates then were incubated in KRBH buffer having 2.8 mM glucose for 60 min at R.T. We collected the supernatants and the aggregates were switched to KRBH buffer with 16.7 mM glucose for 60 min. Supernatants were collected once more. After the experiment, cell aggregates were dissociated into single cells and the cell numbers were calculated to normalize the GSIS. Mercodia Human Insulin ELISA kit (Mercodia, Cat# 10-1113-01) and Human C-peptide ELISA kit (Mercodia, Cat# 10-1141-01) were used to measure the insulin and C-peptide content in supernatant samples according to manufacturer's protocols.

3.14   Western Blotting

PP1 Cells were harvested and lysed using cell lysis buffer with protease inhibitors. A 10% SDS-PAGE gel was used to separate to the samples. The samples were transferred to PVDF Pre-Cut Blotting Membranes and blocked with 5% milk in Tris-based saline with Tween 20 (0.1% TBST) buffer for 1 h at R.T. The PVDF membrane was incubated with 1ST primary antibodies overnight at 4°C and later on incubated with HRP conjugated secondary antibodies at R.T. for 1 h. ECL western blotting detection reagent (Bio-rad, Cat #1705061) was used to detect the protein bands. The following antibodies were used for experiments: goat anti-PDX1 (R&D, Cat #AF2419, 1:1,000), mouse anti-ACTB (b-Actin) (Cell Signaling Technology, Cat #3700S, 1:10,000).

3.15   Statistics

One-way ANOVA followed by Bonferroni's multiple comparisons test in GraphPad Prism (version 8.0.0 for Windows, GraphPad Software, San Diego, California USA) was used, and statistical significance was determined. A value of $P<0.05$ was regarded as statistically significant.

3.16   Accession Numbers

Microarray, RNA-seq and ChIP-seq data has been submitted to the GEO database at NCBI under the accession number GSE125770 and GSE106950.

4 Publications for dissertation

Wang X, Sterr M, Burtscher I, Chen S, Hieronimus A, Machicao F, Staiger H, Häring HU, Lederer G, Meitinger T, Cernilogar FM, Schotta G, Irmler M, Beckers J, Hrabě de Angelis M, Ray M, Wright CVE, Bakhti M, Lickert H. **Genome-wide analysis of PDX1 target genes in human pancreatic progenitors.** Mol Metab. 2018 Mar;9:57-68. doi: 10.1016/j.molmet.2018.01.011. Epub 2018 Jan 31.

Summary of this paper: This paper provides a clear map of stage-specific target genes of PDX1 during *in vitro* differentiation of stem cells into pancreatic progenitors and shows that regions carrying T2D-associated SNPs are active in pancreatic progenitors.

Declaration of my contribution: Xianming Wang performed and analyzed the iPSC experiments and wrote the manuscript.

Wang X, Sterr M, Ansarullah, Burtscher I, Böttcher A, Beckenbauer J, Siehler J, Meitinger T, Häring HU, Staiger H, Cernilogar FM, Schotta G, Irmler M, Beckers J, Wright CVE, Bakhti M, Lickert H. **Point mutations in the PDX1 transactivation domain impair human β-cell development and function.** Mol Metab. 2019 Jun;24:80-97. doi: 10.1016/j.molmet.2019.03.006. Epub 2019 Mar 20.

Summary of this paper: This paper reveals mechanistic details of how common coding mutations in *PDX1* affect human pancreatic endocrine lineage formation and β-cell function.

Declaration of my contribution: Xianming Wang performed and analyzed the iPSC experiments and wrote the manuscript.

Wang X, Chen S, Burtscher I, Sterr M, Hieronimus A, Machicao F, Staiger H, Häring HU, Lederer G, Meitinger T, Lickert H. **Generation of a human induced pluripotent stem cell (iPSC) line from a patient with family history of diabetes carrying a C18R mutation in the *PDX1* gene.** Stem Cell Res. 2016 Sep;17(2):292-295. doi: 10.1016/j.scr.2016.08.005. Epub 2016 Aug 6.

Summary of this paper: The generation of a human integration-free iPSC line from patients carrying a C18R mutation in the *PDX1* gene.

Declaration of my contribution: Xianming Wang performed and analyzed the iPSC experiments and wrote the manuscript.

Wang X, Chen S, Burtscher I, Sterr M, Hieronimus A, Machicao F, Staiger H, Häring HU, Lederer G, Meitinger T, Lickert H. **Generation of a human induced pluripotent stem cell**

(iPSC) line from a patient carrying a P33T mutation in the *PDX1* gene. Stem Cell Res. 2016 Sep;17(2):273-276. doi: 10.1016/j.scr.2016.08.004. Epub 2016 Aug 5.

Summary of this paper: The generation of a human integration-free iPSC line from patients carrying a P33T mutation in the *PDX1* gene.

Declaration of my contribution: Xianming Wang performed and analyzed the iPSC experiments and wrote the manuscript.

5 References

Ahlgren, U., Jonsson, J., and Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. Development *122*, 1409–1416.

Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., and Edlund, H. (1998). beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. Genes Dev. *12*, 1763–1768.

Ahnfelt-Rønne, J., Ravassard, P., Pardanaud-Glavieux, C., Scharfmann, R., and Serup, P. (2010). Mesenchymal bone morphogenetic protein signaling is required for normal pancreas development. Diabetes *59*, 1948–1956.

Ait-Lounis, A., Bonal, C., Seguín-Estévez, Q., Schmid, C.D., Bucher, P., Herrera, P.L., Durand, B., Meda, P., and Reith, W. (2010). The transcription factor Rfx3 regulates β-cell differentiation, function, and glucokinase expression. Diabetes *59*, 1674–1685.

Akerman, I., Tu, Z., Beucher, A., Rolando, D.M.Y., Sauty-Colace, C., Benazra, M., Nakic, N., Yang, J., Wang, H., Pasquali, L., et al. (2017). Human Pancreatic β Cell lncRNAs Control Cell-Specific Regulatory Networks. Cell Metab. *25*, 400–411.

Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P.J., Epstein, J.A., et al. (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell *8*, 376–388.

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

Artner, I., Le Lay, J., Hang, Y., Elghazi, L., Schisler, J.C., Henderson, E., Sosa-Pineda, B., and Stein, R. (2006). MafB: an activator of the glucagon gene expressed in developing islet alpha- and beta-cells. Diabetes *55*, 297–304.

Artner, I., Blanchi, B., Raum, J.C., Guo, M., Kaneko, T., Cordes, S., Sieweke, M., and Stein, R. (2007). MafB is required for islet beta cell maturation. Proc. Natl. Acad. Sci. *104*, 3853–3858.

Artner, I., Hang, Y., Mazur, M., Yamamoto, T., Guo, M., Lindner, J., Magnuson, M.A., and Stein, R. (2010). MafA and MafB regulate genes critical to β-cells in a unique temporal manner. Diabetes *59*, 2530–2539.

Assady, S., Maor, G., Amit, M., Itskovitz-Eldor, J., Skorecki, K.L., and Tzukerman, M. (2001). Insulin Production by Human Embryonic Stem Cells. Diabetes *50*, 1691–1697.

Attali, M., Stetsyuk, V., Basmaciogullari, A., Aiello, V., Zanta-Boussif, M.A., Duvillie, B., and Scharfmann, R. (2007). Control of β-cell differentiation by the pancreatic mesenchyme. Diabetes *56*, 1248–1258.

Avior, Y., Sagi, I., and Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Mol. Cell Biol. *17*, 170–182.

Bader, E., Migliorini, A., Gegg, M., Moruzzi, N., Gerdes, J., Roscioni, S.S., Bakhti, M., Brandl, E., Irmler, M., Beckers, J., et al. (2016). Identification of proliferative and mature β-cells in the islets of langerhans. Nature *535*, 430–434.

Bakhti, M., Böttcher, A., and Lickert, H. (2019). Modelling the endocrine pancreas in health and disease. Nat. Rev. Endocrinol. *15*, 155–171.

Balboa, D., Saarimäki-Vire, J., Borshagovski, D., Survila, M., Lindholm, P., Galli, E., Eurola, S., Ustinov, J., Grym, H., Huopio, H., et al. (2018). Insulin mutations impair beta-cell development in a patient-derived iPSC model of neonatal diabetes. Elife *7*.

Bastidas-Ponce, A., Roscioni, S.S., Burtscher, I., Bader, E., Sterr, M., Bakhti, M., and Lickert, H. (2017). Foxa2 and Pdx1 cooperatively regulate postnatal maturation of pancreatic β-cells. Mol. Metab. *6*, 524–534.

Benthuysen, J.R., Carrano, A.C., and Sander, M. (2016). Advances in β cell replacement and regeneration strategies for treating diabetes. J. Clin. Invest. *126*, 3651–3660.

Beucher, A., Martín, M., Spenle, C., Poulet, M., Collin, C., and Gradwohl, G. (2012). Competence of failed endocrine progenitors to give rise to acinar but not ductal cells is restricted to early pancreas development. Dev. Biol. *361*, 277–285.

Bhushan, A., Itoh, N., Kato, S., Thiery, J.P., Czernichow, P., Bellusci, S., and Scharfmann, R. (2001). Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. Development *128*, 5109–5117.

Blum, B., Hrvatin, S., Schutz, C., Bonal, C., Rezaniaand, A., and Melton, D.A. (2012). Functional β-cells maturation is marked by an increase in the glucose threshold for insulin secretion and by expression of urocortin3. Nat. Biotechnol. *30*, 261–264.

Boyer, L.A., Tong, I.L., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Briggs, R., and King, T.J. (1952). Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. Proc. Natl. Acad. Sci. U. S. A. *38*, 455–463.

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

Cabrera, O., Berman, D.M., Kenyon, N.S., Ricordi, C., Berggren, P.-O., and Caicedo, A. (2006). The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. Proc. Natl. Acad. Sci. U. S. A. *103*, 2334–2339.

Carrasco, M., Delgado, I., Soria, B., Martín, F., and Rojas, A. (2012). GATA4 and GATA6 control mouse pancreas organogenesis. J. Clin. Invest. *122*, 3504–3515.

Cereghini, S., Haumaitre, C., Heliot, C., De Vas, M.G., Sander, M., and Kopp, J.L. (2015). Hnf1b controls pancreas morphogenesis and the generation of Ngn3+ endocrine progenitors. Development.

Charpentier, E., and Doudna, J.A. (2013). Rewriting a genome. Nature *495*, 50–51.

Chollet, M.B., and Pettitt, D.J. (2006). Treatment of gestational diabetes mellitus. Clin. Diabetes *24*, 35–36.

Cleaver, O., and Dor, Y. (2012). Vascular instruction of pancreas development. Development *139*, 2833–2843.

Collombat, P., Mansouri, A., Hecksher-Sørensen, J., Serup, P., Krull, J., Gradwohl, G., and Gruss, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. Genes Dev. *17*, 2591–2603.

Conrad, E., Dai, C., Spaeth, J., Guo, M., Cyphert, H.A., Scoville, D., Carroll, J., Yu, W.-M., Goodrich, L. V., Harlan, D.M., et al. (2015). The MAFB transcription factor impacts islet α-cell function in rodents and represents a unique signature of primate islet β-cells. Am. J. Physiol. Metab. *310*, E91–E102.

Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Developmental Biology: Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science (80-.). *309*, 1369–1373.

D'Amour, K.A., Bang, A.G., Eliazer, S., Kelly, O.G., Agulnick, A.D., Smart, N.G., Moorman, M.A., Kroon, E., Carpenter, M.K., and Baetge, E.E. (2006). Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. Nat. Biotechnol. *24*, 1392–1401.

Delous, M., Yin, C., Shin, D., Ninov, N., Debrito Carten, J., Pan, L., Ma, T.P., Farber, S.A., Moens, C.B., and Stainier, D.Y.R. (2012). sox9b is a key regulator of pancreaticobiliary ductal system development. PLoS Genet. *8*.

DiMeglio, L.A., Evans-Molina, C., and Oram, R.A. (2018). Type 1 diabetes. Lancet *391*, 2449–2462.

Fajans, S.S., Bell, G.I., and Polonsky, K.S. (2002). Molecular Mechanisms and Clinical Pathophysiology of Maturity-Onset Diabetes of the Young. N. Engl. J. Med. *345*, 971–980.

Fiedorek, F.T., and Kay, E.S. (1995). Mapping of the Insulin Promoter Factor I Gene (Ipf1) to Distal Mouse Chromosome 5. Genomics *28*, 581–584.

Flanagan, S.E., De Franco, E., Lango Allen, H., Zerah, M., Abdul-Rasoul, M.M., Edge, J.A., Stewart, H., Alamiri, E., Hussain, K., Wallis, S., et al. (2014). Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man. Cell Metab. *19*, 146–154.

Gannon, M., Tweedie Ables, E., Crawford, L., Lowe, D., Offield, M.F., Magnuson, M.A., and Wright, C.V.E. (2008). pdx-1 function is specifically required in embryonic β cells to generate appropriate numbers of endocrine cell types and maintain glucose homeostasis. Dev. Biol. *314*, 406–417.

Gao, N., LeLay, J., Vatamaniuk, M.Z., Rieck, S., Friedman, J.R., and Kaestner, K.H. (2008). Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. Genes Dev. *22*, 3435–3448.

Gerrish, K., Cissell, M.A., and Stein, R. (2001). The Role of Hepatic Nuclear Factor 1α and PDX-1 in Transcriptional Regulation of the pdx-1 Gene. J. Biol. Chem. *276*, 47775–47784.

Gerrish K., V.V.J.C.S.R., Gerrish, K., Van Velkinburgh, J.C., and Stein, R. (2004). Conserved transcriptional regulatory domains of the pdx-1 gene. Mol. Endocrinol. *18*, 533–548.

Golosow, N., and Grobstein, C. (1962). Epitheliomesenchymal interaction in pancreatic morphogenesis. Dev. Biol. *4*, 242–255.

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

Gragnoli, C., Stanojevic, V., Gorini, A., Von Preussenthal, G.M., Thomas, M.K., and Habener, J.F. (2005). IPF-1/MODY4 gene missense mutation in an Italian family with type 2 and gestational diabetes. Metabolism *54*, 983–988.

Grarup, N., Moltke, I., Andersen, M.K., Bjerregaard, P., Larsen, C.V.L., Dahl-Petersen, I.K., Jørsboe, E., Tiwari, H.K., Hopkins, S.E., Wiener, H.W., et al. (2018). Identification of novel high-impact recessively inherited type 2 diabetes risk variants in the Greenlandic population. Diabetologia *61*, 2005–2015.

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

GURDON, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. *10*, 622–640.

Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C. V, and Teitelman, G. (1995). Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. Development *121*, 11–18.

Hang, Y., and Stein, R. (2011). MafA and MafB activity in pancreatic β cells. Trends Endocrinol. Metab. *22*, 364–373.

Hani, E.H., Stoffers, D.A., Chèvre, J.C., Durand, E., Stanojevic, V., Dina, C., Habener, J.F., and Froguel, P. (1999). Defective mutations in the insulin promoter factor-1 (IPF-1) gene in late-onset type 2 diabetes mellitus. J. Clin. Invest. *104*.

Hattersley, A.T., and Ashcroft, F.M. (2005). Activating mutations in Kir6.2 and neonatal diabetes: New clinical syndromes, new scientific insights, and new therapy. Diabetes *54*, 2503–2513.

Hattersley, A.T., and Patel, K.A. (2017). Precision diabetes: learning from monogenic diabetes. Diabetologia *60*, 769–777.

Hendriks, W.T., Warren, C.R., and Cowan, C.A. (2016). Genome Editing in Human Pluripotent Stem Cells: Approaches, Pitfalls, and Solutions. Cell Stem Cell *18*, 53–65.

Henseleit, K.D. (2005). NKX6 transcription factor activity is required for α- and β-cell development in the pancreas. Development *132*, 3139–3149.

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

Heuvel-Borsboom, H., de Valk, H.W., Losekoot, M., and Westerink, J. (2016). Maturity onset diabetes of the young: Seek and you will find. Neth. J. Med. *74*, 193–200.

Holland, A.M., Góñez, L.J., Naselli, G., MacDonald, R.J., and Harrison, L.C. (2005). Conditional expression demonstrates the role of the homeodomain transcription factor Pdx1 in maintenance and regeneration of beta-cells in the adult pancreas. Diabetes *54*, 2586–2595.

Holmstrom, S.R., Deering, T., Swift, G.H., Poelwijk, F.J., Mangelsdorf, D.J., Kliewer, S.A., and Macdonald, R.J. (2011). LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. Genes Dev. *25*, 1674–1679.

Horikawa, Y., Iwasaki, N., Hara, M., Furuta, H., Hinokio, Y., Cockburn, B.N., Lindner, T., Yamagata, K., Ogata, M., Tomonaga, O., et al. (1997). Mutation in hepatocyte nuclear factor-1ß gene (TCF2) associated with MODY. Nat. Genet. *17*, 384–385.

Horn, S., Kobberup, S., Jorgensen, M.C., Kalisz, M., Klein, T., Kageyama, R., Gegg, M., Lickert, H., Lindner, J., Magnuson, M.A., et al. (2012). Mind bomb 1 is required for pancreatic β-cell formation. Proc. Natl. Acad. Sci. *109*, 7356–7361.

Huypens, P., Ling, Z., Pipeleers, D., and Schuit, F. (2000). Glucagon receptors on human islet cells contribute to glucose competence of insulin release. Diabetologia *43*, 1012–1019.

Inoue, H., Riggs, A.C., Tanizawa, Y., Ueda, K., Kuwano, A., Liu, L., Donis-Keller, H., and Permutt, M.A. (1996). Isolation, characterization, and chromosomal mapping of the human insulin promoter factor 1 (IPF-1) gene. Diabetes *45*, 789–794.

International Diabetes Foundation (2017). IDF Diabetes Atlas -8th Edition 2017.

James, D. (2005). TGF /activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273–1282.

Jennings, R.E., Berry, A.A., Strutt, J.P., Gerrard, D.T., and Hanley, N.A. (2015). Human pancreas development. Development *142*, 3126–3137.

Jensen, J., Pedersen, E.E., Galante, P., Hald, J., Heller, R.S., Ishibashi, M., Kageyama, R., Guillemot, F., Serup, P., and Madsen, O.D. (2000). Control of endodermal endocrine development by Hes-1. Nat. Genet. *24*, 36–44.

Johnson, J.D., Ahmed, N.T., Luciani, D.S., Han, Z., Tran, H., Fujita, J., Misler, S., Edlund, H., and Polonsky, K.S. (2003). Increased islet apoptosis in Pdx1+/- mice. J. Clin. Invest. *111*, 1147–1160.

Jonckheere, N., Mayes, E., Shih, H.P., Li, B., Lioubinski, O., Dai, X., and Sander, M. (2008). Analysis of mPygo2 mutant mice suggests a requirement for mesenchymal Wnt signaling in pancreatic growth and differentiation. Dev. Biol. *318*, 224–235.

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

Kageyama, R., Ohtsuka, T., and Kobayashi, T. (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. Development *134*, 1243–1251.

Kahn, S.E., Cooper, M.E., and Del Prato, S. (2014). Pathophysiology and treatment of type 2 diabetes: Perspectives on the past, present, and future. Lancet *383*, 1068–1083.

Kameswaran, V., Bramswig, N.C., McKenna, L.B., Penn, M., Schug, J., Hand, N.J., Chen, Y., Choi, I., Vourekas, A., Won, K.J., et al. (2014). Epigenetic regulation of the DLK1-MEG3 MicroRNA cluster in human type 2 diabetic islets. Cell Metab. *19*, 135–145.

Kang, H.S., Kim, Y.-S., ZeRuth, G., Beak, J.Y., Gerrish, K., Kilic, G., Sosa-Pineda, B., Jensen, J., Pierreux, C.E., Lemaigre, F.P., et al. (2010). Transcription Factor Glis3, a Novel Critical Player in the Regulation of Pancreatic β-Cell Development and Insulin Gene Expression. Mol. Cell. Biol. *30*, 1864–1864.

Kesavan, G., Sand, F.W., Greiner, T.U., Johansson, J.K., Kobberup, S., Wu, X., Brakebusch, C., and Semb, H. (2009). Cdc42-Mediated Tubulogenesis Controls Cell Specification. Cell *139*, 791–801.

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

Kim, D., Kim, C.-H., Moon, J.-I., Chung, Y.-G., Chang, M.-Y., Han, B.-S., Ko, S., Yang, E., Cha, K.Y., Lanza, R., et al. (2009). Generation of Human Induced Pluripotent Stem Cells by Direct Delivery of Reprogramming Proteins. Cell Stem Cell *4*, 472–476.

Kim, W., Shin, Y.K., Kim, B.J., and Egan, J.M. (2010). Notch signaling in pancreatic endocrine cell and diabetes. Biochem. Biophys. Res. Commun. *392*, 247–251.

Kobayashi, Y., Okada, Y., Itakura, G., Iwai, H., Nishimura, S., Yasuda, A., Nori, S., Hikishima, K., Konomi, T., Fujiyoshi, K., et al. (2012). Pre-Evaluated Safe Human iPSC-Derived Neural Stem Cells Promote Functional Recovery after Spinal Cord Injury in Common Marmoset without Tumorigenicity. PLoS One *7*.

Kondo, Y., Toyoda, T., Inagaki, N., and Osafune, K. (2018). iPSC technology-based regenerative therapy for diabetes. J. Diabetes Investig. *9*, 234–243.

Krapp, A., Knöfler, M., Ledermann, B., Bürki, K., Berney, C., Zoerkler, N., Hagenbüchle, O., and Wellauer, P.K. (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*,

3752–3763.

Leonardo, T.R., Schultheisz, H.L., Loring, J.F., and Laurent, L.C. (2012). The functions of microRNAs in pluripotency and reprogramming. Nat. Cell Biol. *14*, 1114–1121.

Lickert, H., Domon, C., Huls, G., Wehrle, C., Duluc, I., Clevers, H., Meyer, B.I., Freund, J.N., and Kemler, R. (2000). Wnt/(beta)-catenin signaling regulates the expression of the homeobox gene Cdx1 in embryonic intestine. Development *127*, 3805–3813.

Locke, J.M., Hysenaj, G., Wood, A.R., Weedon, M.N., and Harries, L.W. (2015). Targeted allelic expression profiling in human islets identifies cis-regulatory effects for multiple variants identified by type 2 diabetes genome-wide association studies. Diabetes *64*, 1484–1491.

Löf-Öhlin, Z.M., Nyeng, P., Bechard, M.E., Hess, K., Bankaitis, E., Greiner, T.U., Ameri, J., Wright, C. V., and Semb, H. (2017). EGFR signalling controls cellular fate and pancreatic organogenesis by regulating apicobasal polarity. Nat. Cell Biol. *19*, 1313–1325.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. Nature *525*, 469–478.

Ma, S., Viola, R., Sui, L., Cherubini, V., Barbetti, F., and Egli, D. (2018). β Cell Replacement after Gene Editing of a Neonatal Diabetes-Causing Mutation at the Insulin Locus. Stem Cell Reports *11*, 1407–1415.

Macfarlane, W.M., Macfarlane, W.M., Frayling, T.M., Frayling, T.M., Ellard, S., Ellard, S., Evans, J.C., Evans, J.C., Allen, L.I., Allen, L.I., et al. (1999). Missense mutations in the insulin promoter factor-1 gene predispose to type 2 diabetes. J. Clin. Invest. *104*, R33-NaN-39.

Mamidi, A., Prawiro, C., Seymour, P.A., de Lichtenberg, K.H., Jackson, A., Serup, P., and Semb, H. (2018). Mechanosignalling via integrins directs fate decisions of pancreatic progenitors. Nature *564*, 114–118.

Mangino, M., Hwang, S.J., Spector, T.D., Hunt, S.C., Kimura, M., Fitzpatrick, A.L., Christiansen, L., Petersen, I., Elbers, C.C., Harris, T., et al. (2012). Genome-wide meta-analysis points to CTC1 and ZNf676 as genes regulating telomere homeostasis in humans. Hum. Mol. Genet. *21*, 5385–5394.

Marrades, M.P., González-Muniesa, P., Martínez, J.A., and Moreno-Aliaga, M.J. (2010). A dysregulation in CES1, APOE and other lipid metabolism-related genes is associated to cardiovascular risk factors linked to obesity. Obes. Facts *3*, 312–318.

Marshak, S., Benshushan, E., Shoshkes, M., Havin, L., Cerasi, E., and Melloul, D. (2000). Functional conservation of regulatory elements in the pdx-1 gene: PDX-1 and hepatocyte nuclear factor 3beta transcription factors mediate beta-cell-specific expression. Mol. Cell. Biol. *20*, 7583–7590.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. *78*, 7634–7638.

Masui, T., Long, Q., Beres, T.M., Magnuson, M.A., and MacDonald, R.J. (2007). Early pancreatic development requires the vertebrate Suppressor of Hairless (RBPJ) in the PTF1 bHLH complex. Genes Dev. *21*, 2629–2643.

Masui, T., Swift, G.H., Deering, T., Shen, C., Coats, W.S., Long, Q., Elsässer, H.P., Magnuson, M.A., and MacDonald, R.J. (2010). Replacement of Rbpj With Rbpjl in the PTF1 Complex Controls the Final Maturation of Pancreatic Acinar Cells. Gastroenterology *139*, 270–280.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. *18*, 4261–4269.

McCarthy, M.I., and Zeggini, E. (2009). Genome-wide association studies in type 2 diabetes. Curr. Diab. Rep. *9*, 164–171.

Meier, J.J., and Bonadonna, R.C. (2013). Role of reduced β-cell mass versus impaired β-cell function in the pathogenesis of type 2 diabetes. Diabetes Care *36*.

Meier, J.J., Butler, A.E., Saisho, Y., Monchamp, T., Galasso, R., Bhushan, A., Rizza, R.A., and Butler, P.C. (2008). β-cell replication is the primary mechanism subserving the postnatal expansion of β-cell mass in humans. Diabetes *57*, 1584–1594.

Millership, S.J., Da Silva Xavier, G., Choudhury, A.I., Bertazzo, S., Chabosseau, P., Pedroni, S.M., Irvine, E.E., Montoya, A., Faull, P., Taylor, W.R., et al. (2018). Neuronatin regulates pancreatic β cell insulin content and secretion. J. Clin. Invest. *128*, 3369–3381.

Murphy, R., Ellard, S., and Hattersley, A.T. (2008). Clinical implications of a molecular genetic classification of monogenic β-cell diabetes. Nat. Clin. Pract. Endocrinol. Metab. *4*, 200–213.

Murtaugh, L.C. (2005). β-Catenin is essential for pancreatic acinar but not islet development. Development *132*, 4663–4674.

Nair, G., and Hebrok, M. (2015). Islet formation in mice and men: Lessons for the generation of functional insulin-producing β-cells from human pluripotent stem cells. Curr. Opin. Genet. Dev. *32*, 171–180.

Nair, G.G., Liu, J.S., Russ, H.A., Tran, S., Saxton, M.S., Chen, R., Juang, C., Li, M. Ian, Nguyen, V.Q., Giacometti, S., et al. (2019). Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells. Nat. Cell Biol. *21*, 263–274.

Nishimura, W., Kondo, T., Salameh, T., El Khattabi, I., Dodge, R., Bonner-Weir, S., and Sharma, A. (2006). A switch from MafB to MafA expression accompanies differentiation to pancreatic β-cells. Dev. Biol. *293*, 526–539.

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

Ohlsson, H., Karlsson, K., and Edlund, T. (1993). IPF1, a homeodomain-containing transactivator of the insulin gene. EMBO J. *12*, 4251–4259.

Oliver-Krasinski, J.M., and Stoffers, D.A. (2008). On the origin of the beta cell. Genes Dev *22*, 1998–2021.

Oliver-Krasinski, J.M., Kasner, M.T., Yang, J., Crutchlow, M.F., Rustgi, A.K., Kaestner, K.H., and Stoffers, D.A. (2009). The diabetes gene Pdx1 regulates the transcriptional network of pancreatic endocrine progenitor cells in mice. J. Clin. Invest. *119*, 1888–1898.

Pagliuca, F.W., Millman, J.R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J.H., Peterson, Q.P., Greiner, D., and Melton, D.A. (2014). Generation of functional human pancreatic β cells in vitro. Cell *159*, 428–439.

Painter, J.N., O'Mara, T.A., Batra, J., Cheng, T., Lose, F.A., Dennis, J., Michailidou, K., Tyrer, J.P., Ahmed, S., Ferguson, K., et al. (2015). Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. Hum. Mol. Genet. *24*, 1478–1492.

Pan, F.C., and Brissova, M. (2014). Pancreas development in humans. Curr. Opin. Endocrinol. Diabetes Obes. *21*, 77–82.

Pan, F.C., and Wright, C. (2011). Pancreas organogenesis: From bud to plexus to gland. Dev. Dyn. *240*, 530–565.

Pan, F.C., Brissova, M., Powers, A.C., Pfaff, S., and Wright, C.V.E. (2015). Inactivating the permanent neonatal diabetes gene Mnx1 switches insulin-producing β-cells to a δ-like fate

and reveals a facultative proliferative capacity in aged β-cells. Development *142*, 3637–3648.

Pasquali, L., Gaulton, K.J., Rodríguez-Seguí, S.A., Mularoni, L., Miguel-Escalada, I., Akerman, I., Tena, J.J., Morán, I., Gómez-Marín, C., Van De Bunt, M., et al. (2014). Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. Nat. Genet. *46*, 136–143.

Pearson, E.R., Starkey, B.J., Powell, R.J., Gribble, F.M., Clark, P.M., and Hattersley, A.T. (2003). Genetic cause of hyperglycaemia and response to treatment in diabetes. Lancet *362*, 1275–1281.

Petri, A., Ahnfelt-Rønne, J., Frederiksen, K.S., Edwards, D.G., Madsen, D., Serup, P., Fleckner, J., and Heller, R.S. (2006). The effect of neurogenin3 deficiency on pancreatic gene expression in embryonic mice. J. Mol. Endocrinol. *37*, 301–316.

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

Quiroga, A.D., Li, L., Trötzmüller, M., Nelson, R., Proctor, S.D., Köfeler, H., and Lehner, R. (2012). Deficiency of carboxylesterase 1/esterase-x results in obesity, hepatic steatosis, and hyperlipidemia. Hepatology *56*, 2188–2198.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Reith, W., Ucla, C., Barras, E., Gaud, a, Durand, B., Herrero-Sanchez, C., Kobr, M., and Mach, B. (1994). RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. Mol. Cell. Biol. *14*, 1230–1244.

Rezania, A., Bruin, J.E., Riedel, M.J., Mojibian, M., Asadi, A., Xu, J., Gauvin, R., Narayan, K., Karanu, F., O'Neil, J.J., et al. (2012). Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. Diabetes *61*, 2016–2029.

Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., O'Dwyer, S., Quiskamp, N., Mojibian, M., Albrecht, T., et al. (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nat. Biotechnol. *32*, 1121–1133.

Robinton, D.A., and Daley, G.Q. (2012). The promise of induced pluripotent stem cells in research and therapy. Nature *481*, 295–305.

Rosa, A., and Ballarino, M. (2016). Long Noncoding RNA Regulation of Pluripotency. Stem Cells Int. *2016*.

Russ, H. a, Parent, A. V, Ringler, J.J., Hennings, T.G., Nair, G.G., Shveygert, M., Guo, T., Puri, S., Haataja, L., Cirulli, V., et al. (2015). Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. EMBO J. *34*, e201591058.

RUTTER, W.J., WESSELLS, N.K., and GROBSTEIN, C. (1964). CONTROL OF SPECIFIC SYNTHESIS IN THE DEVELOPING PANCREAS. Natl. Cancer Inst. Monogr. *13*, 51–65.

Saarimäki-Vire, J., Balboa, D., Russell, M.A., Saarikettu, J., Kinnunen, M., Keskitalo, S., Malhi, A., Valensisi, C., Andrus, C., Eurola, S., et al. (2017). An Activating STAT3 Mutation Causes Neonatal Diabetes through Premature Induction of Pancreatic Differentiation. Cell Rep. *19*, 281–294.

Sánchez-Arévalo Lobo, V.J., Fernández, L.C., Carrillo-De-Santa-Pau, E., Richart, L., Cobo, I., Cendrowski, J., Moreno, U., Del Pozo, N., Megías, D., Bréant, B., et al. (2018). C-Myc downregulation is required for preacinar to acinar maturation and pancreatic homeostasis. Gut *67*, 707–718.

Schäfer, S.A., Tschritter, O., Machicao, F., Thamer, C., Stefan, N., Gallwitz, B., Holst, J.J., Dekker, J.M., T'Hart, L.M., Nijpels, G., et al. (2007). Impaired glucagon-like peptide-1-induced insulin secretion in carriers of transcription factor 7-like 2 (TCF7L2) gene polymorphisms. Diabetologia *50*, 2443–2450.

Schaffer, A.E., Freude, K.K., Nelson, S.B., and Sander, M. (2010). Nkx6 transcription factors and Ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. Dev. Cell *18*, 1022–1029.

Seymour, P.A., Freude, K.K., Tran, M.N., Mayes, E.E., Jensen, J., Kist, R., Scherer, G., and Sander, M. (2007). SOX9 is required for maintenance of the pancreatic progenitor cell pool. Proc. Natl. Acad. Sci. *104*, 1865–1870.

Seymour, P.A., Shih, H.P., Patel, N.A., Freude, K.K., Xie, R., Lim, C.J., and Sander, M. (2012). A Sox9/Fgf feed-forward loop maintains pancreatic organ identity. Development *139*, 3363–3372.

Shang, L., Hua, H., Foo, K., Martinez, H., Watanabe, K., Zimmer, M., Kahler, D.J., Freeby, M., Chung, W., LeDuc, C., et al. (2014). β-cell dysfunction due to increased ER stress in a stem cell model of wolfram syndrome. Diabetes *63*, 923–933.

Shastry, B.S. (2002). SNP alleles in human disease and evolution. J Hum Genet *47*, 561–

566.

Shi, Z.D., Lee, K., Yang, D., Amin, S., Verma, N., Li, Q. V., Zhu, Z., Soh, C.L., Kumar, R., Evans, T., et al. (2017). Genome Editing in hPSCs Reveals GATA6 Haploinsufficiency and a Genetic Interaction with GATA4 in Human Pancreatic Development. Cell Stem Cell *20*, 675–688.e6.

Shih, H.P., Kopp, J.L., Sandhu, M., Dubois, C.L., Seymour, P.A., Grapin-Botton, A., and Sander, M. (2012). A Notch-dependent molecular circuitry initiates pancreatic endocrine and ductal cell differentiation. Development *139*, 2488–2499.

Shih, H.P., Wang, A., and Sander, M. (2013). Pancreas Organogenesis: From Lineage Determination to Morphogenesis. Annu. Rev. Cell Dev. Biol. *29*, 81–105.

Shu, L., Sauter, N.S., Schulthess, F.T., Matveyenko, A. V, Oberholzer, J., and Maedler, K. (2008). Transcription factor 7-like 2 regulates beta-cell survival and function in human pancreatic islets. Diabetes *57*, 645–653.

da Silva Xavier, G., Mondragon, A., Sun, G., Chen, L., McGinty, J.A., French, P.M., and Rutter, G.A. (2012). Abnormal glucose tolerance and insulin secretion in pancreas-specific Tcf7l2-null mice. Diabetologia *55*, 2667–2676.

Smid, J.K., Faulkes, S., and Rudnicki, M.A. (2015). Periostin induces pancreatic regeneration. Endocrinology *156*, 824–836.

Staffers, D.A., Ferrer, J., Clarke, W.L., and Habener, J.F. (1997). Early-onset type-ll diabetes mellitus (MODY4) linked to IPF1. Nat. Genet. *17*, 138–139.

Stoffel, M., Stein, R., Wright, C.V.F., Espinosa, R., Le Beau, M.M., and Bell, G.I. (1995). Localization of Human Homeodomain Transcription Factor Insulin Promoter Factor I (IPF1) to Chromosome Band 13q12.1. Genomics *28*, 125–126.

Stoffers, D.A., Zinkin, N.T., Stanojevic, V., Clarke, W.L., and Habener, J.F. (1997). Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. Nat. Genet. *15*, 106–110.

Stoffers, D.A., Stanojevic, V., and Habener, J.F. (1998). Insulin promoter factor-1 gene mutation linked to early-onset type 2 diabetes mellitus directs expression of a dominant negative isoprotein. J. Clin. Invest. *102*, 232–241.

Svensson, P., Williams, C., Lundeberg, J., Rydén, P., Bergqvist, I., and Edlund, H. (2007). Gene array identification of Ipf1/Pdx1-/- regulated genes in pancreatic progenitor cells. BMC Dev. Biol. *7*.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takahashi, K., and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. Nat. Rev. Mol. Cell Biol. *17*, 183–193.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872.

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

Tattersall, R.B. (1974). Mild familial diabetes with dominant inheritance. QJM *43*, 339–357.

Teo, A.K.K., Tsuneyoshi, N., Hoon, S., Tan, E.K., Stanton, L.W., Wright, C.V.E., and Dunn, N.R. (2015). PDX1 Binds and Represses Hepatic Genes to Ensure Robust Pancreatic Commitment in Differentiating Human Embryonic Stem Cells. Stem Cell Reports *4*, 578–590.

Teo, A.K.K., Lau, H.H., Valdez, I.A., Dirice, E., Tjora, E., Raeder, H., and Kulkarni, R.N. (2016). Early Developmental Perturbations in a Human Stem Cell Model of MODY5/HNF1B Pancreatic Hypoplasia. Stem Cell Reports *6*, 357–367.

Thomson, J.A. (1998). Embryonic stem cell lines derived from human blastocysts. Science (80-.). *282*, 1145–1147.

Thorens, B. (2014). Neural regulation of pancreatic islet cell mass and function. Diabetes, Obes. Metab. *16*, 87–95.

Thorens, B. (2015). GLUT2, glucose sensing and glucose homeostasis. Diabetologia *58*, 221–232.

Tiyaboonchai, A., Cardenas-Diaz, F.L., Ying, L., Maguire, J.A., Sim, X., Jobaliya, C., Gagne, A.L., Kishore, S., Stanescu, D.E., Hughes, N., et al. (2017). GATA6 Plays an Important Role in the Induction of Human Definitive Endoderm, Development of the Pancreas, and Functionality of Pancreatic β Cells. Stem Cell Reports *8*, 589–604.

Tulachan, S.S., Doi, R., Hirai, Y., Kawaguchi, Y., Koizumi, M., Hembree, M., Tei, E., Crowley, A., Yew, H., McFall, C., et al. (2006). Mesenchymal epimorphin is important for pancreatic duct morphogenesis. Dev. Growth Differ. *48*, 65–72.

Vallier, L. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell Sci. *118*, 4495–4509.

Velazco-Cruz, L., Song, J., Maxwell, K.G., Goedegebuure, M.M., Augsornworawat, P., Hogrebe, N.J., and Millman, J.R. (2019). Acquisition of Dynamic Function in Human Stem Cell-Derived β Cells. Stem Cell Reports *12*, 351–365.

Veres, A., Faust, A.L., Bushnell, H.L., Engquist, E.N., Kenty, J.H.R., Harb, G., Poh, Y.C., Sintov, E., Gürtler, M., Pagliuca, F.W., et al. (2019). Charting cellular identity during human in vitro β-cell differentiation. Nature *569*, 368–373.

Villasenor, A., Chong, D.C., Henkemeyer, M., and Cleaver, O. (2010). Epithelial dynamics of pancreatic branching morphogenesis. Development *137*, 4295–4305.

Vrang, N., Meyre, D., Froguel, P., Jelsing, J., Tang-Christensen, M., Vatin, V., Mikkelsen, J.D., Thirstrup, K., Larsen, L.K., Cullberg, K.B., et al. (2010). The imprinted gene neuronatin is regulated by metabolic status and associated with obesity. Obesity *18*, 1289–1296.

Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N.A., Muth, K., Palmer, J., Qiu, Y., Wang, J., et al. (2015). Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. Cell Stem Cell *16*, 386–399.

Wang, S., Yan, J., Anderson, D.A., Xu, Y., Kanal, M.C., Cao, Z., Wright, C.V.E., and Gu, G. (2010). Neurog3 gene dosage regulates allocation of endocrine and exocrine cell fates in the developing mouse pancreas. Dev. Biol. *339*, 26–37.

Wang, X., Sterr, M., Ansarullah, Burtscher, I., Böttcher, A., Beckenbauer, J., Siehler, J., Meitinger, T., Häring, H.-U., Staiger, H., et al. (2019). Point mutations in the PDX1 transactivation domain impair human β-cell development and function. Mol. Metab.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., et al. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell *7*, 618–630.

Wells, J.M., Esni, F., Boivin, G.P., Aronow, B.J., Stuart, W., Combs, C., Sklenka, A., Leach, S.D., and Lowy, A.M. (2007). Wnt/β-catenin signaling is required for development of the exocrine pancreas. BMC Dev. Biol. *7*.

Wessells, N.K., and Cohen, J.H. (1967). Early pancreas organogenesis: Morphogenesis, tissue interactions, and mass effects. Dev. Biol. *15*, 237–270.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H.S. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810–813.

Wojtusciszyn, A., Armanet, M., Morel, P., Berney, T., and Bosco, D. (2008). Insulin secretion

from human beta cells is heterogeneous and dependent on cell-to-cell contacts. Diabetologia *51*, 1843–1852.

Xu, X., Browning, V.L., and Odorico, J.S. (2011). Activin, BMP and FGF pathways cooperate to promote endoderm and pancreatic lineage cell differentiation from human embryonic stem cells. Mech. Dev. *128*, 412–427.

Yan, W., and Shao, R. (2006). Transduction of a mesenchyme-specific gene periostin into 293T cells induces cell invasive activity through epithelial-mesenchymal transformation. J. Biol. Chem. *281*, 19700–19709.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281–292.

You, L., Wang, N., Yin, D., Wang, L., Jin, F., Zhu, Y., Yuan, Q., and De, W. (2016). Downregulation of Long Noncoding RNA Meg3 Affects Insulin Synthesis and Secretion in Mouse Pancreatic Beta Cells. J. Cell. Physiol. *231*, 852–862.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science (80-.). *318*, 1917–1920.

Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J., and Jiang, Y. (2011). Efficient feeder-free episomal reprogramming with small molecules. PLoS One *6*.

Zeng, H., Guo, M., Zhou, T., Tan, L., Chong, C.N., Zhang, T., Dong, X., Xiang, J.Z., Yu, A.S., Yue, L., et al. (2016). An Isogenic Human ESC Platform for Functional Evaluation of Genome-wide-Association-Study-Identified Diabetes Genes and Drug Discovery. Cell Stem Cell *19*, 326–340.

Zhang, X., Zhang, J., Wang, T., Esteban, M.A., and Pei, D. (2008). Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. J. Biol. Chem. *283*, 35825–35833.

Zhou, Q., and Melton, D.A. (2018). Pancreas regeneration. Nature *557*, 351–358.

Zhou, Q., Law, A.C., Rajagopal, J., Anderson, W.J., Gray, P.A., and Melton, D.A. (2007). A Multipotent Progenitor Domain Guides Pancreatic Organogenesis. Dev. Cell *13*, 103–114.

Zhu, Z., Li, Q. V., Lee, K., Rosen, B.P., González, F., Soh, C.L., and Huangfu, D. (2016). Genome Editing of Lineage Determinants in Human Pluripotent Stem Cells Reveals Mechanisms of Pancreatic Development and Diabetes. Cell Stem Cell *18*, 755–768.

6 Acknowledgements

7 Publications

Wang X, Sterr M, Ansarullah, Burtscher I, Böttcher A, Beckenbauer J, Siehler J, Meitinger T, Häring HU, Staiger H, Cernilogar FM, Schotta G, Irmler M, Beckers J, Wright CVE, Bakhti M, Lickert H. Point mutations in the PDX1 transactivation domain impair human β-cell development and function. Mol Metab. 2019 Mar 20. pii: S2212-8778(19)30134-6. doi: 10.1016/j.molmet.2019.03.006. Epub 2019 Mar 20.

Wang X, Sterr M, Burtscher I, Chen S, Hieronimus A, Machicao F, Staiger H, Häring HU, Lederer G, Meitinger T, Cernilogar FM, Schotta G, Irmler M, Beckers J, Hrabě de Angelis M, Ray M, Wright CVE, Bakhti M, Lickert H. Genome-wide analysis of PDX1 target genes in human pancreatic progenitors. Mol Metab. 2018 Mar;9:57-68. doi: 10.1016/j.molmet.2018.01.011. Epub 2018 Jan 31.

Wang X, Chen S, Burtscher I, Sterr M, Hieronimus A, Machicao F, Staiger H, Häring HU, Lederer G, Meitinger T, Lickert H. Generation of a human induced pluripotent stem cell (iPSC) line from a patient with family history of diabetes carrying a C18R mutation in the PDX1 gene. Stem Cell Res. 2016 Sep;17(2):292-295. doi: 10.1016/j.scr.2016.08.005. Epub 2016 Aug 6.

Wang X, Chen S, Burtscher I, Sterr M, Hieronimus A, Machicao F, Staiger H, Häring HU, Lederer G, Meitinger T, Lickert H. Generation of a human induced pluripotent stem cell (iPSC) line from a patient carrying a P33T mutation in the PDX1 gene. Stem Cell Res. 2016 Sep;17(2):273-276. doi: 10.1016/j.scr.2016.08.004. Epub 2016 Aug 5.

Benda C, Zhou T, **Wang X**, Tian W, Grillari J, Tse HF, Grillari-Voglauer R, Pei D, Esteban MA. Urine as a source of stem cells. Adv Biochem Eng Biotechnol. 2013;129:19-32. doi: 10.1007/10_2012_157.

Li W, **Wang X**, Fan W, Zhao P, Chan YC, Chen S, Zhang S, Guo X, Zhang Y, Li Y, Cai J, Qin D, Li X, Yang J, Peng T, Zychlinski D, Hoffmann D, Zhang R, Deng K, Ng KM, Menten B, Zhong M, Wu J, Li Z, Chen Y, Schambach A, Tse HF, Pei D, Esteban MA. Modeling Abnormal Early Development with Induced Pluripotent Stem Cells from Aneuploid Syndromes. Hum Mol Genet. 2012 Jan 1;21(1):32-45. Epub 2011 Sep 23. **(Co-first author)**

Yang J, Cai J, Zhang Y, **Wang X**, Li W, Xu J, Li F, Guo X, Deng K, Zhong M, Chen Y, Lai L, Pei D, Esteban MA. Induced Pluripotent Stem Cells Can Be Used to Model the Genomic Imprinting Disorder Prader-Willi Syndrome. J Biol Chem. 2010 Dec 17;285(51):40303-11. Epub 2010 Oct 18.

8 Abbreviations

ChIP-seq	ChIP-sequencing
CRISPR	clustered regularly interspaced short palindromic repeats
DE	definitive endoderm
EN	endocrine precursors
FACS	fluorescent activated cell sorting
gRNA	guide RNA
GSIS	glucose-stimulated insulin secretion
hESC	human embryonic stem cell
iPSC	induced pluripotent stem cell
mESC	mouse embryonic stem cell
MODY	maturity-onset diabetes of the young
PDX1	pancreas/duodenum homeobox protein 1
PGT	primitive gut tube
PP1	pancreatic progenitor stage 1/ posterior foregut
PP2	pancreatic progenitor stage 2/ pancreatic endoderm
qPCR	quantitative polymerase chain reaction
RNA-seq	RNA sequencing
SC-β	stem cell derived-β cell
SCNT	somatic cell nuclear transfer
SNP	single-nucleotide polymorphism
ssDNA	single-stranded DNA
TALEN	transcription activator-like effector nuclease
ZFNs	zinc-finger nucleases

Genome-wide analysis of PDX1 target genes in human pancreatic progenitors

Xianming Wang [1,2,9,15], Michael Sterr [1,2,9,15], Ingo Burtscher [1,2], Shen Chen [3], Anja Hieronimus [4,5,10], Fausto Machicao [6,10], Harald Staiger [4,7,10], Hans-Ulrich Häring [4,5,10], Gabriele Lederer [8], Thomas Meitinger [8], Filippo M. Cernilogar [11], Gunnar Schotta [11], Martin Irmler [6], Johannes Beckers [6,10,12], Martin Hrabě de Angelis [6,10,12], Michael Ray [13,14], Christopher V.E. Wright [13,14], Mostafa Bakhti [1,2,10], Heiko Lickert [1,2,9,10,*]

ABSTRACT

Objective: Homozygous loss-of-function mutations in the gene coding for the homeobox transcription factor (TF) PDX1 leads to pancreatic agenesis, whereas heterozygous mutations can cause Maturity-Onset Diabetes of the Young 4 (MODY4). Although the function of Pdx1 is well studied in pre-clinical models during insulin-producing β-cell development and homeostasis, it remains elusive how this TF controls human pancreas development by regulating a downstream transcriptional program. Also, comparative studies of PDX1 binding patterns in pancreatic progenitors and adult β-cells have not been conducted so far. Furthermore, many studies reported the association between single nucleotide polymorphisms (SNPs) and T2DM, and it has been shown that islet enhancers are enriched in T2DM-associated SNPs. Whether regions, harboring T2DM-associated SNPs are PDX1 bound and active at the pancreatic progenitor stage has not been reported so far.

Methods: In this study, we have generated a novel induced pluripotent stem cell (iPSC) line that efficiently differentiates into human pancreatic progenitors (PPs). Furthermore, PDX1 and H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) was used to identify PDX1 transcriptional targets and active enhancer and promoter regions. To address potential differences in the function of PDX1 during development and adulthood, we compared PDX1 binding profiles from PPs and adult islets. Moreover, combining ChIP-seq and GWAS meta-analysis data we identified T2DM-associated SNPs in PDX1 binding sites and active chromatin regions.

Results: ChIP-seq for PDX1 revealed a total of 8088 PDX1-bound regions that map to 5664 genes in iPSC-derived PPs. The PDX1 target regions include important pancreatic TFs, such as *PDX1* itself, *RFX6*, *HNF1B*, and *MEIS1*, which were activated during the differentiation process as revealed by the active chromatin mark H3K27ac and mRNA expression profiling, suggesting that auto-regulatory feedback regulation maintains *PDX1* expression and initiates a pancreatic TF program. Remarkably, we identified several PDX1 target genes that have not been reported in the literature in human so far, including *RFX3,* required for ciliogenesis and endocrine differentiation in mouse, and the ligand of the Notch receptor *DLL1*, which is important for endocrine induction and tip-trunk patterning. The comparison of PDX1 profiles from PPs and adult human islets identified sets of stage-specific target genes, associated with early pancreas development and adult β-cell function, respectively. Furthermore, we found an enrichment of T2DM-associated SNPs in active chromatin regions from iPSC-derived PPs. Two of these SNPs fall into PDX1 occupied sites that are located in the intronic regions of *TCF7L2* and *HNF1B*. Both of these genes are key transcriptional regulators of endocrine induction and mutations in cis-regulatory regions predispose to diabetes.

MOLECULAR METABOLISM 9 (2018) 57–68 57
www.molecularmetabolism.com

Original Article

Conclusions: Our data provide stage-specific target genes of PDX1 during *in vitro* differentiation of stem cells into pancreatic progenitors that could be useful to identify pathways and molecular targets that predispose for diabetes. In addition, we show that T2DM-associated SNPs are enriched in active chromatin regions at the pancreatic progenitor stage, suggesting that the susceptibility to T2DM might originate from imperfect execution of a β-cell developmental program.

Keywords iPSC; T2DM; ChIP-seq; PDX1; SNPs; PP; GWAS

1. INTRODUCTION

Diabetes mellitus is a group of heterogeneous disorders characterized by hyperglycemia and the progressive loss or dysfunction of the insulin-producing β-cells in the pancreas. Today, over 415 million people worldwide have been diagnosed with diabetes and the number is expected to rise to 642 million by 2040 [1]. There are four main types of diabetes mellitus: Type 1 Diabetes Mellitus (T1DM), Type 2 Diabetes Mellitus (T2DM), Gestational Diabetes and Maturity Onset Diabetes of the Young (MODY).

The most common form of diabetes is T2DM, a multifactorial disease that is caused by a complex interplay between genetic, epigenetic, and environmental factors. This form predominantly occurs in adulthood due to the failure of β-cells to produce sufficient amounts of insulin for maintenance of normoglycemia. For decades, analyses of natural genetic variation have been used to understand the pathogenesis of T2DM and to improve diagnosis and treatment for the people who suffer from this disease [1,2].

MODY is inherited in an autosomal dominant fashion and presents early in life due to defects in developing enough β-cells or failure in function. There are at least 13 types of MODY that are diagnosed in 1—2% of diabetes patients, which are caused by mutations in several essential genes, such as *GCK*, *HNF1A*, *HNF1B*, and *PDX1* [3].

Among these genes, *PDX1* encodes one key TF, regulating β-cell development and function [4,5]. In humans, the *PDX1* gene is located on chromosome 13q12.1 and encodes for a protein of 283 amino acids. Typically for a TF it contains a transactivation domain and a homeodomain that binds to DNA. In mouse, the expression of Pdx1 is first evident at embryonic day (E) 8.5—9.0 and becomes restricted to β- and δ-cells in adult islets [6—9]. Homozygous Pdx1 knockout mice form pancreatic buds but fail to develop a pancreas [10]. On the contrary, heterozygous Pdx1 knockout mice develop a pancreas but become diabetic in adulthood and β-cells increasingly undergo apoptosis [11—13]. In humans, PDX1 is expressed in the developing pancreas and heterozygous mutations in the *PDX1* gene cause a strong form of monogenic diabetes, called MODY4 [14,15]. Contrary to the numerous studies highlighting the importance of Pdx1 during mouse pancreas development, little is known about the role of this TF in human β-cell development, homeostasis and function. Specifically, it is important to unravel the PDX1 target gene program to understand its cell-type specific function during development and its contribution to MODY and T2DM in adulthood. Genome-wide association studies have identified multiple loci associated with the susceptibility to T2DM, including *TCF7L2*, *SLC30A8*, *HHEX*, *CDKAL1*, *CDKN2A/B*, and *IGF2BP2* [16,17]. It has been shown that islet enhancers are enriched in T2DM-associated SNPs [18]; however, it is unknown if these SNPs affect PDX1 bound cis-regulatory regions and play a role during pancreas development.

Here, we generated a novel iPSC line from a healthy female donor and confirmed its efficient *in vitro* pancreatic differentiation. We performed transcriptome analysis combined with ChIP-seq profiling of active H3K27ac histone modifications and PDX1 binding sites in PPs and compared these to adult islets to investigate stage-specific functions of

PDX1 in progenitors and adult β-cells. Furthermore, through screening for T2DM-associated SNPs in active chromatin regions of PPs, we suggest that some SNPs might increase the diabetes risk by affecting pancreas and β-cell development.

2. MATERIALS AND METHODS

2.1. Ethics statement

The choice of appropriate human donors and the procedures for skin biopsy, isolation, and characterization of dermal fibroblasts were performed in accordance with study protocols approved by the Ethics Committee of the Medical Faculty of the Eberhard Karls University, Tübingen. The study design followed the principles of the Declaration of Helsinki. All study participants gave informed consent prior to entry into the study. All mice were housed in the facilities at the Helmholtz Zentrum München — German Research Center for Environmental Health (HMGU) and treated in accordance with the German animal welfare legislation and acknowledged guidelines of the Society of Laboratory Animals (GV-SOLAS) and of the Federation of Laboratory Animal Science Associations (FELASA). The teratoma generation procedure was approved by the Institutional Animal Care and Use Committee (IACUC) of HMGU and notified to the local regulatory supervisory authority.

2.2. Skin biopsy, isolation, and characterization of dermal fibroblasts

A full-thickness skin specimen was taken by punch biopsy from the upper arm in the deltoid muscle region. After removal of adipose tissue remnants and visible blood vessels, the sample was digested overnight at 4 °C with 10 U/mL dispase II (Roche Diagnostics, Mannheim, Germany) in 50 mM HEPES pH 7.4, 150 mM NaCl. Thereafter, the digest was incubated for 30 min at 37 °C under continuous shaking (1200 rpm). Using forceps, the dermis was separated from the epidermal layer, and fibroblasts were isolated from the dermis by digestion with 0.2% collagenase CLS I (Biochrom, Berlin, Germany) in DMEM, 10% BSA for 45 min at 37 °C under continuous shaking (1200 rpm). For purification of the fibroblasts, the digest was filtered through a 70 μm mesh and centrifuged. The pelleted cells were resuspended, grown for three days in DMEM, 10% FCS, and subsequently further expanded in Medium 106 supplemented with low serum growth supplement (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Samples of dermal fibroblasts were tested for the presence of viruses, pathogenic to humans, i.e., HBV, HCV, and HIV, with Genesig PCR-based detection kits from Primerdesign Ltd. (Chandler's Ford, UK) and for the presence of mycoplasma with a PCR test kit from PanReac AppliChem (Darmstadt, Germany) and, in parallel, by DNA staining with DAPI. All cultures were found to be negative for the tested contaminants.

2.3. XM001 iPSCs generation

Primary fibroblasts from a healthy female donor were reprogrammed into pluripotent stem cells using a non-integrating Episomal iPSC Reprogramming Kit (Invitrogen, Cat. no. A14703). This kit contains a

mixture of three vectors, which have the oriP/EBNA-1 (Epstein—Barr nuclear antigen-1) backbone that delivers six reprogramming factors: OCT4, SOX2, NANOG, LIN28, KLF4, and L-MYC [20]. Human fibroblasts, at 75—90% confluency, were transfected using the Amaxa 4D-Nucleofector transfection system and a nucleofector kit for human dermal fibroblasts (Lonza, Cat. no. VPD-1001), plated onto geltrex-coated culture dishes and incubated in supplemented fibroblast medium. The fibroblast medium contained knockout DMEM/F-12 (Life Technologies), 10% ESC-qualified FBS (Life Technologies), 1% MEM non-essential amino acids (Life Technologies), 10 μM HA-100 (Santa Cruz) and 4 ng/mL bFGF (Life Technologies). At 24 h after transfection, the medium was exchanged with N2B27 medium supplemented with 0.5 μM PD0325901 (Stemgent), 3 μM CHIR99021 (Stemgent), 0.5 μM A-83-01 (Stemgent), 10 μM HA-100 (Santa Cruz), 10 ng/mL hLIF (Life Technologies) and 100 ng/mL bFGF (Life Technologies). The basic N2B27 medium contained DMEM/F12 with HEPES (Life Technologies), $1 \times$ N2 supplement (Life Technologies), $1 \times$ B27 supplement (Life Technologies), 1% MEM non-essential amino acids, $1 \times$ Glutamax (Life Technologies) and $1 \times$ β-Mercaptoethanol (Life Technologies). On day 15 after transfection, the medium was changed to Essential 8 medium and monitored for the emergence of iPSC colonies. Around 3 weeks after transfection, undifferentiated iPSC colonies were picked and transferred onto fresh geltrex-coated culture dishes for expansion.

2.4. XM001 iPSCs characterization
DNA was extracted from iPSCs using standard procedure. Markers for the episomal backbone were amplified by semi-quantitative PCR to exclude transgene integration. Primers were as follows: oriP forward: TTCCAC-GAGGCTAGTGAACC. oriP reverse: TCGGGGGTGTTAGAGACAAC; EBNA-1 forward: ATCGTCAAAGCTGCACACAG. EBNA-1 reverse: CCCAGGAGTCC-CAGTAGTCA. For karyotype analysis, we used the cells growing in logarithmic phase. These were fed with fresh medium the day before adding colcemid for 2 h. Cells were then trypsinized, treated with hypotonic solution (0.075 M KCl) for 20 min, and fixed with methanol:acetic acid (3:1). Metaphases were spread on microscope slides, and chromosomes were classified according to the International System for Human Cytogenic Nomenclature using the standard G banding technique. At least 20 metaphases were counted per cell line, and the final karyotype was stated if it was present in more than 85% of them. For teratomas, 2×10^6 iPSCs were injected into the right hind leg of immunocompromised NOD/SCID mice. Tumors were excised after 8 weeks, fixed, embedded in paraffin, sectioned and stained with hematoxylin/eosin [51].

2.5. XM001 iPSCs culture
XM001 iPSCs were cultured on 1:100 diluted Matrigel (BD Biosciences) in mTeSR™1 medium (Stem Cell Technologies). At ~70—80% confluency, cultures were rinsed with $1 \times$ DPBS without Mg^{2+} and Ca^{2+} (Invitrogen) followed by incubation with $1 \times$ TrypLE Select Enzyme (Life Technology) for 3—5 min at 37 °C. Single cells were rinsed with mTeSR™1 medium, and spun at 1,000 rpm for 3 min. The resulting cell pellet was resuspended in mTeSR™1 medium supplemented with 10 μM Y-27632 (Sigma—Aldrich) and the single cell suspension was seeded at ~1.5×10^5 cells/cm^2 on Matrigel-coated surfaces. Cultures were fed every day with mTeSR™1 medium and differentiation was initiated 48 h following seeding, resulting in ~90% starting confluency.

2.6. Pancreatic progenitor differentiation

2.6.1. S1: Definitive endoderm (3 d)
XM001 iPSCs plated on 1:100 diluted Matrigel were first rinsed with $1 \times$ DPBS without Mg^{2+} and Ca^{2+} (Invitrogen) and then cultured in

RPMI 1640 medium (Invitrogen) further supplemented with 1.2 g/L sodium bicarbonate (Sigma), 0.2% ESC-qualified FBS (Life Technologies), 100 ng/mL Activin-A (R&D Systems), and 20 ng/mL of Wnt3A (R&D Systems) for day 1 only. For the next 2 days, cells were cultured in RPMI with 0.5% FBS, 1.2 g/L sodium bicarbonate and 100 ng/mL Activin-A.

2.6.2. S2: Primitive gut tube (3 d)
Cells were rinsed with $1 \times$ DPBS (without Mg^{2+} and Ca^{2+}) once and then exposed to DMEM-F12 (Life Technologies) medium further supplemented with 2 g/L sodium bicarbonate, 2% FBS and 50 ng/mL of FGF7 (R&D Systems) for 3 days.

2.6.3. S3: Pancreatic progenitors (PP)/posterior foregut (4 d)
Cultures were maintained for 4 days in DMEM-HG medium (Life Technologies) supplemented with 0.25 μM SANT-1 (Sigma—Aldrich), 2 μM retinoic acid (RA; Sigma—Aldrich), 100 ng/mL of Noggin (R&D Systems), and 1% B27 (Invitrogen).

2.7. Immunofluorescence imaging
Cells were fixed with 4% paraformaldehyde for 30 min and then permeabilized in PBS containing 0.2% Triton X-100. Cells were blocked with PBS containing 3% BSA, and incubated with primary antibodies overnight at 4 °C. Then secondary antibodies were incubated for 1 h at room temperature after washing with PBS. Images were acquired on a TCS SP5 laser-scanning microscope (Leica). The following antibodies and dilutions were used: goat anti-OCT-3/4 (1:500, #sc-8628, Santa Cruz), goat anti-SOX2 (1:500, #sc-17320, Santa Cruz), mouse anti-TRA-1-60 (1:1000, #4746, Cell Signaling), mouse anti-TRA-1-81 (1:50, MAB4381, Millipore), mouse anti-SSEA4 (1:500, #4755, Cell Signaling), rabbit anti-FOXA2 (1:250, #8186, Cell Signaling), goat anti-SOX17 antibody (1:500, #GT15094, Acris/Novus), goat anti-PDX1 antibody (1:500, #AF2419, R&D Systems).

2.8. Flow cytometry
Cells were dissociated using $1 \times$ TrypLE Select Enzyme and washed with cold FACS buffer (5% FBS in $1 \times$ DPBS). Cells were fixed with 4% paraformaldehyde and permeabilized with donkey block solution (0.1% tween-20, 10% FBS, 0.1% BSA and 3% donkey serum) containing 0.5% saponin. The cells were incubated with goat anti-SOX17 antibody (1:500, #GT15094, Acris/Novus) and goat anti-PDX1 antibody (1:500, #AF2419, R&D Systems) for 30 min at room temperature and then stained with Alexa Fluor 555-conjugated donkey antibody directed against goat (1:500, #A21432, Invitrogen) for 30 min at room temperature. Flow cytometry was performed using FACS-Aria III (BD Bioscience). FACS data were analyzed using FlowJo.

2.9. RNA isolation and qPCR
Total RNA was extracted from cells with the miRNeasy mini kit (Qiagen). cDNA synthesis was performed with a high-capacity RNA-to-cDNA kit (Applied Biosystems). TaqMan qPCR was performed under standard conditions using ViiA7 (Applied Biosystems) and TaqMan Fast Advanced Master Mix (Applied Biosystems). Samples were normalized to housekeeping genes 18S ribosomal RNA (RN18S) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Taqman probes (Applied Biosystems): PDX1, Hs00236830_m1; SOX9, Hs01001343_g1; PTF1A, Hs00603586_g1; HNF1B, Hs01001602_m1; OCT4, Hs00999632_g1; NANOG, Hs04260366_g1; GAPDH, Hs02758991_g1; 18S, Hs99999901_s1; NKX6.1, Hs01055914_m1; NKX2.2, Hs00159-616_m1.

SYBR qPCR was performed under standard conditions using ViiA7 (Applied Biosystems) and SYBR™ Green PCR Master Mix (Applied Biosystems). Samples were normalized to the housekeeping gene *ACTB*. Primers were as follows:

HHEX forward: ACGGTGAACGACTACACGC.
HHEX reverse: CTTCTCCAGCTCGATGGTCT;
MEIS1 forward: GGGCATGGATGGAGTAGGC.
MEIS1 reverse: GGGTACTGATGCGAGTGCAG;
ONECUT1 forward: GAACATGGGAAGGATAGAGGCA.
ONECUT1 reverse: GTAGAGTTCGACGCTGGACAT;
RFX6 forward: AAGCAGCGGATCAATACCTGT.
RFX6 reverse: ACCGTGGTAAGCAAACTCCTT;
ACTB forward: CCCAGAGCAAGAGAGG.
ACTB reverse: GTCCAGACGCAGGATG.

2.10. ChIP-seq

XM001 PP cells (2×10^6 cells) were cross-linked in 1% formaldehyde in culture medium for 10 min at room temperature. The cross-linking reaction was stopped by the addition of glycine to a final concentration of 125 mM. For chromatin fragmentation, cells were resuspended in lysis buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 0.5% SDS) and sonicated in a Covaris S220 sonicator with a duty cycle of 2%, a peak incident power of 105 W and 200 cycles per burst for 20 min. The fragmented chromatin was diluted 1:5 in IP-Buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-Desoxycholate, 140 mM NaCl, H_2O, Protease Inhibitors) and directly used for immunoprecipitation. For PDX1 ChIP, 60% of the chromatin (equivalent of 1.2×10^6 cells) was processed in two parallel ChIPs using magnetic beads, preloaded with 3 µl goat anti PDX1 antibody (kindly provided by C. Wright), for each ChIP. For H3K27ac ChIP, 40% of the chromatin (equivalent 0.4×10^6 cells) was processed in three parallel IPs using magnetic beads, preloaded with 3 µg anti H3K27ac antibody (Diagenode, C15410174). For all ChIPs, antibody incubation was performed at 4 °C for 5 h. Beads were then washed 5 times using (1.) IP-Buffer, (2.) Washing-Buffer 1 (500 mM NaCl, 50 mM Tris-HCl (pH 8.0), 0.1% SDS, 1% NP-40), (3.) Washing-Buffer 2 (250 mM LiCl, 50 mM Tris-HCl (pH 8.0), 0.5% Na-Deoxycholate, 1% NP-40) and (4. & 5.) Washing-Buffer 3 (10 mM Tris-HCl (pH 8.0), 10 mM EDTA). Subsequently, protein-DNA complexes were eluted from the beads in Elution-Buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% SDS) at 65 °C for 20 min. Cross-links were reversed at 65 °C overnight, and DNA was purified for library construction with the MicroPlex kit (Diagenode, C05010010). Libraries were sequenced on an Illumina HiSeq 1500 (50 bp, single end).

2.11. ChIP-seq data analysis

Raw reads from PDX1 and H3K27ac ChIP-seq were processed with Trimmomatic (0.35) to remove low quality bases and potential adapter contamination. Next, reads were aligned to hg19 genome using bowtie2 (2.2.6) with very-sensitive option and duplicate reads were removed using samtools (1.3). Binding sites for PDX1 were then called using GEM [52] and filtered, after visual inspection, using a Q-value cut-off of 10^{-4} and overlapping regions were merged. Regions of H3K27ac enrichment were called using BCP-HM [53] with default parameters. PDX1 binding sites and H3K27ac enriched regions were further filtered by removing blacklist regions [54]. Regions and binding sites within 20 kb of a TSS or within a gene body were annotated with the respective gene using bedtools (2.26.0). Motif analysis, annotations with genomic features, and pathway enrichment analysis were

performed using HOMER [55]. For publicly available data sets, raw reads were downloaded from online repositories and processed as described above (GSE54471 [22], GSE58686 [24], E-MTAB-1143 [28], E-MTAB-1919 [18]). For clustering of H3K27ac profiles, first the reads that fall in the intersection of enriched regions from all samples to be compared were counted and normalization factors were calculated using DESeq2. Next, the union of all identified regions from all samples to be compared was merged and reads falling into these regions were counted. Reads were then processed with DESeq2 using normalization factors calculated before and normalized read counts were used to calculate Pearson correlation coefficients that were subsequently subjected to hierarchical clustering. To calculate enrichment of H3K27ac at PDX1-bound sites, three sets of random sites were generated by shuffling all 8088 PDX1 binding sites using bedtools (2.26.0) shuffle. Then, H3K27ac reads were counted and the ratio of the mean read count at PDX1 sites over the mean read count at shuffled sites was calculated. For visualization, BIGWIG files, normalized to 1 × sequencing depth (Reads Per Genomic Content) using deepTools bamCoverage, were used. H3K27ac and PDX1 tracks for adult islets were generated from merged BAM files containing reads from all available experiments. To generate high confidence PDX1 binding sites, the overlap of PDX1 sites from progenitor profiles (XM001, [22,24]) or islet profiles [18,28] was calculated, and sites found in two out of three progenitor data sets or three out of four islet data sets were considered as high confidence binding sites.

2.12. Enrichment of T2DM SNPs

To obtain positions and disease association P-values of T2DM-associated SNPs, DIAGRAM 1000G GWAS meta-analysis Stage 1 Summary statistics [17] was downloaded and filtered to remove all SNPs with an association P-value below 5×10^8. For islet H3K27ac, the union of enriched regions from all 3 available experiments [18,22] was generated and subsequently, regions closer than 1000 bp were merged. Enrichment P-values were calculated using bedtools (2.26.0) fisher and the SNPs overlapping regions of interest were identified using bedtools (2.26.0) intersect. SNPs within 20 kb of a TSS or within a gene body were annotated with the respective gene using bedtools (2.26.0).

2.13. Affymetrix microarray

For gene profiling, total RNA was extracted using miRNeasy Mini kit (Qiagen, #217004), RNA integrity was checked using Agilent 2100 Bioanalyzer (Agilent RNA 6000 Pico Kit), and cDNA was amplified with the Ovation PicoSL WTA System V2 (Nugen, 3312) in combination with the Encore Biotin Module (Nugen, USA). Amplified cDNA was hybridized on GeneChip™ Human Gene 2.0 ST arrays (Affymetrix, 902113). Expression console (v.1.3.0.187, Affymetrix) was used for quality control. All subsequent computational analysis was performed in R using Bioconductor packages. Expression data were RMA normalized using the oligo package (version 1.38.0) and probe sets were annotated using the package hugene20sttranscriptcluster.db (version 8.5.0). Differential expression analyses were performed using the limma package (version 3.30.7) and P-values were adjusted for multiple testing by Benjamini-Hochberg correction. A gene was considered as differentially expressed if the adjusted p-value (FDR) was below a threshold of 0.05 and the fold-change was greater than or equal to 2. Functional enrichments were conducted using HOMER [55]. Gene set enrichment analysis was performed using GSEA 3.0 [56,57] with genes ranked by log2 ratio between XM001 PPs and XM001 iPSCs.
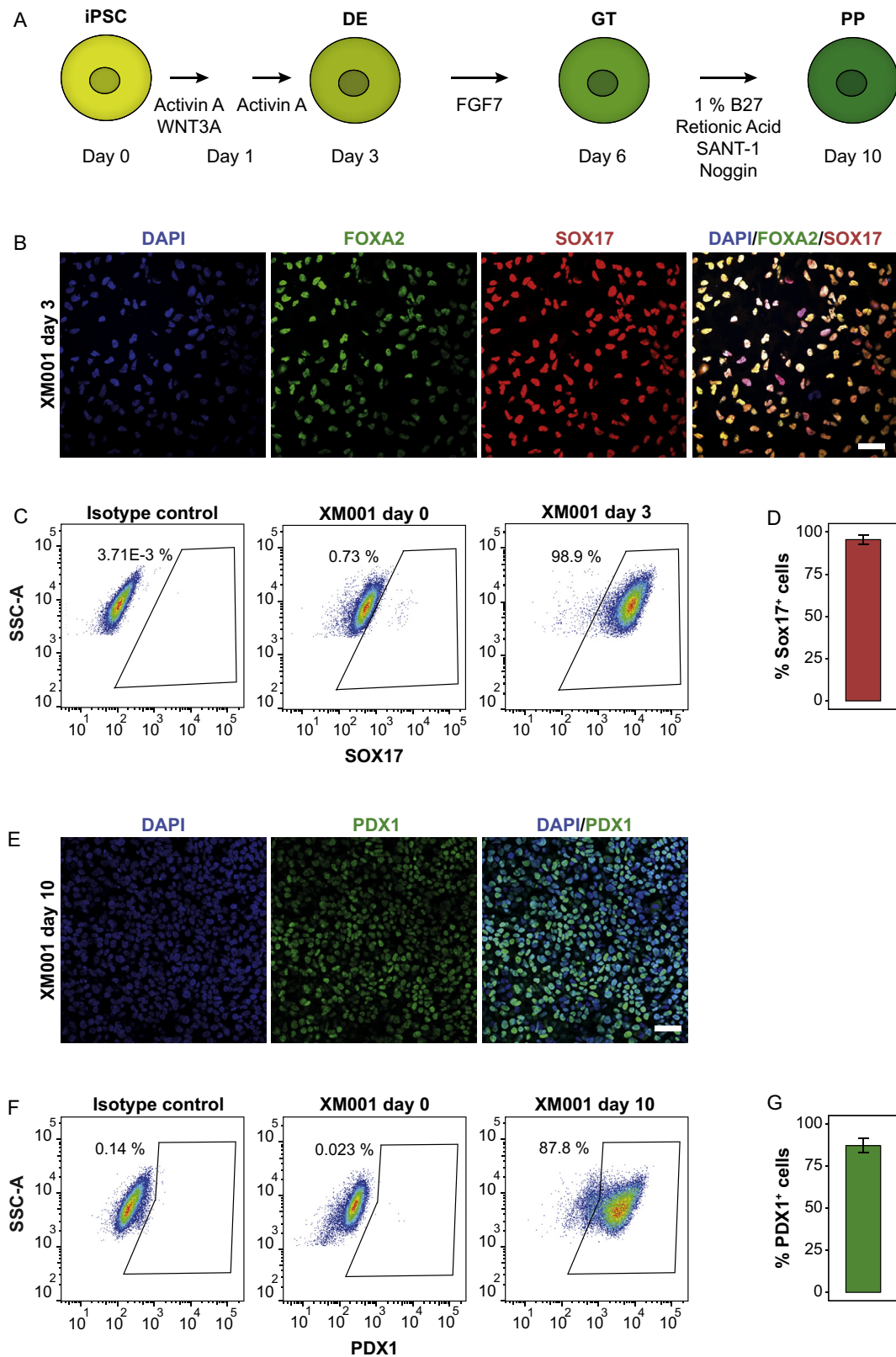
Figure 1: Efficient differentiation of XM001 iPSCs into pancreatic progenitors. (A) Schematic of iPSC-derived pancreatic progenitor differentiation protocol. (B) Immuno-staining for FOXA2 and SOX17 on day 3. Scale bar indicates 50 μm. (C) Representative FACS plot of SOX17$^+$ cells at DE stage. A differentiated sample stained with only the secondary antibody and XM001 iPSCs were used as negative controls. (D) FACS quantification of the percentage of SOX17$^+$ cells at DE stage (n = 3). (E) Immunostaining for PDX1 on day 10. Scale bar indicates 50 μm. (F) Representative FACS plot of PDX1$^+$ cells at the PP stage. A differentiated sample stained with only the secondary antibody and XM001 iPSCs were used as negative controls. (G) FACS quantification of the percentage of PDX1$^+$ cells at PP stage (n = 3).

2.14. Accession number

Microarray and ChIP-seq data have been submitted to the GEO database at NCBI under the accession number GSE106950.

3. RESULTS

3.1. Efficient differentiation of XM001 iPSCs into pancreatic progenitors

To generate the XM001 induced pluripotent stem cell (iPSC) line, primary adult skin fibroblasts from a healthy female donor were reprogrammed into iPSCs by transfection with three episomal plasmids encoding human *OCT4*, *SOX2*, *NANOG*, *LIN28*, *KLF4*, and *L-MYC*. (Suppl. Figure 1A) [19,20]. We established the reprogramming protocol including three stages: transfection, reprogramming and expansion (Suppl. Figure 1B). Colonies with stem cell-like characteristics appeared around day 15 post-infection and were handpicked by day 21 and subjected to quality controls. XM001 iPSCs were episomal vector integration-free and of normal karyotype (Suppl. Figure 1C—D). Next, we confirmed the expression of several pluripotency markers (TRA-1-60 and TRA-1-81, SSEA-3, and SSEA-4) and pluripotency TFs (OCT4 and SOX2) by immunostainings in XM001 iPSCs (Suppl. Figure 2). To rigorously demonstrate multi-lineage potency of the generated iPSC line, we injected 2×10^6 iPSCs subcutaneously into None-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice to generate teratomas. The histological analysis of the induced teratomas revealed that tissues from all three germ layers were differentiated, confirming pluripotency of our XM001 iPSCs *in vivo* (Suppl. Figure 3). These data show successful reprogramming of fibroblasts into pluripotent iPSCs that stably display human embryonic stem cell (hESC)-like characteristics and multi-lineage potency.

Next, we differentiated XM001 iPSCs into pancreatic progenitors (PP) using a previously established protocol recapitulating pancreas development through three main stages: definitive endoderm (DE), primitive gut tube (GT), and PPs of the posterior foregut (Figure 1A)

[21]. Specific markers, corresponding to the DE (FOXA2 and SOX17) and PP (PDX1) stages were verified by immunostaining and flow cytometry at day 3 and 10, respectively. At the DE stage, cells co-expressed FOXA2 and SOX17 (Figure 1B) and flow-cytometry analysis showed that approximately 95% of the cells were positive for the DE marker SOX17 (Figure 1C—D). At the PP stage, the number of PDX1-positive cells was around 85% as measured by flow cytometry and shown by immunostaining (Figure 1E—G).

3.2. mRNA profiles of XM001 PPs show expression of pancreas specific genes

To characterize global gene expression changes during PP differentiation, we performed Affymetrix microarray analysis on XM001 cells at the pluripotency and PP stages. In total, we identified 2658 differentially expressed genes between iPSCs and PPs (FC ≥ 2, FDR ≤ 5%), including important pancreatic genes such as *HNF1B*, *SOX9*, *ONECUT1*, *HHEX*, *FOXA2* and *RFX6* that are upregulated in PPs, as well as pluripotency-associated genes including *NANOG* and *SOX2* that are specifically expressed in iPSCs and downregulated in PPs (Figure 2A). Moreover, gene set enrichment analysis as well as unbiased gene ontology (GO) term and pathway analysis showed significant enrichment of terms associated with pancreas development in PP cells, while terms associated with embryonic stem cells, cell cycle and proliferation were downregulated during the differentiation process (Figure 2B—C). To verify the microarray results, we validated expression changes of important genes involved in pancreas development (*HHEX*, *HNF1B*, *MEIS1*, *ONECUT1*, *PDX1*, *PTF1A*, *RFX6*, and *SOX9*), endocrine formation (*NKX2-2*, *NKX6-1*) and stem cell maintenance (*NANOG*, *OCT4*) by quantitative PCR (qPCR) (Figure 2D—F). Taken together, the XM001 iPSCs were efficiently differentiated into PPs after 10 days of *in vitro* differentiation characterized by the expression of lineage-specific genes and the downregulation of pluripotency genes.
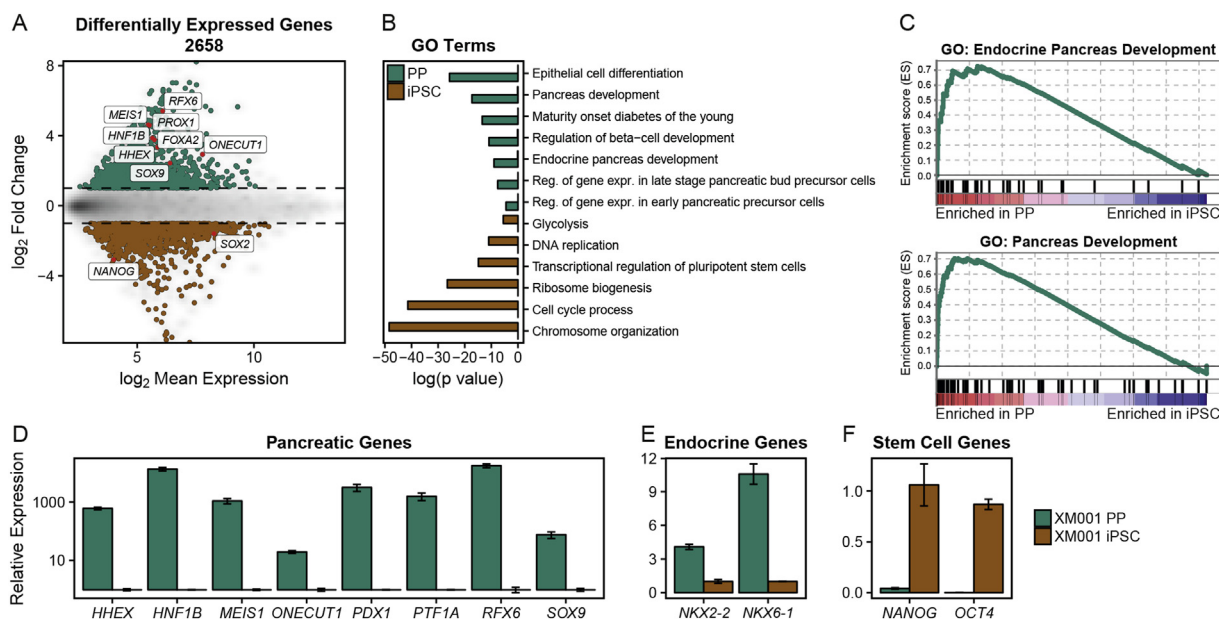

Figure 2: mRNA profiles of XM001 PPs show expression of pancreas specific genes. (A) MA plot showing the mean log2 expression against the log2-fold change of the microarray data obtained from XM001 iPSCs and PPs. Genes with significantly different expression (log2 fold change ≥ 1 and adjusted p-value ≤ 0.05) are drawn in color. Green depicts increased expression in PPs, whereas brown depicts increased expression in iPSCs. (B) Bar graph of p-values from selected GO terms and KEGG and Reactome pathways from differentially expressed genes, shown in (A). (C) Gene set enrichment analysis of pancreas related GO terms. (D—F) qPCR validation of the expression of pancreas (D), endocrine (E) and stem cell (F) related genes in XM001 iPSCs and PPs (n = 3).

Correct specification of tissue-specific enhancers is crucial for the differentiation of individual cell types [22]. In this context, the histone modification H3K27ac is strongly associated with active enhancers and promoters [23]. Therefore, we analyzed this histone modification in XM001 cells after 10 days of differentiation using chromatin immunoprecipitation followed by next generation sequencing (ChIP-seq) and found that key pancreatic genes are marked by an active chromatin configuration. Wang et al. [22] thoroughly characterized the active H3K27ac chromatin mark at several stages of pancreatic endoderm differentiation (embryonic stem cells (ESCs), definitive endoderm (DE), gut tube (GT), posterior foregut (FG; corresponds to PP stage of this study) and late pancreatic progenitors/pancreatic endoderm (PE)) and showed that the chromatin state at lineage-specific enhancers is vital for developmental competence [22]. We used H3K27ac profiles from hESC *in vitro* differentiations into pancreatic progenitors [22] to compare with our H3K27ac profile from differentiated XM001 iPSCs. Unsupervised hierarchical clustering showed that after 10 days of differentiation, the active chromatin landscape in XM001 iPSCs was reminiscent of that in FG and PE stage derived from hESCs, indicating that our iPSC-derived PPs represent an early stage of pancreas development, shortly after PDX1 induction (Suppl. Figure 4).

3.3. Characterization of PDX1 binding in XM001 PP cells

Due to the crucial role of PDX1 in the specification and differentiation of the pancreatic lineages, characterization of its transcriptional targets is important to understand human pancreas development. To do so, we employed ChIP-seq analysis of PDX1 to profile genome-wide PDX1 binding sites in XM001 iPSCs after 10 days of differentiation, a time point at which PDX1 is robustly expressed in approximately 85% of differentiated cells (Figure 1G). We identified a total of 8088 PDX1-bound regions that are associated with 5664 genes which have PDX1 binding sites within 20 kb of their TSS or within their gene body. These include *PDX1* itself and other pancreatic genes, such as *RFX6*,

MEIS1, and the MODY gene *HNF1B* (Figure 3A). Moreover, the active histone modification H3K27ac was enriched at the PDX1-bound sites (Figure 3B and Suppl. Figure 5), which were predominantly found in intergenic (44.2%) and intronic (40.4%) regions. However, with 11.3% of the binding sites occurring in promoter regions, PDX1 was enriched at the transcription start sites (TSS) of its target genes (Figure 3C—D). *De novo* motif analysis was used to discover enriched motifs within peak regions. The most significantly enriched motif was found in around 62% of all peak sequences and matched the PDX1 consensus motif, confirming the efficiency and specificity of the PDX1 ChIP-seq approach (Figure 3E). We also compared our data to previous published datasets of PDX1 ChIP-seq from hESC *in vitro*-derived PPs [22,24] and found that the profiles of PDX1 binding show obvious differences that might reflect hESC and iPSC differences, genetic background variations and differences in the *in vitro* differentiation stage or protocol (Suppl. Figure 6).

3.4. Functional characterization of PDX1-bound genes

Next, we examined the expression levels and function of PDX1-bound genes in XM001 PP cells. From the 5664 PDX1-bound genes, 4523 were covered by the microarray. In XM001 PP cells, 755 PDX1-bound genes were differentially expressed between the iPSC and PP stage, which accounts for 28.4% of the total differentially expressed genes (Figure 4A—B). Notably, we found upregulation of *RFX3* and *DLL1* among PDX1-bound genes during the differentiation process. The mouse orthologues of these genes are known to be involved in endocrine differentiation [25—27], suggesting a potential role of these genes also during human pancreas development. GO term analysis revealed an association of PDX1-bound genes enriched in XM001 PP cells with pancreas, endocrine and β-cell development as well as Notch and Hippo signaling, suggesting that PDX1 induces a pancreatic program and primes for endocrine differentiation (Figure 4C).
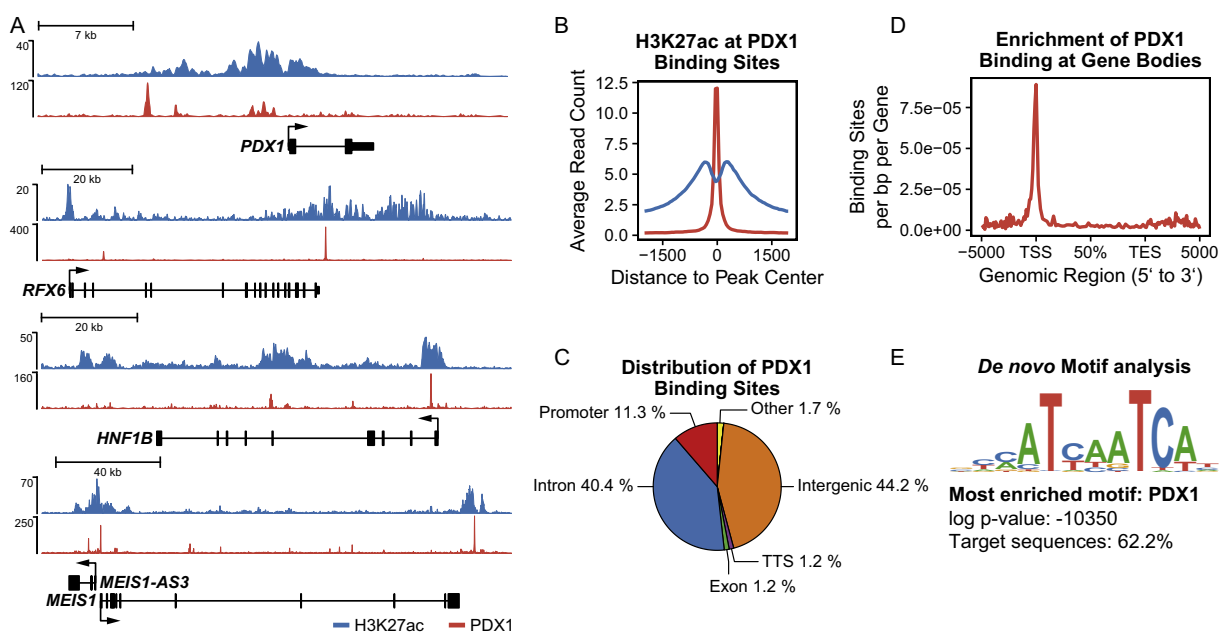

Figure 3: Characterization of PDX1 binding in XM001 PP cells. (A) ChIP-seq data tracks showing the enrichment of H3K27ac (blue) and PDX1 (red) at the loci of important pancreatic genes. (B) Average ChIP-seq Signal of H3K27ac (blue) and PDX1 (red) at PDX1 binding sites shows enrichment of H3K27ac at PDX1-bound sites. (C) Distribution of PDX1 binding sites among genomic features. PDX1 binds predominantly to intergenic, intronic, and promoter regions. (D) Meta-genomic plot of the enrichment of PDX1 at the transcriptional start sites (TSS) of its target genes displayed as binding sites per base pair (bp) per gene over the genomic regions of all RefSeq genes. (E) Most enriched motif discovered by *de novo* motif analysis resembles the known PDX1 consensus sequence and is identified in 62.2% of all PDX1-bound sequences.
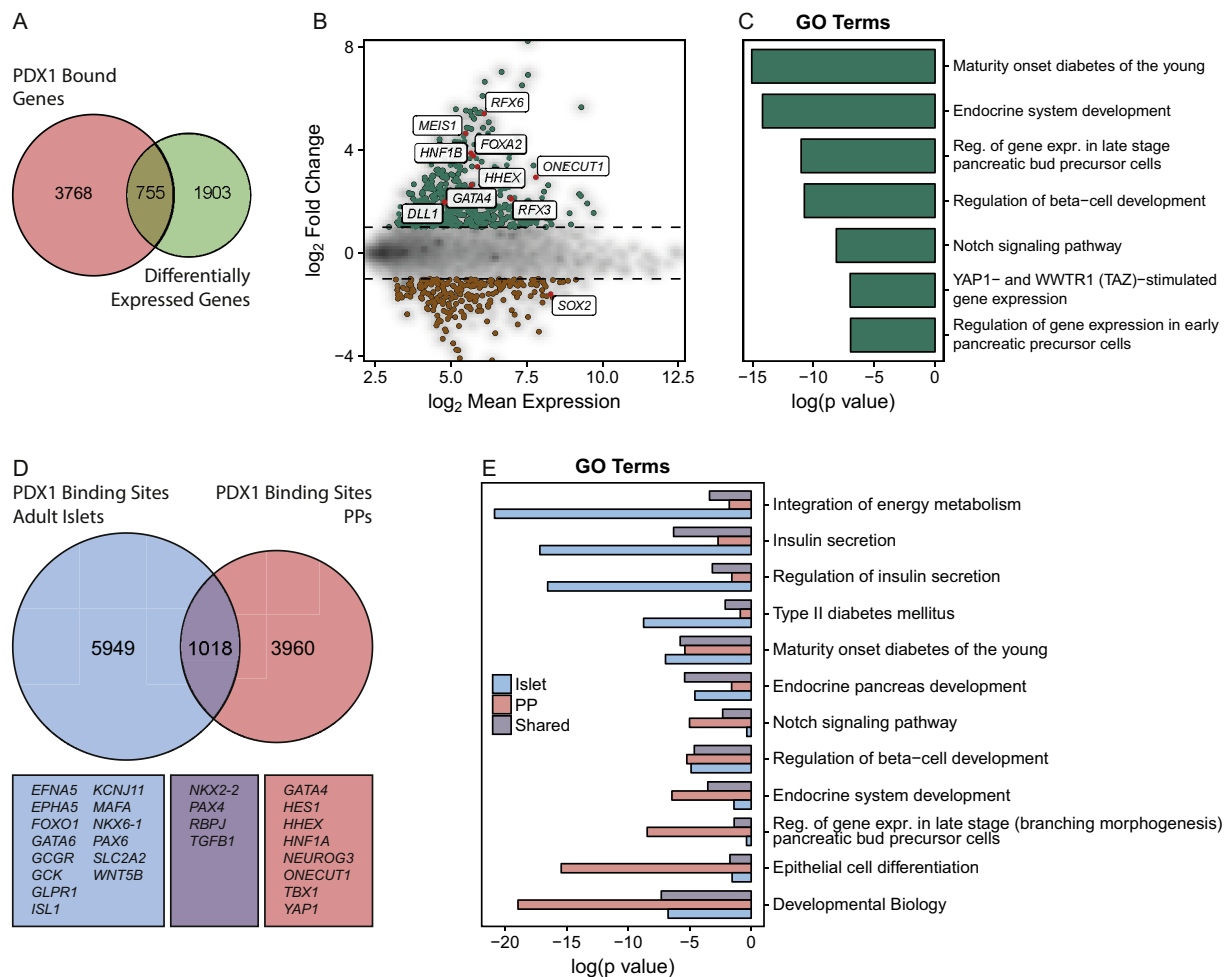
Figure 4: Functional characterization of PDX1-bound genes. (A) Venn diagram depicting the overlap of PDX1-bound genes and the differentially expressed genes identified by microarray analysis. (B) MA plot showing the log2 fold change over the mean log2 expression of PDX1-bound genes. Differentially expressed genes are displayed in color. Green indicates enrichment in PPs, whereas brown indicates enrichment in iPSCs. (C) Bar chart of log10 p-values from enriched GO terms and KEGG and Reactome pathways of PDX1-bound genes upregulated in PPs. (D) Venn diagram showing the overlap of high confidence PDX1 binding sites from adult human islets and PPs and some annotated genes. (E) Bar chart of log10 p-values from enriched gene ontology and KEGG/Reactome pathways from genes bound in adult islets and/or PPs.

Based on the tempo-spatial pattern of Pdx1 expression in mouse that shows ubiquitous expression in all pancreatic progenitors and restricted expression in mature β- and δ-cells, we hypothesized that the molecular programs controlled by Pdx1 might be stage specific. To understand if PDX1 carries out different functions during development and in adult β-cells, we compared PDX1 binding sites from PPs and adult human islets. To that end, we derived a set of high confidence PDX1 binding sites for PPs and adult islets, comprised of PDX1 sites, consistently bound in multiple datasets from this study and previous studies from hESC-derived PPs and islets cells [18,22,24,28]. Of the 4978 high confidence binding sites in PP cells, 1018 (20.4%) were identified in the islet binding data (Figure 4D). This indicates that in PP cells and adult islets only a minority of PDX1 binding sites are conserved, while the majority of sites change during the differentiation and maturation processes. Correspondingly, among the genes near PP-specific high confidence binding sites, we found important genes for pancreas development, such as *GATA4*, *HES1*, *HHEX*, *HNF1A*, *NEUROG3*, *ONEUCUT1*, *TBX1*, and *YAP1*, while genes near islet-specific binding sites, including *GLPR1*, *KCNJ11*, *SLC2A2*, *NKX6-1*, and *MAFA*, are associated with mature β-cell functions, e.g. insulin secretion, glucose sensing and β-cell maturation and identity

(Figure 4D—E). Genes in the vicinity of shared binding sites are implicated in endocrine development and include *NKX2-2*, *PAX4*, *RBPJ*, and *TGFB1*. Of note, we also observed genes that show stage-specific binding of PDX1 that can occur in combination with common binding sites. These include *HNF1B*, *RFX3*, and *RFX6*, *MNX1*, *FOXA2*, the *MEG3* locus, and other important pancreatic genes (Figure 5C and Suppl. Figure 7). Together, these results suggest that PDX1 binding sites change during the process of pancreas and endocrine cell development to meet the functional requirements of the different stages during development and homeostasis.

3.5. Analysis of T2DM SNPs in XM001 PPs and adult islets
In a worldwide effort, genome-wide association studies (GWAS) identified more than 300 genes that are linked to T2DM [29]. Of these genes, 104 (32%) were bound by PDX1 in XM001 PPs. While harsh mutations in MODY genes lead to monogenic forms of diabetes, an increased susceptibility to T2DM is linked to SNPs occurring mostly in non-coding regions. Moreover, it has been shown that islet enhancers are enriched in T2DM-associated SNPs [18,30]. Therefore, we analyzed whether XM001 PP active chromatin regions, as determined by enrichment of H3K27ac, are also enriched in T2DM-associated
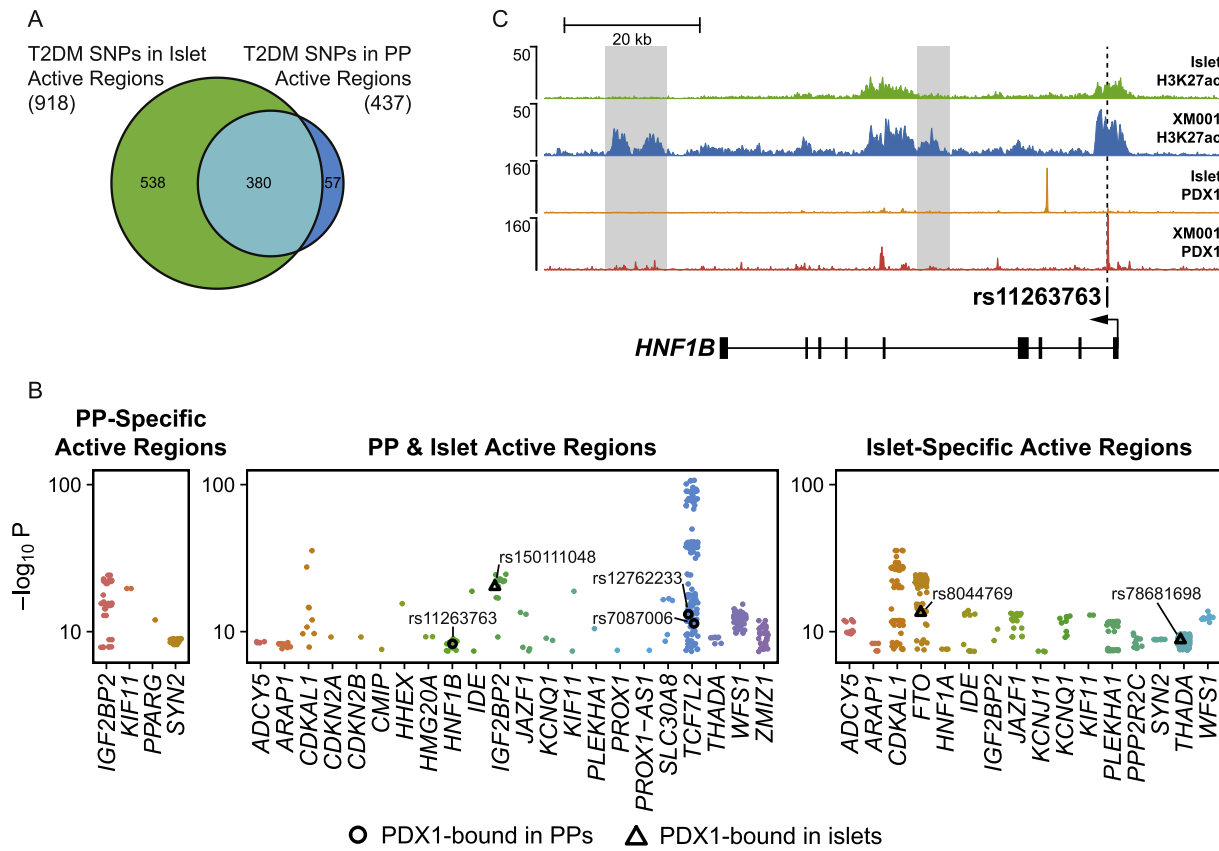
Figure 5: Analysis of T2DM SNPs in XM001 PPs and adult islets. (A) Venn diagram depicting the overlap of T2DM SNPs found in active regions of adult islets and XM001 PPs. (B) P-values of T2DM SNPs near known T2DM-associated genes found in active regions of adult islets [18,22] and XM001 PPs. P-values of SNPs bound by PDX1 in PPs and islets are shown as black circles and triangles, respectively. (C) ChIP-seq data tracks showing H3K27ac and PDX1 from PP and islet [18,22,28] cells at the *HNF1B* locus. The SNP rs11263763 is located in a PP specific PDX1 binding site in the first intron of *HNF1B*. The PP specific enhancer regions are shaded in gray.

SNPs and if there are different variants in XM001 PPs and adult islets. To this end, we used the DIAGRAM 1000G GWAS meta-analysis data [17] and identified SNPs with a high association significance (1788 SNPs, $P \leq 5 \times 10^{8}$) that overlap with active chromatin marks in XM001 PPs and adult islets [18,22]. We found a significant enrichment of T2DM-associated SNPs in active chromatin of both PPs (437 SNPs, $P = 1.0156 \times 10^{-106}$) and islets (918 SNPs, $P = 8.2266 \times 10^{-263}$). Closer examination of the enriched SNPs showed that of the 918 SNPs found in islet active chromatin, 380 fall in regions that are also enriched in H3K27ac in the XM001 PP cells. However, we could also identify 57 SNPs in regions, active specifically in PPs (Figure 5A). When we mapped the SNPs to genes, we identified 857 SNPs in the vicinity of known genes (20 kb up-/downstream of or within a gene body), and these included mostly known T2DM-associated genes, e.g. *TCF7L2*, *IGF2BP2*, *HNF1B*, and *FTO* (Figure 5B). While most of these genes were associated with SNPs in regulatory regions active in PP and islets, some genes were associated with SNPs in cell type-specific active regions. SNPs in islet-specific regions were found near the genes *FTO*, *HNF1A*, *KCNJ11*, *PPP2R2C*, and *SYN2* and SNPs in PP-specific regions near *SYN2*, *IGF2BP2*, *KIF11*, and *PPARG*. Of note, the *PPARG* locus is the only T2DM locus were a T2DM-associated SNP falls exclusively into a PP-specific active region and that does not have any T2DM-associated SNP in a commonly active or islet-specific active site. The enrichment of T2DM SNPs in active regulatory regions of PP cells raises the possibility that the susceptibility to T2DM is, at least in

parts, already established during pancreatic development, possibly through miss-regulation of key genes and developmental programs. Interestingly, when overlapping the disease-associated SNPs with PDX1 binding sites in PP and islet cells, we found only six SNPs that lie in the PDX1-bound sites (3 SNPs in PP PDX1 sites and 3 SNPs in islet PDX1 sites) (Figure 5B). The SNP rs8044769, SNP rs78681698 and SNP rs150111048 are in PDX1-bound regions in islets and are near *FTO*, *THADA*, and *IGF2BP2*. Another 3 SNPs near *HNF1B* and *TCF7L2* are found in PDX1-bound regions of PP cells. At the *HNF1B* locus, the SNP rs11263763 is located in a PP-specific PDX1 binding site in the first intron of *HNF1B* (Figure 5C). In addition, there are PP-specific active enhancer regions, in the 4th intron and downstream of *HNF1B*, suggesting that *HNF1B* might be regulated in different ways during development and adulthood. At the *TCF7L2* locus, the SNPs rs12762233 and rs7087006 are located in PP-specific PDX1 binding sites that may regulate *TCF7L2* expression during development. SNPs near the *TCF7L2* locus have the highest association to T2DM and mutations of *HNF1B* can cause MODY5. Thus, mutations in cis-regulatory regions of these genes may predispose to diabetes. This further supports the notion that the risk for T2DM might be linked to early pancreas development.

4. DISCUSSION

Even though, a recent study confirmed that the key events of pancreatic endocrine cell formation *in vivo* are mimicked during *in vitro*

differentiation [31], the efficient generation of functional and mature β-cells for cell-replacement therapy of diabetes still remains a challenge [32—34]. To produce such functional cells fulfilling the physiological properties of mature islets, it is important to understand the molecular mechanisms and gene regulatory networks governing the specification of early human pancreatic lineages that later differentiate into functional β-cells. In this context, understanding the temporal progression of key pancreatic TF target genes during development is valuable information that helps to better recapitulate the developmental steps required for the differentiation of functional β-cells *in vitro*. However, the access to primary human material of early pancreas development is very limited and thus, *in vitro* approaches have to be employed to study these early steps in a human system.

Although it is well known that Pdx1 plays a crucial role during mouse pancreas development, its stage-specific role in human pancreatic progenitors and adult β-cells has not been thoroughly studied. To address this question, we combined integrated analysis of genome-wide PDX1 binding sites and H3K27ac together with mRNA profiling in PPs derived from a novel human iPSC line.

Global characterization of PDX1 binding sites in iPSC-derived PPs identified numerous target sites and associated genes, including developmental regulatory factors, signaling molecules and factors important for mature β-cell function as well as many of the genes, differentially expressed between iPSCs and PPs. To our knowledge, this is the first genome-wide characterization of PDX1 binding in iPSC-derived PPs and comparison with previous PDX1 datasets from hESC-derived PPs of similar stages shows significant differences in PDX1 binding [22,24]. There are many possible explanations for this discrepancy in PDX1 binding, including differences in the efficiency of differentiation, factors used in the *in vitro* differentiation protocol, and the exact developmental stage at the time of the experiments. This could explain many of the differences between the XM001 data from this study and the data from Wang et al. [22], as these cells are slightly further in development and already express NKX6.1. Moreover, the developmental competence of a given cell line has implications on the differentiation of the cells and might favor a set of target sites over another. Despite the differences, however, sites bound by PDX1 in multiple datasets are useful to identify high confidence stage-specific target genes.

Among the PDX1-bound genes were *RFX3* and *DLL1* that are important for mouse pancreas development [25—27]. *RFX3* is a member of the regulatory factor × family of transcription factors that contain a highly conserved winged helix DNA-binding domain [35]. It has been shown that *Rfx3* is required for the differentiation and function of mature β-cells and regulates the expression of the glucokinase gene in mice [25]. *DLL1* is a member of the delta/serrate/jagged family of Notch ligands. Mice deficient for the Delta-like 1 gene (*Dll1*) show increased expression of Ngn3, which, in turn, causes an accelerated and increased commitment to the endocrine lineage and depletes the progenitor pool, ultimately leading to pancreatic hypoplasia [26,27]. Whether these genes play important roles in human pancreas development has not been characterized. Therefore, our analysis suggests a possible evolutionary conserved function of *RFX3* and *DLL1* during human pancreas development. In contrast, other PDX1-bound genes such as *TIMP2*, *MAFB*, and *SIX2,* are known to be specifically express in human, but not in mouse β-cells [36,37]. To address the function of newly identified PDX1 target genes during development, the XM001 iPSCs has been modified to express CRISPR/Cas9 under the control of an inducible promoter, which will allow gene functional studies in the future [38].

PDX1 is a master regulator of early pancreatic development and is expressed in all PPs *in vivo* but becomes restricted to β-and δ-cells during later stages of endocrine differentiation. Therefore, it is likely, that PDX1 has different functions and targets at different stages of development. When we compared genome-wide PDX1 binding sites from iPSC-derived PPs and adult human islets, only 20.4% of the identified PDX1 targets were common in both cell populations. Interestingly, the function of genes near common and specific binding sites differed markedly and reflects cellular function. PDX1-bound genes in islets were related to β-cell-specific functions. For example, the K$^+$ channel *KCNJ11* as well as the glucose transporter *SLC2A2* are essential for adult β-cell function, whereas *NKX6-1* and *MAFA* are important for β-cell identity and maturation. On the other hand, in PP cells PDX1-bound genes, such as *GATA4* and *ONECUT1*, are associated with processes and pathways in early pancreas development. This shows that, depending on the developmental stage, PDX1 target sites change to activate molecular programs that are required for stage-specific functions.

The presence of H3K27ac in PP cells at regions carrying T2DM-associated SNPs indicates that these regions participate in the regulation of gene activity. Variation in these regions potentially interferes with the regulation of T2DM-relevant genes at the PP stage and might therefore impair developmental programs and lead to reduce β-cell mass or function at birth that contributes to an increased susceptibility to T2DM. It has been shown that β-cell mass in children is highly variable between individuals [39] and diabetic patients show decreased β-cell mass [40]. One of the SNPs that we detected in a PP-specific PDX1 binding site, rs11263763, lies in the first intron of *HNF1B* and has been shown to reduce *HNF1B* expression [41]. This gene has several SNPs in the first and second intron that are linked to T2DM. During mouse pancreatic development, Hnf1β is involved in controlling proliferation and survival of multipotent pancreatic progenitors and deletion of the gene causes pancreatic hypoplasia. Hnf1β is also important for the specification of endocrine progenitors through regulation of Ngn3 expression [42]. In humans, some mutations in HNF1B cause MODY5 [43]. We also identified the SNPs rs12762233 and rs7087006 within PP-specific PDX1 binding sites at the *TCF7L2* locus. This gene, also known as *Tcf4*, is a member of the HMG-box containing T-cell factor (Tiff)/Lymphoid enhancer factor (Lef) transcription factor family of DNA-binding proteins and acts downstream of the canonical Wnt pathway [44]. GWAS studies have identified *TCF7L2* as the locus conveying the highest risk for developing T2DM in European [45]. In mouse, Tcf7l2 is an important regulator of β-cell development [46], whereas in human it regulates β-cell survival and function [47]. Genetic and metabolic studies show that people carrying *TCF7L2* polymorphisms have insulin secretion defects in the presence of normal incretin plasma levels [48]. *HNF1B* and *TCF7L2* are therefore important developmental regulatory factors and SNPs in PDX1-bound regions in these loci might impair early developmental processes that result in a higher risk to develop T2DM later in adulthood. The three SNPs we found in islet-specific binding sites are near the *FTO*, *THADA*, and *IGF2BP2* genes, suggesting that these SNPs might modulate adult β-cell function.

In summary, our study provides a PDX1 regulated human pancreas developmental program that can be used to investigate the effects of *PDX1* mutations, e.g. in point mutations in the transactivation domain found in T2DM patients, on pancreas developmental and islet PDX1 program [49,50]. Also, we find more than 430 T2DM-associated SNPs in active regulatory regions and 32% of T2DM genes to be bound by PDX1 in XM001 PPs, demonstrating that PDX1 occupancy is an efficient way to identify important pancreatic developmental and disease genes. This provides the base for further disease modeling *in vitro* using stem cell and gene editing technologies to understand the causal pathomechanisms underlying diabetes.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data related to this article can be found at https://doi.org/10.1016/j.molmet.2018.01.011.

REFERENCES

[1] International Diabetes Federation, 2015. IDF diabetes atlas.

[2] Prasad, R.B., Groop, L., 2015. Genetics of type 2 diabetes—pitfalls and possibilities. Genes 6(1):87—123. https://doi.org/10.3390/genes6010087.

[3] Heuvel-Borsboom, H., de Valk, H.W., Losekoot, M., Westerink, J., 2016. Maturity onset diabetes of the young: seek and you will find. Netherlands Journal of Medicine, 193—200.

[4] Jennings, R.E., Berry, A.A., Strutt, J.P., Gerrard, D.T., Hanley, N.A., 2015. Human pancreas development. Development 142(18):3126—3137. https://doi.org/10.1242/dev.120063.

[5] Bastidas-Ponce, A., Scheibner, K., Lickert, H., Bakhti, M., 2017. Cellular and molecular mechanisms coordinating pancreas development. Development 144(16):2873—2888. https://doi.org/10.1242/dev.140756.

[6] 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 (Cambridge, England) 122(5):1409—1416. https://doi.org/10.1016/0092-8674(88)90391-1.

[7] Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C.V., et al., 1995. Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. Development (Cambridge, England) 121(1):11—18.

[8] Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., et al., 1996. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development (Cambridge, England) 122(3):983—995. https://doi.org/10.1016/0076-6879(93)25031-v.

[9] Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., Edlund, H., 1998. Beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. Genes & Development 12(12):1763—1768. https://doi.org/10.1101/gad.12.12.1763.

[10] Jonsson, J., Carlsson, L., Edlund, T., Edlund, H., 1994. Insulin-promoter-factor 1 is required for pancreas development in mice. Nature, 606—609. https://doi.org/10.1038/371606a0.

[11] Brissova, M., Shiota, M., Nicholson, W.E., Gannon, M., Knobel, S.M., Piston, D.W., et al., 2002. Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion. Journal of Biological Chemistry 277(13):11225—11232. https://doi.org/10.1074/jbc.M111272200.

[12] Johnson, J.D., Ahmed, N.T., Luciani, D.S., Han, Z., Tran, H., Fujita, J., et al., 2003. Increased islet apoptosis in Pdx1+/- mice. Journal of Clinical Investigation 111(8):1147—1160. https://doi.org/10.1172/JCl200316537.

[13] Holland, A.M., Góñez, L.J., Naselli, G., MacDonald, R.J., Harrison, L.C., 2005. Conditional expression demonstrates the role of the homeodomain transcription factor Pdx1 in maintenance and regeneration of β-cells in the adult pancreas. Diabetes 54(9):2586—2595. https://doi.org/10.2337/diabetes.54.9.2586.

[14] Staffers, D.A., Ferrer, J., Clarke, W.L., Habener, J.F., 1997. Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. Nature Genetics 17(2):138—139. https://doi.org/10.1038/ng1097-138.

[15] Stoffers, D.A., Stanojevic, V., Habener, J.F., 1998. Insulin promoter factor-1 gene mutation linked to early-onset type 2 diabetes mellitus directs expression of a dominant negative isoprotein. Journal of Clinical Investigation 102(1):232—241. https://doi.org/10.1172/JCl2242.

[16] Wu, Y., Li, H., Loos, R.J.F., Yu, Z., Ye, X., Chen, L., et al., 2008. Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. Diabetes 57(10):2834—2842. https://doi.org/10.2337/db08-0047.

[17] Scott, R.A., Scott, L.J., Mägi, R., Marullo, L., Gaulton, K.J., Kaakinen, M., et al., 2017. An expanded genome-wide association study of type 2 diabetes in Europeans. Diabetes, db161253. https://doi.org/10.2337/db16-1253.

[18] Pasquali, L., Gaulton, K.J., Rodríguez-Seguí, S.A., Mularoni, L., Miguel-Escalada, I., Akerman, İ., et al., 2014. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. Nature Genetics 46(2):136—143. https://doi.org/10.1038/ng.2870.

[19] Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 107(5):861—872. https://doi.org/10.1016/j.cell.2007.11.019.

[20] Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J., Jiang, Y., 2011. Efficient feeder-free episomal reprogramming with small molecules. PLoS One 6(3). https://doi.org/10.1371/journal.pone.0017557.

[21] Rezania, A., Bruin, J.E., Riedel, M.J., Mojibian, M., Asadi, A., Xu, J., et al., 2012. Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. Diabetes 61(8):2016—2029. https://doi.org/10.2337/db11-1711.

[22] Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N.A., et al., 2015. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. Cell Stem Cell 16(4):386—399. https://doi.org/10.1016/j.stem.2015.02.013.

[23] Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., et al., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences 107(50):21931—21936. https://doi.org/10.1073/pnas.1016071107.

[24] Teo, A.K.K., Tsuneyoshi, N., Hoon, S., Tan, E.K., Stanton, L.W., Wright, C.V.E., et al., 2015. PDX1 binds and represses hepatic genes to ensure robust pancreatic commitment in differentiating human embryonic stem cells. Stem Cell Reports 4(4):578—590. https://doi.org/10.1016/j.stemcr.2015.02.015.

[25] Ait-Lounis, A., Bonal, C., Seguin-Estevez, Q., Schmid, C.D., Bucher, P., Herrera, P.L., et al., 2010. The transcription factor Rfx3 regulates -cell differentiation, function, and glucokinase expression. Diabetes 59(7):1674–1685. https://doi.org/10.2337/db09-0986.

[26] Kim, W., Shin, Y.-K., Kim, B.-J., Egan, J.M., 2010. Notch signaling in pancreatic endocrine cell and diabetes. Biochemical and Biophysical Research Communications 392(3):247–251. https://doi.org/10.1016/j.bbrc.2009.12.115.

[27] Apelqvist, A., Li, H., Sommer, L., Beatus, P., Anderson, D.J., Honjo, T., et al., 1999. Notch signalling controls pancreatic cell differentiation. Nature 400(6747):877–881. https://doi.org/10.1038/23716.

[28] Khoo, C., Yang, J., Weinrott, S.A., Kaestner, K.H., Naji, A., Schug, J., et al., 2012. Research resource: the pdx1 cistrome of pancreatic islets. Molecular Endocrinology (Baltimore, Md.) 26(3):521–533. https://doi.org/10.1210/me.2011-1231.

[29] MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., et al., 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Research 45(D1):D896–D901. https://doi.org/10.1093/nar/gkw1133.

[30] Corradin, O., Scacheri, P.C., 2014. Enhancer variants: evaluating functions in common disease. Genome Medicine 6(10):85. https://doi.org/10.1186/s13073-014-0085-3.

[31] Ramond, C., Glaser, N., Berthault, C., Ameri, J., Kirkegaard, J.S., Hansson, M., et al., 2017. Reconstructing human pancreatic differentiation by mapping specific cell populations during development. eLife 6. https://doi.org/10.7554/eLife.27564.

[32] Pagliuca, F.W., Millman, J.R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J.H., et al., 2014. Generation of functional human pancreatic β cells in vitro. Cell 159(2):428–439. https://doi.org/10.1016/j.cell.2014.09.040.

[33] Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., et al., 2014. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nature Biotechnology 32(11):1121–1133. https://doi.org/10.1038/nbt.3033.

[34] Russ, H. a, Parent, A.V., Ringler, J.J., Hennings, T.G., Nair, G.G., Shveygert, M., et al., 2015. Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. The EMBO Journal 34(13): e201591058. https://doi.org/10.15252/embj.201591058.

[35] Reith, W., Ucla, C., Barras, E., Gaud, a., Durand, B., Herrero-Sanchez, C., et al., 1994. RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. Molecular and Cellular Biology 14(2):1230–1244. https://doi.org/10.1128/MCB.14.2.1230.Updated.

[36] Segerstolpe, Å., Palasantza, A., Eliasson, P., Andersson, E.-M., Andréasson, A.-C., Sun, X., et al., 2016. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. Cell Metabolism 24(4): 593–607. https://doi.org/10.1016/j.cmet.2016.08.020.

[37] Xin, Y., Kim, J., Okamoto, H., Yancopoulos, G.D., Lin, C., Correspondence, J.G., 2016. RNA sequencing of single human islet cells reveals type 2 diabetes genes. Cell Metabolism 24:608–615. https://doi.org/10.1016/j.cmet.2016.08.018.

[38] Yumlu, S., Stumm, J., Bashir, S., Dreyer, A.-K., Lisowski, P., Danner, E., et al., 2017. Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9. Methods 121–122:29–44. https://doi.org/10.1016/j.ymeth.2017.05.009.

[39] Meier, J.J., Butler, A.E., Saisho, Y., Monchamp, T., Galasso, R., Bhushan, A., et al., 2008. β-cell replication is the primary mechanism subserving the postnatal expansion of β-cell mass in humans. Diabetes 57(6):1584–1594. https://doi.org/10.2337/db07-1369.

[40] Butler, A.E., Janson, J., Bonner-Weir, S., Ritzel, R., Rizza, R.A., Butler, P.C., 2003. Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. Diabetes 52:102–110. https://doi.org/10.2337/diabetes.52.9.2304 (January).

[41] Painter, J.N., O'Mara, T.A., Batra, J., Cheng, T., Lose, F.A., Dennis, J., et al., 2015. Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. Human Molecular Genetics 24(5):1478–1492. https://doi.org/10.1093/hmg/ddu552.

[42] De Vas, M.G., Kopp, J.L., Heliot, C., Sander, M., Cereghini, S., Haumaitre, C., 2015. Hnf1b controls pancreas morphogenesis and the generation of Ngn3+ endocrine progenitors. Development 142(5):871–882. https://doi.org/10.1242/dev.110759.

[43] Horikawa, Y., Iwasaki, N., Hara, M., Furuta, H., Hinokio, Y., Cockburn, B.N., et al., 1997. Mutation in hepatocyte nuclear factor−1β gene (TCF2) associated with MODY. Nature Genetics 17(4):384–385. https://doi.org/10.1038/ng1297-384.

[44] Lickert, H., Domon, C., Huls, G., Wehrle, C., Duluc, I., Clevers, H., et al., 2000. Wnt/(beta)-catenin signaling regulates the expression of the homeobox gene Cdx1 in embryonic intestine. Development (Cambridge, England) 127(17):3805–3813.

[45] McCarthy, M.I., Zeggini, E., 2009. Genome-wide association studies in type 2 diabetes. Current Diabetes Reports, 164–171. https://doi.org/10.1007/s11892-009-0027-4.

[46] da Silva Xavier, G., Mondragon, A., Sun, G., Chen, L., McGinty, J.A., French, P.M., et al., 2012. Abnormal glucose tolerance and insulin secretion in pancreas-specific Tcf7l2-null mice. Diabetologia 55(10):2667–2676. https://doi.org/10.1007/s00125-012-2600-7.

[47] Shu, L., Sauter, N.S., Schulthess, F.T., Matveyenko, A.V., Oberholzer, J., Maedler, K., 2008. Transcription factor 7-like 2 regulates beta-cell survival and function in human pancreatic islets. Diabetes 57(3):645–653. https://doi.org/10.2337/db07-0847.

[48] Schäfer, S.A., Tschritter, O., Machicao, F., Thamer, C., Stefan, N., Gallwitz, B., et al., 2007. Impaired glucagon-like peptide-1-induced insulin secretion in carriers of transcription factor 7-like 2 (TCF7L2) gene polymorphisms. Diabetologia 50(12):2443–2450. https://doi.org/10.1007/s00125-007-0753-6.

[49] Wang, X., Chen, S., Burtscher, I., Sterr, M., Hieronimus, A., Machicao, F., et al., 2016. Generation of a human induced pluripotent stem cell (iPSC) line from a patient carrying a P33T mutation in the PDX1 gene. Stem Cell Research 17(2):273–276. https://doi.org/10.1016/j.scr.2016.08.004.

[50] Wang, X., Chen, S., Burtscher, I., Sterr, M., Hieronimus, A., Machicao, F., et al., 2016. Generation of a human induced pluripotent stem cell (iPSC) line from a patient with family history of diabetes carrying a C18R mutation in the PDX1 gene. Stem Cell Research 17(2):292–295. https://doi.org/10.1016/j.scr.2016.08.005.

[51] Li, W., Wang, X., Fan, W., Zhao, P., Chan, Y.C., Chen, S., et al., 2012. Modeling abnormal early development with induced pluripotent stem cells from aneuploid syndromes. Human Molecular Genetics 21(1):32–45. https://doi.org/10.1093/hmg/ddr435.

[52] Guo, Y., Mahony, S., Gifford, D.K., 2012. High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. PLoS Computational Biology 8(8). https://doi.org/10.1371/journal.pcbi.1002638.

[53] Xing, H., Mo, Y., Liao, W., Zhang, M.Q., 2012. Genome-wide localization of protein-DNA binding and histone modification by a bayesian change-point method with ChIP-seq data. PLoS Computational Biology 8(7). https://doi.org/10.1371/journal.pcbi.1002613.

[54] ENCODE Project Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489(7414):57–74. https://doi.org/10.1038/nature11247.

[55] Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., et al., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for Macrophage and B Cell identities. Molecular Cell 38(4):576–589. https://doi.org/10.1016/j.molcel.2010.05.004.

[56] Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., et al., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences 102(43):15545–15550. https://doi.org/10.1073/pnas.0506580102.

[57] Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al., 2003. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature Genetics 34(3):267–273. https://doi.org/10.1038/ng1180.

Original Article

Point mutations in the PDX1 transactivation domain impair human β-cell development and function

Xianming Wang [1,2,3], Michael Sterr [1,2,3], Ansarullah [1,2], Ingo Burtscher [1,2], Anika Böttcher [1,2], Julia Beckenbauer [1,2], Johanna Siehler [1,2,3], Thomas Meitinger [4], Hans-Ulrich Häring [5,6,8], Harald Staiger [5,7,8], Filippo M. Cernilogar [9], Gunnar Schotta [9], Martin Irmler [10], Johannes Beckers [8,10,11], Christopher V.E. Wright [12,13], Mostafa Bakhti [1,2,8], Heiko Lickert [1,2,3,8,*]

ABSTRACT

Objective: Hundreds of missense mutations in the coding region of *PDX1* exist; however, if these mutations predispose to diabetes mellitus is unknown.

Methods: In this study, we screened a large cohort of subjects with increased risk for diabetes and identified two subjects with impaired glucose tolerance carrying common, heterozygous, missense mutations in the *PDX1* coding region leading to single amino acid exchanges (P33T, C18R) in its transactivation domain. We generated iPSCs from patients with heterozygous $PDX1^{P33T/+}$, $PDX1^{C18R/+}$ mutations and engineered isogenic cell lines carrying homozygous $PDX1^{P33T/P33T}$, $PDX1^{C18R/C18R}$ mutations and a heterozygous *PDX1* loss-of-function mutation ($PDX1^{+/-}$).

Results: Using an *in vitro* β-cell differentiation protocol, we demonstrated that both, heterozygous $PDX1^{P33T/+}$, $PDX1^{C18R/+}$ and homozygous $PDX1^{P33T/P33T}$, $PDX1^{C18R/C18R}$ mutations impair β-cell differentiation and function. Furthermore, $PDX1^{+/-}$ and $PDX1^{P33T/P33T}$ mutations reduced differentiation efficiency of pancreatic progenitors (PPs), due to downregulation of PDX1-bound genes, including transcription factors *MNX1* and *PDX1* as well as insulin resistance gene *CES1*. Additionally, both $PDX1^{P33T/+}$ and $PDX1^{P33T/P33T}$ mutations in PPs reduced the expression of PDX1-bound genes including the long-noncoding RNA, *MEG3* and the imprinted gene *NNAT*, both involved in insulin synthesis and secretion.

Conclusions: Our results reveal mechanistic details of how common coding mutations in *PDX1* impair human pancreatic endocrine lineage formation and β-cell function and contribute to the predisposition for diabetes.

© 2019 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords PDX1; Transactivation domain; β-Cell differentiation; Insulin secretion; PDX1-Bound genes

1. INTRODUCTION

Diabetes mellitus is a group of heterogeneous disorders characterized by high blood glucose levels that are caused by the loss and/or dysfunction of the pancreatic insulin-producing β-cells. The mechanisms leading to the onset of this multifactorial disease still need to be identified, and advancements will aid in improved diagnostics and personalized therapy [1]. The major forms of diabetes include type 1 (T1D) and type 2 diabetes (T2D), gestational diabetes, and maturity-onset diabetes of the young (MODY) [2]. Among these, MODYs are

clinically heterogeneous monogenic forms with autosomal-dominant inheritance that result from loss-of-function mutations in key genes involved in β-cell development or function [3]. In humans 13 different MODYs have been characterized so far including MODY4, which is associated with haploinsufficiency of the transcription factor (TF) pancreas/duodenum homeobox protein 1 (PDX1) [4,5], one of the master regulators of pancreas development and β-cell function. The *PDX1* gene encodes a TF containing a transactivation domain and DNA binding homeodomain. In mouse, Pdx1 not only is important for induction and growth of the embryonic pancreas but also plays a

[1]Institute of Diabetes and Regeneration Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany [2]Institute of Stem Cell Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany [3]Technische Universität München, Ismaningerstraße 22, 81675 München, Germany [4]Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany [5]Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Zentrum München at the University of Tübingen, 72076 Tübingen, Germany [6]Department of Internal Medicine, Division of Endocrinology, Diabetology, Vascular Medicine, Nephrology and Clinical Chemistry, University of Tübingen, 72076 Tübingen, Germany [7]Institute of Pharmaceutical Sciences, Department of Pharmacy and Biochemistry, Eberhard Karls University Tübingen, 72076 Tübingen, Germany [8]German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany [9]Biomedical Center and Center for Integrated Protein Science Munich, Ludwig-Maximilians-University, 82152 Planegg-Martinsried, Germany [10]Institute of Experimental Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany [11]Chair of Experimental Genetics, School of Life Sciences Weihenstephan, Technische Universität München, 85354 Freising, Germany [12]Vanderbilt University Program in Developmental Biology, Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN 37232, USA [13]Vanderbilt Center for Stem Cell Biology, Vanderbilt University, Nashville, TN 37232, USA

*Corresponding author. Helmholtz Zentrum München, Ingolstaedter Landstraße 1, 85764 Neuherberg, Germany. Fax: +49 89 31873761. E-mail: heiko.lickert@helmholtz-muenchen.de (H. Lickert).

crucial role during insulin-producing β- and somatostatin-producing δ-cell development and function in the adult organ [6—9]. Homozygous Pdx1-deficient mice fail to generate a pancreas [10], while heterozygous animals develop a pancreas but become diabetic in adulthood due to β-cell apoptosis [11—13]. In humans, several missense coding mutations in *PDX1* gene such as the P33T and C18R mutations in the transactivation domain have been associated with and increased risk for diabetes of the carrier individuals [14—16]. Currently, there are more than 150 missense coding mutations described for *PDX1* among which mutations at amino acid position 18 and 33 are rather common (gnomad.broadinstitute.org); however, causal link to increased risk for type 2 diabetes is still missing for most mutations [17]. In contrast, *PDX1*$^{P33T/+}$ and *PDX1*$^{C18R/+}$ mutations have been shown to perturb the activity of the PDX1 protein and reduce the expression of insulin gene in INS-1 and NES2Y cell lines [15,16] although the exact mechanisms by which these mutations contribute to diabetes predisposition are not understood. Moreover, whether these mutations exert their effects through impairment in developmental programs, regulating β-cell differentiation or adult β-cell function remains unclear. Although several studies have shed light on the developmental impacts of other coding mutations of pancreatic TFs such as *PDX1*, *HNF1B*, *GATA4*, and *GATA6* [18—20], how common mutations in the *PDX1* gene affect human pancreatic progenitors and β-cells still needs to be addressed. The major obstacle is a lack of appropriate modeling systems to investigate the effect of loss-of-function or point mutations in certain genes on human pancreas development. One of the appealing approaches is the generation of induced pluripotent stem cells (iPSCs) from somatic cells from diabetic patients [21,22]. In such a system, patient-derived somatic cells are reprogrammed to generate patient-specific stem cells, which can be further differentiated into the endocrine lineage cells, mimicking human β-cell development in a culture dish [19,23,24]. Alternatively, advancements in CRISPR-Cas9 gene-editing technology offer targeting of specific mutations in the genes of interest to generate disease-specific cells and investigate the corresponding consequences [20,25].

Previously, we identified the genome-wide target gene profile of PDX1 in human pancreatic progenitors [26]. However, how PDX1 coordinates human pancreatic cell development is not understood in detail. To address this, we investigated the impact of *PDX1* coding mutations as well as its haploinsufficiency (*PDX1*$^{+/−}$) on differentiation of human pancreatic cells from progenitor stage to β-cells. First, we screened a large cohort of human subjects with high risk to develop diabetes and identified individuals with insulin secretion deficiency associated with heterozygous *PDX1*$^{P33T/+}$ and *PDX1*$^{C18R/+}$ missense mutations. Using patient-derived iPSCs, we found that both heterozygous mutations impair *in vitro* β-cell differentiation and function. To further exclude genetic background variations in the human population and investigate dose-dependent effects, we generated isogenic iPSC lines carrying homozygous *PDX1*$^{P33T/P33T}$ and *PDX1*$^{C18R/C18R}$ point mutations. Our results indicate that homozygous point mutations in the PDX1 transactivation domain do not only impact pancreatic endocrine lineage development, but also impair glucose-responsive function of β-cells through misregulation of several PDX1 target genes involved in β-cell development, maturation, and function. Altogether, our data provide novel insight into the mechanisms by which common point mutations in the PDX1 transactivation domain impair human pancreatic β-cell formation and function and contribute to increased risk for diabetes in the general population.

2. MATERIALS AND METHODS

2.1. Ethics statement

The choice of appropriate human donors, the procedures for skin biopsy, isolation of dermal fibroblasts, generation of iPSCs, and their use in further scientific investigations were performed under the positive vote of the Ethics Committee of the Medical Faculty of the Eberhard Karls University, Tübingen. The study design followed the principles of the Declaration of Helsinki. All study participants gave informed consent prior to entry into the study.

2.2. Cell culture

hiPSCs were cultured on 1:100 diluted Matrigel (BD Biosciences, CA, Cat #354277) in mTeSR™1 medium (STEMCELL technologies, Cat #85850). At ∼70—80% confluency, cultures were rinsed with 1 × DPBS without Mg^{2+} and Ca^{2+} (Invitrogen, Cat #14190) followed by incubation with TrypLE Select Enzyme (1 ×) (Life Technologies, Cat #12563011) for 3—5 min at 37 °C. Single cells were rinsed with mTeSR™1 medium, and spun at 1,000 rpm for 3 min. The resulting cell pellet was resuspended in mTeSR™1 medium supplemented with Y-27632 (10 μM; Sigma—Aldrich; MO, Cat #Y0503), and the single cell suspension was seeded at ∼0.75 × 10^5 cells/cm^2 on Matrigel-coated surfaces. Cultures were fed every day with mTeSR™1 medium and differentiation was initiated 72 h following seeding with ∼90% starting confluency. All the cell lines were confirmed mycoplasma-free by using the Lonza MycoAlert Mycoplasma Detection Kit (Lonza, Cat #LT07-418).

2.3. 1st pancreatic lineage differentiation protocol

2.3.1. S1: definitive endoderm (3 d)

Cells were first rinsed with 1 × DPBS without Mg^{2+} and Ca^{2+} and then exposed to MCDB 131 medium (Life Technologies, Cat #10372-019) further supplemented with 1.5 g/l sodium bicarbonate (Sigma, MO, Cat #S6297), 1 × Glutamax (Life Technologies, Cat #35050-079), 10 mM final glucose (Sigma, Cat #G8769) concentration, 0.5% bovine serum albumin fraction V, fatty acid free (Sigma, Cat # 10775835001), 100 ng/ml Activin-A (R&D Systems Inc, Cat #338-AC-050/CF), and 3 μM or 5 μM of CHIR-99021 (GSK3β inhibitor, SelleckChem, Cat #S2924) for day 1. For day 2, cells were cultured in MCDB 131 medium with 0.5% BSA, 1.5 g/l sodium bicarbonate, 1 × Glutamax, 10 mM glucose, 100 ng/ml Activin-A and 0.3 μM CHIR-99021. On day 3, cells were cultured in MCDB 131 with 0.5% BSA, 1.5 g/l sodium bicarbonate, 1 × Glutamax, 10 mM glucose and 100 ng/ml Activin-A.

2.3.2. S2: primitive gut tube (2 d)

Cells were rinsed with 1 × DPBS without Mg^{2+} and Ca^{2+} and then exposed to MCDB 131 medium further supplemented with 1.5 g/l sodium bicarbonate, 1 × Glutamax, 10 mM final glucose concentration, 0.5% BSA, 0.25 mM ascorbic acid (Sigma, Cat #A4544), 50 ng/ml FGF7 (R & D Systems, Cat #251-KG-010/CF) or/and 1.25 μM IWP-2 (Tocris Bioscience, Cat #3533) for 2 d.

2.3.3. S3: posterior foregut (2 d)

Cells were then added for 2 d in MCDB 131 medium supplemented with 2.5 g/l sodium bicarbonate, 1 × Glutamax, 10 mM glucose concentration, 2% BSA, 0.25 mM ascorbic acid, 50 ng/ml FGF7, 0.25 μM SANT-1 (Sigma, Cat #S4572), 1 μM retinoic acid (RA; Sigma, Cat #R2625), 100 nM LDN193189 (LDN; BMP receptor inhibitor, Stemgent, CA, Cat #04-0019), 1:200 ITS-X (Life Technologies, Cat

#51500056), and 200 nM TPB (PKC activator, custom synthesis, ChemPartner).

2.3.4. S4: pancreatic endoderm (3 d)
Cells were exposed to MCDB 131 medium supplemented with 2.5 g/l sodium bicarbonate, 1 × Glutamax, 10 mM final glucose concentration, 2% BSA, 0.25 mM ascorbic acid, 2 ng/ml of FGF7, 0.25 μM SANT-1, 0.1 μM retinoic acid, 200 nM LDN193189, 1:200 ITS-X, and 100 nM TPB for 3 d.

After 3 d of culture, the S4 cells were treated for 4 h with 10 μM Y-27632. Cells were then rinsed with 1 × DPBS without Mg^{2+} and Ca^{2+} and exposed to TrypLE Select Enzyme (1 ×) (Life Technologies, Cat #12563011) for 3—5 min at 37 °C. The released cells were washed with basal MCDB 131 medium and spun at 1,000 rpm for 3 min. Cell pellets were resuspended in S5-7 media with S5 chemical supplements at ∼50 million cells/ml and spotted onto Transwell insert filters (6-well plate: Corning 3414) for culture in air-liquid interface at 10 μl/spot and ∼10 spots per well. S5-7 media with supplements were added to the bottom of each insert: ∼1.5 ml per well.

2.3.5. S5: pancreatic endocrine precursors (3 d)
Cells were further cultured in MCDB 131 medium supplemented with 1.5 g/l sodium bicarbonate, 1 × Glutamax, 20 mM glucose, 2% BSA, 0.25 μM SANT-1, 0.05 μM retinoic acid, 100 nM LDN193189, 1:200 ITS-X, 1 μM T3 (3,3′,5-Triiodo-L-thyronine sodium salt, Sigma, T6397), 10 μM ALK5 inhibitor ll (Enzo Life Sciences, NY, Cat #ALX-270-445), 10 μM zinc sulfate (Sigma, Z0251) and 10 μg/ml of heparin (Sigma, H3149).

2.3.6. S6: immature β-like cells (7 d)
Cells were exposed to the MCDB 131 medium further supplemented with 1.5 g/l sodium bicarbonate, 1 × Glutamax, 20 mM glucose, 2% BSA, 100 nM LDN193189, 1:200 ITS-X, 1 μM T3, 10 μM ALK5 inhibitor ll, 10 μM zinc sulfate, 100 nM gamma secretase inhibitor XX (EMD MilliPore, MA, Cat #565789) and 10 μg/ml of heparin for 7 d.

2.3.7. S7: mature β-like cells (14 d)
Cells were further cultured in MCDB 131 medium supplemented with 1.5 g/l sodium bicarbonate, 1 × Glutamax, 20 mM final glucose concentration, 2% BSA, 1:200 ITS-X, 1 μM T3, 10 μM ALK5 inhibitor ll, 10 μM zinc sulfate, 1 mM N-acetyl cysteine (N-Cys, Sigma, Cat #A9165), 10 μM Trolox (Vitamin E analogue, EMD, Cat #648471), 2 μM R428 (AXL inhibitor, SelleckChem, Cat #S2841) and 10 μg/ml of heparin for 14 d. For all stages, the cultures were fed every day.

2.4. 2nd generation pancreatic progenitor differentiation protocol

2.4.1. S1: definitive endoderm (3 d)
iPSCs plated on 1:100 diluted Matrigel were first rinsed with 1 × DPBS without Mg^{2+} and Ca^{2+} (Invitrogen) and then cultured in RPMI 1640 medium (Invitrogen, Cat #21875-034) further supplemented with 1.2 g/L sodium bicarbonate (Sigma), 0.2% ESC-qualified FBS (Life Technologies, Cat #16141-079), 100 ng/mL Activin-A (R&D Systems), and 20 ng/mL of Wnt3A (R&D Systems, Cat #5036-WN-010) for day 1 only. For the next 2 days, cells were cultured in RPMI with 0.5% FBS, 1.2 g/L sodium bicarbonate and 100 ng/mL Activin-A.

2.4.2. S2: primitive gut tube (3 d)
Cells were rinsed with 1 × DPBS (without Mg^{2+} and Ca^{2+}) once and then exposed to DMEM-F12 (Life Technologies, Cat #21041025) medium further supplemented with 2 g/L sodium bicarbonate, 2% FBS and 50 ng/mL of FGF7 for 3 days.

2.4.3. S3: posterior foregut (4 d)
Cultures were maintained for 4 days in DMEM-HG medium (Life Technologies, Cat #41966029) supplemented with 0.25 μM SANT-1 (Sigma—Aldrich), 2 μM retinoic acid (RA; Sigma—Aldrich), 100 ng/mL of Noggin (R&D Systems, Cat #6057-NG-025), and 1% B27 (Invitrogen, Cat #17504044).

2.5. Generation of clonal hiPSC mutant lines using gRNA or gRNA/ssDNA transfection in hiPSCs
All mutant lines were generated using a previously established protocol [27]. The plasmid for transfection (pU6-(BbsI) sgRNA_CAG-venus-bpA, Addgene ID86985) was a gift from from Dr. Ralf Kühn and contained BbsI site for single or multiple gRNAs, the *Venus* gene and the *Cas9* gene. Human iPSCs were cultured for two days and then dissociated using TrypLE select, replated onto Matrigel-coated plates and transfected in suspension with gRNAs or gRNA/ssDNA using Lipofectamine RNAiMAX (Thermo Fisher Scientific, Cat #13778-150) following manufacturer's instructions. For one well of 6-well plates, 0.75 × 10^6 cells were seeded. For the *PDX1$^{+/-}$* knockout cell line, 2.5 ug plasmid of gRNA was added for one transfection. For the point mutations, 2.5 ug plasmid with the gRNA and 30 pmol ssDNA were added in one transfection. gRNAs or gRNA/ssDNA and Lipofectamine RNAiMAX were diluted separately in Opti-MEM (Thermo Fisher Scientific, Cat # 31985070), mixed together, incubated for 5 min at room temperature (R.T.), and added dropwise to the hiPSCs before plating.

2.6. Establishment of clonal hiPSC mutant lines
Two days after the gRNA or gRNA/ssDNA transfection, hiPSCs were dissociated into single cells and the transfected positive cells characterized by Venus expression were sorted by FACS. After that, the sorted cells were replated at a low density (500-1,000 cells per 10 cm dish). The remaining cells were collected and genomic DNA was extracted. T7 endonuclease I assay was then performed to assess the CRISPR mutagenesis efficiency. The seeded cells were allowed to grow and form colonies from single cells. Medium was changed every 2 days. 10 days later, individual colonies were picked, mechanically disaggregated and replated in mTeSR™1 medium into individual wells of 96-well plates coated with Matrigel. Clonal lines were expanded. A portion of the cells was lysed to gain genomic DNA. Then PCR was performed using Herculase II Fusion DNA Polymerase (Agilent Technologies, Cat #600679) followed by Sanger sequencing to identify mutant clones. Primers for *PDX1* are as follows: forward: *TACCTGGGCCTAGCCTCTTAGTG*, reverse: *TGAGAACCGGAAAGGA-GAAAGGG*. Clonal cell lines carrying desired mutations were further expanded and frozen down.

2.7. Immunofluorescence imaging
Cells were fixed with 4% paraformaldehyde for 30 min and then permeabilized in PBS containing 0.2% Triton X-100. Cells were blocked with PBS containing 3% BSA, and incubated with primary antibodies overnight at 4 °C. Then the cells were incubated with the secondary antibodies for 1 h at room temperature after washing with PBS. Images were acquired on a TCS SP5 laser-scanning microscope (Leica). The following antibodies and dilutions were used: goat anti-OCT-3/4 (1:500, Cat #sc-8628, Santa Cruz), goat anti-SOX2 (1:500, Cat #sc-17320, Santa Cruz), mouse anti-TRA-1-81 (1:50, Cat #MAB4381, Millipore), mouse anti-SSEA4 (1:500, Cat #4755, Cell Signaling), rabbit anti-FOXA2 (1:250, Cat #8186, Cell Signaling), goat anti-SOX17 (1:500, Cat #GT15094, Acris/Novus), goat anti-PDX1 (1:500, Cat #AF2419, R&D Systems), rabbit anti-NKX6.1 (1:300, Cat

#NBP1-82553, Acris/Novus), guinea pig anti-C-Peptide (1:100, Cat #ab30477, Abcam), mouse anti-Glucagon (1:500, Cat #G2654, Sigma), rat anti-Somatostatin (1:300, Cat #MA5-16987, Invitrogen).

2.8. Flow cytometry

Cells were dissociated using $1 \times$ TrypLE Select Enzyme and collected. After that, cells were washed with cold FACS buffer (5% FBS in $1 \times$ DPBS). Cells were fixed with 4% paraformaldehyde and permeabilized with donkey block solution (0.1% tween-20, 10% FBS, 0.1% BSA and 3% donkey serum) containing 0.5% saponin. The cells were incubated with rabbit anti-FOXA2 (1:200, #8186, Cell Signaling), goat anti-SOX17 antibody (1:200, Cat #GT15094, Acris/Novus), goat anti-PDX1 antibody (1:100, Cat #AF2419, R&D Systems), rabbit anti-NKX6.1 antibody (1:200, Cat #NBP1-82553, Acris/Novus), guinea pig anti-C-Peptide antibody (1:100, Cat #ab30477, Abcam), mouse anti-Glucagon (1:500, Cat #G2654, Sigma), for 30 min at room temperature and then stained with appropriate AlexaFluor-555 and -647 secondary antibody for 30 min at room temperature. Flow cytometry was performed using FACS-Aria III (BD Bioscience). FACS data were analyzed using FlowJo. For quantification of median fluorescence intensity (MFI) for PDX1, the PDX1 positive population was first gated and then MFI was calculated using BD FACS software.

2.9. RNA isolation and qPCR

Total RNA was extracted from cells with the miRNeasy mini kit (Qiagen). cDNA synthesis was performed with a high-capacity RNA-to-cDNA kit (Applied Biosystems). TaqMan qPCR was performed under standard conditions using ViiA7 (Applied Biosystems) and TaqMan Fast Advanced Master Mix (Applied Biosystems). Samples were normalized to the housekeeping genes 18S ribosomal RNA (*RNA18S*) and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). Taqman probes (Applied Biosystems): *INS*, Hs02741908_m1; *GCG*, Hs01031536_m1; *SST*, Hs00356144_m1; *PDX1*, Hs00236830_m1; *GAPDH*, Hs02758991_g1; *18S*, Hs99999901_s1; *NKX6.1*, Hs01055914_m1; *MAFA*, Hs01651425_s1; *NEUROD1*, Hs01922995_s1; *UCN3*, Hs00846499_s1; *PAX6*, Hs01088114_m1; *ABCC8*, Hs01093752_m1; *KCNJ11*, Hs00265026_s1; *SLC30A8*, Hs00545183_m1.

2.10. Affymetrix microarray

For microarray analysis, total RNA from $PDX1^{P33T/+}$ and XM001 PPs, produced with the 2nd differentiation protocol, was extracted using the miRNeasy Mini kit (Qiagen, Cat #217004). RNA integrity was checked using the Agilent 2100 Bioanalyzer (Agilent RNA 6000 Pico Kit) and cDNA was amplified with the Ovation PicoSL WTA System V2 (Nugen, 3312) in combination with the Encore Biotin Module (Nugen, USA). Amplified cDNA was hybridized on GeneChip™ Human Gene 2.0 ST arrays (Affymetrix, 902113). Expression console (v.1.3.0.187, Affymetrix) was used for quality control. All subsequent computational analysis was performed in R using Bioconductor packages. Expression data were RMA normalized using the oligo package (version 1.38.0) and probe sets were annotated using the package hugene20sttranscriptcluster.db (version 8.5.0). Differential expression analyses were performed on prefiltered data, containing the 30562 probe sets with highest expression values, using the limma package (version 3.30.7) and p-values were adjusted for multiple testing by Benjamini-Hochberg correction. A gene was considered differentially expressed if the adjusted p-value (FDR) was below a threshold of 0.1 and the fold-change was greater than or equal to 2. Functional enrichments were conducted using HOMER [28]. Functional annotations were based on term affiliations provided by HOMER and literature research.

2.11. ChIP-seq

All samples for ChIP-seq were processed as described in [26]. $PDX1^{P33T/+}$ PP cells (2×10^6 cells) were cross-linked in 1% formaldehyde in culture medium for 10 min at room temperature. The cross-linking reaction was stopped by the addition of glycine to a final concentration of 125 mM. For chromatin fragmentation, cells were resuspended in lysis buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 0.5% SDS) and sonicated in a Covaris S220 sonicator with a duty cycle of 2%, a peak incident power of 105 W and 200 cycles per burst for 20 min. The fragmented chromatin was diluted 1:5 in IP-Buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-Desoxycholate, 140 mM NaCl, H_2O, Protease Inhibitors) and directly used for immunoprecipitation. For PDX1 ChIP, 60% of the chromatin (equivalent of 1.2×10^6 cells) was processed in two parallel ChIPs using magnetic beads, preloaded with 3 µl goat anti PDX1 antibody (kindly provided by C. Wright), for each ChIP. For H3K27ac ChIP, 40% of the chromatin (equivalent 0.4×10^6 cells) was processed in three parallel IPs using magnetic beads, preloaded with 3 µg anti H3K27ac antibody (Diagenode, C15410174). For all ChIPs, antibody incubation was performed at 4 °C for 5 h. Beads were then washed 5 times using (1.) IP-Buffer, (2.) Washing-Buffer 1 (500 mM NaCl, 50 mM Tris-HCl (pH 8.0), 0.1% SDS, 1% NP-40), (3.) Washing-Buffer 2 (250 mM LiCl, 50 mM Tris-HCl (pH 8.0), 0.5% Na-Deoxycholate, 1% NP-40) and (4. & 5.) Washing-Buffer 3 (10 mM Tris-HCl (pH 8.0), 10 mM EDTA). Subsequently, protein-DNA complexes were eluted from the beads in Elution-Buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% SDS) at 65 °C for 20 min. Cross-links were reversed at 65 °C overnight and DNA was purified for library construction using the MicroPlex kit (Diagenode, C05010010).

2.12. ChIP-seq data analysis

For all processing and analysis steps, $PDX1^{P33T/+}$ data from this study was used together with data from XM001 cells, previously published by our group (GSE106950) [26]. Raw reads from PDX1 and H3K27ac ChIP-seq were processed with Trimmomatic (0.35) to remove low quality bases and potential adapter contamination. Next, reads were aligned to hg19 genome using bowtie2 (2.2.6) with very-sensitive option and duplicate reads were removed using samtools (1.3). BAM files from $PDX1^{P33T/+}$ cells were subsampled to match the read depth of the BAM files from XM001 cells using samtools. Binding sites of PDX1 in XM001 and $PDX1^{P33T/+}$ cells were then called using GEM (3.2) in multi condition mode and filtered, after visual inspection, using a Q-value cut-off of 10^{-4} (XM001) and $10^{-4.05}$ ($PDX1^{P33T/+}$). Subsequently, overlapping regions were merged. Regions of H3K27ac enrichment were called using HOMER (4.10) on the merged data from XM001 and $PDX1^{P33T/+}$ cells using parameters -style histone -size 500 -minDist 2000. PDX1 binding sites and H3K27ac enriched regions were further filtered by removing blacklist regions. H3K27ac read counts from XM001 and $PDX1^{P33T/+}$ cells were normalized using a set of promoters from housekeeping genes [29]. Housekeeping genes were derived from the microarray data and defined by an absolute log2 fold change ($PDX1^{P33T/+}$ vs XM001) < log2 (1.1) and a linear coefficient of variation < 0.15. Reads were counted in a region from 1000 bp upstream to 100 bp downstream of the TSS of the housekeeping genes and DESeq2 (1.20) was used to estimate size factors for normalization. Normalized read counts where used when indicated in the figures. Regions and binding sites within 20 kb of a TSS or within a gene body were annotated with the respective gene using bedtools (2.26.0). Motif analysis, annotations with genomic features and pathway enrichment analysis were performed using HOMER. To calculate enrichment of H3K27ac at PDX1-bound sites, three sets of

random sites were generated by shuffling all PDX1 binding sites from $PDX1^{P33T/+}$ cells using bedtools (2.26.0) shuffle. Then, H3K27ac reads were counted and ratio of the mean read count at PDX1 sites over the mean read count at shuffled sites was calculated. For visualization, BIGWIG files were generated from downsampled BAM files and normalized using the size factors previously determined by DESeq2 (H3K27ac only) using deeptools bamCoverage (3.1.0).

2.13. RNA-seq

RNA samples from $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, $PDX1^{C18R/C18R}$ and XM001 lines were collected at the PP1 stage were produced by the 1st differentiation protocol. Total RNA from $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, $PDX1^{C18R/C18R}$, and XM001 lines was extracted using miRNeasy Mini kit (Qiagen, #217004) and RNA integrity was checked using Agilent 2100 Bioanalyzer (Agilent RNA 6000 Pico Kit). Libraries were prepared using the TruSeq Stranded mRNA Library Prep (Illumina).

2.14. RNA-seq analysis

RNA-seq data was quantified using salmon (0.11.3) quant with options — validateMappings — rangeFactorizationBins 4 — numBootstraps 100 — seqBias — gcBias. Quantification results were imported into R using tximport. DESeq2 (1.20) [30] was used for differential expression analysis as follows. Genes with less than 6 reads were discarded and DESEq2 was run with default parameters. Results were calculated with independent filtering and independent hypothesis weighting [31]. For the comparison PPs against iPSCs, all data sets were used and all PP samples (i.e. XM001, $PDX1^{C18R/C18R}$, $PDX1^{P33T/P33T}$, and $PDX1^{+/-}$) were compared with XM001 iPSCs using the lfcThreshold = 1.5 option in the results function. For the comparison of the different PP cells, only data from PPs were loaded into DESeq2. Results for these comparisons were generated without the lfcThreshold option. For all analyses, fold changes were shrunken using the apeglm method [32] using the results from independent hypothesis weighting. Genes were considered as differentially expressed when the FDR was >0.05 (PP vs iPSC) or 0.1 (PDX1 mutations vs control) and absolute shrunken log2 fold change was >2 (PP vs iPSC) or 1 (PDX1 mutations vs control). GO-term and pathway analysis were performed using HOMER.

2.15. Static glucose stimulation insulin secretion

Static glucose stimulated insulin secretion (GSIS) of the generated β-like cells was performed based on previous protocols [18,20]. Briefly, 5 aggregates were picked and rinsed three times with KRBH buffer (129 mM NaCl, 4.8 mM KCl, 2.5 mM CaCl2, 1.2 mM MgSO4, 1 mM Na2HPO4, 1.2 mM KH2PO4, 5 mM NaHCO3, 10 mM HEPES and 0.1% BSA in deionized water and sterile filtered) and then equilibrated in KRBH buffer at 37 °C for 1 h. Aggregates then were incubated in KRBH buffer with 2.8 mM glucose for 60 min at R.T. Supernatants were collected, and the aggregates were transferred to KRBH buffer with 16.7 mM glucose for 60 min. Supernatants were collected again. At the end of the experiment, cell aggregates were dissociated into single cells and the cell numbers were counted to normalize the GSIS. Mercodia Human Insulin ELISA kit (Mercodia, Cat #10-1113-01) and Human C-peptide ELISA kit (Mercodia, Cat #10-1141-01) was used to measure the insulin and C-peptide content in supernatant samples following manufacturer's protocols.

2.16. Western blotting

Cells were harvested at PP1 stages and lysed using cell lysis buffer containing protease inhibitors. Samples were separated on a 10% SDS-PAGE gel and transferred to PVDF Pre-Cut Blotting Membranes followed by blocking with 5% milk in Tris-based saline with Tween 20

(0.1% TBST) buffer for 1 h at R.T. The membrane was incubated with primary antibodies overnight at 4 °C, followed by incubation with HRP conjugated secondary antibodies at R.T. for 1 h. ECL western blotting detection reagent (Bio-rad, Cat #1705061) was used to visualize the protein bands. The following antibodies were used with the dilution ratio noted: goat anti-PDX1 (R&D, Cat #AF2419, 1:1,000), mouse anti-ACTB (b-Actin) (Cell Signaling Technology, Cat #3700S, 1:10,000).

2.17. Statistics

Statistical significance was determined using one-way ANOVA followed by Bonferroni's multiple comparisons test using GraphPad Prism (version 8.0.0 for Windows, GraphPad Software, San Diego, California USA). A value of $P < 0.05$ was considered statistically significant.

2.18. Accession numbers

Microarray, RNA-seq and ChIP-seq data has been submitted to the GEO database at NCBI under the accession number GSE125770.

3. RESULTS

3.1. Generation of β-like cells from patients carrying $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ mutations

To obtain samples from subjects with PDX1 mutations, 2547 participants of the Tübingen Family Study for T2D (TÜF study, currently more than 3000 participants, inclusion criteria given in Table 1) were screened by mass spectrometry for known rare mutations in the MODY4 gene PDX1 [33]. In this cohort, we identified six heterozygous P33T (Pro-Thr) carriers (frequency ∼1:425) and one heterozygous C18R (Cys-Arg) carrier (frequency ∼1:2500). Of note, both of the mutations occur at highly conserved amino acids in the transactivation domain of PDX1 (Suppl. Figure 1A,B). One of the PDX1 P33T carriers and the PDX1 C18R carrier were recruited for full-thickness skin biopsy, and primary dermal fibroblasts were isolated from the skin specimens. The clinical characteristics of the two mutation carriers compared to the average of the whole TÜF cohort showed decreased insulin secretion (Table 1). We reprogramed the primary fibroblasts into iPSCs using nucleofection with three episomal plasmids encoding human OCT4, SOX2, NANOG, LIN28, KLF4, and L-MYC (Suppl. Figure 1C) and several colonies were picked and expanded. The established control cell line XM001 as well as the P33T ($PDX1^{P33T/+}$) and C18R ($PDX1^{C18R/+}$) iPSC lines were reported previously [26,34,35].

To investigate the disease-relevant phenotype of these mutations during β-cell development, we differentiated patient-derived iPSCs into β-like cells using a differentiation protocol which is based on Rezania et al., 2014 (Figure 1A) [36]. Fluorescent activated cell sorting (FACS) analysis revealed approximately 90% positive cells for the definitive endoderm (DE) marker SOX17 (Suppl. Figure 2A-B). As expected, there was no difference between the patients and the control iPSC lines in endoderm differentiation, since PDX1 is not expressed until PP1 stage (Suppl. Figure 2C). Moreover, we detected ∼85% PDX1-positive cells for all rounds of iPSC differentiation using FACS analysis at PP1 stage (Suppl. Figure 3A-C), while the protein levels of PDX1 were the same in cells derived from the patients (Suppl. Figure 3D-E). A similar differentiation efficiency was also observed for late stage PP2 where we detected ∼40% PDX1/NKX6.1 double-positive cells derived from all iPSC cell lines by immunostaining and FACS analysis (Suppl. Figure 4A-C). Altogether, these data indicate that $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ mutations do not significantly impair early pancreatic progenitor differentiation.

Trait	$PDX1^{P33T/+}$	$PDX1^{C18R/+}$	TÜF overall (N = 3029)	
Gender	Female	Female	64.2% Females	
Age (yrs)	40	33	42	31—54
BMI (kg/m^2)	23.0[a]	20.6[a]	28.7	24.3—35.3
Body fat content (%)	29	24	33	24—45
Waist-hip ratio	0.819	0.766[a]	0.869	0.813—0.941
Glucose$_0$ (mmol/L)	5.10	5.06	5.17	4.88—5.56
Glucose$_{120}$ (mmol/L)	5.83	4.90[a]	6.22	5.28—7.33
HbA1c (%)	4.7[a]	4.4[a,b]	5.5	5.2—5.8
Insulin$_0$ (pmol/L)	39[a]	40[a]	68	42—112
Insulin$_{30}$ (pmol/L)	419	268[a]	522	326—836
C-peptide$_0$ (pmol/L)	643	427	554	415—777
C-peptide$_{30}$ (pmol/L)	2372	1487	1824	1397—2407
Proinsulin$_0$ (pmol/L)	3	4	3	2—6
Proinsulin$_{30}$ (pmol/L)	12	16[a]	8	4—13
AUC insulin$_{0-30}$/AUC glucose$_{0-30}$ (10^{-9})	28.8	25.7[a]	43.6	28.5—66.7
AUC C-peptide$_{0-120}$/AUC glucose$_{0-120}$ (10^{-9})	340	266	288	237—356
HOMA-IR (10^{-6} mol*U*L^{-2})	1.47[a]	1.50[a]	2.61	1.57—4.47
ISI OGTT (10^{19} L^2*mol^{-2})	13.6	26.5[a]	9.8	5.8—16.0
FFA (μmol/L)	519	618	566	434—720
Leukocytes (μl^{-1})	5140	2520[a]	6080	5118—7380
CRP (mg/dL)	0.11	0.14	0.16	0.07—0.48
GOT (U/L)	18	18	21	17—27
GGT (U/L)	14	23	19	12—31
Triglycerides (mg/dL)	61[a]	59[a]	101	72—144
Total cholesterol (mg/dL)	164[a]	190	191	167—217
HDL-cholesterol (mg/dL)	40[a,b]	66[a]	52	43—62
LDL-cholesterol (mg/dL)	106	122	113	94—137
TÜF inclusion criteria fulfilled	Previous GDM	FHD (one second-degree relative)	IFG and/or BMI ≥27 kg/m^2 and/or FHD and/or previous GDM	

Table 1 — Clinical characteristics of mutation carriers compared to the average of the TÜF study.

TÜF data represent proportion (gender) or medians and interquartile ranges (all other traits). Subscript numbers indicate time-points (in minutes) of the oral glucose tolerance test.
[a] Outside TÜF interquartile range.
[b] Outside clinical reference range (central laboratory data). AUC — area under the curve; BMI — body-mass index; CRP — C-reactive protein; FFA — free fatty acids; FHD — family history of diabetes; GDM — gestational diabetes mellitus; GGT — γ-glutamyl transferase; GOT — glutamate-oxaloacetate transaminase; HbA1c — glycated hemoglobin A1c; HDL — high-density lipoprotein; HOMA-IR — homeostasis model assessment of insulin resistance; IFG — impaired fasting glycemia; ISI — insulin sensitivity index; LDL — low-density lipoprotein; OGTT — oral glucose tolerance test; TÜF — Tübingen Family study for type-2 diabetes.

To assess whether these two point mutations in *PDX1* affect late stages of β-cell development, we differentiated patient-derived iPSCs towards β-like cells. Immunostaining analysis revealed that both patients and control iPSCs generate predominantly cells, co-expressing C-peptide/NKX6.1 and C-peptide/PDX1, and minor populations expressing the α-cell hormone glucagon or the δ-cell hormone somatostatin (Figure 1B and Suppl. Figure 5A). Quantification of β-like cell differentiation efficiency showed ∼25% C-peptide-positive cells for the control but only ∼8% C-peptide-positive cells for the patient samples (Figure 1C,D). The transcript level of *INS* was lower in cells derived from the patients as revealed by q-PCR analysis (Figure 1E). The impairment in β-cell differentiation was also reflected by a weaker response to a glucose challenge in patient iPSCs-derived β-like cells compared to those derived from the control iPSCs, as assessed by a glucose-stimulated insulin secretion (GSIS) assay (Figure 1F and Suppl. Figure 5B). This reduced GSIS might be due to impaired β-cell glucose-responsive machinery components or delayed β-cell differentiation programs. Altogether, these data indicate that PDX1 heterozygous point mutations in the transactivation domain impair the differentiation of β-like cells and consequently impact their glucose-responsive function.

3.2. Characterization of $PDX1^{+/-}$, $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations at the early pancreatic stage (PP1)

Although we analyzed two different patient-derived iPSC lines, it is possible that the observed phenotypes result from differences in the genetic background of control and the patient iPSC lines and are not due to the point mutations in the PDX1 transactivation domain. To

exclude this possibility, we generated cell lines carrying homozygous *PDX1 P33T* ($PDX1^{P33T/P33T}$) and *C18R* ($PDX1^{C18R/C18R}$) point mutations from the control XM001 iPSC line. This approach also allowed us to analyze dose-dependent effects of the common *PDX1* mutations (Suppl. Figure 6A, C-D). Using targeted genetic modifications by the CRISPR/Cas9 system, we generated XM001 iPSC lines carrying $PDX1^{P33T/P33T}$ or $PDX1^{C18R/C18R}$ mutations. Furthermore, to study the effects of PDX1 haploinsufficiency, we generated a $PDX1^{+/-}$ XM001 iPSC line (Suppl. Figure 6B) by targeting the transactivation domain to generate a frame-shift mutation. Isogenic cells with *PDX1* haploinsufficiency or the point mutations on both alleles, displayed normal karyotype (Suppl. Figure 7A) and expressed pluripotency markers (TRA-1-81, SSEA-4, OCT3/4 and SOX2; Suppl. Figure 7B).

To address whether *PDX1* heterozygosity and homozygous point mutations affect pancreas development, we differentiated the isogenic control and mutant iPSCs towards the pancreatic lineage using an improved differentiation protocol by increasing the concentration of CHIR-99021 during S1 stage and adding IWP-2 during S2 stage. We found that all iPSC lines differentiated normally into endoderm cells co-expressing SOX17 and FOXA2 (∼90%) as determined by immunostaining and FACS analysis (Suppl. Figure 8A-C). At the PP1 stage, all control and mutant iPSC lines generated pancreatic progenitors with similar efficiency (∼90%) (Figure 2A—C), however, PDX1 protein levels (Figure 2D) as measured by FACS and immunofluorescence (Figure 2E,F) were substantially lower in PP1 cells differentiated from the $PDX1^{P33T/P33T}$ and $PDX1^{+/-}$ iPSCs as compared to the control and $PDX1^{C18R/C18R}$ iPSCs.
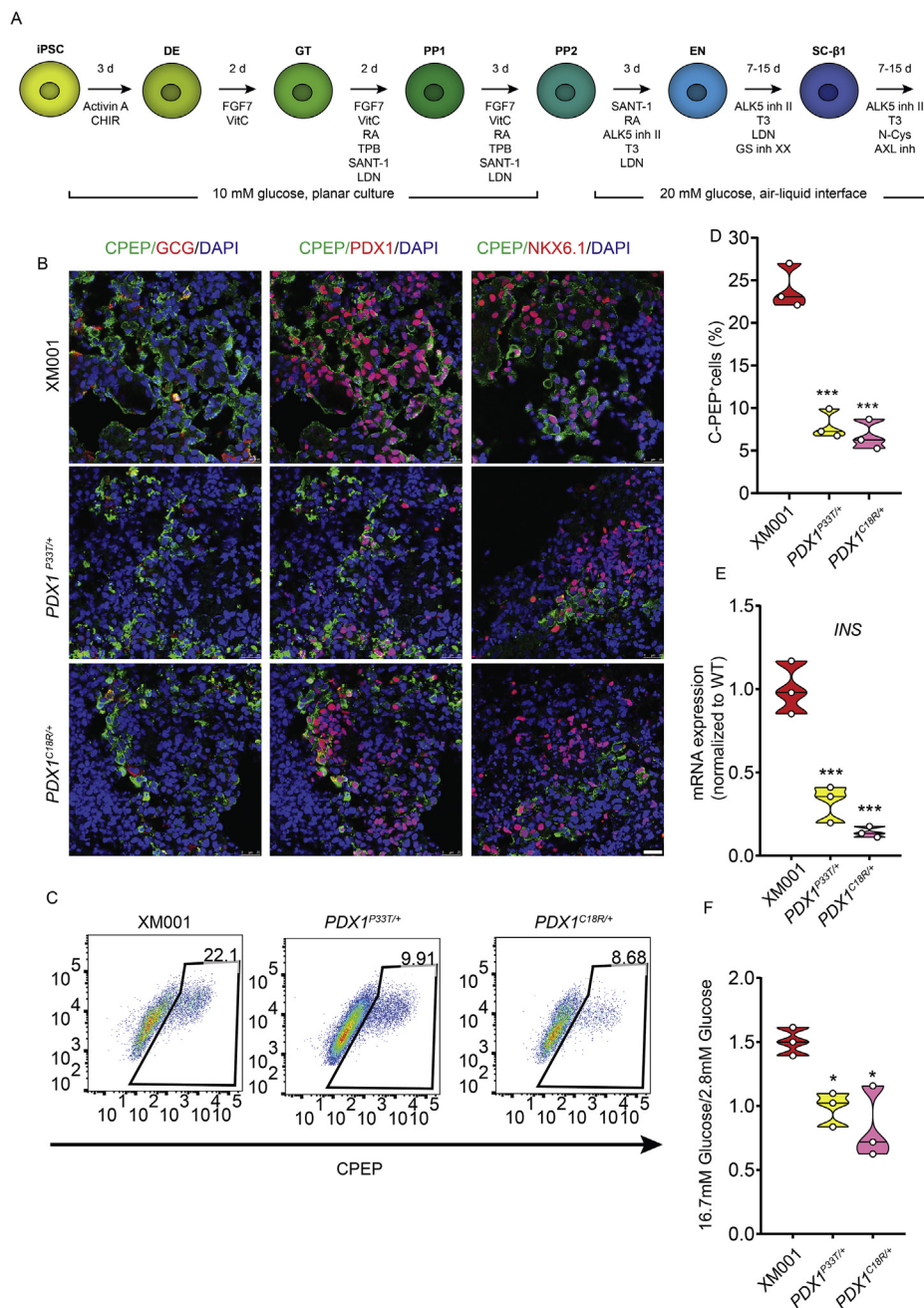
Figure 1: Generation of β-like cells from patients carrying $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ mutations. (A) Schematic of iPSC-derived β-like cells protocol. (B) Immunostaining for CPEP, GCG, PDX1, and NKX6.1 in the XM001, $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ cells at the S7 stage. Scale bar indicates 25 μm. (C) Representative FACS plots of CPEP$^+$ cells in the XM001, $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ cells at the S7 stage. (D) FACS quantification of the percentage of CPEP$^+$ cells in the XM001, $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ cells at the S7 stage (n = 3). (E) RT-qPCR analysis of INS gene expression in the XM001, $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ cells at the S7 stage (n = 3). (F) Glucose-stimulated insulin secretion assay for the XM001, $PDX1^{P33T/+}$, and $PDX1^{C18R/+}$ cells at the S7 stage. The fold change of insulin secretion with high glucose (16.7 mM) relative to low glucose (2.8 mM) treatment is shown (n = 3).

3.3. Characterization of $PDX1^{+/-}$, $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations at the late pancreatic stage (PP2)

Although we detected PDX1 and NKX6.1 double-positive cells in all isogenic cell lines, the number of double-positive cells were significantly lower in $PDX1^{P33T/P33T}$ and $PDX1^{+/-}$ iPSC lines, as compared to $PDX1^{C18R/C18R}$ and control iPSC lines (Figure 3A—C), clearly demonstrating that the expression of $NKX6.1$ and pancreatic differentiation into PP2 cells were impaired.

3.4. Characterization of the impact of $PDX1^{+/-}$, $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations on β-like cell differentiation

To investigate the functional consequences of haploinsufficient PDX1 levels or homozygous $PDX1$ point mutations in β-cell development, we differentiated isogenic iPSCs towards β-like cells. Similar to cells derived from the patient iPSCs, a majority of cells co-expressed C-peptide/NKX6.1 and C-peptide/PDX1 and only few cells expressed glucagon and somatostatin (Figure 4A and Suppl. Figure 9A). However,
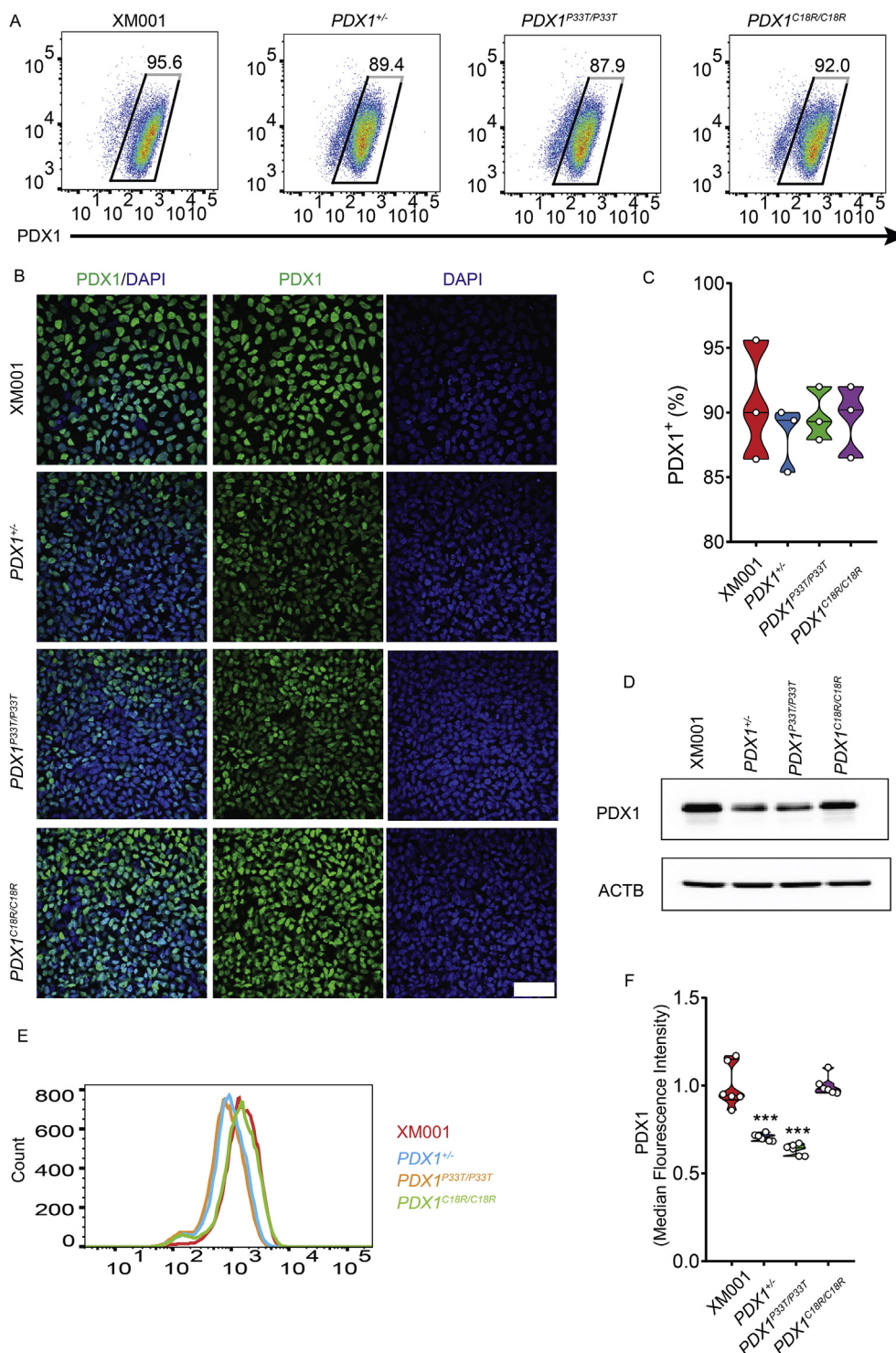
 www.molecularmetabolism.com

Figure 2: Characterization of *PDX1*⁺/⁻, *PDX1*^P33T/P33T, and *PDX1*^C18R/C18R mutations at the early pancreatic stage (PP1). (A) Representative FACS plots of PDX1⁺ cells in XM001, *PDX1*⁺/⁻, *PDX1*^P33T/P33T, and *PDX1*^C18R/C18R cells at the early pancreatic stage. (B) Representative immunofluorescence staining of PDX1 in XM001, *PDX1*⁺/⁻, *PDX1*^P33T/P33T, and *PDX1*^C18R/C18R. Scale bar indicates 50 μm. (C) FACS quantification of the percentage of PDX1⁺ cells in XM001, *PDX1*⁺/⁻, *PDX1*^P33T/P33T and *PDX1*^C18R/C18R cells at the PP1 stage (n = 3). (D) Representative immunoblot of PDX1 expression from XM001, *PDX1*⁺/⁻, *PDX1*^P33T/P33T, and *PDX1*^C18R/C18R cells at the PP1 stage. (E) Representative FACS histograms comparing the differentiation efficiencies towards the PDX1⁺ cells in XM001, *PDX1*⁺/⁻, *PDX1*^P33T/P33T, and *PDX1*^C18R/C18R cells. (F) Median Fluorescence Intensity (MFI) quantification for XM001, *PDX1*⁺/⁻, *PDX1*^P33T/P33T, and *PDX1*^C18R/C18R cells at the PP1 stage stained with PDX1 antibody (n = 6).
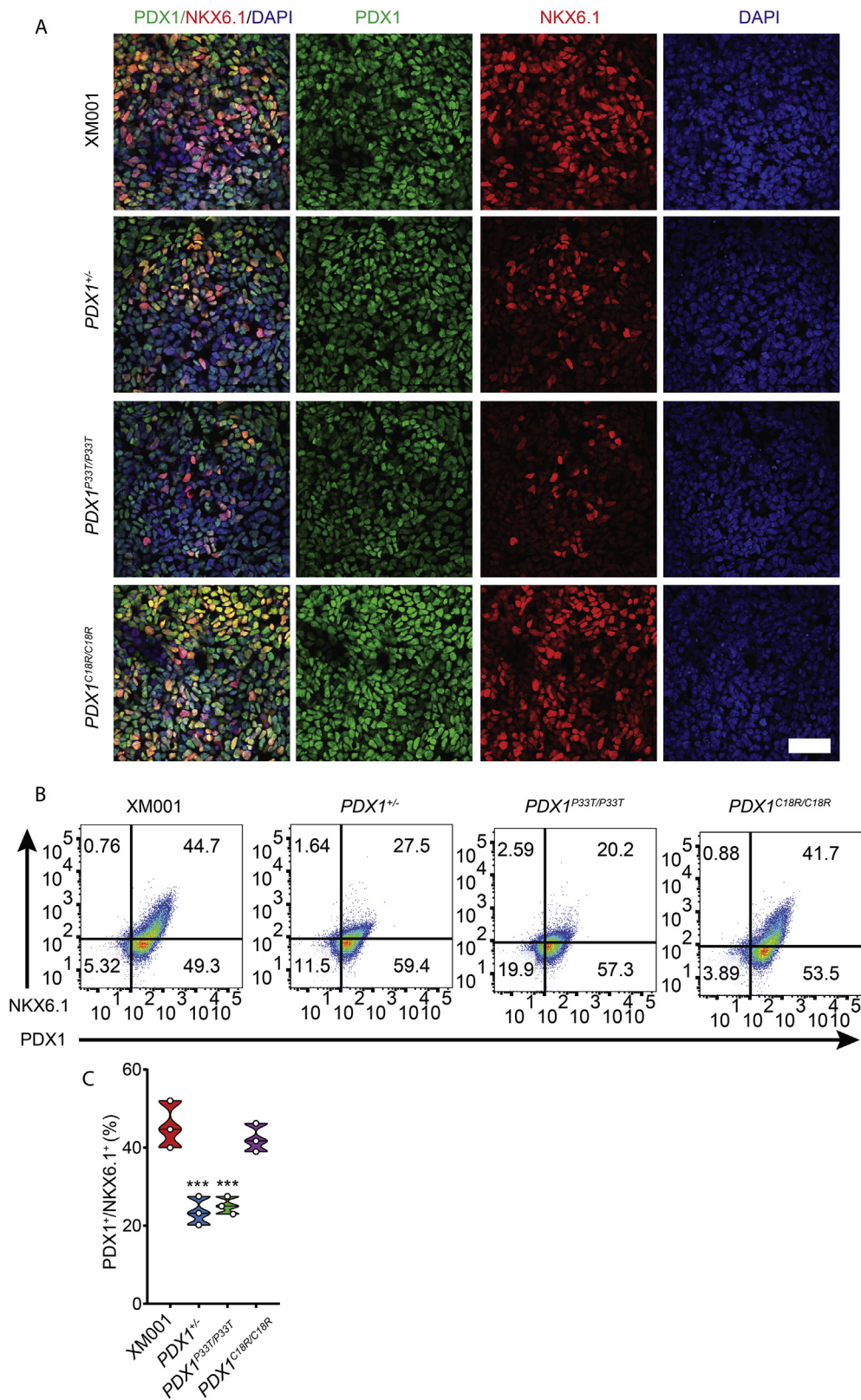
Figure 3: Characterization of *PDX1*$^{+/-}$, *PDX1*$^{P33T/P33T}$ and *PDX1*$^{C18R/C18R}$ mutations at the late pancreatic stage (PP2). (A) Representative immunofluorescence staining of PDX1 and NKX6.1 in XM001, *PDX1*$^{+/-}$, *PDX1*$^{P33T/P33T}$, and *PDX1*$^{C18R/C18R}$ cells. Scale bar indicates 50 μm. (B) Representative FACS plots of PDX1$^+$ and NKX6.1$^+$cells in XM001, *PDX1*$^{+/-}$, *PDX1*$^{P33T/P33T}$, and *PDX1*$^{C18R/C18R}$ cells at the PP2 stage. (C) FACS quantification of the percentage of PDX1$^+$ and NKX6.1$^+$cells in XM001, *PDX1*$^{+/-}$, *PDX1*$^{P33T/P33T}$, and *PDX1*$^{C18R/C18R}$ cells at the PP2 stage (n = 3).

the number of total C-peptide-positive cells, monohormonal C-peptide-positive cells and cells expressing both C-peptide and NKX6.1 derived from $PDX1^{+/-}$ and $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutant isogenic iPSC lines were less compared to the control (Figure 4B—D and Figure 5A,B).

3.5. *PDX1* mutations reduce glucose-responsive function of β-like cells

Furthermore, we performed functional analysis of isogenic iPSC-derived β-like cells using GSIS that revealed a significant impairment in the response to glucose in cells derived from the isogenic mutants compared to the control cells (Figure 5C and

Suppl. Figure 9B). To explain how *PDX1* mutations cause defects in β-like cell function, we performed q-PCR analysis of differentiated endocrine cells. The transcript levels of *INS* and *SST* but not *GCG* were lower in cells derived from $PDX1^{+/-}$ and $PDX1^{C18R/C18R}$ compared to the control and the transcript levels of *INS*, *SST*, and *GCG* were lower in cells derived from $PDX1^{P33T/P33T}$ (Figure 5D). We also found lower expression levels of *NEUROD1*, *ISLI*, *ABCC8*, and *UCN3* in cells derived from all isogenic mutations. Importantly, the expression levels of *PDX1*, *NKX6-1*, *MAFA*, *PAX6*, *SLC30A8*, and *KCNJ11* were less in the $PDX1^{+/-}$ and $PDX1^{P33T/P33T}$ iPSCs-derived β-like cells but not in those originated from the $PDX1^{C18R/C18R}$ (Figure 5D), highlighting distinct effects on gene expression of different mutations in *PDX1*. Of note, with the
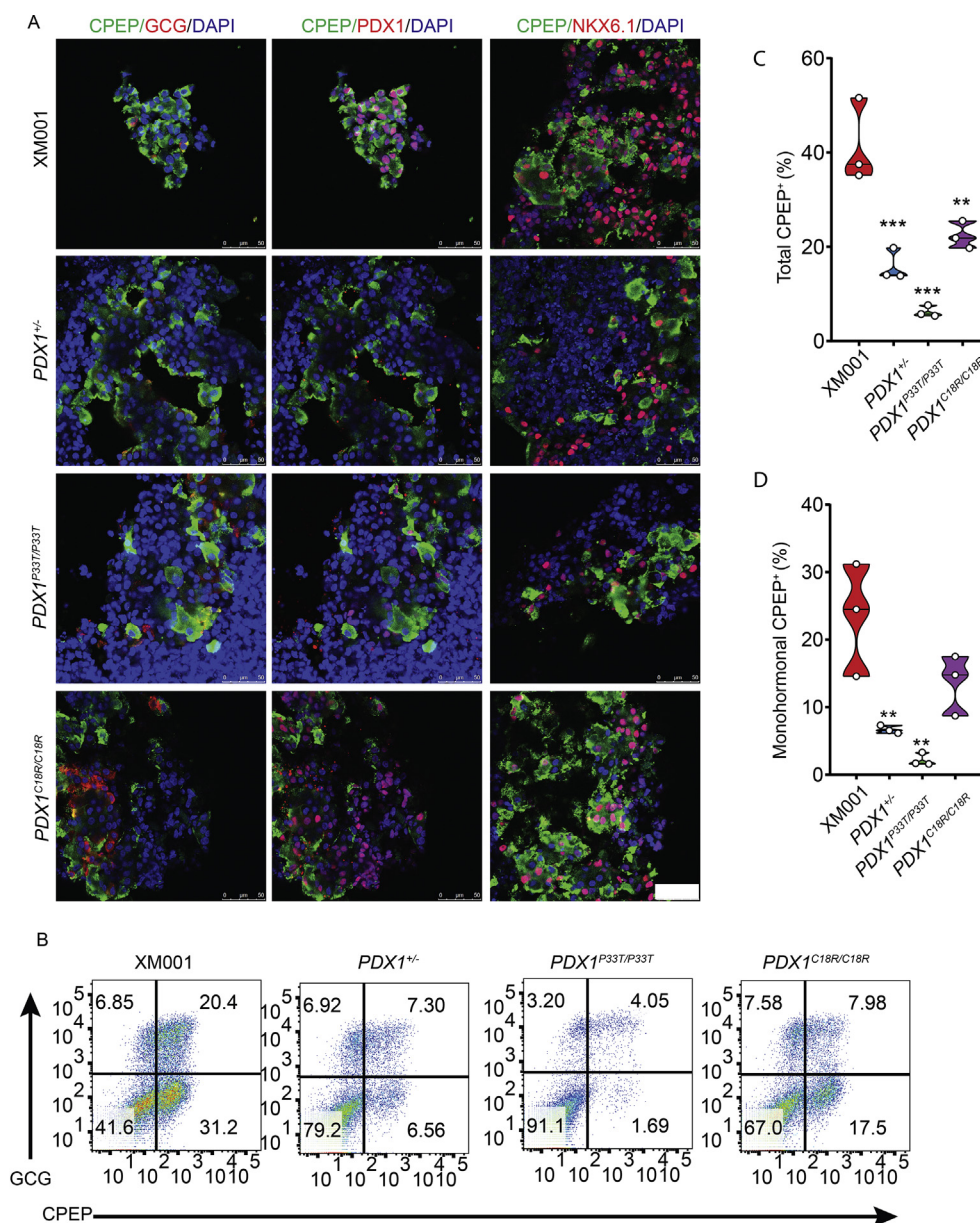

Figure 4: Characterization of the impact of $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ mutations on β-like cell differentiation. (A) Representative immunofluorescence staining for C-peptide, Glucagon, PDX1 and NKX6.1 in XM001, $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ cells at the S7 stage. Scale bar indicates 50 μm. (B) Representative FACS plots of CPEP$^+$ and GCG$^+$ cells in XM001, $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ cells at the S7 stage. (C) FACS quantification of the percentage of total CPEP$^+$ cells in the XM001, $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ cells at the S7 stage (n = 3). (D) FACS quantification of the percentage of mono-hormonal CPEP$^+$ cells in the XM001, $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ cells at the S7 stage (n = 3).

† Bound by PDX1 in XM001 PPs
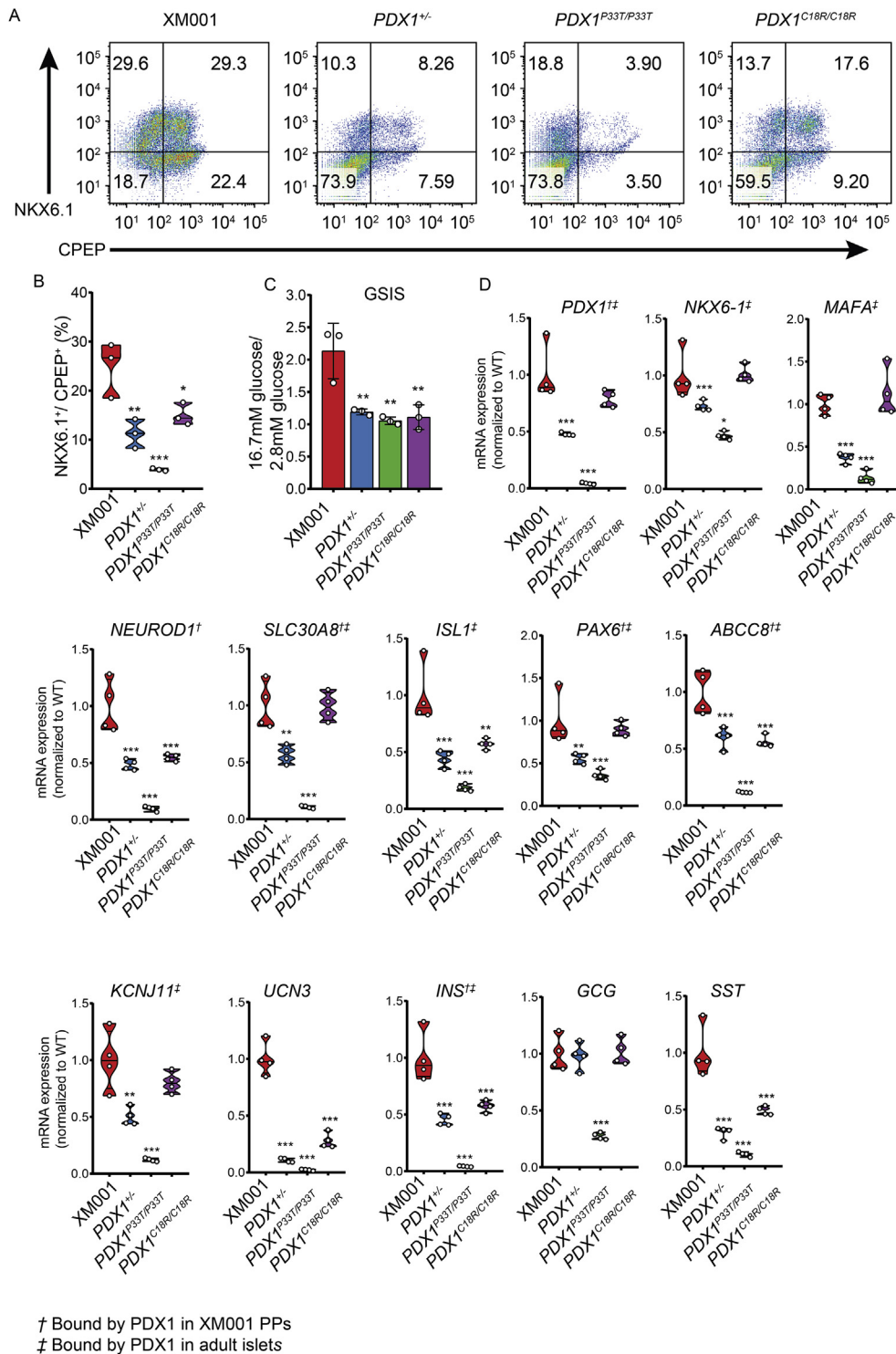‡ Bound by PDX1 in adult islets

Figure 5: *PDX1* mutations reduce glucose-responsive function of β-like cells. (A) Representative FACS plots for C-peptide⁺ and NKX6.1⁺ in XM001, *PDX1*^+/−^, *PDX1*^P33T/P33T^ and *PDX1*^C18R/C18R^ cells at the S7 stage. (B) FACS quantification of the percentage of total CPEP⁺ and NKX6.1⁺ cells in the XM001, *PDX1*^+/−^, *PDX1*^P33T/P33T^, and *PDX1*^C18R/C18R^ cells at the S7 stage (n = 3). (C) GSIS assay for the XM001, *PDX1*^+/−^, *PDX1*^P33T/P33T^, and *PDX1*^C18R/C18R^ cells at the S7 stage. The fold change of insulin secretion with high glucose (16.7 mM) relative to low glucose (2.8 mM) treatment is shown (n = 3). (D) RT-qPCR analysis of expression of β-cell transcription factors, hormonal markers and β-cell functional markers at the S7 stage (n = 4).

exception of *UCN3*, *GCG*, and *SST*, all tested genes are bound by PDX1 in either adult islets (high confidence PDX1 binding sites [26]), XM001 PP1 cells [26] or both (Figure 5D). These data indicate that point

mutations in the transactivation domain reduce the potential of the PDX1 TF to activate the expression of its target genes during β-like cell development and maturation. Furthermore, the *PDX1*^P33T/P33T^ mutation

causes greater impact on β-like cell formation and function than the $PDX1^{C18R/C18R}$ mutation.

3.6. RNA-seq profiling of pancreatic progenitors (PP1) from $PDX1^{+/-}$, $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ iPSC lines

To identify the immediate early PDX1 target genes, which react to haploinsufficiency and the homozygous point mutations in the transactivation domain and which could explain the impaired PP2 differentiation, we performed RNA-seq at the PP1 stage. As quality control, we compared control iPSCs at pluripotency to the average of all other tested cell lines at PP1 stage and, as expected, identified upregulation of a pancreatic gene program (*PDX1*, *HNF1B*, *SOX9*, *ONECUT1*, *HHEX*, *FOXA2*, and *RFX6*) as well as downregulation of a pluripotency program (Figure 6A,B). Furthermore, principal component analysis (PCA)

showed two distinct clusters along the first principal component, corresponding to control iPSCs and the different PPs. Moreover, the second principal component separates the $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ PPs from XM001 control PPs and shows that $PDX1^{+/-}$ and $PDX1^{P33T/P33T}$ are more similar to each other compared to $PDX1^{C18R/C18R}$ and control cells (Figure 6C). Next, we compared mRNA profiles from PP1 cells with isogenic *PDX1* mutation to control PP1 cells. When compared to the control PPs, we found 112, 196 and 199 deregulated genes in $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ PPs, respectively. Among these, 20, 41, and 54 genes were bound by PDX1, respectively (Figure 6D−F). Interestingly, we found the diabetes risk genes *LARGE1* and *ANPEP* and insulin resistance genes *MEG3* and *CES1* to be consistently downregulated among all PPs with *PDX1* mutations (Figure 6G). Moreover, genes such as the metabolic disease
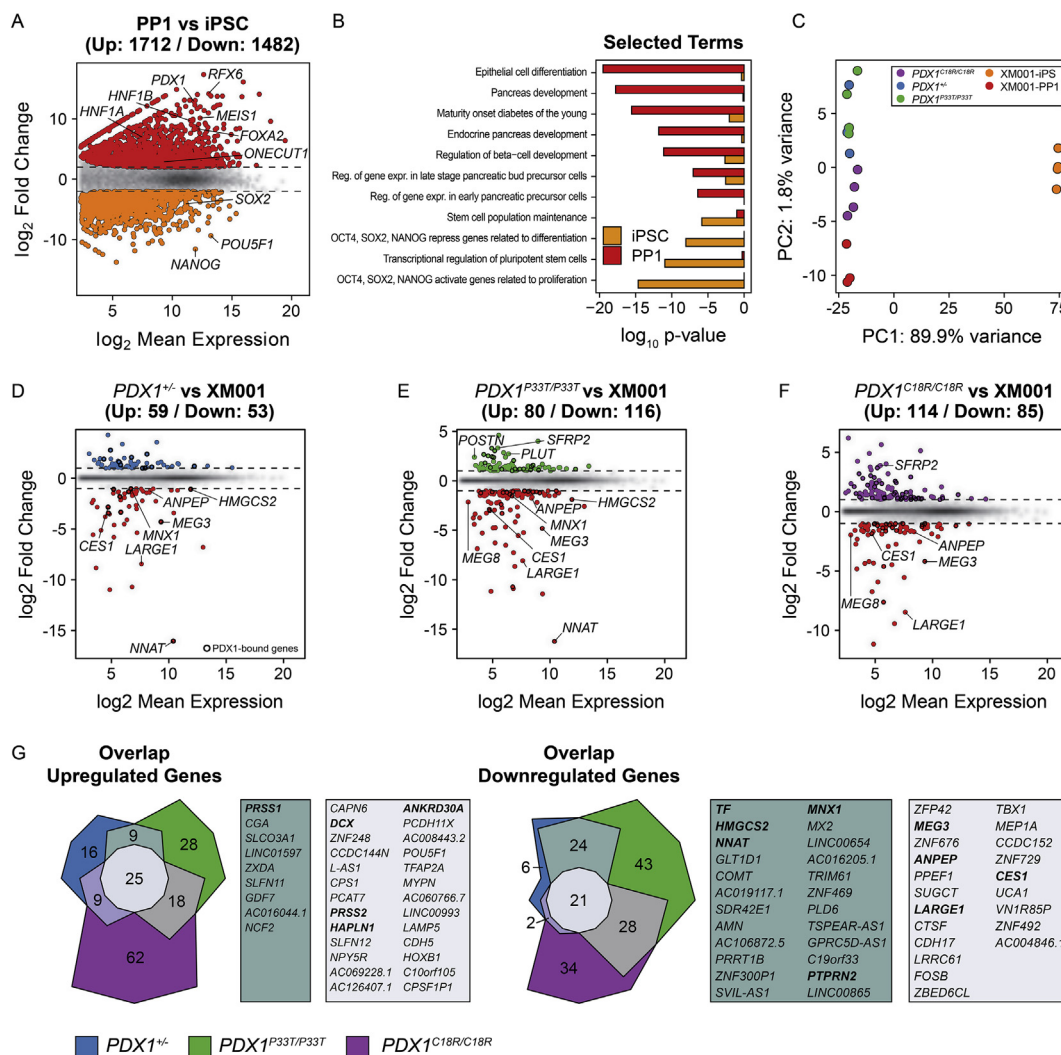

Figure 6: RNA-seq profiling of pancreatic progenitors (PP1) from $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ iPSC lines. (A) MA plot showing the mean $\log_2$ expression against the $\log_2$-fold change of the RNA-Seq data obtained from XM001 iPSCs and PPs from XM001 and isogenic *PDX1* mutants. Genes with significantly different expression ($\log_2$ fold change ≥ 2 and adjusted p-value ≤ 0.05) are drawn in color. Red depicts increased expression in PPs, whereas orange depicts increased expression in iPSCs. (B) Bar graph of p-values from pathway enrichment analysis, showing selected GO terms and KEGG and Reactome pathways from differentially expressed genes. (C) Principal component analysis (PCA) of iPSCs and PP1 cells. (D−F) MA plots showing the mean $\log_2$ expression against the $\log_2$ fold change of the RNA-Seq data obtained from $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ compared to XM001. Genes with significantly different expression ($\log_2$ fold change ≥ 1 and adjusted p-value ≤ 0.1) are drawn in color. Red depicts increased expression in XM001 control PPs, blue, green and purple depict increased expression in $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$, respectively. Black circles genes with PDX1 binding sites in XM001 PPs. (G) Venn diagrams showing the overlap of up- and downregulated genes from $PDX1^{+/-}$, $PDX1^{P33T/P33T}$, and $PDX1^{C18R/C18R}$ PPs compared to XM001 PPs. Some relevant pancreatic and disease associated genes are set in bold.

associated genes *HAPLN1*, the β-cell hallmark gene *DCX*, the fibro-calculous pancreatic diabetes associated genes *PRSS1/2*, which are the main contributors to the difference between PPs from mutant and control isogenic cell lines, were consistently deregulated (Suppl. Figure 10). Since PDX1 levels are reduced in *PDX1^{P33T/P33T}* PPs, we compared the downregulated genes with those from *PDX1^{+/−}* PPs. Among these genes, we found the transcription factor *MNX1* and the proteolipid *NNAT*, as well as genes that seem to distinguish *PDX1^{P33T/P33T}* and *PDX1^{+/−}* from *PDX1^{C18R/C18R}* and XM001 PPs like *TF* and *METTL9* (Figure 6D,E and G and Suppl. Figure 10).

3.7. Characterization of PDX1 binding in patient-derived *PDX1^{P33T/+}* pancreatic progenitors (PP1)

Since the *PDX1^{P33T/P33T}* mutation in isogenic cells severely reduced PP2 cell differentiation and the *PDX1^{P33T/+}* patient was diagnosed with gestational diabetes, we sought to identify genes that might be differentially regulated due to the heterozygous expression of the P33T allele. To this end, we first differentiated iPSCs from the *PDX1^{P33T/+}* patient and the control iPSC into PP1 using the 2nd differentiation protocol. To understand whether the heterozygous P33T mutation affects PDX1 DNA-binding activity, we employed ChIP-seq analysis to profile genome-wide PDX1 binding sites in control [26] and *PDX1^{P33T/+}* patient iPSC-derived PPs. We identified 8970 PDX1 binding sites in the patient iPSC-derived PP cells that were predominantly found in intergenic and intronic regions and showed a clear enrichment at promoters and 5' UTRs (Figure 7A,B and Suppl. Figure 11A-B). Motif analysis showed that 71% of the binding sites harbor a PDX1 consensus sequence, confirming the high quality of our data (Figure 7C). Moreover, ChIP-seq of the active histone modification H3K27ac in control [26] and patient-derived PPs showed a ~5 fold enrichment of H3K27ac at the PDX1-bound sites (Figure 7D and Suppl. Figure 11C). Compared to the ChIP-seq data from control PP
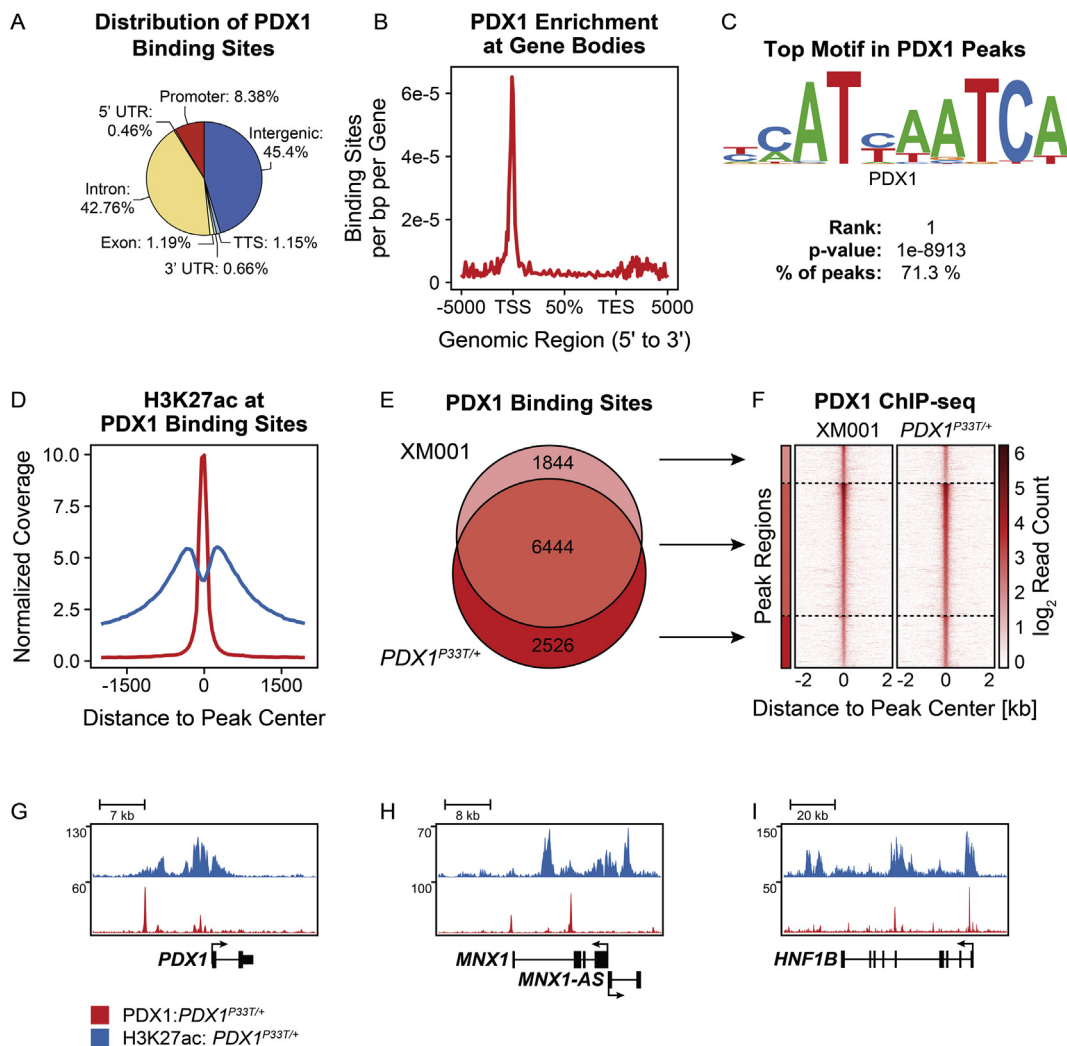

Figure 7: Characterization of PDX1 binding in patient-derived *PDX1^{P33T/+}* pancreatic progenitors (PP1). (A) Distribution of PDX1 binding sites among genomic features. PDX1^{P33T/+} binds predominantly to intergenic, intronic and promoter regions. (B) Meta-genomic plot showing the enrichment of PDX1 at the transcriptional start sites (TSS) of its target genes displayed as binding sites per base pair (bp) per gene over the genomic regions of all RefSeq genes. (C) Most enriched motif detected by motif analysis resembles the known PDX1 consensus sequence and is identified in 71.3% of all PDX1-bound sequences. (D) Average ChIP-seq Signal of H3K27ac (blue) and PDX1 (red) at PDX1 binding sites shows enrichment of H3K27ac at PDX1-bound sites. (E) Venn diagram showing the overlap of PDX1-binding sites in XM001 PPs and *PDX1^{P33T/+}* PPs. (F) Heatmap of PDX1 ChIP-seq signal at all PDX1 binding sites in XM001 and *PDX1^{P33T/+}* PPs, showing high resemblance of PDX1 binding in these cells. (G) ChIP-seq data tracks showing the enrichment of H3K27ac (blue) and PDX1 (red) at the loci of important pancreatic genes.

cells [26], we did not find pronounced differences in PDX1 binding or H3K27ac enrichment in cells derived from the patient compared to the control (Figure 7E,F and Suppl. Figure 11D-K). Of the 8970 PDX1 binding sites in the patient, 6444 are shared with the control PPs (Figure 7E). Next, we mapped PDX1 binding sites, shared between the control and $PDX1^{P33T/+}$ PPs, to genes and identified 3978 potential target genes. Among those, we found important pancreatic genes such as $PDX1$, $MNX1$, or $HNF1B$ (Figure 7G—I).

3.8. mRNA profiles of patient-derived $PDX1^{P33T/+}$ pancreatic progenitors (PP1)

Furthermore, microarray analysis of control [26] and patient-derived $PDX1^{P33T/+}$ cells was performed at the pluripotency and PP stages. A total of 2370 differentially expressed genes were identified between

iPSC and PP (FC $\geq$ 2, FDR $\leq$ 5%). Important pancreatic genes including $HNF1B$, $SOX9$, $ONECUT1$, and $HHEX$ were found to be upregulated in PPs. In contrast, pluripotency associated genes such as $NANOG$ and $LIN28A$ were specifically expressed in the iPSC stage (Figure 8A). Functional analysis of deregulated genes by means of GO term and pathway enrichment showed that genes upregulated in PPs were associated with pancreas development while downregulated genes were linked to embryonic stem cell, cell cycle and proliferation (Figure 8B). Comparing the transcriptional profile of PPs from $PDX1^{P33T/+}$ with controls, we identified 88 deregulated genes (FC $\geq$ 2, FDR $\leq$ 5%), among which 21 genes possessed PDX1 binding sites (Figure 8C). Remarkably, we identified downregulation of $MEG3$ and $NNAT$ genes, which have PDX1 binding sites and are involved in pancreas development and insulin secretion. Importantly, we also
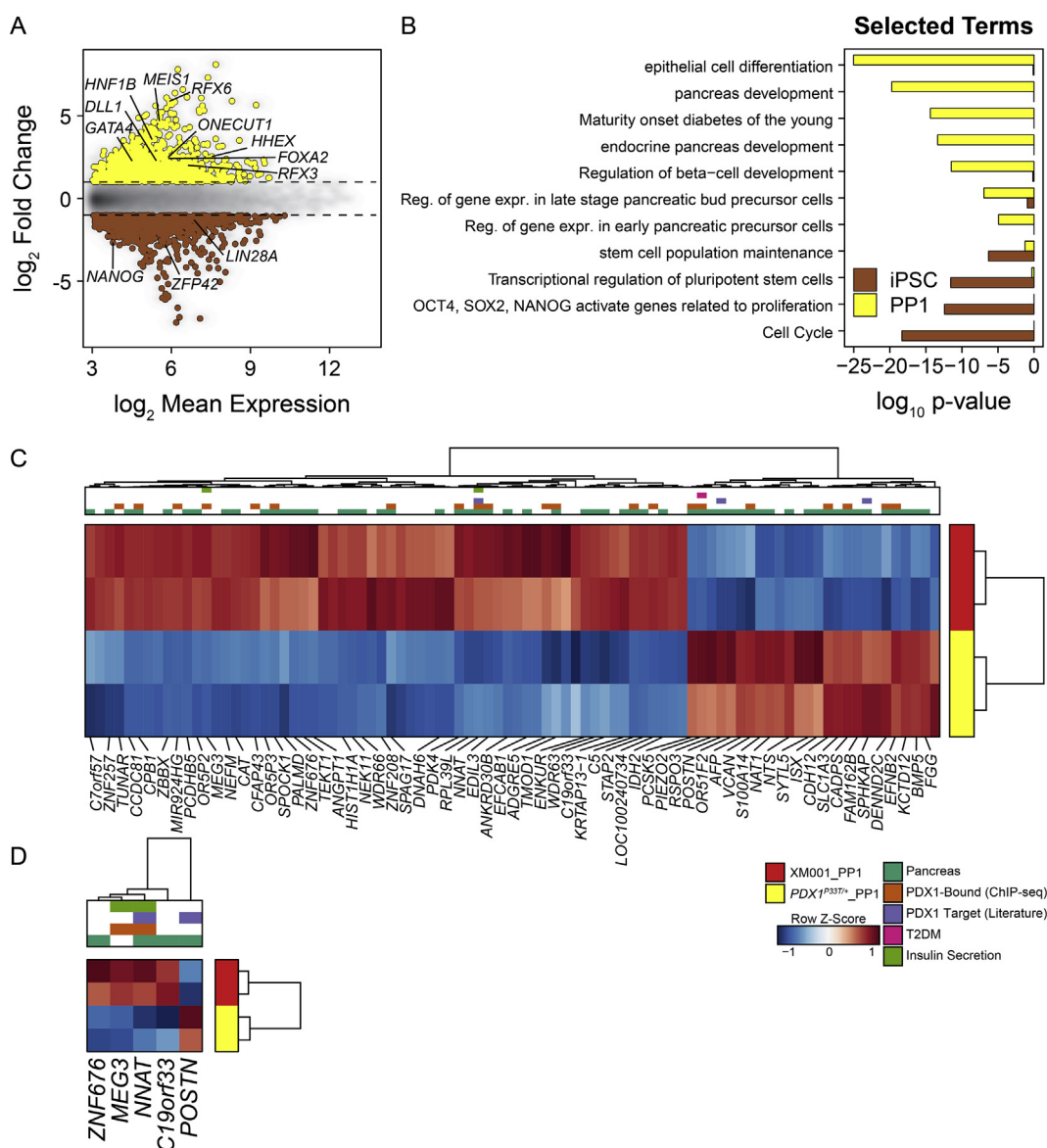

Figure 8: mRNA profiles of patient-derived $PDX1^{P33T/+}$ pancreatic progenitors (PP1). (A) MA plot showing the mean log₂ expression against the log₂-fold change of the microarray data obtained from XM001 iPSCs and PPs. Genes with significantly different expression (log₂ fold change ≥ 1 and adjusted p-value ≤ 0.1) are drawn in color. Yellow depicts increased expression in PPs, whereas brown depicts increased expression in iPSCs. (B) Bar graph of p-values from pathway enrichment analysis, showing selected GO terms and KEGG and Reactome pathways from differentially expressed genes. (C) Heatmap of differentially expressed genes between patient-derived $PDX1^{P33T/+}$ and XM001 control PPs. (D) Heat map and clustering showing consistently deregulated genes in patients $PDX1^{P33T/+}$ and isogenic $PDX1^{P33T/P33T}$ PPs.

identified a set of genes showing consistent deregulation between $PDX1^{P33T/P33T}$ PPs and patient derived $PDX1^{P33T/+}$ PPs that include *NNAT*, *MEG3*, and *POSTN* (Figure 8D). Intriguingly, Periostin, which is encoded by the *POSTN* gene, is involved in pancreas and β-cell regeneration [37], and both *NNAT* and *MEG3* are associated with insulin secretion.

4. DISCUSSION

The translation of findings from successful preclinical reports from animal models to human often fail due to the evolutionary diversity in organ development and function among species. In this context, discovering novel treatments for diabetes mellitus necessitates comprehensive understanding of the mechanisms underlying human endocrine cell formation, function and failure. Due to the limited access to human primary samples and the unfeasibility of performing longitudinal analyses in humans, the establishment of novel modeling systems to study human pancreas homeostasis is essential [38]. Along these lines, *in vitro* differentiation of β-cells from stem cells provides a unique platform to study human endocrine cell development and to model human disease [36,39,40]. In particular, this system enables the study of defects in developmental programs, leading to an increased predisposition to diabetes, by mimicking the phenotypes of mutations in specific genes such as those associated with MODY. Uncovering the consequence of such mutations will aid to identify the carrier individuals and design personalized prevention and treatment approaches, as has been shown for sulfonylurea treatment for patients carrying *KCNJ11* mutation in the potassium channel [41].

Here, using iPSCs endocrine differentiation culture, we investigated the impact of two common heterologous mutations (P33T, C18R) in the *PDX1* gene on human β-cell development and function. Using iPSCs from glucose intolerant patients as well as newly generated isogenic cell lines, we found that these mutations impair endocrine progenitor and β-cell development. However, our data revealed $PDX1^{P33T/+}$ and $PDX1^{C18R/+}$ mutations in patient-derived cells do not affect early stages of pancreatic differentiation. In comparison, $PDX1^{P33T/P33T}$ impaired the induction of endocrine fate at the PP2 stage, suggesting a dose-dependent effect of the P33T mutation on the endocrine lineage decision. Importantly, the levels of PDX1 protein were reduced at the PP1 stage in cells carrying a homozygous P33T mutation. In mouse, it has been shown that Pdx1 can regulate its own expression through an auto-regulatory positive feedback mechanism [42–44]. As the P33T mutation lies in the transactivation domain of PDX1, it is tempting to speculate that this mutation weakens protein–protein interactions and impairs the recruitment of transcriptional co-activators required for the induction of target gene expression and *PDX1* itself. Interestingly, in the $PDX1^{P33T/P33T}$ mutant PP cells, the lncRNA *PLUTO* is upregulated. *PLUTO*, which is downregulated in T2D, regulates *PDX1* expression by promoting interactions between the *PDX1* promoter and an upstream enhancer cluster [45], suggesting that its upregulation in $PDX1^{P33T/P33T}$ mutant PPs is due to a compensatory mechanism aiming to increase *PDX1* expression. The reduction in PDX1 levels at PP1 stage coincided with the impaired PP2 differentiation. During endocrinogenesis, the maintenance of high levels of PDX1 is essential for induction and differentiation of endocrine progenitors [46]. Therefore, the reduced levels of PDX1 likely hamper endocrine fate specification in human $PDX1^{P33T/P33T}$ mutant progenitors, supporting an evolutionary conserved mechanism of PDX1 function during endocrine cell differentiation. Interestingly, $PDX1^{C18R/C18R}$ mutation reduced neither PDX1 levels at PP1 stage nor at PP2 cell development. This difference between the effects of the two mutations might be due to the more severe changes in protein

structure, induced by the replacement of proline with threonine in P33T mutation. Proline is the major amino acid generating tight turns in protein sequences and its replacement with any other amino acid might result in significant changes in transactivation domain structure and protein–protein interactions. Therefore, it is likely that changing proline to threonine in PDX1 reduces its interaction with co-binding transcription factors. This is also supported by the similar phenotype of cells carrying $PDX1^{P33T/P33T}$ mutation and those carrying $PDX1^{+/-}$, suggesting a significant loss-of function upon this point mutation. However, at least in the $PDX1^{P33T/P33T}$ and $PDX1^{+/-}$ isogenic cell lines, the reduced expression of PDX1 might cause binding patterns to change, since it is possible that the remaining amount of PDX1 is not sufficient to activate low affinity binding sites.

Gene expression analysis from PPs derived from patient and isogenic cells carrying $PDX1^{P33T/+}$ and $PDX1^{P33T/P33T}$, identified downregulation of several PDX1 target genes including *MEG3* and *NEURONATIN* (*NNAT*). Maternally expressed gene 3 (*Meg3*) is an imprinted gene, coding for a long noncoding RNA, that resides in the distal region of mouse chromosome 12, and its human homolog is located in the syntenic region on chromosome 14. In mouse, the downregulation of Meg3 affects insulin synthesis and secretion in pancreatic β-cells [47]. Furthermore, this gene is downregulated in islets from T2D, suggesting an essential role for β-cell development and/or function [48]. *NNAT* is an imprinted gene coding for the proteolipid *NEURONATIN*, is expressed in neuroendocrine and metabolic tissues in a hormone- and glucose-sensitive manner. Rare single-nucleotide polymorphisms (SNPs) at the human *NNAT* locus are associated with extreme childhood obesity [49]. Mice lacking *Nnat* specifically in β-cells, exhibit reduced insulin content and impaired GSIS, demonstrating its involvement in regulation of glucose homeostasis and insulin secretion [50]. However, a function of this protein in pancreatic developmental has not been reported so far; therefore, it will be exciting to explore whether *NNAT* influences human β-cell development. Moreover, PPs derived from all isogenic mutants exhibited downregulation of common genes, i.e. *MEG3*, *LARGE1*, *ANPEP*, and *CES1*. It has been shown that mutations in *LARGE1* might increase the risk of type 2 diabetes through accumulation of visceral fat and increased insulin resistance [51]. In the *ANPEP* locus, SNPs influencing the expression of the *ANPEP* gene, coding for the aminopeptidase N, are associated with an increased risk of T2D [52]. Furthermore, global $Ces1^{-/-}$ mice have shown elevated plasma levels of triglycerides, free cholesterol, FFAs, and insulin, and displayed insulin resistance [53]. However, whether *LARGE1*, *ANPEP*, and *CES1* play a role during human β-cell development and/or function needs further investigation. In addition, we detected the downregulation of *MNX1* in $PDX1^{P33T/P33T}$ PPs. In mouse, this TF promotes and maintains the β-cell program during embryonic and postnatal stages, respectively, thereby regulating endocrine lineage allocation and β-cell fate maintenance [54]. Although homozygous *MNX1* mutation causes permanent neonatal diabetes in humans [55], the mechanisms by which this protein impacts human β-cell development still need to be addressed.

Both $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations deteriorated β-cell development and function. We found the downregulation of several key TFs regulating β-cell differentiation such as *NEUROD1* and *ISLI* upon homozygous single base-pair mutations. These data indicate that PDX1 fails to activate these target genes upon point mutations and further highlight the essential role of this TF for human β-cell development. Of note, we found significant reduction in the expression levels of *UCN3* in cells expressing homozygous PDX1 mutants, underlining the importance of proper PDX1 function for β-cell maturation, similar to rodents. Recently, we have shown that impairment in Pdx1 TF activity in double

knock-in reporter mice reduces β-cell maturation in male mice and leads to diabetes, whereas female animals are protected until pregnancy when they develop gestational diabetes [56,57]. Due to the report on gestational diabetes in the $PDX1^{P33T/+}$ patient and the reduction in $UCN3$ levels in β-like cells carrying this mutation, it is likely that β-cells from this patient function sub-optimally, which becomes evident upon metabolic demand similar to the above described mouse model. These results suggest an evolutionary conserved function for PDX1 in regulating β-cell maturation and function.

The differentiated β-cells carrying $PDX1$ point mutations were less functional compared to WT cells in term of glucose responsiveness. In mouse, Pdx1 is known to regulate GSIS. This TF not only directly regulates insulin synthesis, but also induces the expression of secretory machinery components required for insulin release [11]. We showed the downregulation of INS, $ABCC8$, $KCNJ11$, and $SLC30A8$ upon $PDX1^{P33T/P33T}$ and $PDX1^{C18R/C18R}$ mutations in β-like cells, highlighting the regulatory function of this protein for insulin synthesis and secretion in human β-cells. This finding is supported by previous studies, reporting decreased binding activity of heterozygous P33T and C18R mutant PDX1 protein to human insulin promotor using electrophoretic mobility shift assay [15,16].

In summary, we report the impact of two common point mutations in the transactivation domain of PDX1 on human β-cell development and function and show that these mutations predispose to diabetes. Characterizing the impact of such mutations will expand our understanding of disease heterogeneity in the general population and might allow for diagnosing people at risk and predicting their susceptibility to diabetes or gestational diabetes [1]. Our previous study has revealed that the reduced level of Pdx1 promotes development of diabetes and gestational diabetes in male and female mice, respectively [56]. Here, we suggest a similar function of PDX1 in humans and propose that developmental and compensary β-cell formation and expansion depends on a fully functional PDX1 TF.

AUTHOR CONTRIBUTION

X.W. performed and analyzed experiments. M.S. performed ChIP-seq experiments and computational analyses. A.B. helped to establish the FACS protocol. J.B. and J.S. helped to generate mutant lines. T.M. performed karyotyping. H.S. and H.-U.H. provided cells for iPSC generation. F.M.C. and G.S. provided infrastructure for ChIP-seq experiments and contributed to ChIP-seq analysis. M.I. and J.B. performed microarray and initial analysis of microarray data. C.V.E.W. provided the goat anti PDX1 antibody, used for ChIP experiments. A, I.B. and H.S. reviewed and edited the manuscript. X.W., M.S. and M.B. wrote the manuscript. H.L. wrote the manuscript and conceived the work. H.L. is the guarantor of this work.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

APPENDIX A. SUPPLEMENTARY DATA

REFERENCES

[1] Ahlqvist, E., Storm, P., Käräjämäki, A., Martinell, M., Dorkhan, M., Carlsson, A., et al., 2018. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. The Lancet Diabetes and Endocrinology 6(5):361—369. https://doi.org/10.1016/S2213-8587(18)30051-2.

[2] Diagnosis and classification of diabetes mellitus. Diabetes Care 37(Supplement_1), 2014:S81—S90. https://doi.org/10.2337/dc14-S081.

[3] Hattersley, A.T., Patel, K.A., 2017. Precision diabetes: learning from monogenic diabetes. Diabetologia, 769—777. https://doi.org/10.1007/s00125-017-4226-2.

[4] Staffers, D.A., Ferrer, J., Clarke, W.L., Habener, J.F., 1997. Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. Nature Genetics 17(2):138—139. https://doi.org/10.1038/ng1097-138.

[5] Heuvel-Borsboom, H., de Valk, H.W., Losekoot, M., Westerink, J., 2016. Maturity onset diabetes of the young: seek and you will find. The Netherlands Journal of Medicine, 193—200.

[6] 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 (Cambridge, England) 122(5):1409—1416. https://doi.org/10.1016/0092-8674(88)90391-1.

[7] Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C.V., et al., 1995. Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. Development (Cambridge, England) 121(1):11—18.

[8] Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., et al., 1996. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development (Cambridge, England) 122(3):983—995. https://doi.org/10.1016/0076-6879(93)25031-v.

[9] Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., Edlund, H., 1998. beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. Genes & Development 12(12):1763—1768. https://doi.org/10.1101/gad.12.12.1763.

[10] Jonsson, J., Carlsson, L., Edlund, T., Edlund, H., 1994. Insulin-promoter-factor 1 is required for pancreas development in mice. Nature, 606—609. https://doi.org/10.1038/371606a0.

[11] Brissova, M., Shiota, M., Nicholson, W.E., Gannon, M., Knobel, S.M., Piston, D.W., et al., 2002. Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion. Journal of Biological Chemistry 277(13):11225—11232. https://doi.org/10.1074/jbc.M111272200.

[12] Johnson, J.D., Ahmed, N.T., Luciani, D.S., Han, Z., Tran, H., Fujita, J., et al., 2003. Increased islet apoptosis in Pdx1+/- mice. Journal of Clinical Investigation 111(8):1147—1160. https://doi.org/10.1172/JCl200316537.

[13] Holland, A.M., Góñez, L.J., Naselli, G., MacDonald, R.J., Harrison, L.C., 2005. Conditional expression demonstrates the role of the homeodomain transcription factor Pdx1 in maintenance and regeneration of beta-cells in the adult pancreas. Diabetes 54(9):2586—2595. https://doi.org/10.2337/diabetes.54.9.2586.

[14] Stoffers, D.A., Stanojevic, V., Habener, J.F., 1998. Insulin promoter factor-1 gene mutation linked to early-onset type 2 diabetes mellitus directs expression of a dominant negative isoprotein. Journal of Clinical Investigation 102(1): 232–241. https://doi.org/10.1172/JCI2242.

[15] Gragnoli, C., Stanojevic, V., Gorini, A., Von Preussenthal, G.M., Thomas, M.K., Habener, J.F., 2005. IPF-1/MODY4 gene missense mutation in an Italian family with type 2 and gestational diabetes. Metabolism 54(8):983–988. https://doi.org/10.1016/j.metabol.2005.01.037.

[16] Macfarlane, W.M., Macfarlane, W.M., Frayling, T.M., Frayling, T.M., Ellard, S., Ellard, S., et al., 1999. Missense mutations in the insulin promoter factor-1 gene predispose to type 2 diabetes. Journal of Clinical Investigation 104. https://doi.org/10.1172/jci7449. R33-NaN-39.

[17] Fuchsberger, C., Flannick, J., Teslovich, T.M., Mahajan, A., Agarwala, V., Gaulton, K.J., et al., 2016. The genetic architecture of type 2 diabetes. Nature 536(7614):41–47. https://doi.org/10.1038/nature18642.

[18] Zhu, Z., Li, Q.V., Lee, K., Rosen, B.P., González, F., Soh, C.L., et al., 2016. Genome editing of lineage determinants in human pluripotent stem cells reveals mechanisms of pancreatic development and diabetes. Cell Stem Cell 18(6):755–768. https://doi.org/10.1016/j.stem.2016.03.015.

[19] Teo, A.K.K., Lau, H.H., Valdez, I.A., Dirice, E., Tjora, E., Raeder, H., et al., 2016. Early developmental perturbations in a human stem cell model of MODY5/HNF1B pancreatic hypoplasia. Stem Cell Reports 6(3):357–367. https://doi.org/10.1016/j.stemcr.2016.01.007.

[20] Shi, Z.D., Lee, K., Yang, D., Amin, S., Verma, N., Li, Q.V., et al., 2017. Genome editing in hPSCs reveals GATA6 haploinsufficiency and a genetic interaction with GATA4 in human pancreatic development. Cell Stem Cell 20(5):675–688. https://doi.org/10.1016/j.stem.2017.01.001 e6.

[21] Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 107(5):861–872. https://doi.org/10.1016/j.cell.2007.11.019.

[22] Li, W., Wang, X., Fan, W., Zhao, P., Chan, Y.C., Chen, S., et al., 2012. Modeling abnormal early development with induced pluripotent stem cells from aneuploid syndromes. Human Molecular Genetics 21(1):32–45. https://doi.org/10.1093/hmg/ddr435.

[23] Millman, J.R., Xie, C., Van Dervort, A., Gürtler, M., Pagliuca, F.W., Melton, D.A., 2016. Generation of stem cell-derived β-cells from patients with type 1 diabetes. Nature Communications 7. https://doi.org/10.1038/ncomms11463.

[24] Shang, L., Hua, H., Foo, K., Martinez, H., Watanabe, K., Zimmer, M., et al., 2014. β-cell dysfunction due to increased ER stress in a stem cell model of wolfram syndrome. Diabetes 63(3):923–933. https://doi.org/10.2337/db13-0717.

[25] Zeng, H., Guo, M., Zhou, T., Tan, L., Chong, C.N., Zhang, T., et al., 2016. An isogenic human ESC platform for functional evaluation of genome-wide-association-study-identified diabetes genes and drug discovery. Cell Stem Cell 19(3):326–340. https://doi.org/10.1016/j.stem.2016.07.002.

[26] Wang, X., Sterr, M., Burtscher, I., Chen, S., Hieronimus, A., Machicao, F., et al., 2018. Genome-wide analysis of PDX1 target genes in human pancreatic progenitors. Molecular Metabolism 9:57–68. https://doi.org/10.1016/j.molmet.2018.01.011.

[27] Yumlu, S., Stumm, J., Bashir, S., Dreyer, A.-K., Lisowski, P., Danner, E., et al., 2017. Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9. Methods 121–122:29–44. https://doi.org/10.1016/j.ymeth.2017.05.009.

[28] Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., et al., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular Cell 38(4):576–589. https://doi.org/10.1016/j.molcel.2010.05.004.

[29] Denny, S.K., Yang, D., Chuang, C.H., Brady, J.J., Lim, J.S.S., Grüner, B.M., et al., 2016. Nfib promotes metastasis through a widespread increase in chromatin accessibility. Cell 166(2):328–342. https://doi.org/10.1016/j.cell.2016.05.052.

[30] Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15(12). https://doi.org/10.1186/s13059-014-0550-8.

[31] Ignatiadis, N., Klaus, B., Zaugg, J.B., Huber, W., 2016. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nature Methods 13(7):577–580. https://doi.org/10.1038/nmeth.3885.

[32] Zhu, A., Ibrahim, J.G., Love, M.I., 2018. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics. https://doi.org/10.1093/bioinformatics/bty895.

[33] Stefan, N., Machicao, F., Staiger, H., Machann, J., Schick, F., Tschritter, O., et al., 2005. Polymorphisms in the gene encoding adiponectin receptor 1 are associated with insulin resistance and high liver fat. Diabetologia 48(11): 2282–2291. https://doi.org/10.1007/s00125-005-1948-3.

[34] Wang, X., Chen, S., Burtscher, I., Sterr, M., Hieronimus, A., Machicao, F., et al., 2016. Generation of a human induced pluripotent stem cell (iPSC) line from a patient carrying a P33T mutation in the PDX1 gene. Stem Cell Research 17(2):273–276. https://doi.org/10.1016/j.scr.2016.08.004.

[35] Wang, X., Chen, S., Burtscher, I., Sterr, M., Hieronimus, A., Machicao, F., et al., 2016. Generation of a human induced pluripotent stem cell (iPSC) line from a patient with family history of diabetes carrying a C18R mutation in the PDX1 gene. Stem Cell Research 17(2):292–295. https://doi.org/10.1016/j.scr.2016.08.005.

[36] Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., et al., 2014. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nature Biotechnology 32(11):1121–1133. https://doi.org/10.1038/nbt.3033.

[37] Smid, J.K., Faulkes, S., Rudnicki, M.A., 2015. Periostin induces pancreatic regeneration. Endocrinology 156(3):824–836. https://doi.org/10.1210/en.2014-1637.

[38] Bakhti, M., Böttcher, A., Lickert, H., 2018. Modelling the endocrine pancreas in health and disease. Nature Reviews Endocrinology. https://doi.org/10.1038/s41574-018-0132-z.

[39] Russ, H. a., Parent, A.V., Ringler, J.J., Hennings, T.G., Nair, G.G., Shveygert, M., et al., 2015. Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. The EMBO Journal 34(13): e201591058. https://doi.org/10.15252/embj.201591058.

[40] Pagliuca, F.W., Millman, J.R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J.H., et al., 2014. Generation of functional human pancreatic β cells in vitro. Cell 159(2):428–439. https://doi.org/10.1016/j.cell.2014.09.040.

[41] Hattersley, A.T., Ashcroft, F.M., 2005. Activating mutations in Kir6.2 and neonatal diabetes: new clinical syndromes, new scientific insights, and new therapy. Diabetes, 2503–2513. https://doi.org/10.2337/diabetes.54.9.2503.

[42] Marshak, S., Benshushan, E., Shoshkes, M., Havin, L., Cerasi, E., Melloul, D., 2000. Functional conservation of regulatory elements in the pdx-1 gene: PDX-1 and hepatocyte nuclear factor 3beta transcription factors mediate beta-cell-specific expression. Molecular and Cellular Biology 20(20):7583–7590. https://doi.org/10.1128/MCB.20.20.7583-7590.2000.

[43] Gerrish, K., Cissell, M.A., Stein, R., 2001. The role of hepatic nuclear factor 1α and PDX-1 in transcriptional regulation of the pdx-1 gene. Journal of Biological Chemistry 276(51):47775–47784. https://doi.org/10.1074/jbc.M109244200.

[44] Gerrish, K., Van Velkinburgh, J.C., Stein, R., 2004. Conserved transcriptional regulatory domains of the pdx-1 gene. Molecular Endocrinology (Baltimore, Md.) 18(3):533–548. https://doi.org/10.1210/me.2003-0371.

[45] Akerman, I., Tu, Z., Beucher, A., Rolando, D.M.Y., Sauty-Colace, C., Benazra, M., et al., 2017. Human pancreatic β cell lncRNAs control cell-specific regulatory networks. Cell Metabolism 25(2):400–411. https://doi.org/10.1016/j.cmet.2016.11.016.

[46] Mamidi, A., Prawiro, C., Seymour, P.A., de Lichtenberg, K.H., Jackson, A., Serup, P., et al., 2018. Mechanosignalling via integrins directs fate decisions

of pancreatic progenitors. Nature 564(7734):114—118. https://doi.org/10.1038/s41586-018-0762-2.

[47] You, L., Wang, N., Yin, D., Wang, L., Jin, F., Zhu, Y., et al., 2016. Down-regulation of long noncoding RNA Meg3 affects insulin synthesis and secretion in mouse pancreatic beta cells. Journal of Cellular Physiology 231(4):852—862. https://doi.org/10.1002/jcp.25175.

[48] Kameswaran, V., Bramswig, N.C., McKenna, L.B., Penn, M., Schug, J., Hand, N.J., et al., 2014. Epigenetic regulation of the DLK1-MEG3 MicroRNA cluster in human type 2 diabetic islets. Cell Metabolism 19(1):135—145. https://doi.org/10.1016/j.cmet.2013.11.016.

[49] Vrang, N., Meyre, D., Froguel, P., Jelsing, J., Tang-Christensen, M., Vatin, V., et al., 2010. The imprinted gene neuronatin is regulated by metabolic status and associated with obesity. Obesity 18(7):1289—1296. https://doi.org/10.1038/oby.2009.361.

[50] Millership, S.J., Da Silva Xavier, G., Choudhury, A.I., Bertazzo, S., Chabosseau, P., Pedroni, S.M.A., et al., 2018. Neuronatin regulates pancreatic β cell insulin content and secretion. Journal of Clinical Investigation 128(8):3369—3381. https://doi.org/10.1172/JCI120115.

[51] Grarup, N., Moltke, I., Andersen, M.K., Bjerregaard, P., Larsen, C.V.L., Dahl-Petersen, I.K., et al., 2018. Identification of novel high-impact recessively inherited type 2 diabetes risk variants in the Greenlandic population. Diabetologia 61(9):2005—2015. https://doi.org/10.1007/s00125-018-4659-2.

[52] Locke, J.M., Hysenaj, G., Wood, G., Weedon, M.N., Harries, L.W., 2015 Apr. Targeted allelic expression profiling in human islets identifies cis-regulatory effects for multiple variants identified by type 2 diabetes genome-wide association studies. Diabetes 64(4):1484—1491. https://doi.org/10.2337/db14-0957.

[53] Quiroga, A.D., Li, L., Trötzmüller, M., Nelson, R., Proctor, S.D., Köfeler, H., et al., 2012. Deficiency of carboxylesterase 1/esterase-x results in obesity, hepatic steatosis, and hyperlipidemia. Hepatology (Baltimore, Md.) 56(6):2188—2198. https://doi.org/10.1002/hep.25961.

[54] Pan, F.C., Brissova, M., Powers, A.C., Pfaff, S., Wright, C.V.E., 2015. Inactivating the permanent neonatal diabetes gene Mnx1 switches insulin-producing β-cells to a δ-like fate and reveals a facultative proliferative capacity in aged β-cells. Development 142(21):3637—3648. https://doi.org/10.1242/dev.126011.

[55] Flanagan, S.E., De Franco, E., Lango Allen, H., Zerah, M., Abdul-Rasoul, M.M., Edge, J.A., et al., 2014. Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man. Cell Metabolism 19(1):146—154. https://doi.org/10.1016/j.cmet.2013.11.021.

[56] Bastidas-Ponce, A., Roscioni, S.S., Burtscher, I., Bader, E., Sterr, M., Bakhti, M., et al., 2017. Foxa2 and Pdx1 cooperatively regulate postnatal maturation of pancreatic β-cells. Molecular Metabolism 6(6):524—534. https://doi.org/10.1016/j.molmet.2017.03.007.

[57] Kleinert, M., Clemmensen, C., Hofmann, S.M., Moore, M.C., Renner, S., Woods, S.C., et al., 2018. Animal models of obesity and diabetes mellitus. Nature Reviews Endocrinology, 140—162. https://doi.org/10.1038/nrendo.2017.161.

Lab resource: Stem Cell Line

Generation of a human induced pluripotent stem cell (iPSC) line from a patient with family history of diabetes carrying a C18R mutation in the *PDX1* gene

CrossMark

Xianming Wang [a,b,i], Shen Chen [c], Ingo Burtscher [a,b], Michael Sterr [a,b], Anja Hieronimus [d,e,j], Fausto Machicao [f,j], Harald Staiger [d,g,j], Hans-Ulrich Häring [d,e,j], Gabriele Lederer [k], Thomas Meitinger [h,k], Heiko Lickert [a,b,i,j,∗]

[a] Institute of Diabetes and Regeneration Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany
[b] Institute of Stem Cell Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany
[c] iPS and Cancer Research Unit, Department of Histology and Embryology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China
[d] Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Zentrum München at the University of Tübingen, 72076 Tübingen, Germany
[e] Department of Internal Medicine, Division of Endocrinology, Diabetology, Vascular Disease, Nephrology and Clinical Chemistry, University of Tübingen, 72076 Tübingen, Germany
[f] Institute of Experimental Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany
[g] Institute of Pharmaceutical Sciences, Department of Pharmacy and Biochemistry, Eberhard Karls University Tübingen, 72076 Tübingen, Germany
[h] Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany
[i] Technische Universität München, Ismaninger straße 22, 81675 München, Germany
[j] German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany
[k] Institute of Human Genetics, Technische Universität München, 81675 München, Germany

ARTICLE INFO

ABSTRACT

Homozygous loss-of-function mutations in the gene coding for the homeobox transcription factor *PDX1* leads to pancreatic agenesis, whereas certain heterozygous point mutations are associated with Maturity-Onset Diabetes of the Young 4 (MODY4) and Type 2 Diabetes Mellitus (T2DM). To understand the pathomechanism of MODY4 and T2DM, we have generated iPSCs from a woman with a C18R heterozygous mutation in the transactivation domain of PDX1. The resulting *PDX1* C18R iPSCs generated by episomal reprogramming are integration-free, have a normal karyotype and are pluripotent *in vitro* and *in vivo*. Taken together, this iPSC line will be useful to study diabetes pathomechanisms.

Resource table

Name of stem cell line	PDX1 C18R iPSC1
Institution	Institute of Diabetes and Regeneration Research
Person who created resource	Xianming Wang
Contact person and email	Heiko Lickert heiko.lickert@helmholtz-muenchen.de
Date archived/stock date	April 2014
Origin	Human dermal fibroblasts
Type of resource	Induced pluripotent stem cells from a woman carrying a PDX1 C18R mutation
Sub-type	Cell line
Key transcription factors	OCT4, SOX2, NANOG, LIN28, KLF4, and L-MYC

(continued)

Name of stem cell line	PDX1 C18R iPSC1
Authentication	Identity and purity of cell line confirmed
Link to related literature	Not available
Information in public databases	Not available
Ethics	Informed written consent obtained

Resource details

The Tübingen Family Study for T2DM (TÜF study, $N = 2500$) was screened for known rare mutations in the MODY4 gene *PDX1*. By mass spectrometry-based genotyping, one heterozygous *PDX1* C18R carrier could be identified and recruited for full-thickness skin biopsy. From the biopsy material, the epidermal layer was separated from the dermis by dispase digestion and fibroblasts were isolated from the dermis by trypsin digestion using an established protocol (see the Materials and

∗ Corresponding author at: Helmholtz Zentrum München, Parkring 11, D-85748 Garching, Germany.
  E-mail address: heiko.lickert@helmholtz-muenchen.de (H. Lickert).

methods section). The *PDX1* C18R fibroblasts were expanded using commercially available growth media and deep-frozen in liquid nitrogen. From this collection of dermal fibroblasts, samples were tested for the presence of viruses pathogenic to humans (HBV, HCV, HIV) as well as for the presence of mycoplasma. All cultures were found to be negative for all of these potential contaminants. *PDX1* C18R primary fibroblasts were reprogrammed into iPSCs using nucleofection with three episomal plasmids encoding human *OCT4, SOX2, NANOG, LIN28,*
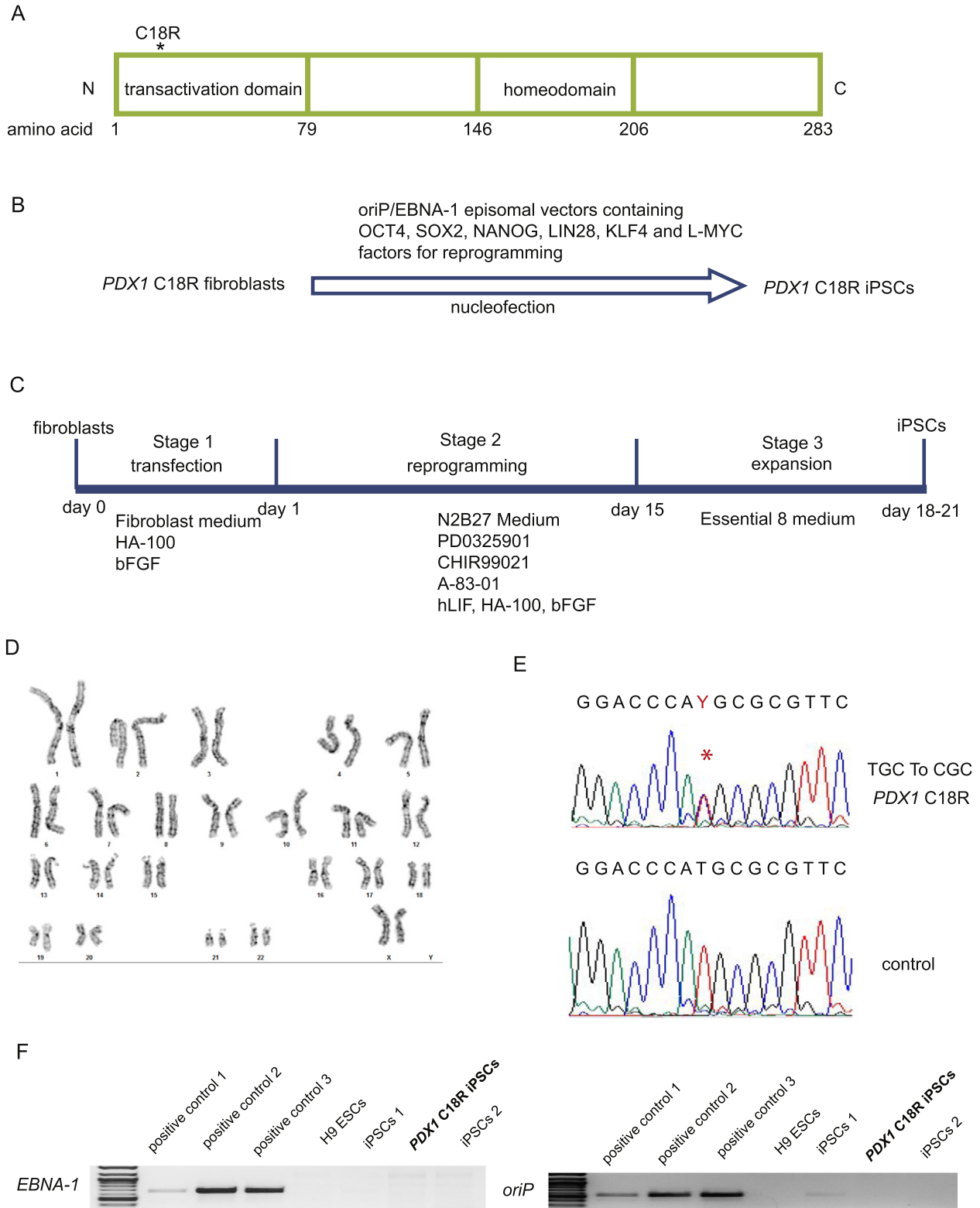
Fig. 1. Generation of *PDX1* C18R iPSCs. (A) An overview of the PDX1 protein structure. (B) Scheme shows the reprogramming factors for the generation of the *PDX1* C18R iPSC line. (C) Reprogramming protocol of skin fibroblasts into integration-free iPSCs using episomal vectors. (D) Normal karyotype (46, XX) of one *PDX1* C18R iPSC clone. (E) Sequencing result shows heterozygous T > C mutation in *PDX1*. (F) Semi-quantitative PCR confirms that episomal vectors did not integrate into the genomic DNA of the selected *PDX1* C18R iPSC clone. H9 ESCs and two independent iPSC clones (#1 and #2) were used as negative controls, whereas transfected fibroblast at day 6 were used as positive controls.

KLF4, and *L-MYC* (Fig. 1B). The reprogramming protocol (transfection, reprogramming and expansion) is shown in Fig. 1C. 3 weeks after transfection, many colonies with human embryonic stem cell (hESC)-like morphology appeared that were hand picked by day 21, expanded and frozen down as individual iPSC lines. The analysis of one of these iPSC lines confirmed that *PDX1* C18R fibroblasts were successfully reprogrammed into iPSCs with a normal karyotype (Fig. 1D) and containing the T > C mutation causing a C18R amino acid exchange in the transactivation domain (Fig. 1A, E). Integration of the exogenous transgenes in the genome of iPSCs was excluded by semi-quantitative PCR (Fig. 1F). To confirm that this iPSC line was pluripotent, we performed immunofluorescence staining for the pluripotency transcription factors OCT4 and SOX2, as well as for the hESC surface markers SSEA-3, SSEA-4 and TRA-1-81, which were all expressed in the *PDX1* C18R iPSC line (Fig. 2A). To further demonstrate pluripotency of the selected iPSC line we injected 2×10^6 cells subcutaneously into NOD/SCID mice to generate teratomas. The histological analysis of the induced teratomas revealed that all tissues from all three germ layers were differentiated confirming pluripotency in this *in vivo* assay (Fig. 2B). Taken together, we have successfully reprogrammed *PDX1* C18R dermal fibroblasts into pluripotent iPSCs that can be used to study a point mutation in the transactivation domain of PDX1 to study the development of diabetes.

Materials and methods

Ethics statement

The study adhered to the Declaration of Helsinki. All participants gave informed written consent, and the study protocols were approved by the local ethics board (Ethics Committee of the Eberhard Karls University Tübingen).

Study participants

The ongoing TÜF study currently includes >3000 non-related individuals at risk for type-2 diabetes mellitus (T2DM), *i.e.*, healthy subjects with a family history of T2DM, a body mass index ≥27 kg/m², impaired fasting glycaemia, and/or previous gestational diabetes (Stefan et al., 2005). From all TÜF participants, the medical history, smoking status, and alcohol consumption habits were queried and documented. Furthermore, all participants underwent physical examination, routine blood tests, bioelectric impedance measurement, and 5-point oral glucose tolerance tests (OGTTs) with insulin and glucose measurements. The current study population comprised 2500 individuals who spent DNA for genotyping, were not on any medication known to influence glucose tolerance, insulin sensitivity, or insulin secretion, and who had complete OGTT data sets.
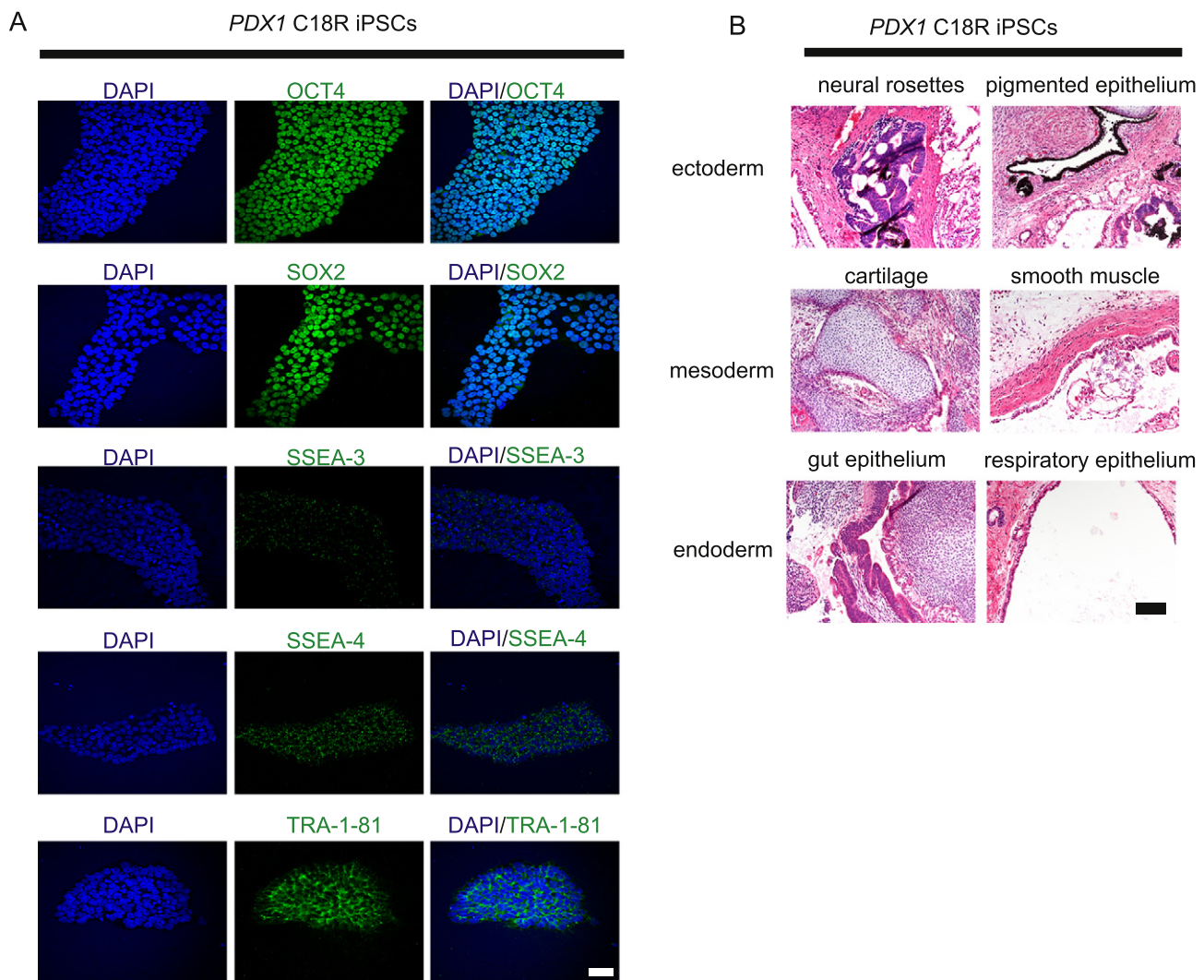

Fig. 2. *PDX1* C18R iPSCs are pluripotent and differentiate into all germ layers *in vivo*. (A) *PDX1* C18R iPSCs display typical morphological ESC-like characteristics and uniformly express several pluripotency markers. Scale bar indicates 50 μm. (B) Hematoxylin/eosin staining of tissue sections of teratomas generated from the *PDX1* C18R iPSC clone. All germ layer-derivatives are clearly detectable *in vivo*. Scale bar indicates 100 μm.

Using the mass spectrometry-based genotyping platform massARRAY from Sequenom (Hamburg, Germany) and the manufacturer's iPLEX software, the 2500 subjects were screened for known rare mutations in maturity onset of diabetes (MODY) genes that were described in the literature to co-segregate with MODY phenotypes in families. We identified seven individuals with heterozygous mutations in *PDX1* (MODY4) (McKinnon and Docherty, 2001). From these cases, one glucose-tolerant female *PDX1* Cys18Arg (C18R) carrier (aged 33 years; BMI 20.6 kg/m^2) with history of diabetes could be recruited for full-thickness skin biopsy.

Skin biopsy, isolation and testing of dermal fibroblasts

A full-thickness skin specimen was taken by punch biopsy from the upper arm in the deltoid muscle region. After removal of adipose tissue remnants and visible blood vessels, the sample was digested overnight at 4 °C with 10 U/ml dispase II (Roche Diagnostics, Mannheim, Germany) in 50 mM Hepes pH 7.4, 150 mM NaCl. Thereafter, the digest was heated for 30 min at 37 °C under continuous shaking (1200 rpm). Using forceps, the dermis was separated from the epidermal layer, and fibroblasts were isolated from the dermis by digestion with 0.2% collagenase CLS I (Biochrom, Berlin, Germany) in DMEM, 10% BSA for 45 min at 37 °C under continuous shaking (1200 rpm). For purification of the fibroblasts, the digest was filtered through a 70 μm mesh and centrifuged. The pelleted cells were resuspended, grown for three days in DMEM, 10% FCS, and subsequently further expanded in Medium 106 supplemented with low serum growth supplement (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Samples of dermal fibroblasts were tested for the presence of viruses pathogenic to humans, *i.e.*, HBV, HCV, and HIV, with Genesig PCR-based detection kits from Primerdesign Ltd. (Chandler's Ford, UK) and for the presence of mycoplasma with a PCR test kit from PanReac AppliChem (Darmstadt, Germany) and, in parallel, by DNA staining with DAPI. All cultures were found to be negative for the tested contaminants.

iPSC generation

Primary fibroblasts from patient were reprogrammed into pluripotent stem cells by using a non-integrating Episomal iPSC Reprogramming Kit (Invitrogen, Cat. no. A14703). This kit contains mixture of three vectors which has the oriP/EBNA-1 (Epstein-Barr unclear antigen-1) backbone that delivers six reprogramming factors: OCT4, SOX2, NANOG, LIN28, KLF4, and L-MYC (Yu et al., 2011). Human fibroblasts at 75–90% confluent were transfected using the Amaxa 4D-Nucleofector transfection system and a nucleofector kit for human dermal fibroblast (Lonza, Cat. no. VPD-1001), plated onto geltrex-coated culture dishes, incubated in supplemented fibroblast medium. This medium contained knockout DMEM/F-12 (Life technologies), 10% FBS of ESC-qualified (Life technologies), 1% MEM non-essential amino acids (Life technologies), 10 μM HA-100 (Santa Cruz) and 4 ng/ml bFGF (Life technologies). At 24 h after transfection, the medium were exchanged with N2B27 medium supplemented with 0.5 μM PD0325901 (Stemgent), 3 μM CHIR99021 (Stemgent), 0.5 μM A-83-01 (Stemgent), 10 μM HA-100 (Santa Cruz), 10 ng/ml hLIF (Life technologies) and 100 ng/ml bFGF. The basic N2B27 medium contained DMEM/F12 with HEPES (Life technologies), 1 × N2 supplement (Life technologies), 1 × B27 supplement (Life technologies), 1% MEM non-essential amino acids, 1 × Glutamax (Life technologies) and 1 × β-Mercaptoethanol (Life technologies). On day 15 after transfection, the medium was exchanged with Essential 8 medium and monitored for the emergence of iPSC colonies. Around 3 weeks after transfection, undifferentiated iPS colonies were picked and transferred onto fresh geltrex-coated culture dishes for expansion.

iPSC characterization

DNA was extracted from iPSCs by using standard procedure. Markers for the episomal backbone were amplified by semi-quantitative PCR to exclude transgene integration. Primers are as follows: *oriP* forward: TTCCACGAGGCTAGTGAACC. *oriP* reverse: TCGGGGGTGTTAGAGACAAC; *EBNA-1* forward: ATCGTCAAAGCTGCACACAG. *EBNA-1* reverse: CCCAGGAGTCCCAGTAGTCA. For karyotype analysis, we used the cells growing in logarithmic phase. They were fed with fresh medium the night before adding colcemid for 2 h. Cells were then trypsinized, treated with hypotonic solution (0.075 M KCl) for 20 min and fixed with methanol:acetic acid (3:1). Metaphases were spread on microscope slides, and chromosomes were classified according to the International System for Human Cytogenic Nomenclature using the standard G banding technique. At least, 20 metaphases were counted per cell line, and the final karyotype was stated if it was present in >85% of them. For teratomas, 2×10^6 iPSCs were injected into the right hind leg of immunocompromised NOD/SCID mice. Tumors were excised after 8 weeks, fixed, embedded in paraffin, sectioned and stained with hematoxylin/eosin (Takahashi et al., 2007).

Mutation analysis

Genomic DNA was isolated from *PDX1* C18R and control iPSCs using standard procedure. PCR amplification with a set of primers flanking the mutation site was performed in those two samples. PCR products were sequenced using the reverse primer by Sanger sequencing on an ABI 3130 genetic analyzer (Applied Biosystems). Primers are as follows: *PDX1* forward: GGAGTGTGCAGCAAACTCAG, *PDX1* reverse: ACGCGTGAGCTTTGGTAGAC.

Immunofluorescence imaging

Cells were fixed with 4% paraformaldehyde for 30 min and then permeabilized in PBS containing 0.2% Triton X-100. Cells were blocked with PBS containing 3% BSA, and incubated with primary antibodies overnight at 4 °C. Then secondary antibodies were incubated for 1 h at room temperature after washing with PBS. Images were acquired on a laser scanning microscope. The following antibodies and dilutions were used: goat anti-OCT3/4 (1:500, Santa Cruz), goat anti-SOX2 (1:500, Santa Cruz), rat anti-SSEA3 (1:50, Invitrogen), mouse anti-SSEA4 (1:500, cell signaling) and mouse anti-TRA-1-81 (1:50, Millipore).

Author disclosure statement

There are no competing financial interests in this study.

References

McKinnon, C.M., Docherty, K., 2001. Pancreatic duodenal homeobox-1, PDX-1, a major regulator of beta cell identity and function. Diabetologia 44, 1203–1214.
Stefan, N., Machicao, F., Staiger, H., Machann, J., Schick, F., Tschritter, O., Spieth, C., Weigert, C., Fritsche, A., Stumvoll, M., Häring, H.U., 2005. Polymorphisms in the gene encoding adiponectin receptor 1 are associated with insulin resistance and high liver fat. Diabetologia 48, 2282–2291.
Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861–872.
Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J., Jiang, Y., 2011. Efficient feeder-free episomal reprogramming with small molecules. PLoS One 6, e17557.
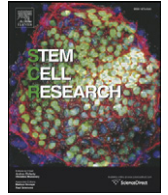
Lab resource: Stem Cell Line

Generation of a human induced pluripotent stem cell (iPSC) line from a patient carrying a P33T mutation in the *PDX1* gene

Xianming Wang [a,b,i], Shen Chen [c], Ingo Burtscher [a,b], Michael Sterr [a,b], Anja Hieronimus [d,e,j], Fausto Machicao [f,j], Harald Staiger [d,g,j], Hans-Ulrich Häring [d,e,j], Gabriele Lederer [k], Thomas Meitinger [h,k], Heiko Lickert [a,b,i,j,*]

[a] *Institute of Diabetes and Regeneration Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany*
[b] *Institute of Stem Cell Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany*
[c] *iPS and Cancer Research Unit, Department of Histology and Embryology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China*
[d] *Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Zentrum München at the University of Tübingen, 72076 Tübingen, Germany*
[e] *Department of Internal Medicine, Division of Endocrinology, Diabetology, Vascular Disease, Nephrology and Clinical Chemistry, University of Tübingen, 72076 Tübingen, Germany*
[f] *Institute of Experimental Genetics of the Helmholtz Zentrum München, 85764 Neuherberg, Germany*
[g] *Institute of Pharmaceutical Sciences, Department of Pharmacy and Biochemistry, Eberhard Karls University Tübingen, 72076 Tübingen, Germany*
[h] *Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany*
[i] *Technische Universität München, Ismaninger straße 22, 81675 München, Germany*
[j] *German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany*
[k] *Institute of Human Genetics, Technische Universität München, 81675 München, Germany*

ARTICLE INFO

ABSTRACT

Homozygous loss-of-function mutations in the gene coding for the homeobox transcription factor *PDX1* leads to pancreatic agenesis, whereas certain heterozygous point mutations are associated with Maturity-Onset Diabetes of the Young 4 (MODY4) and Type 2 Diabetes Mellitus (T2DM). To understand the pathomechanism of MODY4 and T2DM, we have generated iPSCs from a woman with a P33T heterozygous mutation in the transactivation domain of PDX1. The resulting *PDX1* P33T iPSCs generated by episomal reprogramming are integration-free, have a normal karyotype and are pluripotent *in vitro* and *in vivo*. Taken together, this iPSC line will be useful to study diabetes pathomechanisms.

1. Resource table

Name of stem cell line	*PDX1* P33T iPSC1
Institution	Institute of Diabetes and Regeneration Research
Person who created resource	Xianming Wang
Contact person and email	Heiko Lickert heiko.lickert@helmholtz-muenchen.de
Date archived/stock date	April 2014
Origin	Human dermal fibroblasts
Type of resource	Induced pluripotent stem cells from a woman carrying a *PDX1* P33T mutation
Sub-type	Cell line
Key transcription factors	OCT4, SOX2, NANOG, LIN28, KLF4, and L-MYC
Authentication	Identity and purity of cell line confirmed
Link to related literature	Not available
Information in public databases	Not available
Ethics	Informed written consent obtained

* Corresponding author at: Helmholtz Zentrum München, Parkring 11, D-85748 Garching, Germany.
E-mail address: heiko.lickert@helmholtz-muenchen.de (H. Lickert).

2. Resource details

The Tübingen Family Study for T2DM (TÜF study, N = 2500) was screened for known rare mutations in the MODY4 gene *PDX1*. By mass spectrometry-based genotyping, one heterozygous *PDX1* P33T carrier could be identified and recruited for full-thickness skin biopsy. From the biopsy material, the epidermal layer was separated from the dermis by dispase digestion and fibroblasts were isolated from the dermis by trypsin digestion using an established protocol (see Materials and methods section). The *PDX1* P33T fibroblasts were expanded using commercially available growth media and deep-frozen in liquid nitrogen. From this collection of dermal fibroblasts, samples were tested for the presence of viruses pathogenic to humans (HBV, HCV, HIV) as well as for the presence of mycoplasma. All cultures were found to be negative for all of these potential contaminants. *PDX1* P33T primary fibroblasts were reprogrammed into iPSCs using nucleofection with three episomal plasmids encoding human *OCT4*, *SOX2*, *NANOG*, *LIN28*, *KLF4*, and *L-MYC* (Fig. 1B). The reprogramming protocol (transfection, reprogramming and expansion) is shown in Fig. 1C. Three weeks after transfection, many colonies with human embryonic stem cell (hESC)-
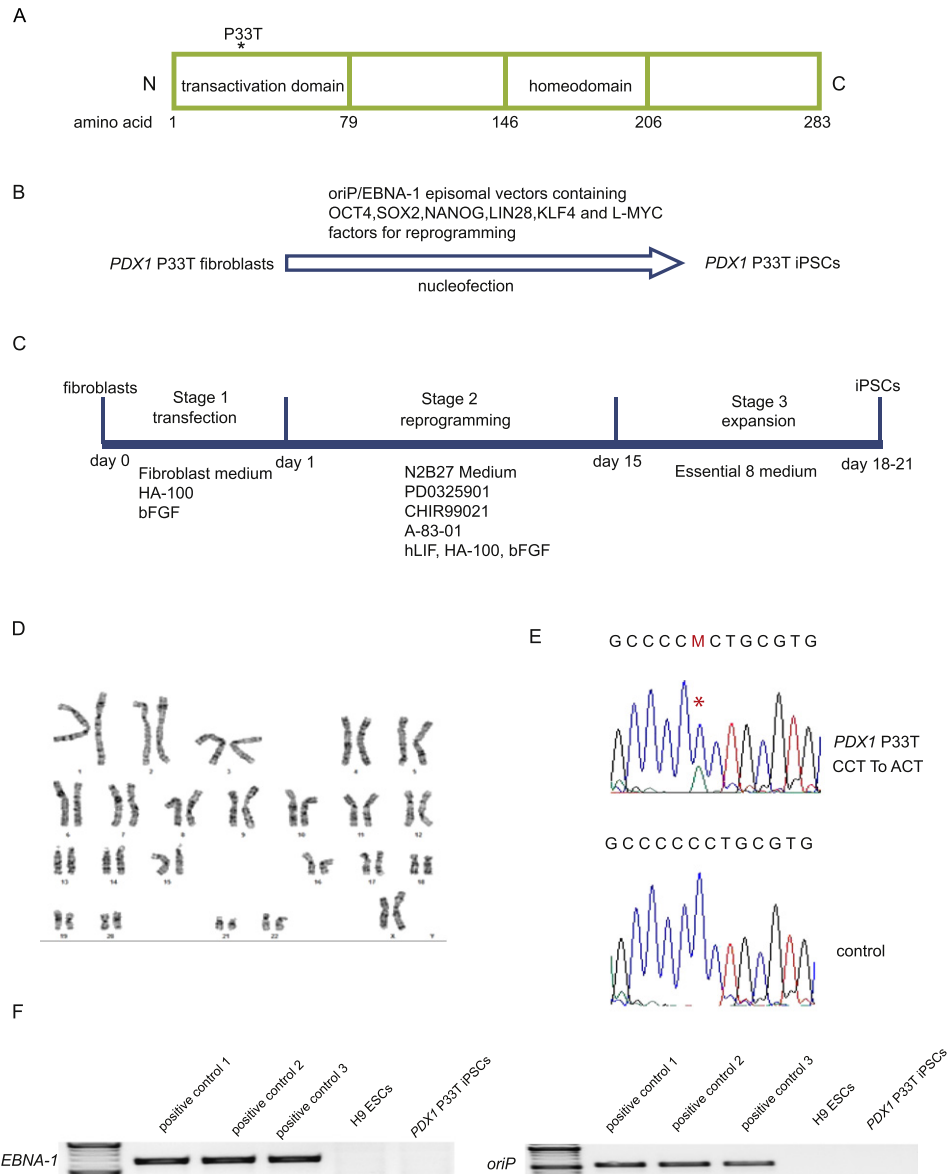
Fig. 1. Generation of *PDX1* P33T iPSCs. (A) An overview of the PDX1 protein structure. (B) Scheme shows the reprogramming factors for the generation of the *PDX1* P33T iPSC line. (C) Reprogramming protocol of skin fibroblasts into integration-free iPSCs using episomal vectors. (D) Normal karyotype (46, XX) of one *PDX1* P33T iPSC clone. (E) Sequencing result shows heterozygous C>A mutation in *PDX1*. (F) Semi-quantitative PCR confirms that episomal vectors did not integrate into the genomic DNA of the selected *PDX1* P33T iPSC clone. H9 ESCs were used as a negative control, whereas transfected fibroblasts at day 6 were used as positive controls.

like morphology appeared that were hand picked by day 21, expanded and frozen down as individual iPSC lines. The analysis of one of these iPSC lines confirmed that *PDX1* P33T fibroblasts were successfully reprogrammed into iPSCs with a normal karyotype (Fig. 1D) and containing the C>A mutation causing a P33T amino acid exchange in the transactivation domain (Fig. 1A, E). Integration of the exogenous transgenes in the genome of iPSCs was excluded by semi-quantitative PCR (Fig. 1F). To confirm that this iPSC line was pluripotent, we performed immunofluorescence staining for the pluripotency transcription factors OCT4 and SOX2, as well as for the hESC surface markers SSEA-3, SSEA-4 and TRA-1-81, which were all expressed in the *PDX1* P33T iPSC line (Fig. 2A). To further demonstrate pluripotency of the selected iPSC line we injected 2×10^6 cells subcutaneously into NOD/SCID mice to generate teratomas. The histological analysis of the induced teratomas revealed that all tissues from all three germ layers were differentiated confirming pluripotency in this *in vivo* assay (Fig. 2B). Taken together, we have successfully reprogrammed *PDX1* P33T dermal fibroblasts

into pluripotent iPSCs that can be used to study a point mutation in the transactivation domain of PDX1 to study the development of diabetes.

3. Materials and methods

3.1. Ethics statement

The study adhered to the Declaration of Helsinki. All participants gave informed written consent, and the study protocols were approved by the local ethics board (Ethics Committee of the Eberhard Karls University Tübingen).

3.2. Study participants

The ongoing TÜF study currently includes >3000 non-related individuals at risk for type-2 diabetes mellitus (T2DM), *i.e.*, healthy subjects with
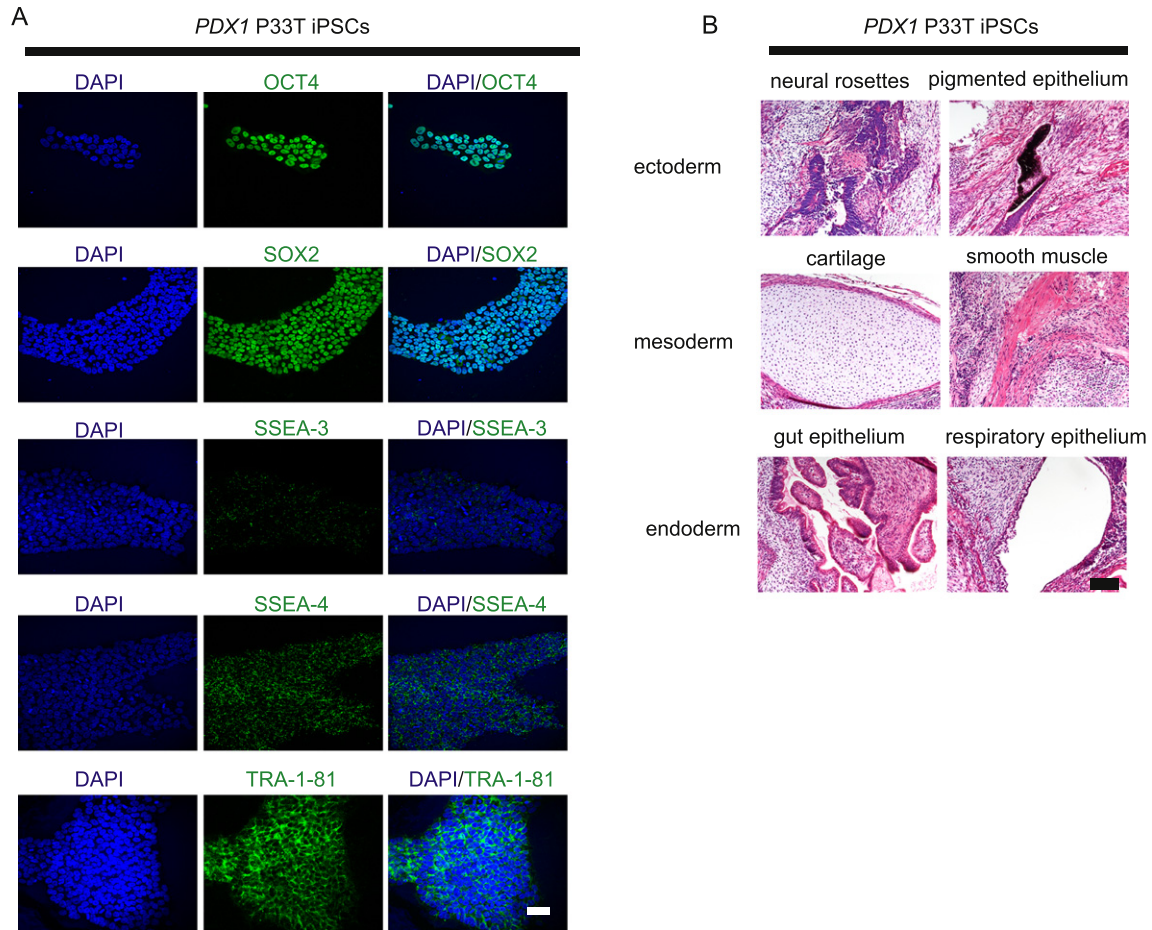
A — PDX1 P33T iPSCs

B — PDX1 P33T iPSCs

Fig. 2. *PDX1* P33T iPSCs are pluripotent and differentiate into all germ layers *in vivo*. (A) *PDX1* P33T iPSCs display typical morphological ESC-like characteristics and uniformly express several pluripotency markers. Scale bar indicates 50 μm. (B) Hematoxylin/eosin staining of tissue sections of teratomas generated from the *PDX1* P33T iPSC clone. All germ layer-derivatives are clearly detectable *in vivo*. Scale bar indicates 100 μm.

a family history of T2DM, a body mass index ≥27 kg/m^2, impaired fasting glycemia, and/or previous gestational diabetes (Stefan N, Machicao F, Staiger H, Machann J, Schick F, Tschritter O, Spieth C, Weigert C, Fritsche A, Stumvoll M, Häring HU. Polymorphisms in the gene encoding adiponectin receptor 1 are associated with insulin resistance and high liver fat. Diabetologia 2005;48:2282-2291). From all TÜF participants, the medical history, smoking status, and alcohol consumption habits were queried and documented. Furthermore, all participants underwent physical examination, routine blood tests, bioelectric impedance measurement, and 5-point oral glucose tolerance tests (OGTTs) with insulin and glucose measurements. The current study population comprised 2500 individuals who spent DNA for genotyping, were not on any medication known to influence glucose tolerance, insulin sensitivity, or insulin secretion, and who had complete OGTT data sets.

Using the mass spectrometry-based genotyping platform MassARRAY from Sequenom (Hamburg, Germany) and the manufacturer's iPLEX software, the 2500 subjects were screened for known rare mutations in maturity onset of diabetes (MODY) genes that were described in the literature to co-segregate with MODY phenotypes in families. We identified seven individuals with heterozygous mutations in *PDX1* (MODY4) (McKinnon and Docherty, 2001). We identified seven individuals with heterozygous mutations in *PDX1* (MODY4). From these cases, one glucose-tolerant female *PDX1* Pro33Thr (P33T) carrier (aged 40 years; BMI 23.0 kg/m^2) could be recruited for full-thickness skin biopsy.

3.3. Skin biopsy, isolation and testing of dermal fibroblasts

A full-thickness skin specimen was taken by punch biopsy from the upper arm in the deltoid muscle region. After removal of adipose tissue remnants and visible blood vessels, the sample was digested overnight at 4 °C with 10 U/ml dispase II (Roche Diagnostics, Mannheim, Germany) in 50 mM Hepes pH .4, 150 mM NaCl. Thereafter, the digest was heated for 30 min at 37 °C under continuous shaking (1200 rpm). Using forceps, the dermis was separated from the epidermal layer, and fibroblasts were isolated from the dermis by digestion with 0.2% collagenase CLS I (Biochrom, Berlin, Germany) in DMEM, 10% BSA for 45 min at 37 °C under continuous shaking (1200 rpm). For purification of the fibroblasts, the digest was filtered through a 70 μm mesh and centrifuged. The pelleted cells were resuspended, grown for three days in DMEM, 10% FCS, and subsequently further expanded in Medium 106 supplemented with low serum growth supplement (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Samples of dermal fibroblasts were tested for the presence of viruses pathogenic to humans, *i.e.*, HBV, HCV, and HIV, with Genesig PCR-based detection kits from Primerdesign Ltd. (Chandler's Ford, UK) and for the presence of mycoplasma with a PCR test kit from PanReac AppliChem (Darmstadt, Germany) and, in parallel, by DNA staining with DAPI. All cultures were found to be negative for the tested contaminants.

3.4. iPSC generation

Primary fibroblasts from patient were reprogrammed into pluripotent stem cells by using a non-integrating Episomal iPSC Reprogramming Kit (Invitrogen, Cat. no. A14703). This kit contains mixture of three vectors which has the oriP/EBNA-1 (Epstein-Barr unclear antigen-1) backbone that delivers six reprogramming factors: *OCT4*, *SOX2*, *NANOG*, *LIN28*, *KLF4*, and *L-MYC* (Yu et al., 2011). Human fibroblasts at 75–90% confluent were transfected using the Amaxa 4D-Nucleofector transfection system and a nucleofector kit for human dermal fibroblast (Lonza, Cat. no.VPD-1001), plated onto geltrex-coated culture dishes, incubated in supplemented fibroblast medium. This medium contained knockout DMEM/F-12 (Life Technologies), 10% FBS of ESC-qualified (Life Technologies), 1% MEM non-essential amino acids (Life Technologies), 10 μM HA-100 (Santa Cruz) and 4 ng/ml bFGF (Life Technologies). At 24 h after transfection, the medium was exchanged with N2B27 medium supplemented with 0.5 μM PD0325901 (Stemgent), 3 μM CHIR99021 (Stemgent), 0.5 μM A-83-01 (Stemgent), 10 μM HA-100 (Santa Cruz), 10 ng/ml hLIF (Life Technologies) and 100 ng/ml bFGF. The basic N2B27 medium contained DMEM/F12 with HEPES (Life Technologies), 1× N2 supplement (Life Technologies), 1× B27 supplement (Life Technologies), 1% MEM non-essential amino acids, 1× Glutamax (Life Technologies) and 1× β-Mercaptoethanol (Life Technologies). On day 15 after transfection, the medium was exchanged with Essential 8 medium and monitored for the emergence of iPSC colonies. Around 3 weeks after transfection, undifferentiated iPS colonies were picked and transferred onto fresh geltrex-coated culture dishes for expansion.

3.5. iPSC characterization

DNA was extracted from iPSCs by using standard procedure. Markers for the episomal backbone were amplified by semi-quantitative PCR to exclude transgene integration. Primers are as follows: *oriP* forward: TTCCACGAGGCTAGTGAACC. *oriP* reverse: TCGGGGGTGTTAGAGACAAC; *EBNA-1* forward: ATCGTCAAAGCTGCACACAG. *EBNA-1* reverse: CCCAGGAGTCCCAGTAGTCA. For karyotype analysis, we used the cells growing in logarithmic phase. They were fed with fresh medium the night before adding colcemid for 2 h. Cells were then trypsinized, treated with hypotonic solution (0.075 m KCl) for 20 min and fixed with methanol:acetic acid (3:1). Metaphases were spread on microscope slides, and chromosomes were classified according to the International System for Human Cytogenic Nomenclature using the standard G banding technique. At least, 20 metaphases were counted per cell line, and the final karyotype was stated if it was present in >85% of them. For teratomas, 2×10^6 iPSCs were injected into the right hind leg of immunocompromised NOD/SCID mice. Tumors were excised after 8 weeks, fixed, embedded in paraffin, sectioned and stained with hematoxylin/eosin (Takahashi et al., 2007).

3.6. Mutation analysis

Genomic DNA was isolated from *PDX1* P33T and control iPSCs using standard procedure. PCR amplification with a set of primers flanking the mutation site was performed in those two samples. PCR products were sequenced using the reverse primer by Sanger sequencing on an ABI 3130 genetic analyzer (Applied Biosystems). Primers are as follows: *PDX1* forward: GGAGTGTGCAGCAAACTCAG, *PDX1* reverse: ACGCGTGAGCTTTGGTAGAC.

3.7. Immunofluorescence imaging

Cells were fixed with 4% paraformaldehyde for 30 min and then permeabilized in PBS containing 0.2% Triton X-100. Cells were blocked with PBS containing 3% BSA, and incubated with primary antibodies overnight at 4 °C. Then secondary antibodies were incubated for 1 h at room temperature after washing with PBS. Images were acquired on a laser scanning microscope. The following antibodies and dilutions were used: goat anti-OCT3/4 (1:500, Santa Cruz), goat anti-SOX2 (1:500, Santa Cruz), rat anti-SSEA3 (1:50, Invitrogen), mouse anti-SSEA4 (1:500, Cell signaling) and mouse anti-TRA-1-81 (1:50, Millipore).

Author disclosure statement

There are no competing financial interests in this study.

References

McKinnon, C.M., Docherty, K., 2001. Pancreatic duodenal homeobox-1, PDX-1, a major regulator of beta cell identity and function. Diabetologia 44, 1203–1214.
Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861–872.
Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J., Jiang, Y., 2011. Efficient feeder-free episomal reprogramming with small molecules. PLoS One 6, e17557.