TECHNISCHE UNIVERSITÄT MÜNCHEN

FAKULTÄT FÜR MEDIZIN

Functional Analysis of *INK4* Locus in Endocrine Development and Function

Alireza Shahryari

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

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitz: Prof. Dr. Roland M. Schmid

Prüfer*innen der Dissertation:

1. Prof. Dr. Heiko Lickert

2. apl. Prof. Dr. Thomas Floss

Die Dissertation wurde am 21.07.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 21.02.2023 angenommen.

Contents

List of figures

Abbreviation

T1D .... Type 1 diabetes

T2D .... Type 2 Diabetes

GD .... Gestational diabetes

MODY .... Maturity-onset diabetes of the young

INK4 ……………….. Inhibitor of CDK4

ORF ………………... Open reading frame

TF …………………. Transcription factor

Pdx1………………... Pancreatic and duodenal homeobox 1

Ptf1a ……………….. Pancreas-specific transcription factor 1a

Ngn3 ………………. Neurogenin 3

Arx ………………… Aristaless-related homeobox

Rb …………………. Retinoblastoma

E2F ………………... Early region 2 transcription factor

RA ………………… Retinoic acid

CRISPR …………… Clustered regularly interspaced short palindromic repeats

tracrRNA ………….. Trans-encoded crRNA

ZFN ………………... Zinc finger nuclease

HDR ……………….. Homology-directed repair

NHEJ ………………. Non-homologous end joining

SSA ………………... Single-strand annealing

BIR ………………..... Breakage-induced replication

Ku80/70 ……………. Ku heterodimer

DNA-PKcs ………… DNA-dependent protein kinase catalytic subunit

XRCC4 ……………. X-ray repair cross-complementing protein 4

Het …………………Heterozygous

Hom ………………. Homozygous

bp …………………. Base pair

DE ………………… Definitive endoderm

GT ………………… Gut tube

PE ………………... Pancreatic endoderm

PP ………………... Pancreatic progenitor

GWAS …………... Genome-wide association study

SNP ……………… Single nucleotide polymorphism

hESC …………….. Human embryonic stem cell

iPSC ……………... Induced pluripotent stem cell

MPCs ……………. Multipotent progenitor cells

NLS ……………… Nuclear localization signal

PCR ……………… Polymerase chain reaction

RNA-seq ………… RNA sequencing

rpm ……………… Rotations per minute

SEM …………….. Standard error of the mean

Abstract

Diabetes is a multifactorial disease, where both genetic disposition and environmental influence play a role in its manifestation. In genome wide association studies (GWAS), an 8 kb intergenic region harboring six single nucleotides polymorphisms (SNP) associated with type 2 diabetes (T2D) was found downstream of the human INK4 locus. The molecular mechanisms of how these SNPs are linked to T2D are still elusive. The human INK4 locus contains three protein-coding genes (*ARF/P14*, *CDKN2B/P15* and *CDKN2A/P16*) and one non-coding RNA gene (*ANRIL*). ARF/P14, CDKN2B/P15, and CDKN2A/P16 play fundamental roles in cell-cycle inhibition via retinoblastoma (RB) and TP53 pathways. However, the *ANRIL* non-coding RNA recruits polycomb repressing complexes (PRCs) 1 and 2 to induce the silencing of *CDKN2A/B* genes.

In silico analyses showed that the 8 kb intergenic region at the INK4 locus harbors several histone marks for active or open chromatin. Furthermore, this region contains binding sites for major transcription factors playing a role in pancreas development and/or *β*-cell function. These indicate that the region might harbor cis-regulatory element. By CRISPR/Cas9 gene editing, we deleted the 8 kb genomic block in induced pluripotent stem cells (iPSC). The resulting knockout (KO) iPSC line (named HMGUi001-A-5) was karyotypically normal, pluripotent, and could differentiate into endoderm, mesoderm, and ectoderm. Following successful characterization, we differentiated this line into pancreatic progenitors, endocrine cells, and β-like cells. The INK4 genes were not expressed at the early stages of differentiation; however, their expression gradually increased from pancreatic progenitor stage onwards and reached a maximum level in β-like cells. *ANRIL* and *CDKN2B/P15* were downregulated at later time points of differentiation in the KO-derived β-like cells. Deleting the T2D risk DNA at the INK4 locus resulted in a diminished proliferation rate in pancreatic progenitors, endocrine cells and β-like cells. The KO-derived β-like cells show less insulin content and reduced glucose-stimulated insulin secretion (GSIS). Altogether, our data represent the potential roles of the T2D risk DNA at the INK4 region in regulating local gene expression and β-cell proliferation and function.

In parallel with the INK4 project, we could design a method to improve the gene editing efficiency for CRISPR/Cas9 system. Upon a double-strand break generated by CRISPR/Cas9 there are two playing pathways: The non-homologous end joining (NHEJ) pathway creates random insertion or deletion of nucleotides, whereas the homology-directed repair (HDR) results in base pair exact modifications. The Cas9 expression vector was modified to increase gene editing efficiency by inserting two different cassettes 1) a short hairpin RNA (shRNA) expression cassette to target *DNAPK* and *XRCC4*, two major actors of NHEJ, or 2) an anti-apoptotic expression construct of miRNA-21 to increase cell survival. For a simple readout in iPSCs, the pluripotency marker *SOX2* was targeted with a *T2A-tdTomato* reporter gene. Downregulating XRCC4 and DNAPK improved the efficiency of *SOX2* knock-in (KI) about twofold. Moreover, ectopic expression of miRNA-21 with Cas9 increased the efficiency of *SOX2* KI about threefold. Totally, our approaches yield an upgrade for CRISPR/Cas9-mediated precise gene integration in human pluripotent stem cells.

Zusammenfassung

Diabetes ist eine multifaktorielle Erkrankung, bei deren Manifestation sowohl genetische Disposition als auch Umwelteinflüsse eine Rolle spielen. In GWAS-Studien wurde ein 8 kb intergene Region im INK4-Lokus gefunden, die sechs Einzelnukleotid-Polymorphismen (SNP) beherbergt, die mit Typ-2-Diabetes (T2D) assoziiert sind. Die molekularen Mechanismen wie diese SNPs mit T2D verbunden sind ist unklar. Der humane INK4-Locus enthält drei Protein-kodierende-Gene (ARF/P14, CDKN2B/P15 und CDKN2A/P16) und ein nicht-kodierendes RNA-Gen (ANRIL). ARF/P14, CDKN2B/P15 und CDKN2A/P16 spielen eine grundlegende Rolle bei der Hemmung des Zellzykluses über den Retinoblastom-(Rb) oder P53-Signalweg. Die nicht kodierende RNA von ANRIL rekrutiert Polycomb-Repressionskomplexe (PRCs) 1 und 2, um die CDKN2A/B-Gene zu hemmen.

In-silico-Analysen zeigten, dass die genomische 8 kb intergene Region in INK4-Lokus mehrere Histonmarkierungen von aktivem oder offenem Chromatin enthält, was darauf hinweist, dass sich dort ein regulatorisches Element befindet. Darüber hinaus enthält diese Region Bindungsstellen für wichtige Transkriptionsfaktoren, die eine Rolle bei der Entwicklung der Bauchspeicheldrüse und/oder der β-Zellfunktion spielen. Durch CRISPR/Cas9-Geneditierung haben wir den 8 kb-Genomblock in induzierten pluripotenten Stammzellen (iPSC) entfernt. Wir bestätigten, dass die Knockout- (KO) oder HMGUi001-A-5-Linie karyotypisch normal und pluripotent war und die Fähigkeit hatte in endodermale, mesodermale oder ektodermale Vorläufer zu differenzieren. Nach erfolgreicher Charakterisierung differenzierten wir diese Linie in Richtung Pankreas-Vorläufer, endokrine Zellen und β-ähnliche Zellen. INK4-Gene wurden in frühen Stadien der Differenzierung nicht exprimiert; ihre Expression stieg jedoch allmählich von pankreatischen Vorläuferstadien an und erreichte ein maximales Niveau in β-ähnlichen Zellen. ANRIL und CDKN2B/P15 wurden zu späteren Zeitpunkten der Differenzierung in den von KO differenzierten β-ähnlichen Zellen herunterreguliert. Die Deletion der T2D-Risikoregion am INK4-Locus führte zu einer verringerten Proliferationsrate in pankreatischen Vorläuferzellen, endokrinen Zellen und β-ähnlichen Zellen. Die von KO differenzierten β-ähnlichen Zellen zeigen einen geringeren Insulingehalt und eine reduzierte Glukose-stimulierte Insulinsekretion (GSIS). Insgesamt zeigen unsere Daten die potenziellen Rollen der T2D-Risiko-Region am INK-Lokus bei der Regulierung der lokalen Genexpression und der Proliferation und Funktion von β-Zellen.

Parallel zum INK4-Projekt wollten wir eine Methode entwickeln, um die Effizienz des Gen-Targetings durch das CRISPR/Cas9-System zu verbessern. Bei einem durch CRISPR/Cas9 erzeugten Doppelstrangbruch gibt es zwei konkurrierende Reperaturmechanismen: Der nicht-homologe Endverbindungsweg (NHEJ) erzeugt eine zufällige Insertion oder Deletion von Nukleotiden, während die homologiegesteuerte Reparatur (HDR) zu basenpaargenauen Modifikationen führt. Der Cas9-Expressionsvektor wurde modifiziert, um die Gen-Editing-Effizienz zu erhöhen, indem zwei verschiedene Kassetten eingefügt wurden: 1) eine Short-Hairpin-RNA (shRNA)-Expressionskassette zur Herunterregulierung von DNAPK und XRCC4, zwei Hauptakteure des NHEJ, oder 2) eine anti-apoptotische Expressionskassette von miRNA-21 zur Steigerung des Zellüberlebens. Für ein einfaches Auslesen in iPSCs wurde der Pluripotenzmarker SOX2 mit einem T2A-tdTomato-Reportergen modifiziert. Das Herunterregulieren von XRCC4 und DNAPK verbesserte die Effizienz von SOX2 KI um etwa das Zweifache. Darüber hinaus erhöhte die ektopische Expression von miRNA-21 mit Cas9 die Effizienz von SOX2 KI um etwa das Dreifache. Insgesamt bieten unsere Ansätze einen einfachen Weg für eine effiziente Genbearbeitung über CRISPR/Cas9 in menschlichen iPSCs.

1. Introduction

1.1 Pancreas Physiology

The pancreas is one of the most important organs of the human body involved in regulating energy consumption and metabolism. It functions as a unique dual gland, consisting of an endocrine and an exocrine compartment in charge of synthesizing and secretion of hormones for glucose regulation and enzymes essential for digestion. The exocrine tissue comprises 95% of the pancreas; however, the endocrine tissue consists of less than 5%. The exocrine compartment includes acinar cells that are connected to ductal cells. The acinar cells yield digestive enzymes, such as proteinases, lipases, and amylases, while ductal cells create channels that transport enzymes into the duodenum and the small intestine to digest carbohydrates, fats, and proteins for absorption. The endocrine section takes part in sustaining glucose homeostasis via secreting several hormones into the blood. Five different hormone-secreting cell types: α-cells, β-cells, δ-cells, pancreatic polypeptide (PP) cells, and ε-cells belong to the endocrine compartment that are embedded within small clusters termed islets of Langerhans. Insulin, secreted by β-cells, and glucagon secreted by α-cells are the central hormones that regulate glucose homeostasis (Muraro et al., 2016) (Kettunen and Tuomi, 2020). δ-cells secrete somatostatin, a hormone that impedes secretion of both hormones of insulin and glucagon. Following feeding, PP-cells secrete pancreatic polypeptide, a 36-amino acid peptide that plays a role as primary feedback inhibitor of pancreatic secretion. It also regulates liver glycogen storage and gastrointestinal secretion. The ε-cells compose < 1% of all islet cells producing the ghrelin hormone that induces appetite (Bakhti et al., 2019) (Zhou and Melton, 2018) (Figure 1.1).

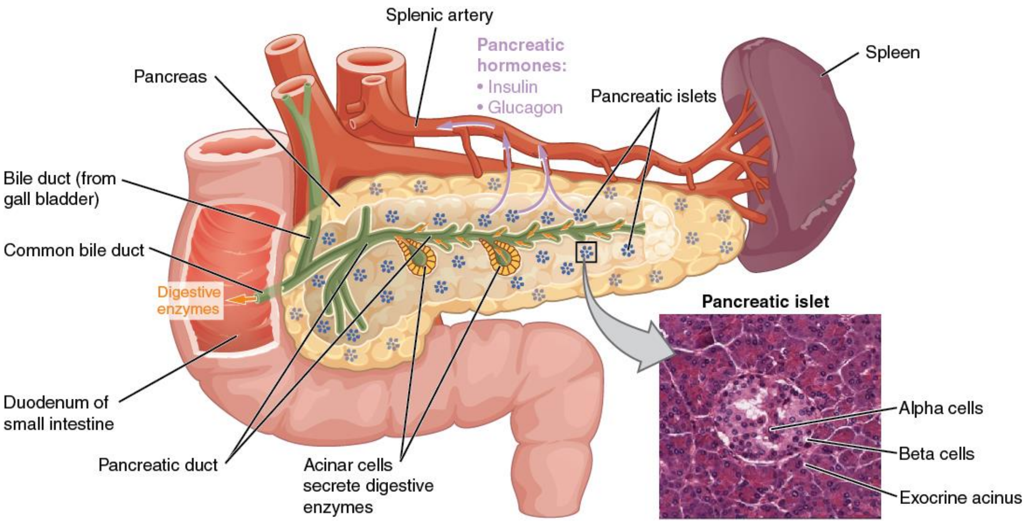


Figure 1-1: The anatomy of the pancreas. The human pancreas is a heterogeneous gland, i.e. it has both an endocrine and a digestive exocrine function. The exocrine compartment encompasses acinar cells and ductal cells while the endocrine compartment comprises α-cells, β-cells, δ-cells, PP-cells, and ε-cells (Roder et al., 2016).

1.2 Glucose Homeostasis

Glucose homeostasis is of essential importance to human health due to its fundamental role as an energy source. Hence, regulation of glucose levels in the blood is mandatory for survival. Glucose homeostasis is carried out by regulating the activities of insulin and glucagon. These hormones mainly act on the liver, fat, and skeletal muscle. The rate of endogenous glucose production and utilization is a major factor in controlling this process. During fasting, 75−85% of endogenous glucose is produced in the liver and the rest in the kidney. In other words, hepatic glucose production is the dominant source of fasting blood glucose levels. Most glucose utilization happens in the skeletal muscle (Keenan et al., 2019) (Salis et al., 2017) (Horie et al., 2018).

Two opposing hormones, glucagon and insulin maintain the balance between glucose production and utilization. In the fed state and response to an increased level of plasma glucose, insulin secretion is hindered from the β-cells. At the same time, glucagon secretion is inhibited from the α-cells. The β-cells react to elevated blood glucose by rising oxidative metabolism. Hence, ATP generation is increased in mitochondria leading to an enhanced ratio of ATP/ADP in the cytoplasm. Following the elevated intracellular ATP/ADP, the ATP-sensitive K^+ channels (K_{ATP}) are closed, resulting in a decreased hyperpolarizing outward K^+

flux. This leads to depolarization of the plasma membrane, and influx of extracellular Ca^{2+} via the voltage-gated Ca^{2+} channels. A quick rise in intracellular Ca^{2+} and operation of protein motors and relevant kinases stimulate exocytosis of vesicles harboring insulin (Matschinsky, 1996) (Maechler et al., 2006) (Rutter, 2001) (Fridlyand et al., 2003). Next, the production of endogenous glucose is inhibited in the liver, while glucose uptake is increased in the adipose, liver, and muscle cells. These tissues play different roles in glucose homeostasis, needing tissue-specific insulin signal transduction pathways. For instance, insulin elevates glucose utilization and storage in skeletal muscle via promoting glucose transport and glycogen synthesis. In the liver, insulin displays three roles 1) stimulates glycogen synthesis, 2) boosts lipogenic gene expression, and 3) reduces gluconeogenic gene expression. In adipocyte tissue, insulin inhibits lipolysis and stimulates glucose transport and lipogenesis.

During fasting or exercise and in response to a decrease in plasma glucose levels, glucagon is released from α-cells, which encompass the β-cells in the pancreas. Glucagon mainly stimulates liver cells and/or muscle cells to break down stored glycogen into glucose. The produced endogenous glucose is later released into the bloodstream, raising the blood glucose level. In other words, both α-cells and β-cells are especially sensitive to glucose concentrations by which they regulate hormone production and release in response to slight alterations in the levels of plasma glucose (Ghasemi and Norouzirad, 2019) (Saltiel, 2016) (Petersen and Shulman, 2018).

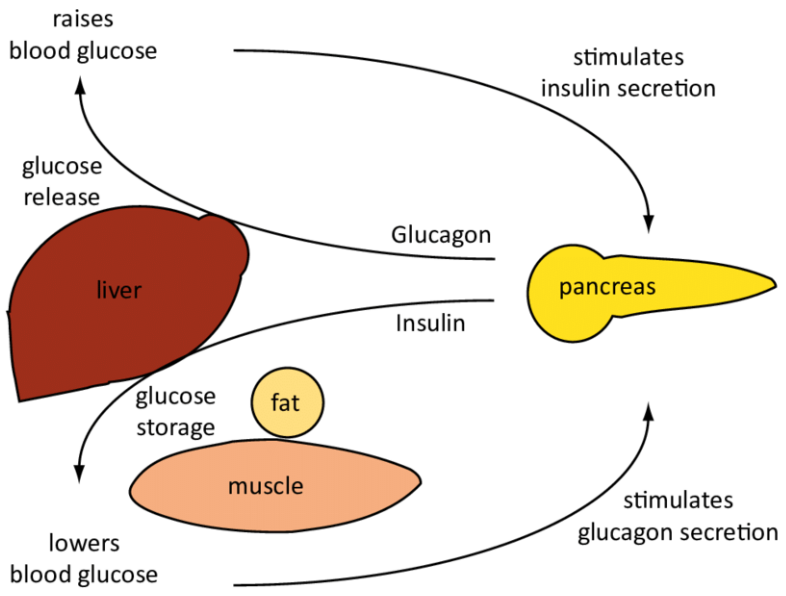


Figure 1-2: Glucose homeostasis. After feeding and when the blood glucose levels are high, insulin is released from β cells of the pancreas. In contrast, when blood glucose levels are low, glucagon is secreted from α cells of the pancreas (Steinbusch et al., 2011).

1.3 Embryonic Development of the Pancreas in Mice

In mice, the pancreas emerges from the foregut endoderm, and its formation occurs via multiple steps of morphological events to produce specified cell types. The first development of this organ can be grouped into two major steps, primary and secondary transitions (Wells and Melton, 1999) (Zorn and Wells, 2009) (Bastidas-Ponce et al., 2017). In the primary transition (from embryonic day (E) 9.0 to E12.5), multipotent progenitor cells (MPCs) are initially emerging and expanding via signals transduction from the notochord, endothelium and mesenchyme that eventually give rise to pancreatic buds (Gittes, 2009) (Lammert et al., 2003) (Larsen and Grapin-Botton, 2017). The initial stage in pancreas development specifies a dorsal and ventral pancreatic bud from the foregut endoderm. These buds come from the MPCs that express two main pancreatic transcription factors (TF) pancreatic and duodenal homeobox 1 (Pdx1) and pancreas-specific transcription factor 1a (Ptf1a) (Burlison et al., 2008). Next, these cells face a higher proliferation rate to form a multilayered epithelium where microlumen structures emerge. After the formation of the pancreatic buds, the next stages of morphogenesis result in a particularly branched, tubular epithelial tree-like network. This extremely regulated event needs epithelial stratification, cell polarization, microlumen generation and fusion and finally develops into a luminal plexus. Next, the plexus is rearranged into a sophisticated tubular web. The formation of exocrine compartment is initiated by epithelial remodeling and/or branching morphogenesis at E11.5 (Marty-Santos and Cleaver, 2015). The presence of mesenchyme is mandatory for the induction of exocrine differentiation. The mesenchyme produces and secretes pro-exocrine factors, e.g. the TGF-β antagonist follistatin inducing exocrine differentiation but inhibits the formation of endocrine cells (Miralles et al., 1998). Moreover, the canonical Wnt signaling pathway balances exocrine cell numbers (Baumgartner et al., 2014) (Wells et al., 2007).

In the time of secondary transition, between E12.5-15.5, the fusion of microlumina forms a central plexus that additionally shapes into an uninterrupted branched epithelial web, separated into bipotent trunk epithelium e.g. tip and trunk domains (Bankaitis et al., 2015) (Kesavan et al., 2009) (Villasenor et al., 2010). The endocrine progenitors originate from the bipotent trunk epithelium. These progenitors temporarily express Ngn3 and generates all types of endocrine cells (Gradwohl et al., 2000) (Gu et al., 2002) (Solar et al., 2009). The morphological changes coincide with the generation of three major cell types including endocrine, exocrine or acinar and ductal cells, demonstrating a tight regulation between morphogenesis and differentiation events throughout the pancreas formation in mice. Following the secondary transition, differentiated endocrine cells exit the ductal epithelium, and move to the neighboring mesenchyme to generate proto-islets. The interactions between these shapes and endothelial, mesenchymal and neuronal cells boost the development of Langerhans islets. These events are highly monitored by the spatiotemporal functions of a number of various signaling pathways and harmonizing cell kinetics and dynamics. The cooperation of extrinsic signals and intrinsic genetic systems could orchestrate the emergence of functional hormone-secreting cells. Additional complex interactions between pancreatic cells and the adjacent mesenchyme, endothelium and neuronal network shape the ultimate anatomy of the adult pancreas (Cleaver and Dor, 2012b) (Thorens, 2014).

1.4 Key Players Controlling Pancreas Development and Cell Type Specification in Mice

Several transcription factors (TFs) monitor pancreas induction and generation from multipotent progenitor cells in the foregut endoderm. Pancreatic duodenal homeobox 1 (Pdx1), Pancreas specific transcription factor 1a (Ptf1a) and SRY-Box Transcription Factor

9 (Sox9) are TFs that play critical roles in the early pancreas development (Ahlgren et al., 1996) (Guz et al., 1995) (Krapp et al., 1998) (Seymour et al., 2007). All cell types derived from the pancreatic endoderm express the TF Pdx1. Its expression is initially started as early as E8.5 in the mice foregut endoderm. Pdx1 is expressed in both the ventral and dorsal buds at E9.5. Its expression is downregulated about E10. Then it is expressed in endocrine cells and adult β-cells. Pdx1 is a fundamental mediator of mesenchymal signaling pathway. This is essential for the branching development involved in forming the ductal network at E10.5. *Pdx1* gene harbors three distinct binding sites for transcription initiation. Several transcription factors such as Foxa2, Hnf6, Ptf1a, Mnx1, Mafa, Hnf, Sp1/3, Usf1/2 and Pdx1 itself can bind to these sites and induce Pdx1 expression in the β-cell (Gao et al., 2008) (Vanhoose et al., 2008) (Offield et al., 1996) (Gu et al., 2002) (Sharma et al., 1996).

Ptf1a is an essential TF regulating exocrine gene transcription. *Ptf1a* gene is expressed in endocrine, exocrine, and ductal cell types. *Ptf1a* is initially expressed as early as E8.0 in the ventral and dorsal pancreatic ducts; however, its expression is limited to acinar precursor cells by E13.5. *Ptf1a* regulates the expression of *Delta-like ligand 1* (*Dll1*), which is mandatory for controlling the early pancreas development mediated by the Notch signaling pathway. Then, activation of *Dll1* within multipotent progenitor cells (MPC) induces proliferation and pancreas growth via sustained expression of hairy and enhancer of split 1 (*HES1*) and *Ptf1a* (Ahnfelt-Ronne et al., 2012) (Beres et al., 2006) (Hald et al., 2008) (Kawaguchi et al., 2002).

Surrounding mesenchyme can regulate the identity of the early pancreas development. The mesenchyme secretes a number of signaling factors including- Fgf10, Egf and Wnt that have been demonstrated to be fundamental for the pancreas formation. Fgf10 can activate the Fgf signaling pathway, which results in propagation of pancreatic progenitors via boosting the expression of *Pdx1* and *Ptf1a*. Fgf10, Sox9 and Fgfr2 comprise a feed-forward cycle throughout the early growth phase of the pancreatic buds. Fgf10 preserves *Sox9* expression, and afterward, Sox9 induces *Fgfr2* expression to activate the Fgf10 signaling pathway. Dysregulation of this cycle can induce a loss of specification in pancreatic epithelia cells (Ahnfelt-Ronne et al., 2012) (Attali et al., 2007) (Bhushan et al., 2001) (Jonckheere et al., 2008) (Kim and Hebrok, 2001) (Tulachan et al., 2006) (Seymour et al., 2012).

Around E12.5, the endocrine, acinar and ductal cells emerge from the microlumina. Later, a cluster of cells, including endocrine cells, endothelial, mesenchymal and neuronal cells generate the islets of Langerhans (Cleaver and Dor, 2012a) (Thorens, 2014). The segregation of MPCs into trunk and tip domains is crucial at this stage. The tip domain differentiates into acinar cell types expressing TFs Ptf1a, c-Myc, Nr5a2 and Cpa1: however, the trunk domains generates endocrine and/or duct progenitor cells expressing TFs Nkx6.1, Sox9, Hnf1b, Nkx2.2 and Pdx1 (Zhou et al., 2007).

Nkx6.1 and Ptf1a particularly regulate the separation into tip or trunk domains. Nkx6.1 triggers the generation of trunk by seizing the tip fate: however, Ptf1a induces the formation of tip domain via inhibiting the trunk formation (Schaffer et al., 2010). Additionally, the formation of tip and trunk is controlled by Notch signaling pathway through modulating the expression of *Nkx6.1* and *Ptf1a*. High expression of *Ptf1a* initiates acinar cell differentiation, which begins from the distal tip epithelium at E13.5, whereas its low expression of *Ptf1a*, *Rbp-jl* and *Nr5a2/LRH-1* maintain the identity of MPC (Holmstrom et al., 2011) (Masui et al., 2007) (Masui et al., 2010). The TFs c-Myc and β-catenin are also involved in differentiation, expansion, and maintenance of the acinar cells (Lobo et al., 2018).

Special cells in the bi-potent trunk progenitor pool express the TF Neurogenin 3 (Ngn3), that has a significant role in endocrine formation and function. It is expressed in all endocrine progenitor cells in two distinct phases. During primary transition *Ngn3* is expressed in mice in a few pancreatic progenitors generating glucagon positive cells. During secondary transition its expression starts at E12 with a high expression level at E15.5, which corresponds to endocrine cell allocation. *Ngn3* is associated with several transcription factors that have functions in endocrine differentiation, cell specification and maintenance, such as *Pdx-1*, *Nkx2.2, Nkx6.1, NeuroD1, Pax4, and Pax6* (Collombat et al., 2005) (Gu et al., 2002) (Villasenor et al., 2008) (Gradwohl et al., 2000) (Schwitzgebel et al., 2000).

The cells in the trunk domain that do not express Ngn3 finally yield the ductal web. Several TFs containing Sox9, Hes1, Hnf1b and Glis3 determine ductal cell destiny (De Vas et al., 2015) (Delous et al., 2012) (Kang et al., 2009) (Shih et al., 2012). Moreover, the Notch signaling pathway plays a major role in ductal cell specification. Notch signaling inactivates Ngn3 inducing ductal cell differentiation (Shih et al., 2012). Furthermore, endocrine cell differentiation is controlled by a number of signaling pathways, such as Notch signaling, Wnt signaling and sphingosine-1-phosphate signaling (Kim and Hebrok, 2001).

The *paired box containing gene 4* (*Pax4*) is an essential TF during early β-cell differentiation. It is initially expressed at E9.5 and is temporarily expressed in all endocrine progenitors in the time of pancreatic formation. *Pax4* is expressed following Ngn3 activation. In the *Pax4* knockout cells, β-cells and δ-cells are not formed, while more α-cells are generated. Also, the depletion of *Pax4* inhibits the expression of *Pdx1* and insulin in β-cell precursors. TF aristaless-related homeobox (Arx) counteracts *Pax4*. *Arx* is initially expressed at E9.5 during mouse pancreatic formation and restricted to mature α-cells. In *Arx* knockout mice do not α-cells do not form. Thus, the stability of *Pax4* and *Arx* is a critical factor for cell fate specification of both α-cells and β- cells (Collombat et al., 2005) (Collombat et al., 2003) (Lin and Vuguin, 2012) (Wang et al., 2004).

Members of the NK homeodomain transcription factor family have a fundamental function in regulating organ formation. *Nkx2.2* and *Nkx6.1* are most important during pancreatic formation since have roles in endocrine cell lineage. Nkx2.2 regulates *insulin* and *Pax4* expressions. At first, the *Nkx2.2* gene is expressed with *Pdx1* at E8.75 in dorsal buds and at E9.5 in ventral buds. However, its expression is restricted to the endocrine cells by E15.5. Loss of *Nkx2.2* resulted in a diminished number of α-cells and PP-cells, and the complete failure of the development of β-cells. In the *Nkx2.2* knockout, insulin and *Nkx6.1* genes are not expressed, demonstrating that *Nkx2.2* is crucial for the identity of the β-cells. The expression of *Nkx6.1* is like that of *Nkx2.2*. *Nkx6.1* is initially expressed in the ventral buds at E8.75. Then, its expression switches to the dorsal buds from E9.0 to E10.5. Then its expression is limited to the central epithelium by E11.5. In adult mice, *Nkx6.1* expression is only observed in the β-cells. It inhibits glucagon expression and controls glucose-stimulated insulin secretion (GSIS). Lack of *Nkx6.1*, affects the late step of pancreas formation, resulting in failure of β-cell development (Sander et al., 2000) (Stanfel et al., 2005) (Cissell et al., 2003) (Jorgensen et al., 2007) (Sussel et al., 1998) (Schisler et al., 2005).

1.5 Pancreas Development in Humans

Human pancreas formation or development is poorly understood due to the restricted access to human tissues. Our knowledge comes from analyses of embryonic and fetal tissue samples.

Basically, the main events and molecular players involved in pancreas development are conserved among mouse and human; however, variabilities in timing, checkpoints, and developmental players have been depicted (Fowden and Hill, 2001) (Bastidas-Ponce et al., 2017).

Like mice, the human pancreas development starts with the rearrangement of the foregut endoderm, at Carnegie step (CS) 9. This eventually generates ventral and dorsal buds at CS13. In contrast to mice, a primary transition is not observed in human pancreas and *NKX2.2* expression is not detectable in pancreatic progenitors (Jennings et al., 2013) (Pan and Brissova, 2014) (Jennings et al., 2015). Marked populations of tip-like and trunk-like domains are observed by CS19. The endocrine progenitors that are $NGN3^{+}$ $SOX9^{-}$reach at the highest point at eight weeks post coitus (wpc), decrease at ~26-28 wpc and are not observed by 35 wpc. Similar to mice, human endocrine progenitors transiently express NGN3; however, *NGN3* is also expressed in recently differentiated human endocrine cells. Surprisingly, *NGN3* homozygous mutations result in developing a mild diabetic phenotype. This can demonstrate that there is NGN3-independent mechanisms for generating endocrine progenitors (Capito et al., 2013) (Salisbury et al., 2014, Jennings et al., 2013) (Lyttle et al., 2008) (Rubio-Cabezas et al., 2014).

In human, the initial fetal insulin-expressing β-cells develop at ~8 wpc. Then, glucagon-producing α-cells emerge at 9 wpc. The endocrine clustering initiates by 10 wpc, and all types of endocrine cells are distinguishable in the emerging islets by 12-13 wpc. Of note, the morphology of human islets changes during development. For example, α-cells are located at the periphery while β-cells are in the core at 14 wpc; however, both cell types are intermixed within the islets by 21 wpc. This change in islet architecture could be essential for the maturation of human endocrine cells (Hanley et al., 2010) (Jennings et al., 2013) (Riedel et al., 2012) (Meier et al., 2010) (Jeon et al., 2009).

1.6 Diabetes

Diabetes is a metabolic disorder highlighted by hyperglycemia, a sustained increase of glucose levels in the blood. This disease is caused by a decrease or partially or complete damage of insulin-producing β-cells. Based on the diagnostic criteria, etiology, and genetic examinations, diabetes can be grouped into four major types: type 1 diabetes (T1D), type 2 diabetes (T2D), gestational diabetes (GD), and monogenic diabetes. T2D accounts for about

90% of individuals, while 10% of cases of the disease primarily belong to the other forms of diabetes. Common to all shapes of diabetes is the elevated glucose level in the blood. Diagnosis of diabetes is carried out via measuring the blood tests such as fasting plasma glucose (FBS), oral glucose tolerance test, or glycated hemoglobin (A1C) (Patterson et al., 2009) (Yang and Chan, 2016). It was reported that diabetes affected 463 million people by 2019, 578 million people by 2030 and will rise to 700 million by 2045. The prevalence rate is lower in rural (7.2%) regions than urban (10.8%) , and in low-income countries (4.0%) than high-income (10.4%) . Furthermore, the global prevalence of diminished glucose tolerance reached 374 million cases in 2019, rising to 454 million by 2030 and 548 million by 2045 (Saeedi et al., 2019).

T1D is an autoimmune disorder leading to lack of insulin production due to an autoimmune response against pancreatic β-cells. T1D is also considered one of the most common chronic diseases that usually manifests at the first decades of life. In T2D, previously known as adult-onset diabetes, insulin is produced; however, its secretion is failed, or muscle or fat tissues are resistant to insulin. T2D is a common form of diabetes that is ended by a complicated cooperation between various risk factors from environment and genetic. Changes in eating habits, overweight, and lack of exercise or a sedentary lifestyle are major risk factors for T2D. The diminished insulin sensitivity is compensated by increased insulin secretion of β-cells which causes β-cell stress and cell death resulting in reduced b-cell mass. GD is defined by glucose intolerance emerging throughout the second or third trimester of pregnancy. T1D, T2D and GD are classified as multifactorial disorders that both genetic and environmental factors trigger the disease (Kahn et al., 2014) (Seely, 2006).

Monogenic diabetes is due to mutations in individual genes involved in β-cell formation and action (Hattersley and Patel, 2017). Maturity-onset diabetes of the young (MODY) and neonatal diabetes mellitus (NDM) are two shapes of monogenic diabetes. MODY shows an autosomal dominant genetic pattern and is observed in adolescents. The diminished number of β-cells or impaired β-cell activity plays major roles in the development of MODY. Mutations in a number of key genes, such as *GCK*, *ABCC8*, *NEUROD1*, *HNF1A*, *HNF1B* and *PDX1*, account for development of MODY. For a therapeutic strategy, sulfonylureas and insulin have been applied for MODY patients having a mutation in *HNF1A*, *HNF4A* and *HNF1B* (Heuvel-Borsboom et al., 2016) (Murphy et al., 2008) (Pearson et al., 2003).

1.7 Genetics of Type 2 Diabetes

As mentioned above, decreased β-cells mass, failure of insulin secretion, and muscle and fat tissue resistance to insulin are common manifestations of T2D. Many genetic factors have roles in T2D. Currently, 70 loci conferring susceptibility to T2D have been reported. These loci have been discovered via three different methods, including linkage studies, candidate gene studies, and genome-wide association studies (GWAS). These findings are beneficial for a better understanding of the pathophysiology of T2D (Meigs, 2019) (Langenberg and Lotta, 2018) (Mishra et al., 2021).

1.8 Linkage Analysis

The linkage concept evaluates genes and genomic markers located close to each other on the same chromosome and inherited together. This linkage analysis shows relatively poor resolution as only a few hundred genetic markers have been genotyped and discovered across the genome. Additionally, the positions recognized by this method could encompass hundreds of genes and millions of DNA base pairs. Hence, linkage analysis is particularly successful for discovering single gene diseases not complex or polygenic ones. Both TF 7-like 2 (TCF7L2) and Calpain 10 (CAPN10) are associated with T2D were identified via the linkage analysis (Ali, 2013).

CAPN10 encodes a cysteine protease that belongs to the Calpain family, a huge family of ubiquitously expressing genes that display critical functions in intracellular rearrangement, post-receptor signaling, and other intracellular activities. This gene is embedded within chromosome 2 and is the first T2D related target gene identified by linkage analysis. Furthermore, CAPN10 is involved in several other activities including cell signaling, apoptosis, exocytosis, mitochondrial metabolism, and cytoskeletal remodeling. Impairment of *CAPN10* expression and function has been reported in diverse pathologies. *CAPN10* gene can be accounted for T2D prevalence, and its single nucleotide polymorphisms (SNP) are associated with an elevated risk of the disease (Hanis et al., 1996) (Panico et al., 2014).

The Gene of *TCF7L2* is embedded within chromosome 10q. At the beginning, it was introduced as a T2D susceptibility gene. Besides the linkage analysis, the association between T2D and several SNPs in the *TCF7L2* gene was approved in a number of Genome-wide association studies (GWAS). Up to now, the *TCF7L2* gene has remained the most observed and particularly associated T2D risk gene. The gene encodes a transcription factor, which exhibits a fundamental role in pancreatic islet development and function. It is considered as a player in WNT signaling pathway. TCF7L2 protein can bind to β-catenin generating heterodimers, which induce the expression of a number of genes, including the *glucagon-like peptide 1* (*GLP*-1) gene, the *insulin* gene, and other candidate genes play roles in processing and exocytosis of insulin granules (Duggirala et al., 1999) (Jin, 2016) (Grant et al., 2006).

1.9 Candidate Gene Analysis

In a candidate gene approach, the target genes that are already supposed to have an impact on the pathogenesis and prevalence of T2D are determined via precise sequencing tests. The method relies on genes that have well-known functions in glucose uptake and metabolism, insulin production and secretion, insulin receptors, post-receptor signaling and lipid metabolism. These genes are *peroxisome proliferator-activated receptor gamma* (*PPARG*), *potassium inwardly rectifying channel*, *subfamily J*, *member 11* (*KCNJ11*), *Wolfram syndrome 1* (*wolframin*) (*WFS1*), *HNF1 homeobox A* (*HNF1A*), *HNF1 homeobox B* (*HNF1B*) and *HNF4A* (Gaulton et al., 2008).

PPARG gene encodes the molecular target of thiazolidenediones, a typical form of anti-diabetic drugs. Converting proline to arginine at position 36 in the PPARG protein increases the risk of diabetes by 20%. PPARG regulates glucose metabolism and fatty acid storage process. The target genes that are switched on by PPARG induce lipid uptake and adipogenesis in the fat cells. *PPARG* knockout mice lack adipose tissue, demonstrating *PPARG* gene is a major regulator of adipocyte differentiation. Furthermore, PPARG increases insulin sensitivity by several molecular mechanisms including induction of the fatty acids storage in the fat cells, stimulating FGF21-mediated stimulating adiponectin release from fat cells, and inducing the production of nicotinic acid adenine dinucleotide phosphate via upregulating the CD38 enzyme (Ruchat et al., 2009) (Cataldi et al., 2021) (Song et al., 2012).

The gene of *Potassium Voltage-Gated Channel Subfamily J Member 11* (*KCNJ11*) produces the Kir6.2 ATP-sensitive K^+ channel that exhibits a major function in the control of insulin release from the ß-cells. Kir6.2 is an integral membrane protein that forms a channel that allows potassium to flow into the cells. This process is monitored by the G-proteins coupled sulfonylurea receptor (SUR), constituting the ATP-sensitive K^+ channel. This channel couples the metabolic status of the cells to their electrical activity and is present in diverse cell types, including brain, cardiac, skeletal, smooth muscle, and pancreatic β-cells. Missense DNA polymorphisms in *KCNJ11* are associated with T2D (Hani et al., 1998) (Pipatpolkai et al., 2020).

WFS-1 gene encodes Wolframin (a protein inside the endoplasmic reticulum (ER) membrane) that is mutated in patients with the Wolfram syndrome. The disease is highlighted by diabetes insipidus, juvenile diabetes, optic atrophy, and deafness, usually presenting in childhood or early adult life. *WFS-1* is particularly expressed in islet β-cells in which it forms a cation-selective ion channel playing fundamental roles. Pancreatic β-cells are selectively eliminated from the islets of wolfram syndrome patients. ER localization reveals that WFS1 channel is involved in physiological events including membrane trafficking, secretion, processing and controlling ER calcium homeostasis. Dysregulations of these processes simulate ER stress responses and eventually apoptosis. Two SNPs in the *WFS-1* gene are significantly associated with T2D (Ueda et al., 2005) (Sandhu et al., 2007).

HNF1A, *HNF1B*, and *HNF4A* are famous as MODY genes involved in liver development, monitoring hepatic metabolic roles, and β-cells formation. However, they are also expressed in various tissues and organs, such as the pancreas and the kidney, regulating development and function. MODY genes carry rare high penetrance mutations that result in a monogenic form of diabetes in juveniles. Risk variants in these genes are linked to reduced insulin release and increased susceptibility to T2D in different human populations (Ma et al., 2016) (Lau et al., 2018).

1.10 Genome Wide Association Studies (GWAS)

Candidate gene and linkage analyses discovered only a little T2D risk genes but have negligible contributions to the hereditary of T2D. Emerging and maturation of high-throughput SNP genotyping technologies and the accessibility of Hapmap database (https://www.sanger.ac.uk/resources/downloads/human/hapmap3.html) lists hundreds of SNPs that have a link to T2D. Most SNPs are located inside and/or outside of coding genes. According to the GWAS studies, *TCF7L2* gene is the most important target gene that has a function in the pathophysiology of T2D. The other most significant candidate genes include hematopoietically expressed homeobox (*HHEX*) *HHEX*, *Solute carrier family 30 member 8* (*SLC30A8*), *insulin-like growth factor 2 mRNA binding protein 2* (*IGF2BP2*), *KCNJ11*, *Peroxisome proliferator-activated receptor gamma* (*PPARG*) and *cyclin-dependent kinase inhibitor 2A/B* (*CDKN2A/B)* genes (Pal and McCarthy, 2013) (Cugino et al., 2012).

1.11 INK4 Locus

According to the GWAS studies, SNPs embedded upstream of the *CDKN2A* and *CDKN2B* genes are associated with the risk of T2D in several huge populations worldwide (Figure 1.1). This locus that contains *CDKN2A* and *CDKN2B* is called *INK4* (Inhibitor of CDK4) as these two are inhibitors of cyclin-dependent kinase 4 and 6 (CDK4 and 6). *INK4* genes are found on chromosome 9p21.3 and are transcribed into several RNA isoforms, *CDKN2A/CDKN2B/ARF* resulting P16, P15 and P14 proteins. The *INK4* locus also encodes a long non-coding RNA termed ANRIL. ANRIL is transcribed from the opposite DNA strand compared to the other INK4 genes. Its gene overlaps the *CDKN2A* gene promoter and all parts of *CDKN2B* gene. Furthermore, it is spliced into several linear or circular transcript variants (Kim and Sharpless, 2006) (Sharpless and Sherr, 2015b) (Figure 1.2) (Figure 1.3 A-C).

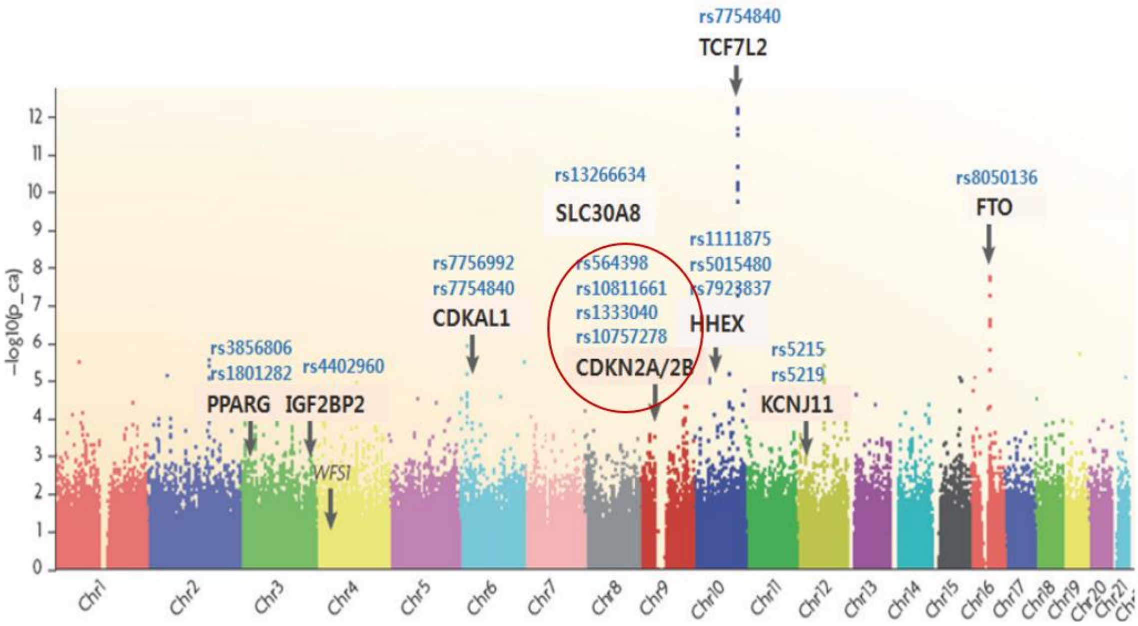


Figure 1-3: GWAS evidence for the INK4 link to T2D. Many GWAS studies showed that there are several SNPs located in the INK4 locus which have an association to the T2D prevalence (Hara et al., 2014a).

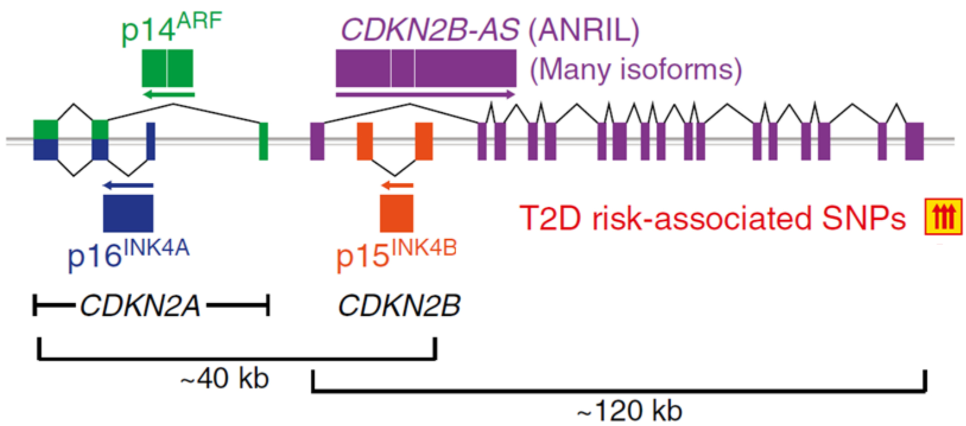


Figure 1-4: INK4 locus genes and structure at human chromosomal region of 9P21.Genomic structure of INK4 locus. Ink4 locus harbors three protein-coding genes including CDKN2A, CDKN2B and ARF. CDKN2A and ARF start with different first exon but share the remaining exons. These three genes are transcribed in the same direction. The locus also harbors a non-coding RNA gene termed ANRIL that is transcribed in from the opposite DNA strand compared to the other INK4 genes (Kong et al., 2016).

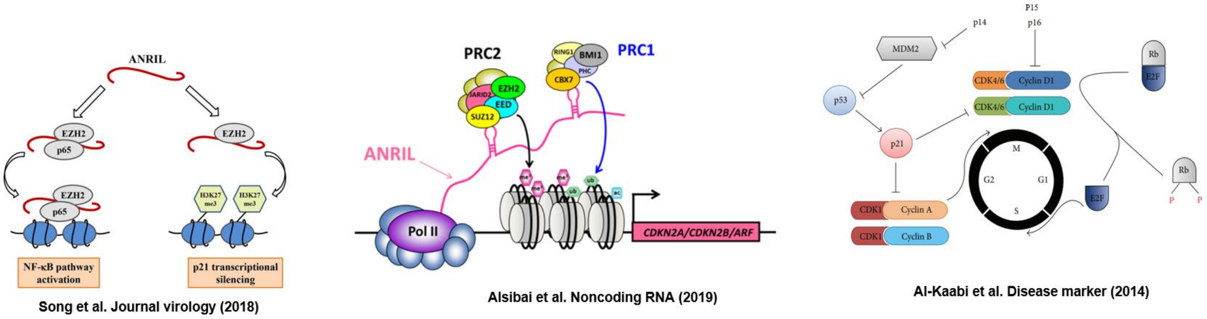


Figure 1-5: Molecular functions of INK4 genes. A: The ANRIL non-coding RNA recruits EZH2, then induces p21 transcriptional silencing and NF-kB activation. **B**: Furthermore, ANRIL interacts with CBX7 (PRC1) and SUZ12 (PRC2) and induces transcriptionally silencing of CDKN2A/CDKN2B/ARF. **C**: CDKN2A/CDKN2B/ARF codes P16/P15/P14 involving in P53 and RB pathways (Song et al., 2018) (Drak Alsibai et al., 2019) (Al-Kaabi et al., 2014).

P16 and P14 are generated by the *CDKN2A* gene. The second and third exons are shared, but the first exon and promoter are different in the two transcripts. P16 and P14 have different open reading frames (ORF), leading to different protein sequences despite the common mRNA sequence. They are famous as inhibitors of CDK4 kinase. *CDKN2B* gene is located about 30 kb upstream of *CDKN2A*, and encodes p15 or INK4B. It generates at least 2 splice variants. *CDKN2A/B* genes control the cell cycle, and are known for their role in tumor suppression (Robertson and Jones, 1999).

INK4 locus was identified to be linked to T2D in several GWAS reports among various human populations. It is predicted that the risk alleles add an odds ratio for the T2D prevalence between 1.2 and 1.5. Of note, the molecular mechanisms of how the risk variants in the INK4 locus increase the diabetes risk is not fully understood, but recent studies point to a failure in insulin release rather than insulin function. These risk alleles are also observed in GWAS for cardiovascular disorders, especially for atherosclerosis; however, the mechanism involving this association is not clear (Nanda et al., 2016) (Quelle et al., 1995) (Poi et al., 2013).

The mouse *Cdkn2a/b* locus is located on chromosome 4, encoding p16Ink4a, p19Arf and p15Ink4b in a close order to the human locus. However, the locus encodes a different lncRNA termed AK148321 in the same location as *ANRIL*. P14, P15 and P16 are cell cycle inhibitors and suppress cell cycle courses affecting tumorigenesis, senescence, and aging events. P16 and P15 block activation of CDK4/6 by cyclin D. Hypophosphorylated retinoblastoma (Rb) mediated by CDK4/6 inhibits the early region 2 transcription factor (E2F), consequently preventing cell cycle entry. In addition, P14 is known as an anti-proliferative factor that stabilize the tumor suppressor P53 by seizing its negative regulator, mouse double minute 2 homologue (MDM2) (Serrano et al., 1993) (Levine et al., 1991).

The INK4 locus is also well-known for its role in metabolic pathways. The most significant adverse effects related to dysregulation of *INK4* genes, mainly P16 are aging and decreased β cell proliferation/mass and regeneration. The *p16INK4A* gene is expressed in human islets in an age-dependent manner fashion and increases with age. Progressive demethylation of the *CDKN2A* locus with age has a potential role in this event. In the human pancreas, nuclear p16 staining is dramatically lower in samples from younger (age 0–9 years) than older (age 10–59 and 60–79 years) patients (Taneera et al., 2013) (Avrahami et al., 2015) (Mizukami et al., 2014).

1.12 Induced Pluripotent Stem Cells (iPSC)

Induced pluripotent stem cells (iPSCs) are pluripotent stem cells that originated from somatic cells. These somatic cells are genetically reprogrammed to embryonic state through the ectopic expression of Yamanaka factors (Oct4, Sox2, Klf4, and c-Myc). These genes are mandatory for sustaining the pluripotent properties of stem cells. The generation of iPSCs allowed scientists to apply pluripotent stem cells for many research areas and clinical applications without the controversial use of embryo-derived stem cells. Indeed, this technology provided beneficial cellular tools to reprogram any kind of somatic cells. Additionally, the differentiated cells that are derived from iPSCs have a similar pattern of gene expression to the cell donor. This is a critical point for drug screening and disease modeling. It is estimated that iPSCs will help scientists uncover novel molecular pathways involved in cellular homeostasis. Furthermore, these valuable cells have a high potential to regenerate damaged cells in human (Singh et al., 2015) (Moradi et al., 2019) (Pfannkuche et al., 2010).

The generation of mouse iPSCs from fibroblasts was initially performed in 2006 by the Yamanaka group. In 2007, human iPSCs were produced from human fibroblasts by two independent Yamanaka's and Thomson's groups. They examined hundreds of genes to find the leading players of the pluripotency network. The reprogramming of somatic cells into iPSCs was initially carried out via ectopic expression of four players, Oct4, Sox2, Klf4, and c-Myc. Also, Oct4, Sox2, Nanog, and LIN28 can reprogram fibroblast cells into the iPSCs. The ectopic expression of these factors eventually gets silenced, and upon the generation of iPSCs, the cell expresses endogenous pluripotency markers involved in the cellular pluripotency network. The iPSCs can be generated from fibroblasts via viral integration of only three factors, Oct4, Sox2, and Klf4. Although these iPSCs revealed reduced tumorigenicity in chimeras and mice; however, the reprogramming rate and its efficiency are significantly reduced. These data demonstrated that the ectopic expression of the three TFs Oct4, Klf4, and Sox2 is mandatory for reprogramming somatic cells into the iPSCs (Takahashi and Yamanaka, 2006) (Takahashi et al., 2007) (Yu et al., 2007).

1.13 iPSCs for Modeling of Diabetes

The generation of insulin-producing β-cells from human stem cells can be a potential therapy for both T1D and T2D. Currently, iPSCs technology has led to a substantial achievement in generating functional β-cells. Human embryonic stem cells and iPSCs can be differentiated to convert into β like-cells in a similar manner. These technologies can be applied for disease modeling and drug screening, as well as for understanding the central pathways involved in β-cells development and function (Mayhew and Wells, 2010) (Nihad et al., 2021).

Tateishi et al. showed that the human iPSCs could be differentiated into insulin-producing islet-like clusters (ILCs) under feeder-free conditions. The iPSCs-derived ILCs not only produced C-peptide and glucagon hormones but also secreted insulin in response to glucose stimuli. Similarly, Zhang et al. developed an efficient protocol to differentiate human iPSCs into mature insulin-producing β-cells in a chemical-dependent culture system. The resulting cells released insulin upon glucose stimulation similar to adult human islets. Interestingly, most iPSCs-derived β cells showed very similar expression patterns to adult β-cells. For example, mature β cell-specific markers including NKX6-1 and PDX1 were expressed in those cells (Tateishi et al., 2008) (Zhang et al., 2009).

As mentioned above, iPSCs can also be applied for diabetes modeling. For example, many loss-of-function mutations in the *PDX1* gene are linked to T2D. To understand the pathomechanism of T2D, our group already generated an iPSC line from a female donor with a mutation (P33T) in the transactivation domain of PDX1 (PDX1$^{P33T/P33T}$). Our group also generated iPSCs from a female donor with another mutation (C18R) in PDX1 (PDX1$^{C18R/C18R}$). Both iPSCs lines could be useful for investigating diabetes pathomechanisms related to *PDX1* mutations. Applying an in vitro β-cell differentiation approach, Wang et al. claimed that both heterozygous PDX1$^{P33T/+}$, PDX1$^{C18R/+}$ and homozygous PDX1$^{P33T/P33T}$, PDX1$^{C18R/C18R}$ mutations demonstrated failure in β-cell development and activity. Additionally, the differentiation efficiency of pancreatic progenitors (PPs) was reduced in iPSCs harboring PDX1$^{+/-}$ and PDX1$^{P33T/P33T}$ mutations. This could be due to reduced activities of PDX1 target genes such as transcription factors PDX1 and MNX1 and insulin resistance gene CES1. Furthermore, the expression of long-noncoding RNA, MEG3 and the imprinted gene NNAT was downregulated in PPs in both PDX1$^{P33T/+}$ and PDX1$^{P33T/P33T}$ mutants. MEG3 and NNAT involve in the processes of insulin synthesis and secretion (Wang et al., 2016a) (Wang et al., 2016b) (Wang et al., 2019).

1.14 Mapping iPSC Differentiation to Pancreas Development

The iPSCs differentiation protocols to generate pancreatic islet-like cells mimic the pancreatic developmental stages, initiating with definitive endoderm and primitive gut tube, followed by limiting the cell fate to pancreatic and EPs, and eventually targeting the final differentiated β-like cells (Rezania et al., 2014) (Hogrebe et al., 2020) (Velazco-Cruz et al., 2019). Small compounds, cytokines and growth factors are applied to steer the pathways essential for the differentiation process, imitating embryonic development. Recently, β-cell differentiation approaches have been carefully developed (Nostro and Keller, 2012) (Velazco-Cruz et al., 2020) (Theis and Lickert, 2019). Before the recent protocols, iPSCs could not differentiate into β cells as efficiently as embryonic stem cells. The relative immaturity of iPSC-derived β-cells could demonstrate an insufficient number or wrong combination of compounds essential for the appropriate pancreas development. Moreover, undesirable cells, including α and δ cells, would have been observed in past protocols; hence, the generated β cell clusters were renamed to islet clusters. In the optimized protocols, the efficiencies for generating iPSC derived-β cells ranged from 17% to 73% C-Peptide^{+}. However, the C-Peptide^{+} cells can be functional (NKX6.1^{+}C-Peptide^{+}) or nonfunctional and immature (GCG^{+}C-Peptide^{+} and SST^{+}C-Peptide^{+}). Of note, the protocols with lower efficiency applied gene reporter iPSC lines to enrich β cells according to the β cell markers such as insulin, NKX6.1, or CD177 (Augsornworawat et al., 2021) (Cosentino et al., 2018, Mahaddalkar et al., 2020) (Russ et al., 2015) .

These advances generated a seven-stage protocol (Figure 1.6) for modelling iPSCs-derived human pancreas development. For example, little was known about the specification of pancreatic endoderm from trials on human embryogenesis. Following this, data from iPSCs differentiation suggested that SHH inhibitors such as SANT can be applied to induce pancreatic endoderm development from its foregut precursor (stage 4 in the protocol) (D'Amour et al., 2005) (Jennings et al., 2013) (Pagliuca et al., 2014) (Rezania et al., 2014).

Noggin or LDN and retinoic acid (RA), as inhibitors of FGFs and BMP,are used to induce foregut differentiation and pancreatic identity (steps 3 and 4). The pancreatic endoderm cells produced by this advanced protocol were optimized for expressing TFs PDX1, NKX6.1 and SOX9. Furthermore, three-dimensional (3D) culture possibly allowed closer mimicry of *in vivo* pancreas formation and elevated the expression of NEUROG3 at stage 5 (McGrath et al., 2015) (Wandzioch and Zaret, 2009) (Pagliuca et al., 2014) (Rezania et al., 2014).

Tri-iodothyronine (T3) is used to induce the maturation of endocrine progenitors into more mature β-cells at stages 6 and 7. Furthermore, adding an inhibitor to the AXL receptor tyrosine kinase, together with T3 and inhibition of TGFβ type 1 receptor kinase (also known as ALK5), induces the expression of MAFA, essential amendatory β-cell transcription factor, promoting β-cells maturation at stage 7 (Rezania et al., 2014) (Vanhoose et al., 2008).

The use of BMP inhibitors during endocrine specification generates bihormonal cells *in vitro*. This demonstrates a normal status of human fetal development since both insulin^{+}/glucagon^{+} cells emerge. However, their percentage is variable, ranging from ∼5% to 92% of endocrine cells *in vitro* differentiation models (Riedel et al., 2012). The iPSC-derived monohormonal (insulin) β-cells expressing PDX1, NKX6.1 and MAFA are ideal resulting cells. But, they are not precisely insulin-releasing cells or not well-responsive to different concentrations of glucose, a major challenge for clinical transplantation (Russ et al., 2015) (Riedel et al., 2012) (Riopel et al., 2014) (Rezania et al., 2014). Altogether, remarkable achievements have been reported towards development of in vitro iPSC-derived mature β-cells. However, still there are more tips to be considered from human development to translate this innovation to the clinics.

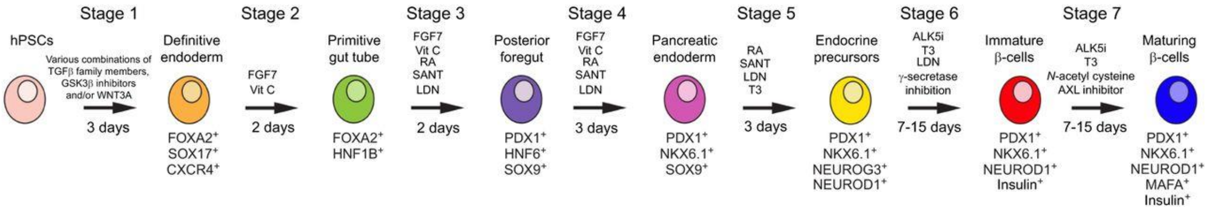


Figure 1-6: *In vitro* differentiation of human PSCs (hPSCs) towards pancreatic β-cells. Schematic representation of the steps of in vitro pancreatic differentiation via recent published, and improved methods. Major molecular players of each step are accentuated, as are the markers used at each step to direct differentiation along the proper course. SANT, Hedgehog signaling antagonist; LDN, BMP type 1 receptor inhibitor; T3, triiodothyronine; ALK5i, ALK5 inhibitor; Vit C, vitamin C (Rezania et al., 2014) (Pagliuca et al., 2014).

1.15 Gene Editing with CRISPR System

Clustered regularly interspaced short palindromic repeats (CRISPR) system belongs to bacteria and archaea to save them against viral pathogens. This system is recognized by palindromic repeats sequences ranging from 21 to 37 base pairs (bp). These sequences are interspaced by spacer segments that come from pathogens' DNA. CRISPR-associated (Cas) genes are close to the CRISPR locus. The CRISPR-Cas system can be categorized into three types (I-III) and 12 subtypes which have their distinct genetic information and structural variations (Barrangou, 2015) (Barrangou, 2015) (Makarova et al., 2015).

Adaptation, biogenesis, and interference stages are observed in the activities of the CRISPR-Cas system. In the adaptation stage, fragments of foreign DNA named protospacers are integrated into the CRISPR array as new spacers. These sequences generate a record of viral infections that will protect the bacteria from the next infection. Throughout the biogenesis stage, the CRISPR array is transcribed, resulting in a single long transcript termed pre-crRNA. Then it is trimmed to generate CRISPR RNAs (crRNAs) with only one spacer sequence. During the interference, the spacers in these crRNAs guide Cas proteins to cut foreign DNAs (Barrangou, 2013) (Wiedenheft et al., 2012) (Marraffini and Sontheimer, 2010).

Cas9 protein belongs to the type II CRISPR-Cas system. It needs two small RNAs, the crRNA and the trans-encoded crRNA (tracrRNA). TracrRNA generates a secondary structure interacting with the cas9 enzyme. Moreover, it has a complementary region which can bind to pre-crRNA. The dsRNA formed between tracrRNA and pre-crRNA is then processed by

RNase III to produce mature crRNA guides (Sapranauskas et al., 2011) (Anders et al., 2014) (Nishimasu et al., 2014).

In DNA editing, crRNA and tracrRNA are synthesized in a single RNA strand termed guide RNA (gRNA). A short DNA element (3-5 bp) named protospacer adjacent motif (PAM) is needed for gene targeting mediated by CRISPR/Cas9. Otherwise, the PAM sequence is mandatory for the activity of the Cas9 enzyme. The PAM element can be a component of the virus DNA or vector. The initial level in target recognition is the temporary binding of Cas9 to PAM DNA. This resulted in the melting of the two DNA strands adjacent to the PAM. The spacer sequence of the crRNA attaches to the opened DNA (6-8 bp in length), and then creates an RNA-DNA heterodimer which stimulates cleavage of the target DNA. Following recognition, the CRISPR-Cas9 generates a crRNA-specific DSB in the target DNA that can be repaired either by homology-directed repair (HDR) or non-homologous end joining (NHEJ) repair pathways (Jinek et al., 2012) (Sternberg et al., 2014) (Szczelkun et al., 2014).

Engineering nucleases such as ZFN, TALEN, and Cas9 are highly used in gene-targeting procedures. These enzymes generate a DSB at a specific position in DNA that can be corrected by NHEJ or HDR pathways. Additionally, there are several alternative error-prone DSB repair pathways including single-strand annealing (SSA) and breakage-induced replication (BIR). In SSA, reconnecting DNA ends with direct sequence repeats that occur without needing a homologous template. BIR repairs one-ended DSBs, an event that is induced by the collapse due to a replication fork (Shahryari et al., 2021a) (Shahryari et al., 2021b) (Jasin and Rothstein, 2013) (Mayle et al., 2015) (Symington, 2014).

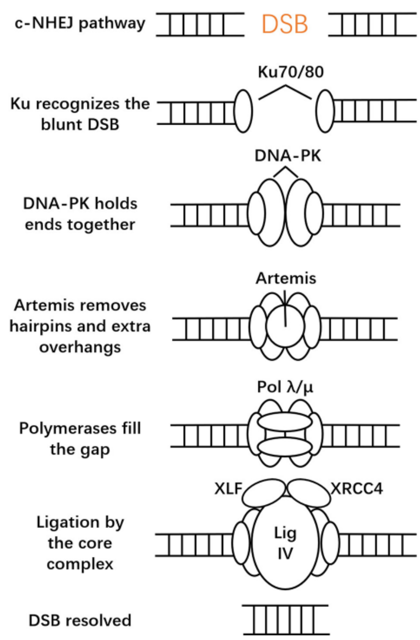


Figure 1-7: Non-homologous end joining repair pathway. Following CRISPR-Cas9 generated a DSB, NHEJ is triggered by the joining of the Ku heterodimeric complex. This eventually generates the major complex recognizing and binding to thebroken ends and holds them together. DNA-PK and Artemis join to this complex later. The ends will then be ligated by as XRCC4 ligase (Liu et al., 2018)

1.16 NHEJ Pathway

After occurring a DSB in DNA, NHEJ is initially activated. Compared to other DNA repair routes, the NHEJ is fast, predominant, and highly flexible (Salsman and Dellaire, 2017). There are two types of NHEJ pathways, canonical NHEJ (c-NHEJ) and alternative NHEJ (alt-NHEJ). c-NHEJ stabilizes the DSB mediated by translocations throughout the cell cycle (Bae et al., 2014). There are several NHEJ complexes that are involved in repairing a DSB. The main complex includes the Ku heterodimer (Ku80/70), the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), DNA ligase IV, and the X-ray repair cross-complementing protein 4 (XRCC4), the XRCC4-like factor (XLF). Ku heterodimer is made of two subunits (70 and 80 kD), which attach to blunt DSBs. In c-NHEJ, Ku interacts with DNA-PKcs resulting in a stable complex at the DSB site. Then, DNA-PK regulates DSB event and attracts Artemis nuclease. Artemis has DNA-PKcs-dependent 5′ and 3′ endonuclease activity on single-stranded overhangs and hairpin structures. The X family DNA polymerases such as pol mu and pol lambda incorporate missing nucleotides at the DSB ends (Moshous et al., 2001) (Weterings et al., 2010) (Daley et al., 2005). Next, the

DSBs will be joined by Ligase IV/XRCC4/XLF, which is controlled by DNA-PK. DNA is often corrected via the c-NHEJ pathway with an efficiency of approximately 90% (Dow et al., 2015).

1.17 HDR Pathway

HDR is a repair mechanism in the cells to correct double-strand DNA lesions. The HDR mechanism can only be applied by the cell in the presence of a homologous fragment of DNA in the nucleus, which mostly occurs throughout G2 and S phases of the cell cycle. The most common shape of HDR is homologous recombination which only occurs in the presence of a homologous DNA template. Following occurring a DSB, the MRE11-RAD50- NBS1 complex recognizes dsDNA generating a nick 15–20 bp from the 5′ -ends of the DSB. SGS1-DNA2 and EXO1 exonucleases finish the resection stage. Then, it proceeds to flanking dsDNA regions and ataxia telangiectasia mutated (ATM) kinase is recruited. ATM interacts with CtIP. The DNA ends are tethered by MRN, facilitating ATM activation (Makharashvili and Paull, 2015) (Kim and Mirkin, 2018) (Dupre et al., 2006).

The MRN complex is made of three subunits, MRE11, SAE2 and RAD50. SAE2 induces the endonuclease function of MRE11 and controls the resection stage in the course of cell cycle. RAD50 is responsible for chromosome maintenance and shows ATPase activity. RAD50 homodimer binds to DNA. Then, MRE11 can link to the ATPase heads of the RAD50 homodimer. RAD50, as the core part of MRN tethers DSB ends during homologous recombination (Hohl et al., 2011) (Cannavo and Cejka, 2014) (Mathiasen and Lisby, 2014). NBS1 and BRCA1 bind to MRE11 and recruit ATM, connecting the central MRN events to DNA failure response players. ATM induces phosphorylation of DDR cascades, including BRCA1, Chk2, and p53 (Lavin, 2008) (Williams et al., 2009). ssDNA can be produced by nuclease resection with the MRN-C-terminal binding protein-interacting protein (CtIP), and EXO1/BLM. BRCA1 is involved in HR by joining to MRN after DNA injury and binds directly to the resection factor CtIP (Sartori et al., 2007) (Mladenov et al., 2016).

BRCA1 supports binding RAD51 to ssDNA by expelling RPA. This can help BRCA2 bind to DSBs through the bridging protein PALB2. BRCA1 also halts the resection suppressor 53BP1. The formation of a RAD51 complex induces homologous screening by locating and pairing the 3′ -overhang with a homologous dsDNA and forming strand invasion termed single-end invasion. The two terminuses of the DSB are same with different functions. One of the ends is known as the "1st end," generates displacement loops (D-loops) structure searching for the homologous sequence and while the 2nd end involves in the next process (Zelensky et al., 2014) (Bunting et al., 2010) (Ma et al., 2017) (Kim and Mirkin, 2018).

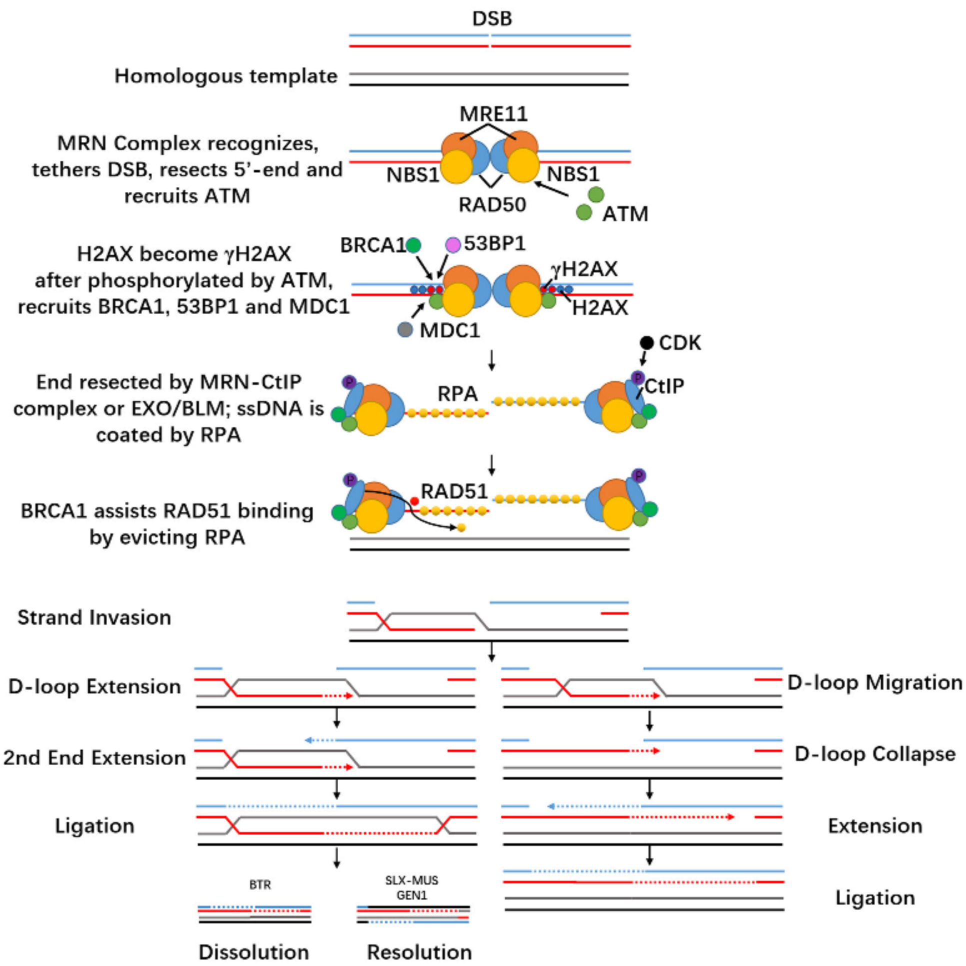


Figure 1-8: Homology-directed repair pathway. When DSB occurs during cell cycle, then DSB can be repaired via the HDR pathway if its ends are resected. Terminuses are occupied with different players and then bind to homologous duplex DNA to generate the D-loop structure. This structure is expanded by DNA synthesis. The second terminus connects to the D-loop and initiates extension. Ligation forms the Holliday junction, which can be cut by HJ resolvases into either crossover or non-crossover products. Following D-loop generation, amplification and branch migration occur that can result in D-loop translocation which collapses simply. Following collapse, the extended first terminus may hybridize to complementary ssDNA in the resected second terminus. Replicative extension of two terminuses and ligation yields non-crossover products (Liu et al., 2018).

Among gene-editing platforms, the CRISPR-Cas9 system is dramatically used for gene manipulations as it is efficient, fast, and easy to run. The CRISPR technology is being applied for different kinds of gene manipulations, including knock-out (KO) and knock-in (KI) aims, and generating short and long genomic DNA editing and/or rearrangements (Cong et al., 2013) (Mali et al., 2013) (Shahryari et al., 2020). It is well-documented that both NHEJ and HDR repair events are functional in all types of cells. But, when a DSB occurs, NHEJ is the major route to fill up the gap. On the other hand, NHEJ is a quick process and basically occurring in less than one hour; however, the HDR pathway is much slower finishing in several hours (Mao et al., 2008) (Jasin, 1996).

As mentioned earlier, the ku70-ku80 heterodimer, DNAPK and XRCC4 are the fundamental players in the NHEJ event. The KU70-KU80 complex binds to DSBs along with DNAPK connecting both DNA ends together. Later, XRCC4 and LIG4 correct the gap. As NHEJ produces indels at DSB, resulting in frameshifts, this can apply to the disruption of genes in loss-of-function (LOF) studies. However, the presence of a donor template is very critical for HDR to be recruited (Guschin et al., 2010) (Davis and Chen, 2013) (Her and Bunting, 2018) (Hug et al., 2016) (Branzei and Foiani, 2008).

The HDR pathway is frequently used for generating precise gene-editing such as gene KI or point mutations. To correct a DSB, HDR reveals less functionally compared to NHEJ. For increasing precise gene editing, several methods have been emerged to downregulate NHEJ and/or boost HDR with small chemical compounds. Accordingly, elevated cellular death due to cytotoxic stress because of lipofectamine transfection is another barrier to efficient gene editing via CRISPR-Cas9 platform. Furthermore, a high level of Cas9 nuclease is toxic for the cell and causes cell death, resulting in decreased efficiency for gene editing (Liu et al., 2018) (Ardehali et al., 2011) (Ihry et al., 2018).

1.18 Aims of the Thesis

1.18.1 Aim 1: Functional Analysis of the T2D Risk Region Upstream of *INK4* for β-cell Proliferation and Insulin Secretion

Six SNPs linked to T2D located upstream of the INk4 locus suggests a potential role of that region in the regulation of the neighboring genes or INK4 genes in a *cis* or *trans* manner. First, we generate *Δ INK4 T2D Risk Region* in iPSCs using CRISPR/Cas9 and delete the whole region in human iPSCs encompassing all six SNPs. After the generation of knockout cells, we performed quality control experiments, i.e., genotyping, karyotyping, pluripotency and differentiation tests. Then, we used validated cells for the next applications, such as generating the insulin-producing β-cells.

Then, upon Upon the differentiation of iPSCs to insulin-producing β-like cells, we measured INK4 genes expression at RNA and protein levels at the early stages of differentiation (progenitor cells) and the final stages (immature β-cells and mature β-cells) of differentiation by qPCR and western blotting respectively. Of special interest was the proliferation rate, determined by staining for proliferation markers like Ki67 and EdU labeling, and quantitative FACS analysis. Finally, insulin secretion was measured in β-like cells by GSIS. The differential expression of RNAs at the early stages of differentiation (progenitor cells) and the final stages (immature β-cells and mature β-cells) of differentiation were also measured by RNA-Seq method.

1.18.2 Aim 2: Design a Strategy for Increasing Gene Editing Efficiency

Here, we designed new procedures to improve DNA editing efficiency for CRISPR/Cas9 system. Like previous approaches, we downregulated NHEJ, and increased cell viability to enhance DNA editing efficiency in human iPSCs. We aimed to target the pluripotency gene of *SOX2* as it is highly expressed in iPSCs, and it is easy to monitor SOX2-reporter expression. The gene of *SOX2* was targeted in two different iPSC lines with the targeting plasmid, Addgene ID89991 (Balboa et al., 2017), to produce SOX2-Thosea asigna virus 2A like peptide-tandem dimer Tomato (SOX2-T2A-tdTomato) reporter iPSC lines. The process had a quantifiable readout as it was easy to quantify the percentage of the cells expressing the reporter by fluorescent activating cell sorting (FACS) machine. To interfere with NHEJ, we incorporated expression constructs of short hairpin RNA (shRNA) into the CRISPR/Cas9 expression vector. Our shRNAs are designed to downregulate XRCC4 and DNAPK, the major players in the NHEJ repair pathway. In parallel, to interfere with stress-induced apoptosis and to increase cell survival, we integrated an expression construct of the miRNA-21 into the Cas9 expression plasmid. Thus, we developed novel approaches to improve gene

KI efficiency for CRISPR-Cas9 by downregulating the activity of the NHEJ pathway and by enhancing cell viability in human stem cells such as iPSCs.

2. Results

2.1 Increasing CRISPR/Cas9 Gene Editing Efficiency

2.1.1 CRISPR/Cas9 Mediated SOX2-T2A-tdTomato KI in Human iPSCs

The NHEJ and HDR repair pathways compete to repair a gap generated by the CRISPR/Cas9. Hence, NHEJ is a barrier to accurate gene integration or editing. Some small compounds such as SCR7 and NU7026 target the important players of NHEJ and can improve the efficiency of CRISPR/Cas9 for precise gene Knock-in (KI). Furthermore, increasing cell survival can improve the DNA editing efficiency of the CRISPR/Cas9 system. Lipofectamine transfection or the overexpressed Cas9 enzyme can lead to cellular stress resulting in massive cell death. This can reduce the genome targeting efficiency of CRISPR/Cas9. Herein, we could add to the efficiency of HDR-mediated gene targeting for CRISPR/Cas9 platform with two simple strategies. First, if main players of NHEJ pathway are downregulated by shRNAs, this can potentially improve the DNA editing efficiency of CRISPR/Cas9. Second, we wondered if improving cell survival via ectopic expression of an anti-apoptotic gene such as miR21 could also increase gene editing efficiency. To evaluate these two hypotheses, we looked for a highly expressed gene in iPSCs. Hence, *SOX2* locus was targeted to produce *SOX2-T2A-tdTomato* reporter iPSC line. As the TF SOX2 is highly expressed in iPSCs we could quickly evaluate gene targeting efficiency by measuring fluorescent activity. To insert *T2A-tdTomato* reporter construct to the *SOX2* locus by the CRISPR/Cas9 system, we used our two developted human iPSC lines (iPSC line I: HMGUi001-A; 46, XX) (Wang et al., 2018) and (iPSC line II: HMGUi002-A; 46, XY) (Wang et al., 2016b). The *T2A-tdTomato* sequence was correctly added before the termination codon of the *SOX2* gene following HDR-mediated gene targeting (Figure 2.1a). We could confirm the homologous recombination at the *SOX2* gene region was by 5′ and 3′ genomic PCR analyses spanning the homologous recombination borders, as shown in Figure 2.1b.

In order to evaluate the random incorporation of *T2A-tdTomato,* we carried out further PCR analyses with primers which anneal to the backbone of the targeting vector and the T2A-tdTomato sequence. The vector backbone is released if the non-digested targeting vector is inserted at the right region at the borders of both homology arms. On other hand, if random insertion happens, most probably the backbone close to homology arms would also be integrated. The PCR output indicated an 1888 bp amplicon particular for the targeting plasmid. The absence of that product in the targeting cells, declined the random integration of the T2A-tdTomato into the genome locus (Figure 2.1c).

In brief, the cells that received CRISPR/Cas9 vectors, expressed Cas9-GFP fusion protein. Next, if HDR happens, the SOX2-T2A-tdTomato fusion protein is expressed in GFP+ cells. Three days after the transfection, GFP+ cells were sorted by FACS, and the percentage of cells expressing T2A-tdTomato was measured (Figure 2.2). The transfections efficiency of iPS cells were measured by the percentage of GFP+ cells using FACS analysis, ranging between 30 and 50% (Figure 2.3a and 3.3b). The T2A-tdTomato+ cells were propagated and cultured in 2D and 3D to produce cells populations for further analyses (Figure 2.4a-c).

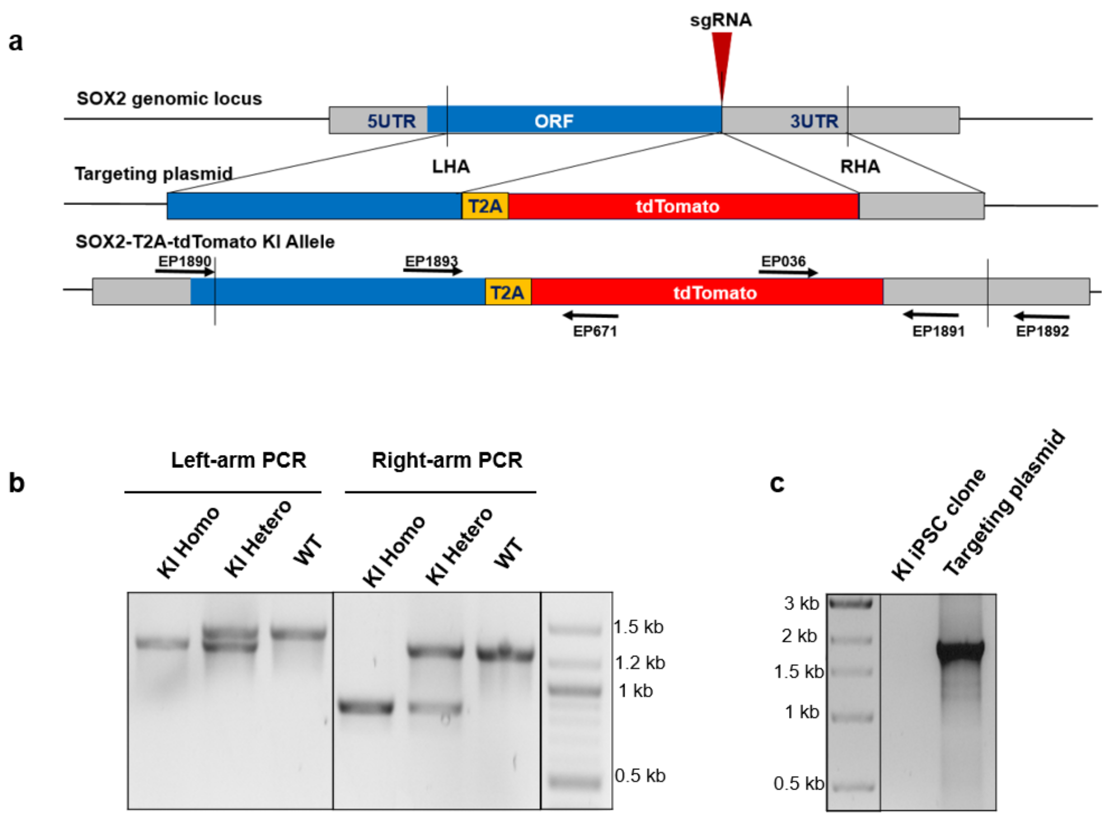


Figure 2-1: A schematic representation of DNA KI method. A: HDR-mediated KI at SOX2 locus in iPSCs. Human SOX2 gene is an individual exon and embedded within chromosomal region of 3q26.*3*-q27. Targeting plasmid carrying LHA, tdTomato reporter and RHA was applied as a template to add the reporter to the 3′ terminus of SOX2 coding region. B: The precise gene integration at the SOX2 gene was approved by 5′ (left) and 3′ (right) genomic DNA PCR test spanning the boundaries of homologous recombination (1401bp and 1310bp bands are representing KI and WT alleles, respectively, in the left-arm PCR while 1294bp and 753bp bands are representing for KI and WT alleles, respectively, in the right-arm PCR). C: PCR analysis for evaluating random insertion (PCR product size: 1888bp).

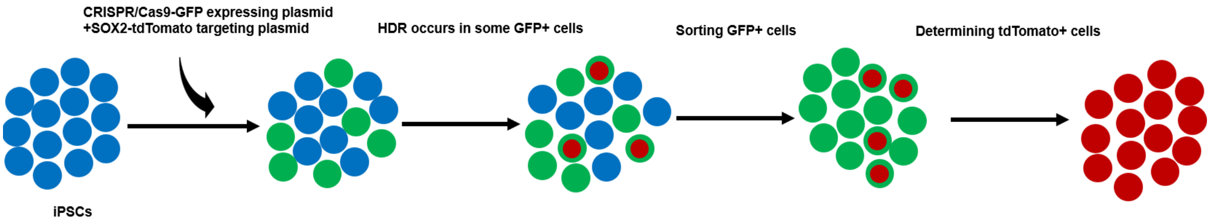


Figure 2-2: Cloning screening and measuring the efficiency of gene KI. Targeting plasmid together with CRISPR/Cas9 expressing vectors were added to the stem cells. The cells received the vector, expressed Cas9 that was fused to GFP protein. In these GFP+ cells, if precise insertion occurs then, the cells express SOX2-T2A-tdTomato fusion protein. Following 72 hours post transfection, GFP+ cells were first sorted via FACS. Then among the sorted cells, T2A-tdTomato cells were measured representing the percentage of gene KI.

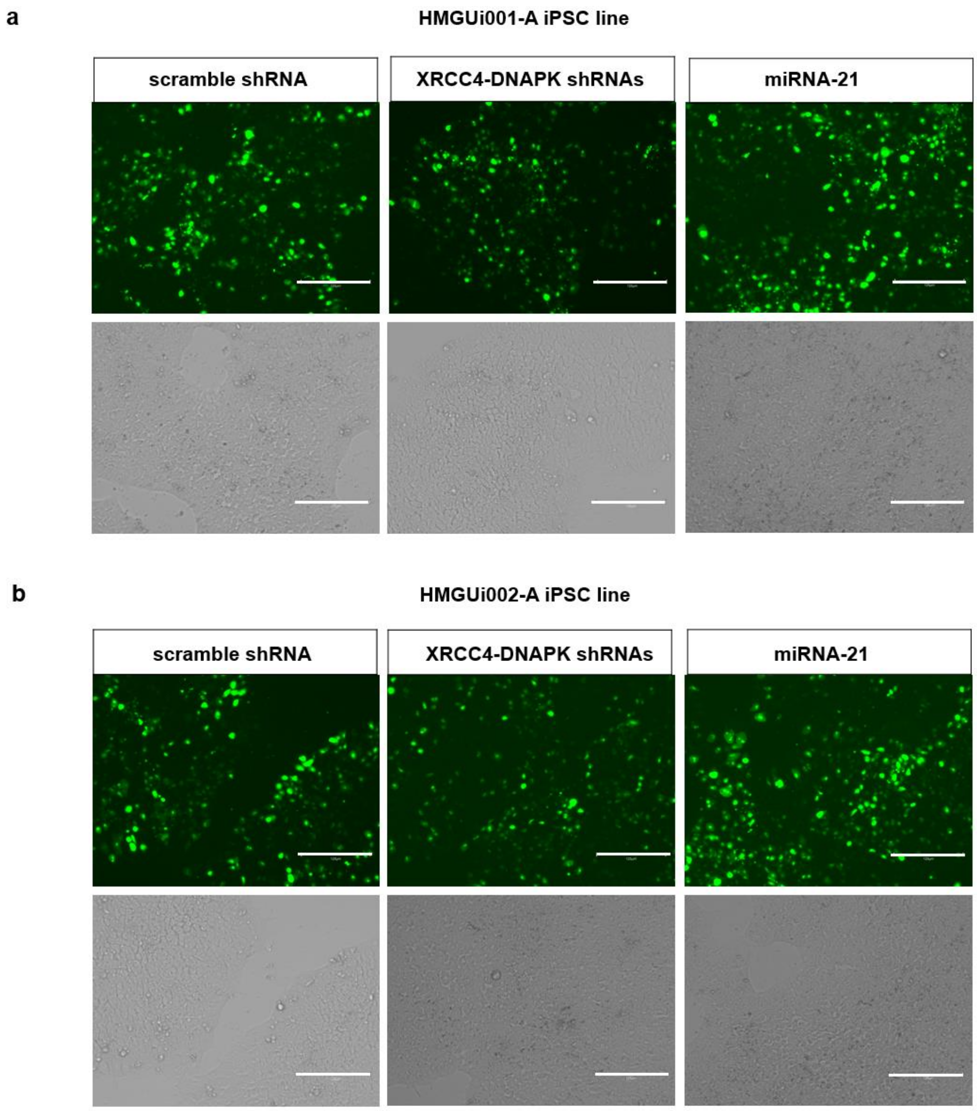


Figure 2-3: Non-quantitative control of transfection efficiency. A and **B:** GFP expression 48 hours after the transfection in iPSC lines I and II (Scale bar: 125µm).

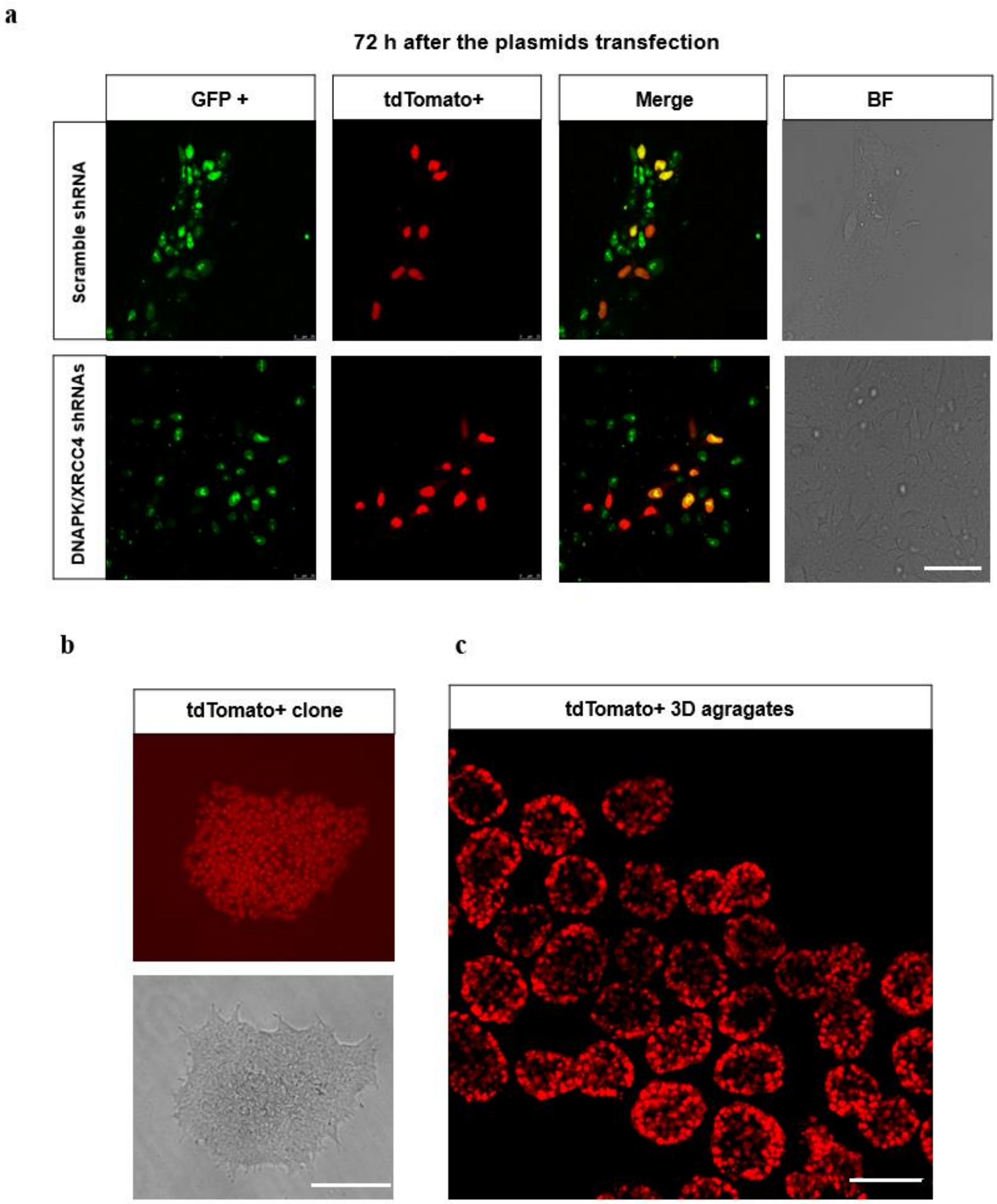


Figure 2-4: Clone expansion. A: 72 hours following the transfection, some of the GFP+ cells express T2A-tdTomato (Scale bar: 50µm). BF stands for bright field. **B:** After the sorting, single cells are cultured and expanded (Scale bar: 200µm). **(c)** The SOX2-T2A-tdTomato reporter iPSCs are cultured in 3D to produce cell clusters (Scale bar: 100µm).

2.1.2 shRNA-mediated Downregulation of NHEJ Increases HDR-mediated SOX2 *T2A-tdTomato* KI

To address the first hypothesis and downregulate NHEJ, we knock-downed *DNAPK* and *XRCC4* genes via shRNA. Instead of using small compounds targeting *DNAPK* and *XRCC4* genes, we generated CRISPR/Cas9 vector expressing their corresponding shRNAs cassettes (Figure 2.5). To this end, shRNAs targeting *DNAPK* and *XRCC4* genes were designed. Initially, the shRNAs efficiency was measured. The downregulation of *DNAPK* and *XRCC4* genes mediated by shRNAs was confirmed by qPCR. The efficiency gene knockdown was 0.15 and 0.11-fold for *DNAPK* and *XRCC4*, respectively (Figure 2.6a). Following determining shRNAs efficiency, the highly efficient ones were added into a single CRISPR/Cas9 vector generating Cas9-GFP/sgRNA-shRNAI-shRNAII construct. At the same time, each small RNA cassette was regulated and monitored by its own U6 promoter. Then, we examined the impact of targeting *DNAPK* and *XRCC4* genes via shRNAs on the efficiency of HDR-mediated *SOX2* KI in two independent iPSC lines. The efficiencies of KI for shRNAI-shRNAII targeting *DNAPK* and *XRCC4* were15.52% in HMGUi001-A line (n = 3 and P-value < 0.001), and 22.48% in HMGUi002-A iPSC line (n = 3 and P-value < 0.01). Moreover, the efficiencies of scramble shRNA were 8.48% and 12.5% in HMGUi001-A and HMGUi002-A iPSC cells, respectively. Altogether, downregulation of *DNAPK* and *XRCC4* by shRNAs co-expressed with the CRISPR vector improved SOX2 KI 1.83-fold in HMGUi001-A and 1.79-fold in HMGUi002-A (almost 2-fold) (Figure 2.6a, Figure 2.6b and Figure 2.8).

2.1.3 Cas9/miR21 Expression System Increases the Efficiency of SOX2 *T2A-tdTomato* KI

To test the second hypothesis and increase cell survival during gene editing, the small RNA miR21 was used as it concomitantly downregulates apoptotic genes, in particular Caspase3. It was estimated that ectopic expression of miR21 could improve cell survival and eventually increase the CRISPR/Cas9 gene targeting efficiency. To this end, miR21 sequence was added to the Cas9 expression plasmid producing the Cas9-GFP-sgRNA-miR21 cassette (Figure 2.5). To calculate DNA editing efficiency, the same transfection and screening procedures mentioned earlier were used e.g. the percentage of SOX2-tdTomato+ cells among the Cas9-GFP+ cells was measured. The efficiency of DNA integration using miR21 was 24.66% (2.9-fold) and 32.4% (2.59-fold) for the HMGUi001-A and HMGUi002-A iPSC lines, respectively (n =3 and *P*-value < 0.0001) (scramble shRNA was used as negative control). On other words, miR21 increased the efficiency of KI almost 3-fold compared with the scramble control (Figure 2.6b, Figure 2.6.c and Figure 2.8). Then, we examined if miR21 can increase cell survival in the HMGUi001-A iPSC line. To address this, the number of alive cells at 8 and 24 hours after the transfection was measured. We reported no significant difference in the survival rate of transfected iPSCs at 8 hours following the transfection (95% vs 93%). However, the survival rate improved at 24 hours (92% survival rate in the presence of miR21 versus 69% in its absence (n =2 and *P*-value < 0.01) (Figure 2.7a and Figure 2.7b). Altogether, the improved viability of transfected cells increased the efficiency of DNA editing for the CRISPR/Cas9 platform.

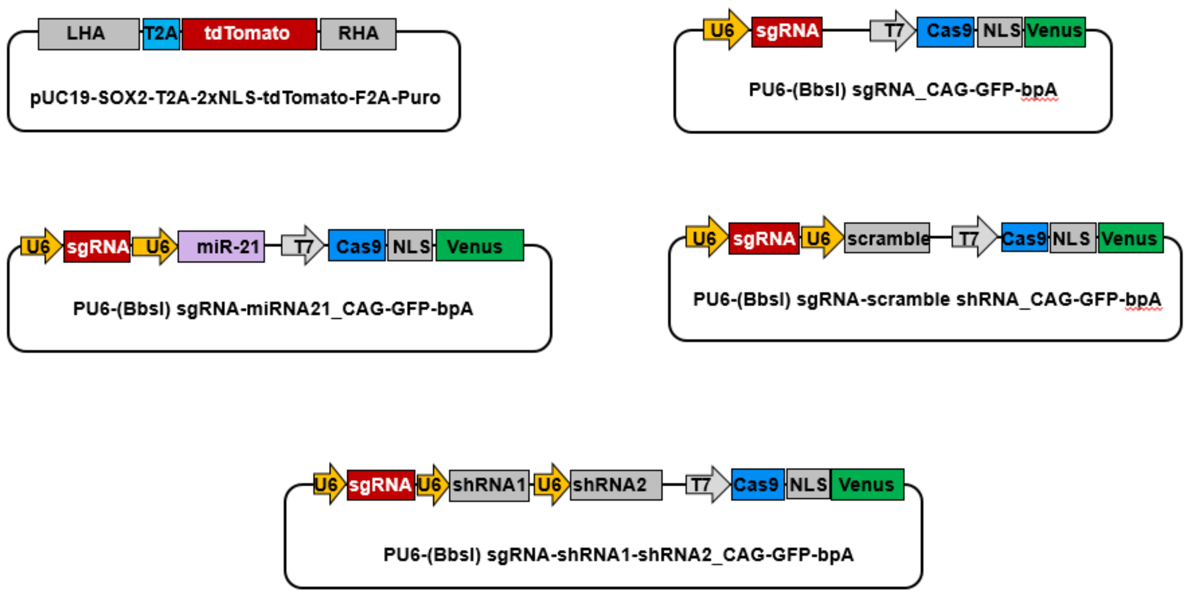


Figure 2-5: CRISPR/Cas9 plasmids. The SOX2 LHA-T2A-SOX2 RHA targeting plasmid and cassettes expressed CRISPR/Cas9 systems (carrying sgRNA, sgRNA-shRNAI-shRNAII, sgRNA-scramble RNA and sgRNA-miR21). All our CRISPR/Cas9 plasmids expressed sgRNA together with shRNA or miR21.

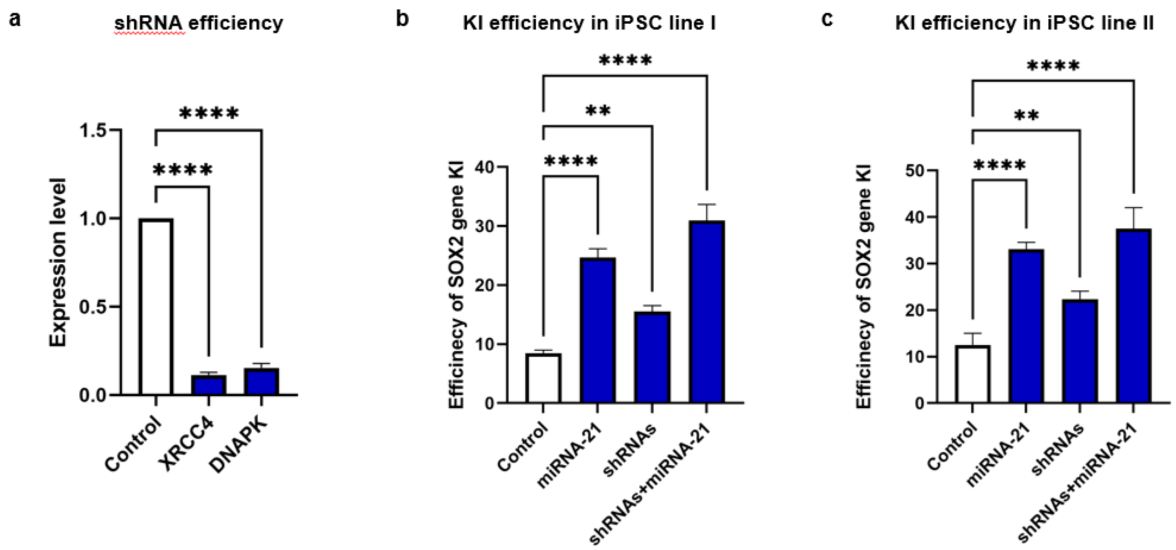


Figure 2-6: Efficiency of HDR mediated Gene editing in iPSCs. A: Measuring shRNA efficiency by qPCR. The data were represented as mean + S.D. (n = 3). B and C: HDR mediated gene targeting efficiencies for production of SOX2-tdTomato reporter in our two iPSC lines, I and II. The iPS cells transfected with scramble shRNA were used as control. The data were shown as mean + S.D. (n = 3). (** $p < 0.01$ and ****$p < 0.0001$).

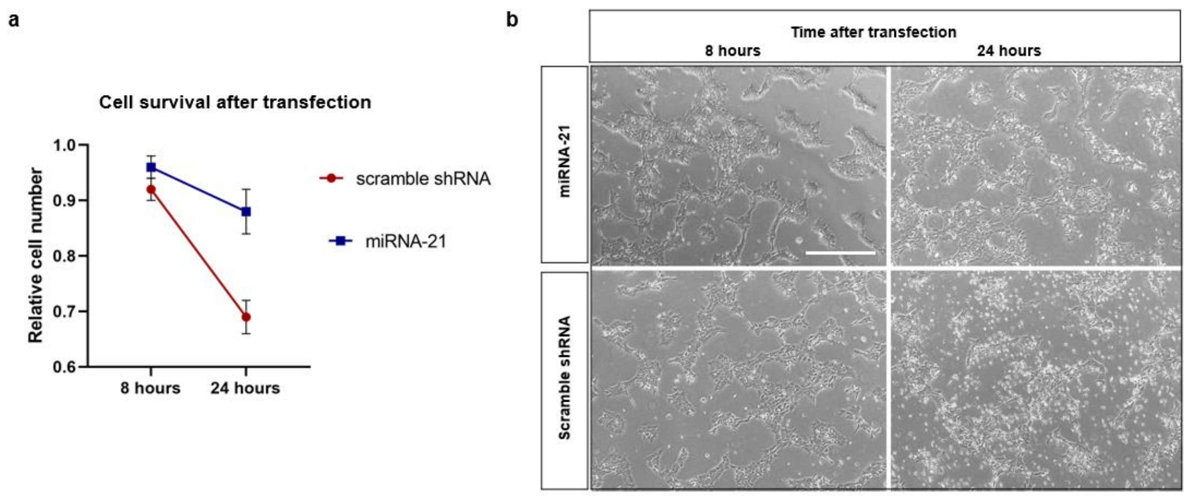


Figure 2-7: Cell survival assay. A: Alterations in relative cell numbers after transfection with DNA editing and/targeting vectors harboring scramble shRNA and miR21. Cell survival was measured at two-time courses, 8 and 24 h following treatment with trypan blue dye. The data were shown as mean + S.D. (n = 2), $p < 0.01$. B: Representative images of human iPSCs at 8- and 24-hours following transfection with scramble shRNA and miR21 (scale bar: 1050μm).

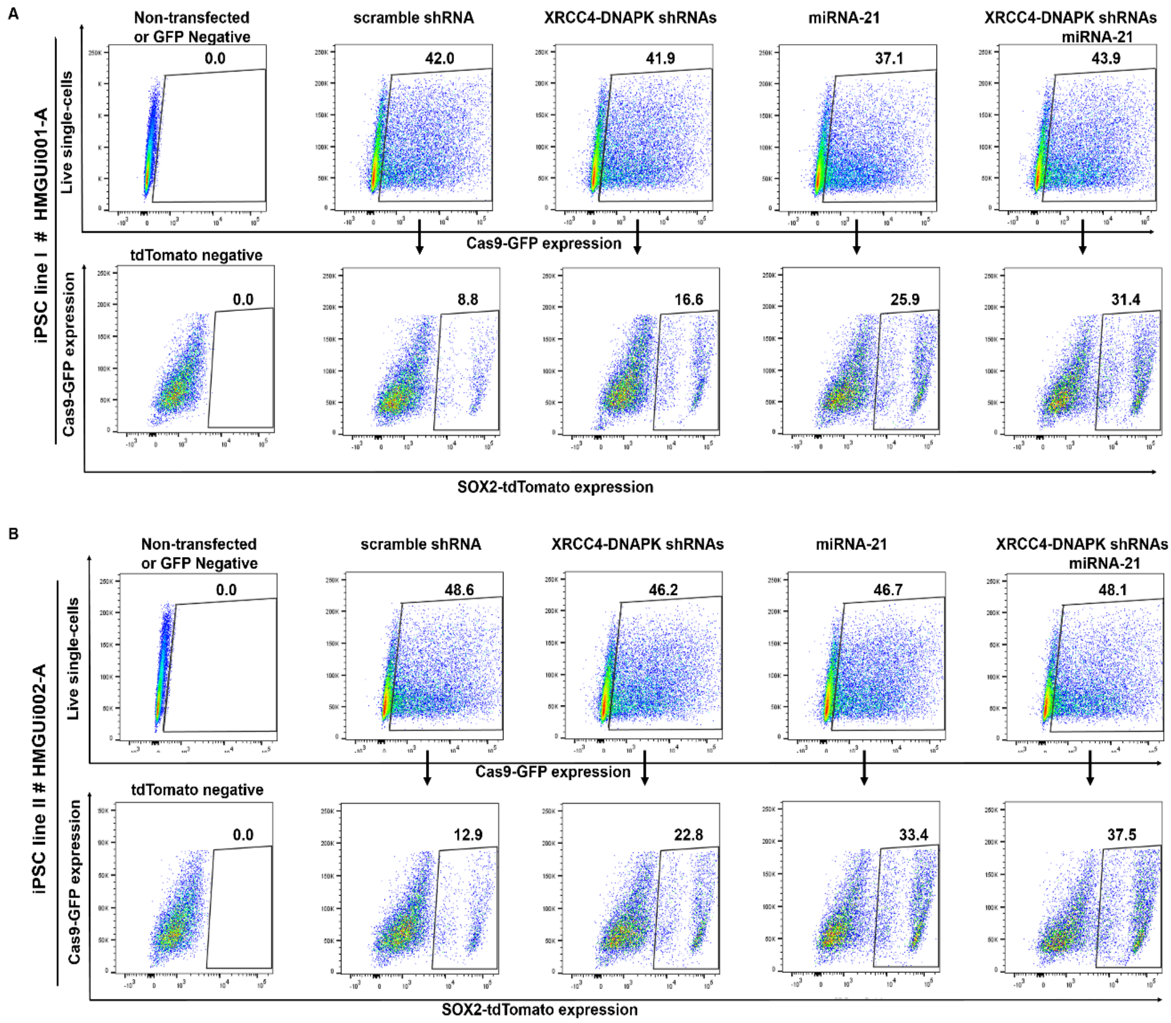


Figure 2-8: Our strategies for measuring transfection efficiency and DNA editing efficiencies for all plasmids conditions. A and **B:** The expression of Cas9-GFP fusion was detected by FACS, determining the transfection efficiency in our both iPSC lines, I and II for plasmids expressed sgRNA-scramble shRNA, sgRNA-miR21, sgRNA-shRNAI-shRNAII and shRNAs-miRNA. The non-transfected iPSCs were considered negative control. Among the GFP+ cells, the number of T2A-tdTomato expressing cells reflecting gene integration efficiency were measured. Importantly, the cells that received Cas9 fused toGFP vector, but not SOX2-T2A-tdTomato targeting plasmid were considered as negative control.

2.1.4 Additive Effect of miR21 and shRNAs on the Gene-editing Efficiency of CRISPR/Cas9 in iPSCs

To evaluate an accumulative impact of increasing cell survival and downregulating NHEJ on gene KI efficiency of CRISPR/Cas9, *DNAPK- XRCC4* shRNAs and miR-21 plasmids were combined. To calculate this, we used a mixture of Cas9-GFP-sgRNA- miR21, Cas9-GFP-sgRNA-shRNAs and SOX2-T2A-tdTomato vectors transfect our two iPSC lines. Using both miRNA and shRNA expressing constructs improved the gene-editing efficiency in transfected cells to 31.4% versus miR21 plasmid alone (25.9%) or shRNA plasmids alone (16.6%). In the second iPSC cell line mixture of the plasmids improved the DNA editing efficiency in transfected cells to 37.5% from 22.8% and 33.4% using shRNA or miRNA plasmid alone (n =3 and *P*-value < 0.0001) (Figure 2.6 and Figure 2.8). The homologous recombination rate marginally enhanced upon combining both approaches.

2.1.5 Differentiation of SOX2-*td*Tomato Reporter iPSC Line into Pancreatic Progenitors

Temporary expression of miR21 could influence pluripotency and/or differentiation capacity of edited iPSCs. To examine this, the protein expression of major pluripotency markers, SOX2 and OCT4, were measured by immunostaining. Moreover, the edited HMGUi001-A iPSCs were differentiated towards pancreatic progenitors (PP) to observe if they differed from control iPSCs. The differentiation process was performed according to an altered Rezania protocol in 3D culture. The aggregates were collected on day 10 following differentiation (at the pancreatic progenitors stage). Then, we stained the aggregates for the two earliest pancreatic TFs, PDX1 and NKX6.1. The miR21 edited iPSCs showed high expression of SOX2 and OCT4; however, they were undetectable in the control iPSCs (Figure 2.9). Moreover, we could efficiently differentiate these cells to pancreatic progenitors as showed by a high expression of PDX1 only cells or PDX1 and NKX6.1 double-positive cells at time point, day 10 of the differentiation course (Figure 2.10).

Our methods showed that co-expression of shRNAs downregulating the leading players of NHEJ, XRCC4, and DNA-PK, together with miRNA-21 and Cas9, increase DNA editing efficiency for the CRISPR system. Our new CRISPR/Cas9 strategy dispenses more plasmids or small molecules and still interferes with the NHEJ event or stress-simulated cellular apoptosis. This increased gene integration diminishes the downstream assignments essential for screening and recognizing the cells having a precise DNA editing mediated by HDR-based genome editing.

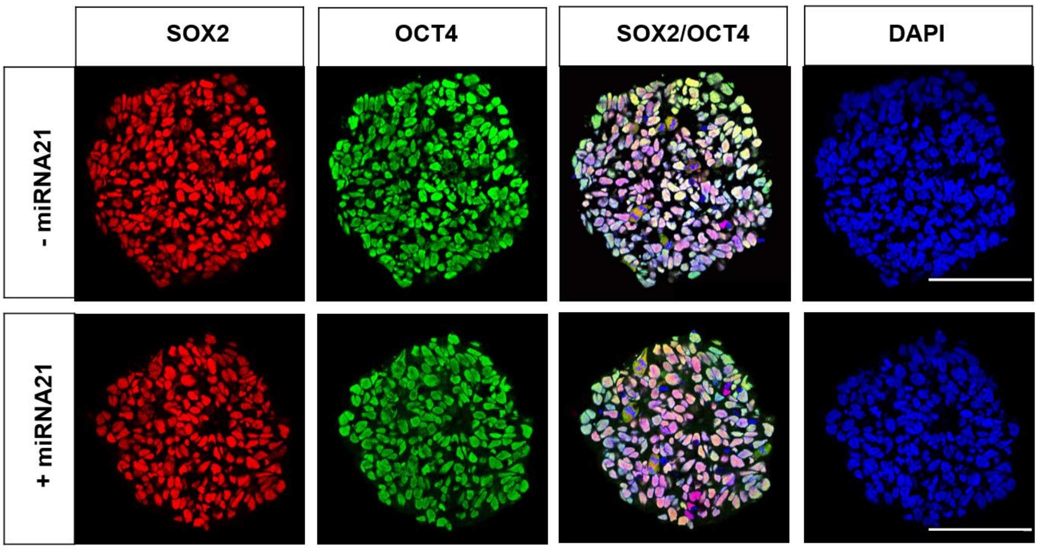


Figure 2-9: Non-quantitative expression of pluripotency markers in miR-21 treated iPS cells. Immunostaining of luripotency markers, SOX2 and OCT4 in the SOX2-*td*Tomato reporter iPSC line or miRNA-21 treated and non-treated iPS cells (scale bar: 100µm).

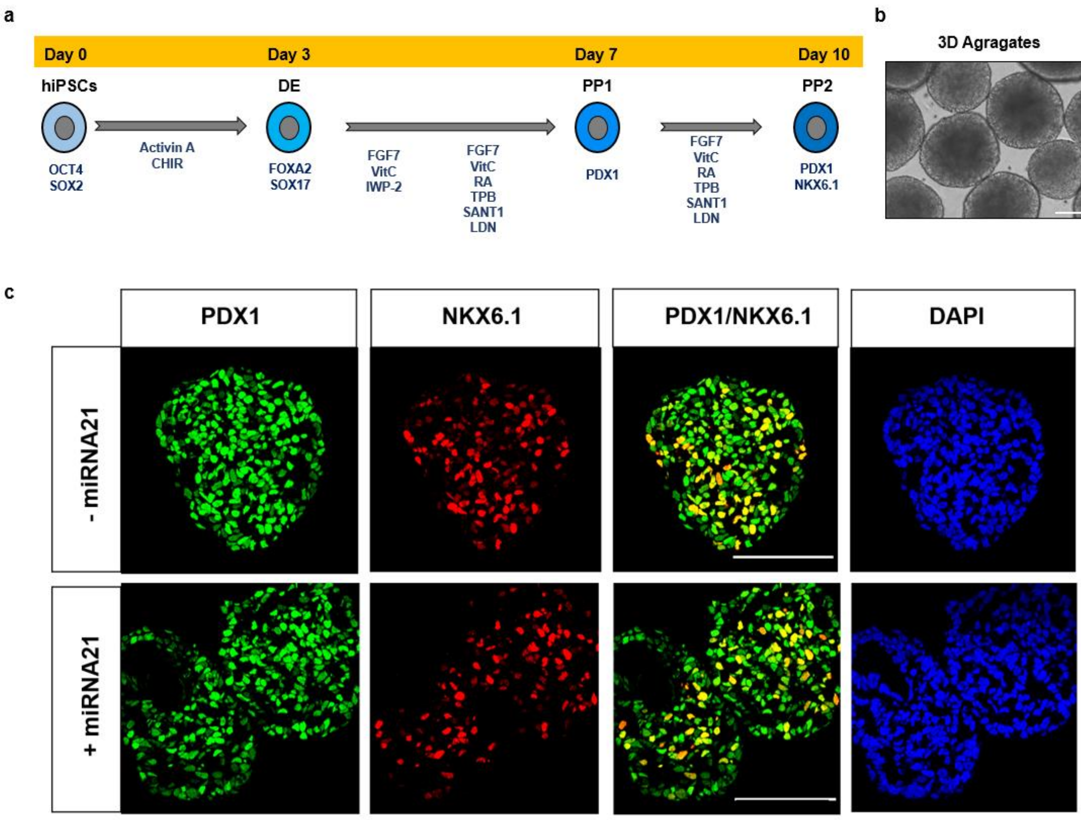


Figure 2-10: Differentiation of the SOX2-tdTomato reporter iPSCs into the pancreatic progenitors. A: Schematic representation of protocol for differentiating the iPSC line I (also named HMGUi001-A) into pancreatic progenitors. *FGF7* (Fibroblast Growth Factor 7), VitC (Vitamin C), RA (Retinoic Acid), CHIR (CHIR99021), IWP-2 (Inhibitor of Wnt Production-2), *SANT-1* (Inhibitor of hedgehog signaling). B: The iPSCs transferred from 2D to 3D culture. Bright-field picture of 3D cell clusters at stage 5 of differentiation (scale bar: 100µm). C: Immunostaining for TFs PDX1 and NKX6.1 in pancreatic progenitors in the presence and absence of miR21 (scale bar: 100µm).

2.2 Functional Analysis of INK4 Locus in Endocrine Development and Function

2.2.1 In silico Analysis for Chromatin State of INK4 T2D Risk Region

There is a T2D risk region (8 kb) inside the human INK4 locus but outside the gene coding region. Interestingly, six SNPs related to T2D are located in that genomic block. The molecular mechanisms of how these SNPs are linked to diabetes prevalence are still questionable. To address this, we decided to study the potential role(s) of that 8 kb genomic block. This block is located about 10 kb downstream of the non-coding RNA *ANRIL* gene.

The SNPs include *rs2383208* (A > G / A > T), *rs10965250* (G > A), *rs7018475* (T > G), *rs1333051* (A > G / A > T), *rs10757283* (C > A / C > T), and *rs10811661* (T > C). Interestingly, according to the published data as shown in Table 2.1 all these SNPs are associated with T2D.

Table 2.1. Association of INK4 SNPs to T2D

Name	Allele	Trait	p-value	Reference
rs2383208	A > G / A > T	Type 2 Diabetes	2E-29	(Takeuchi et al., 2009)
		Type 2 Diabetes	3E-17	(Li et al., 2013)
		Type 2 Diabetes	3E-6	(Tabassum et al., 2013)
rs10965250	G > A	Type 2 Diabetes	1E-10	(Voight et al., 2011)
rs7018475	T > G	Type 2 Diabetes	3E-8	(Huang et al., 2012)
		Glucose homeostasis traits	5E-6	(Palmer et al., 2015)
rs1333051	A > G / A > T	Type 2 Diabetes	6E-10	(Parra et al., 2011)
rs10757283	C > A / C > T	Type 2 Diabetes	5E-3	(Cheng et al., 2011)
rs10811661	T > C	Type 2 Diabetes	5E-8	(Saxena et al., 2007)
		Type 2 Diabetes	8E-15	(Scott et al., 2007)
		Type 2 Diabetes	5E-6	(Zeggini, 2007)
		Type 2 Diabetes	7E-7	(Timpson et al., 2009)
		Type 2 Diabetes	7E-6	(Manning et al., 2012)
		Type 2 Diabetes	1E-18	(Hara et al., 2014b)
		Fasting glucose-related traits	1E-27	(Mahajan et al., 2014)
		Fasting plasma glucose	9E-12	(Hwang et al., 2015)

Our first aim was to understand the chromatin status of this genomic block to check if the region harbors binding sites for well-known TFs or contains active histone or DNA marks. To this end, we used several web tools such as UCSC Genome Brower (https://genome.ucsc.edu), HaploReg https://pubs.broadinstitute.org/mammals/haploreg/haploreg_v4.php) and Islet Regulome Browser (http://pasqualilab.upf.edu/app/isletregulome). According to UCSC Genome Brower, the data analysis on the 8 kb genomic block shows high enrichment for the binding sites of major regulators involved in pancreas development, such as PDX1 and NKX2.2 in adult islets (Figure 2.11A). Other TFs such as FOXA2, NKX26.1 and MAFB showed high enrichment for binding at INK4-T2D risk region in adult islets (Figure 2.11B). Furthermore, as illustrated in Figure 2.11C, active histone marks, H3K27ac, H3K4me1, H3K4me3 and H3K436me3 showed high chromatin enrichment in the T2D risk region at the INK4 genomic region in adult islets.

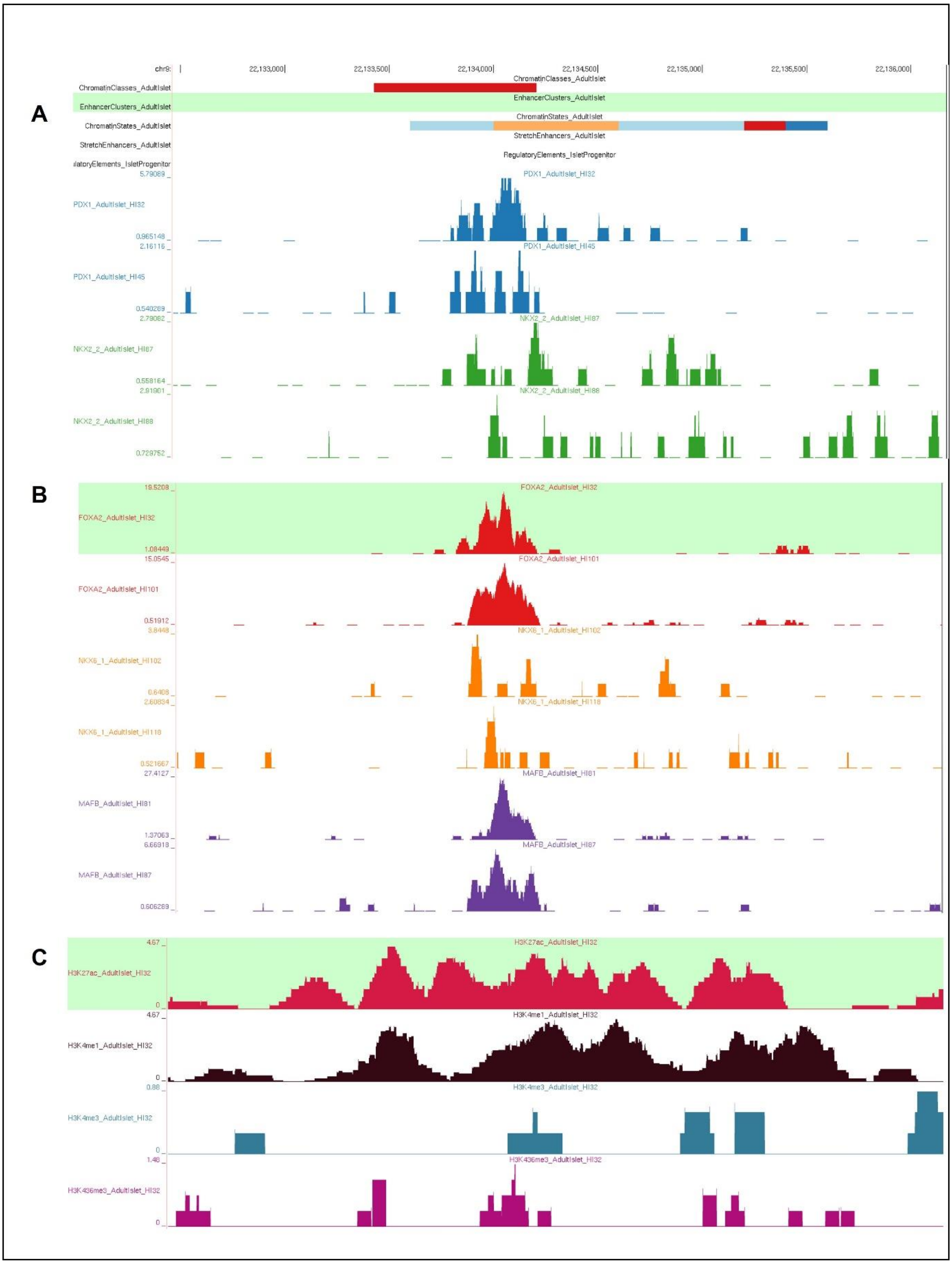


Figure 2-11: Chromatin state and chromatin enrichment for binding site of major regulators in pancreas development. A: Chromatin enrichment for binding sites of PDX1 and NKX2.2 at INK4-T2D risk region in adult islets. **B:** Chromatin enrichment for binding sites of FOXA2, NKX26.1 and MAFB at INK4-T2D risk region in adult islets. **C:** Chromatin enrichment for active histone marks, H3K27ac, H3K4me1, H3K4me3 and H3K436me3 at INK4-T2D risk region in adult islets. All these data are derived from UCSC genome browser.

Further analysis of the 8 kb genomic block using HaploReg web tool showed that the chromatin contains an H3K4me1 histone mark which indicates/hints at the presence of an active enhancer in the pancreatic islets (Figure 2.12A). In line with this, analysis using the Islet Regulome Browser also revealed that the chromatin is open and contains a potential active enhancer. This tool also showed that transcription factors FOXA2 and MAFB have binding sites inside the enhancer in the adult islets (Figure 2.12B).

Figure 2-12: Histone marks and evidence of binding TFs at the INK4-T2D risk region. A: Chromatin at INK4-T2D risk region represents H3K4me1 histone marks indicating a potential enhancer in pancreatic islets. The data is derived from HaploReg v4.1 browser. B: The INK4-T2D

risk region shows open chromatin and encompasses an active enhancer that has binding sites for FOXA2 and MAFB in adult islets. The data is derived from Islet Regulome browser.

ChIP-seq data from previous studies in our lab (Wang et al., 2018) (XM001 Rezania protocol; stage 6 of differentiation) were analyzed to see the status of chromatin modifications at the INK4 locus. We focused on the 8 kb genomic block to check the chromatin status of those six SNPs. The H3K27ac histone mark is observed in positions of the two SNPs, *rs10811661* and *rs10757283*. Otherwise, these two SNPs are located in an active chromosomal region; however, according to this analysis, most SNPs are embedded in an inactive chromatin region (Figure. 2.13).

Furthermore, Chromatin interactions at the INK4 locus and ChIP data were evaluated together. As illustrated in Figure. 2.14, several proximal and distal chromatin elements interact with the promoter and regulatory regions of the non-coding RNA *ANRIL*, *CDKN2B* and *CDKN2A*. As mentioned earlier, the INK4 T2D risk region is located about 10 kb downstream of the *ANRIL* gene. The diagram clearly represents that the INK4 T2D risk region has interactions with INK4 genes promoters, particularly *ANRIL* and *CDKN2B* (Figure 2.14).

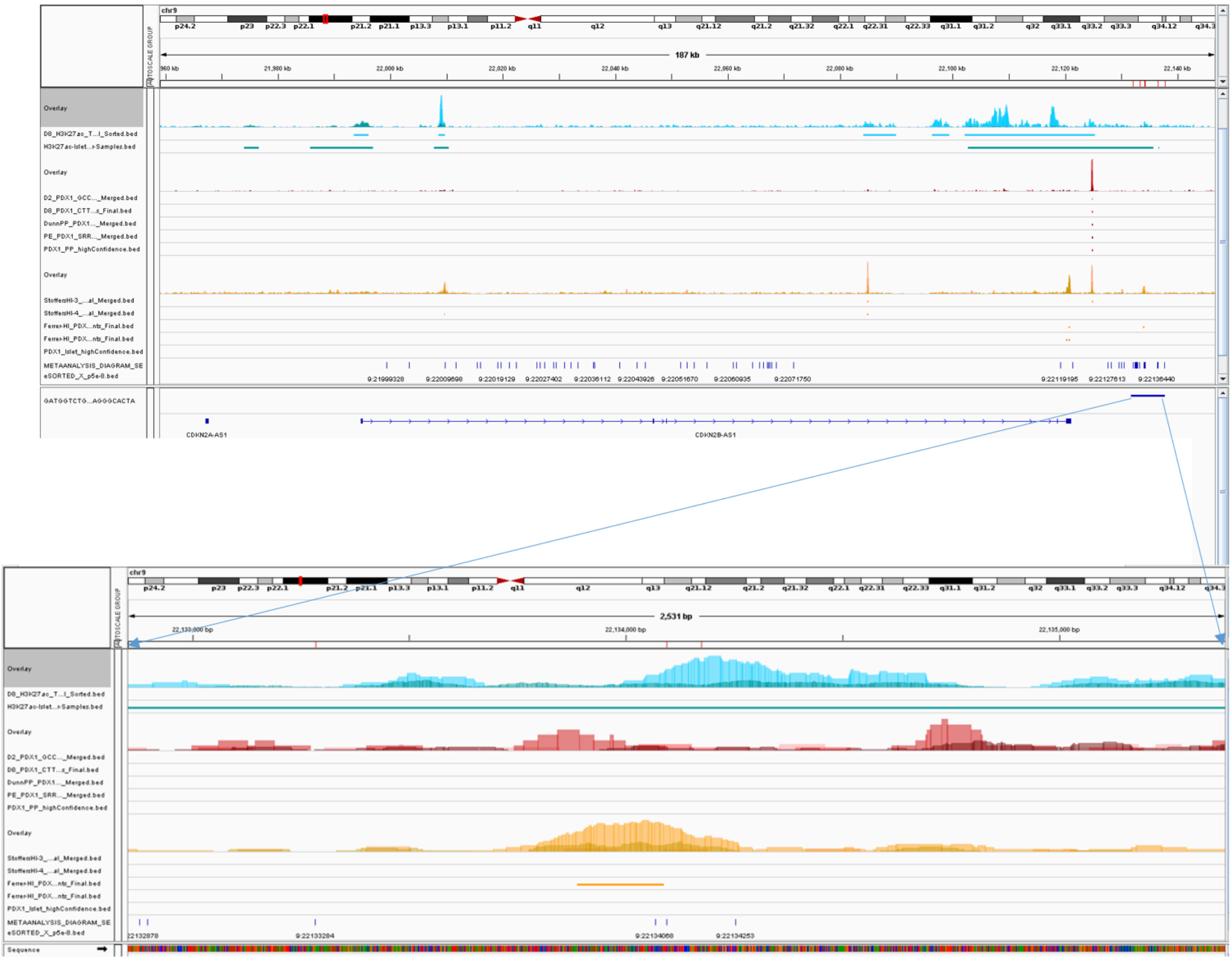


Figure 2-13: Mapping T2D related SNPs with the histone marks at the INK-T2D risk region. According to the ChIP-seq data, most of the SNPs are in inactive regions. (Red lines on top are SNPs). Two of the SNPs, *rs10811661* and *rs10757283* are in a region with H3K27ac histone marks representing an active chromatin region/ state.

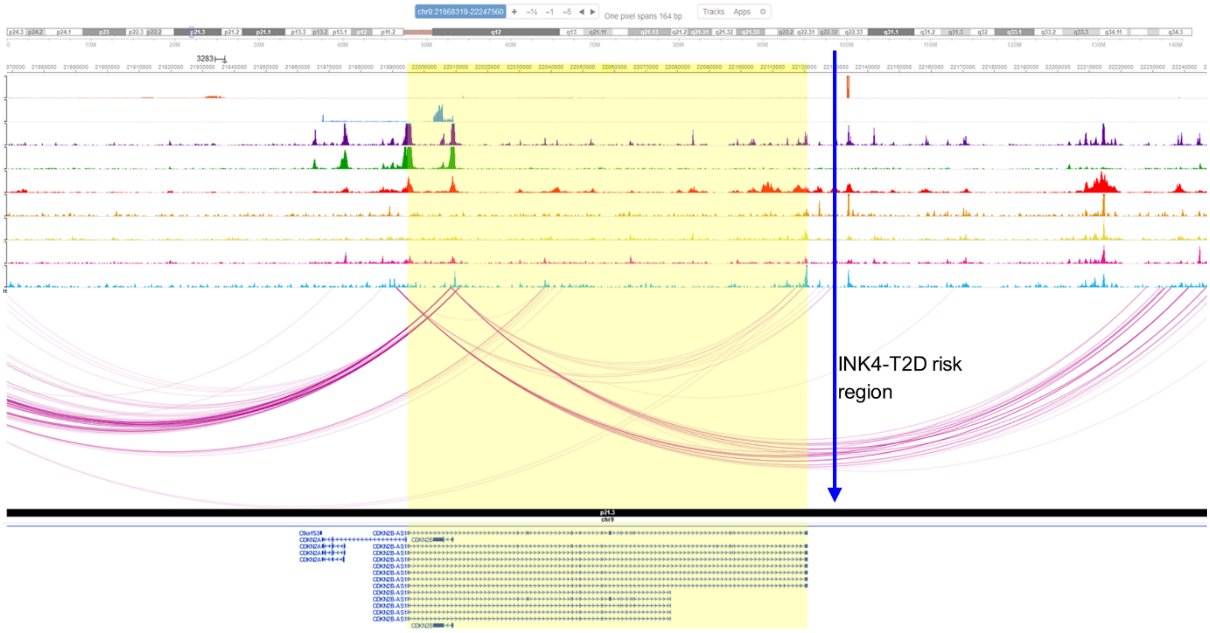


Figure 2-14: Chromatin interactions of proximal and distal regulatory elements at the human INK4 locus. The diagram clearly represents that INK4 T2D risk region interacts with INK4 genes promoters (INK4 T2D risk region is marked by blue line).

2.2.2 Generation of an iPSC Line Lacking INK4 T2D Risk Region

As mentioned above, the INK4 locus contains six T2D-associated SNPs embedded within an 8 kb genomic block (we called the INK4 T2D risk region) near the ANRIL gene. Yet, the mechanisms of how these SNPs are linked to T2D is not known. To have a cell model for functional studies of the T2D-associated SNPs of the INK4 region, we decided to establish an iPSC line lacking genomic region of the INK4 T2D risk fragment e.g., null allele. To this end, we used our previous developed iPSC line (HMGUi001-A-1) (Wang et al., 2018) for gene editing with the CRISPR/Cas9 platfrom.

2.2.2.1 Strategy for Genetic Manipulation of iPSCs via CRISPR/Cas9 System

To target the HMGUi001-A-1 iPSC line, we followed a 7-step strategy as illustrated in Figure. 2.15. These seven steps include 1) Cloning; preparation of vectors expressing sgRNA and Cas9 enzyme 2) Transfection; the process of introducing the generated vectors into the target cells 3) FACS analysis; Cas9 positive cells are sorted by FACS 4) Singularization; in order to obtain single cells, the sorted cells are diluted and cultured 6) Expansion; single cell-derived clones are grown and expanded 6) Analysis; each expanded single-cell clone is analyzed for desired genetic manipulation either insertion or deletion 7) Cell line; following analysis, the cell lines are established and cryo-conserved.

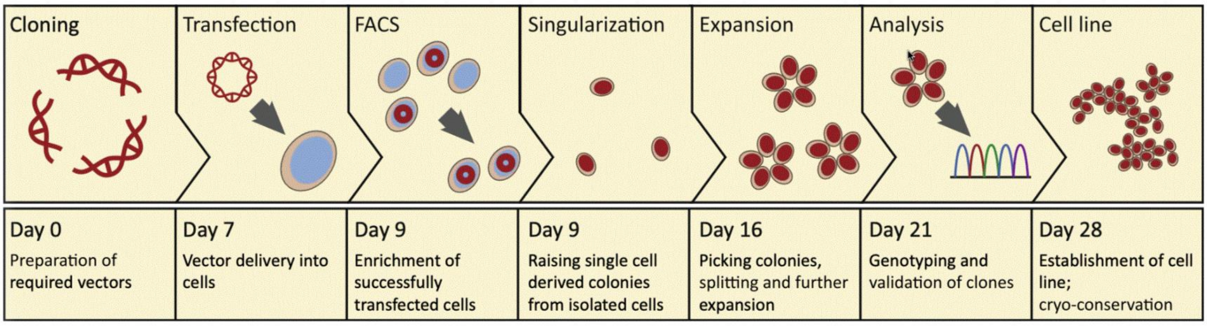


Figure 2-15: General outlines for gene targeting of iPSC with CRISPR/Cas9 genome engineering.

2.2.2.2 Design, and Cloning sgRNA Expression Plasmid for Targeting the INK4 Locus

Initially, the CRISPR target sites were subjected to sequencing e.g. upstream (region A) and downstream (region B) of the INK4-T2D risk region since they are located in a low conserved genomic DNA. Hence, we checked if the target sites contain any variation and differ from the reference sequences. Hence, two pairs of primers (FA, RA; FB, RB) were designed to amplify the border sites of sgRNAs cut regions. Later, two pairs of sgRNAs with high specificity scores were designed by two web-tools CRISPETA and CRISPR. The details of designed sgRNAs are summarized in Figures 2.16 A and 3.16 B. Following synthesizing the oligos, they were subjected to annealing and cloning, into the CRISPR/Cas9 plasmid. The vector expresses Cas9 fused with GFP. By Gibson assembly cloning we produced a dual-sgRNA CRISPR/Cas9-GFP plasmid, which co-expressed two sgRNAs targeting both ends of the 8 kb genomic DNA (Figure 2.17). As illustrated in Figure 2.18, the right cloning of single and dual sgRNA plasmids was confirmed via Sanger sequencing.

a

b

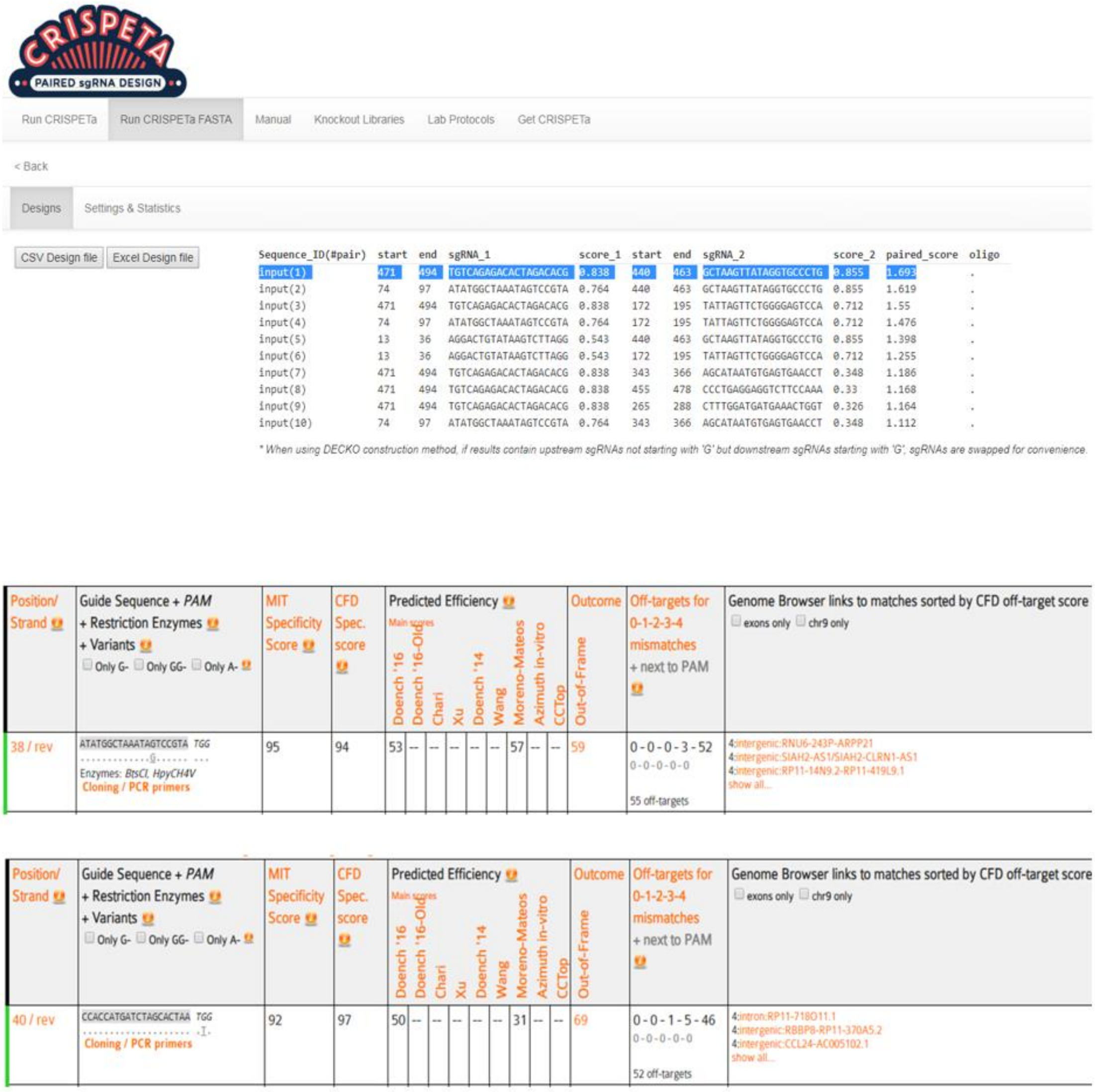


Figure 2-16: A schematic representation of CRISPR design platforms. A: The detailed information for first pair of sgRNAs designed by CRISPETA web-tool and B: the second pair that was designed via CRISPOR web-tool.

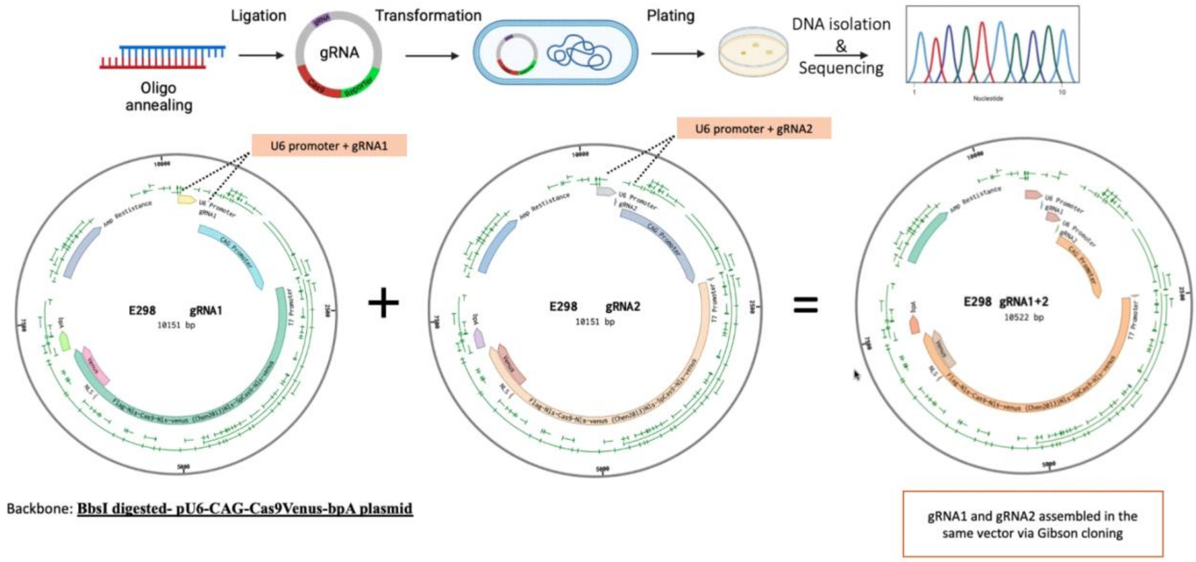


Figure 2-17. Cloning sgRNA expression vectors. The single sgRNAs including sgRNA1 and sgRNA2 were initially cloned into the BbsI digested-pU6-CAG-Cas9Venus-bpA plasmid. Thenthe dual sgRNA expression vector were derived from single sgRNAs vectors using Gibson cloning method.

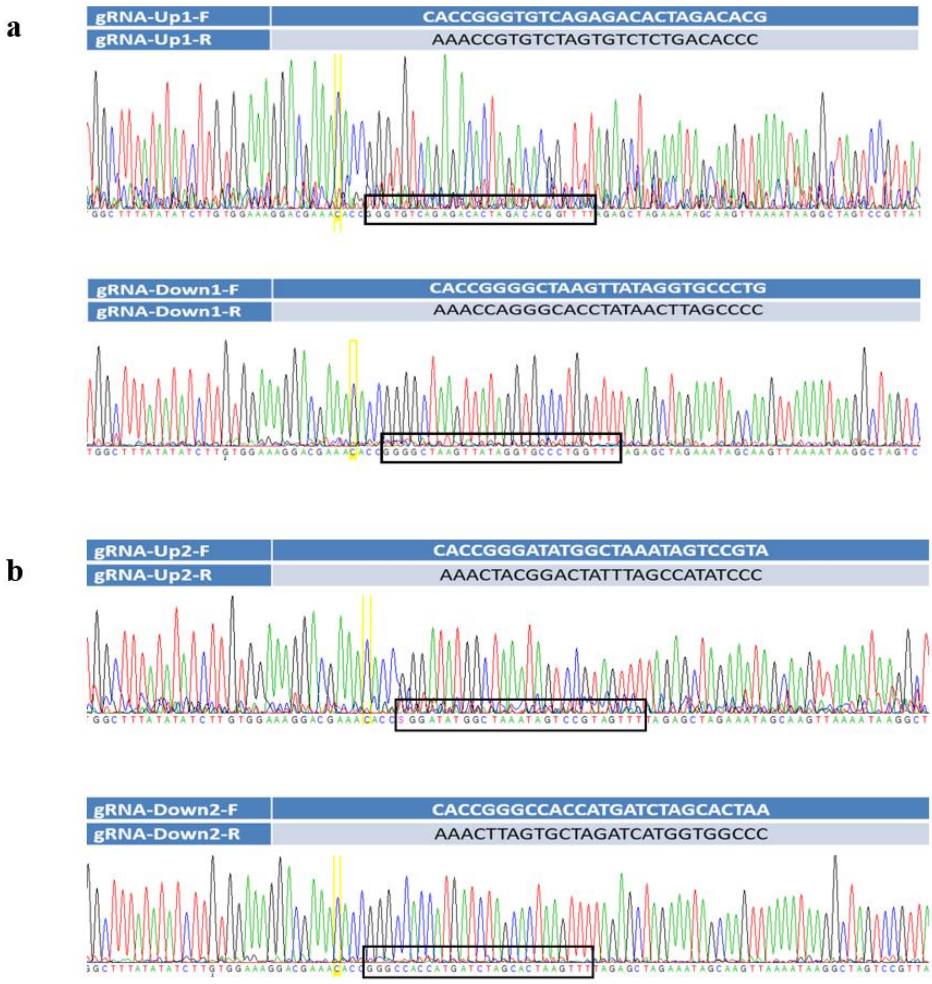


Figure 2-18. Sanger sequencing result. A and **B:** Following cloning of all four oligo nucleotides (two pairs of sgRNAs), they were subjected to the Sanger sequencing. The readout approved the accuracy of the cloning of sgRNAs.

2.2.2.3 Screening the Cas9 Positive iPSC Cells for the Deletion

Additionally, a pair of primers (FC, RC) was designed for screening the deletion (Figure 2.19A). Two days after the transfection, the GFP+ cells were sorted using FACS. Highly GFP expressed cells that accounted for 7% population (5500 events) were sorted by FACS. Following plating and expansion of the GFP positive iPSCs, genomic DNA was extracted from a total of 234 single-cell clones. Genotyping for the 8 kb deletion using FA-RB and FC-RC primers (Figure 2.19A) was performed, yielding a 1910 bp or 750 bp PCR product for the wild type or deleted allele respectively (Figure 2.19A/B and Figure2.20). After screening

234 clones, one homozygous and 19 heterozygous clones were identified. The average efficiency of deletion was 5%; however, this parameter for the generation of homozygous deletion was 0.4% (Table 2.2). To evaluate the accuracy of deletion at the target region, the PCR fragments were cloned in TA plasmid and sequenced via Sanger sequencing. The output confirmed the linking of the deletion borders in both alleles (Figure 2.19C). Eventually, we could successfully delete the 8 kb region in both alleles in the hiPSC clone. The resulting iPSC clone demonstrated normal iPSC morphology with no signature of unforced differentiation (Figure 2.19D).

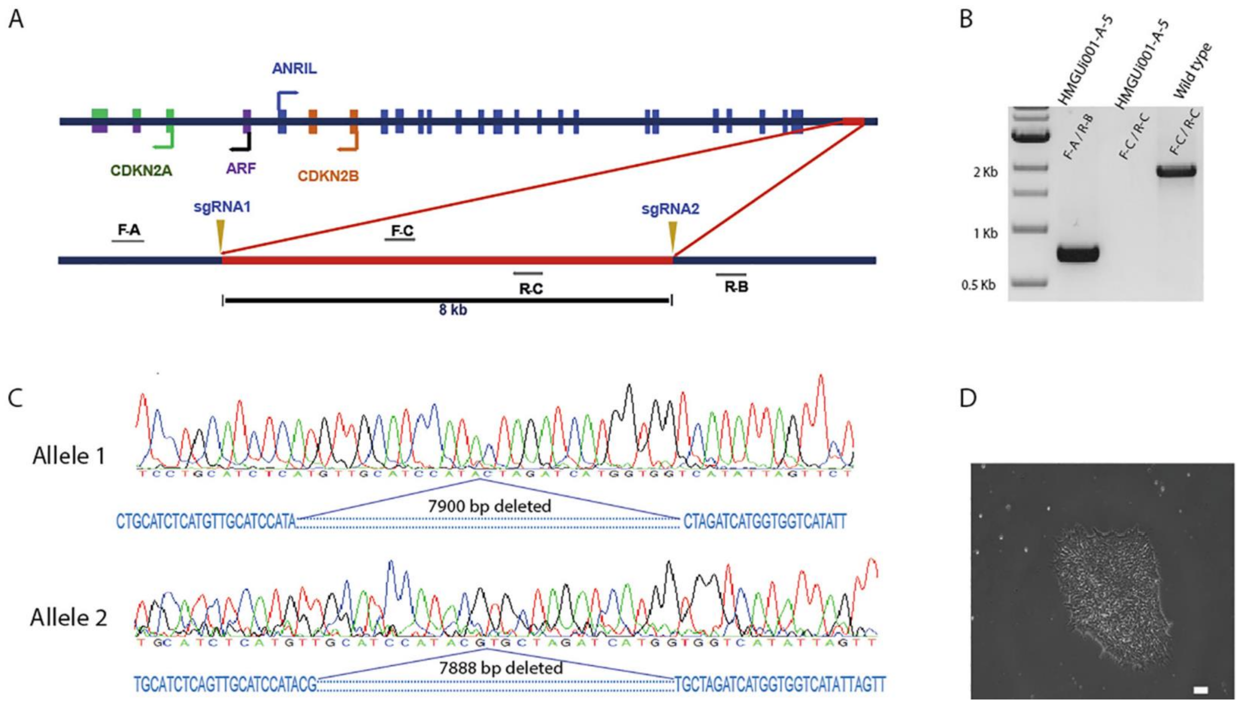


Figure 2-19. A schematic representation of gene targeting and screening. A: Schematic representation of the human INK4 genomic region at 9p21.3 region and the INK4 related T2D risk segment. The finalized sgRNA pair for the deletion of T2D risk region are represented as sgRNA I and sgRNA II at the terminal points of the target. The oligo primers applied for sequencing and screening are presented. FA-RB and FC-RC primer pairs were applied to identify the deletion products. **B:** The 750 bp and 1910 bp PCR amplicons reflect biallelic deletion and wild type respectively. **C:** Sanger sequencing approved the joint of the two terminal ends of the deletion parts. Deleted DNA bases are shown with dash marks. **D:** Bright field image of the selected colony at pluripotency step (Scale bar: 100 μm).

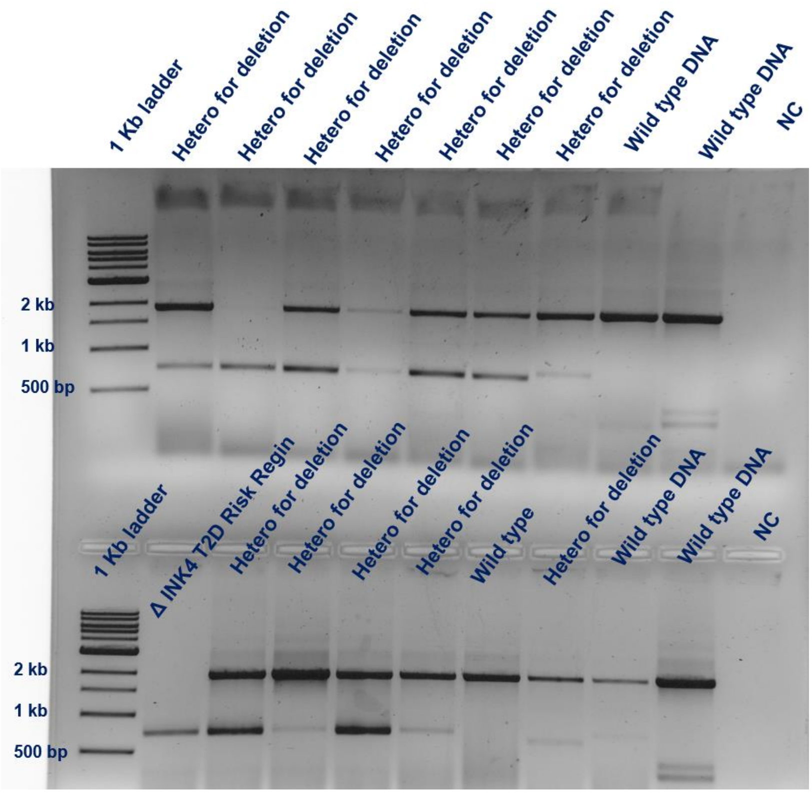


Figure 2-20: PCR Screening of the clones. The 750 bp and 1910 bp PCR amplicons are indicating biallelic deletion and wild type respectively. First well at the down arrow is the only homozygous DNA for deletion of the 8 kb.

Table 2.2: Statistics of screening and Efficiency of DNA editing

Gene targeting details	**Amount**
Transfection efficiency	20%
Selection of Highest GFP Positive cells (7%)	55000 Events
Number of picked clones	234
Number of wild type clones	214
Number of heterozygous cells for deletion	19
Number of homozygous cells for deletion	1
Efficiency of deletion (homozygous cells for 8 kb deletion)	0.4 %
Total efficiency of deletion	5 %

2.2.2.4 Characterization of the Edited iPSC Line (HMGUi001-A-5)

The only homozygous iPSC clone for the deletion of theINK4 T2D risk region was further characterized. Metaphase chromosomes were analyzed by using the standard G banding method and showed a normal karyotype (46, XX) of the iPSC clone (Figure 2.21A). The data from short tandem repeat (STR) analysis (for 16 sites) confirmed that the clone was derived from its parental iPSC line HMGU001-A (Table 2.3). Then, the expression of pluripotency factors OCT4 and SOX2 were examined by immunostaining and FACS analyses (Figure 2.21B and Figure 2.21C). Finally, the *ΔINK4 T2D risk region* iPSC line was differentiated towards the three germ layers, endoderm, mesoderm, and ectoderm (Figure 2.22). The results indicate the well establishment of the ΔINK4 T2D risk region iPSC line that shows multi-lineage potency. Furthermore, we checked three intergenic and/or intragenic genomic positions with the highest off- targeting degree for each sgRNA. The sequencing result showed no mutations at these regions (Figure 2.23A). Moreover, the iPSC clone does not contain mycoplasma contamination (Figure 2.23B). Altogether, the newly established iPSC line or HMGUi001-A-5 provides a beneficial cellular tool to investigate the potential causal association of INK4 SNPs to diabetes prevalence during the development of endocrine lineage or other stages of β-like cell development.

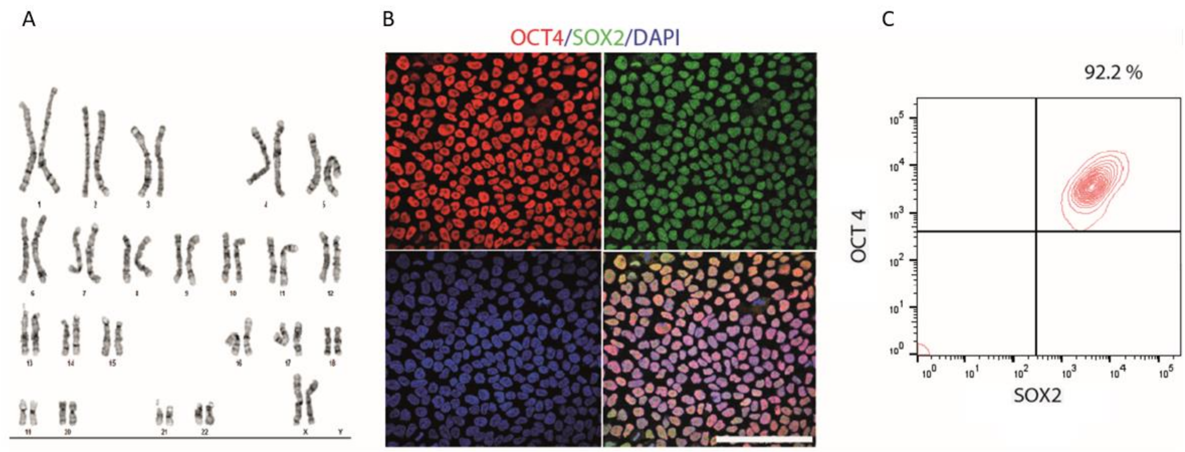


Figure 2-21. Characterizing HMGUi001-A-5 iPSC line for Karyotyping and pluripotency markers. A: Chromosome karyotype of the HMGUi001-A-5 iPSC line showed normal karyotype (46, XX). **B:** ICC staining illustrates the expression of pluripotency markers, SOX2 and OCT4 in the ΔINK4 T2D risk region iPSC line at

the maintenance step (Scale bar: 100 µm). **C:** Representative FACS plots of double positive cells for expression of OCT4 and SOX2 c in the HMGUi001-A-5 iPSC line.

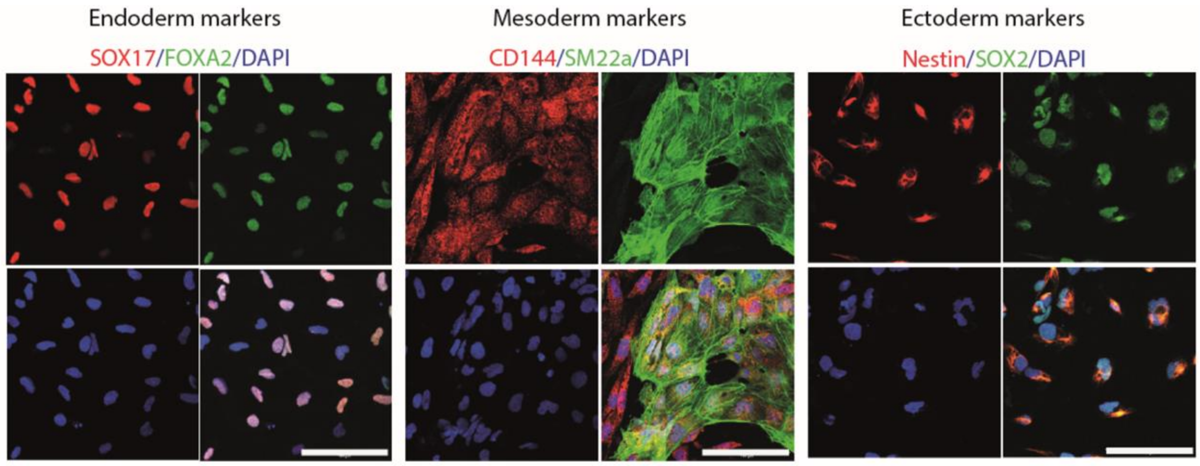
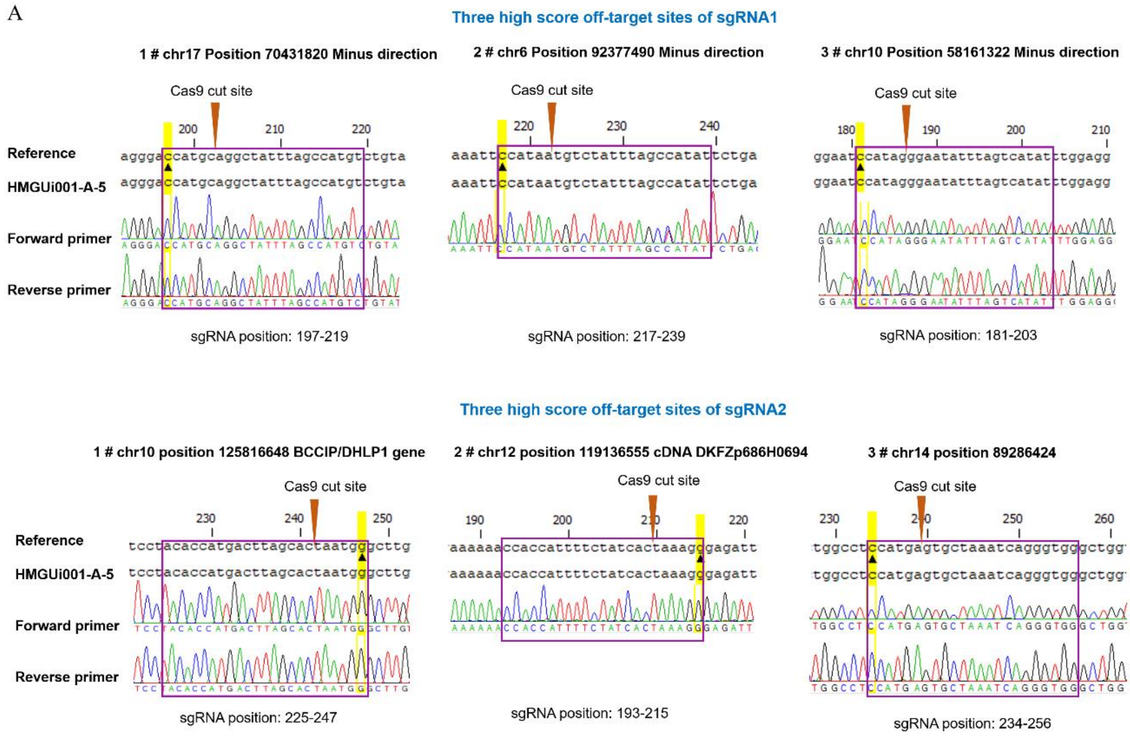


Figure 2-22. Evaluating multi-lineage differentiation of HMGUi001-A-5 iPSC line. Using multi-lineage potency assay, the HMGUi001-A-5 iPSC line was differentiated towards the three germ layers, endoderm, mesoderm, and ectoderm and were immunostained for SOX17/FOXA2, CD144/SM22a and Nestin/SOX2, respectively (Scale bars: 100 µm).

A



B

Mycoplasma results
Cell line: *ΔINK4 T2D Risk Region* hiPSC line (Passage number: P31)

Test method: Biochemical luminescence MycoAlert™ plus Mycoplasma Detection Kit, Lonza

Reading A: 32385 Reading B: 11400 Ratio (B/A): 0.359

Comments:
Ratios < 0.9 are negative for Mycoplasma
Ratios 0.9 – 1.2 are borderline
Ratios >1.2 are positive for Mycoplasma

Result: Negative

Figure 2-23: Evaluating the off-target sites using Sanger segueing. A: The Sanger sequencing output of three different intergenic and/or intragenic genomic regions with the highest off- target degree for each sgRNA showed no scars. **B:** The HMGUi001-A-5 iPSC line was negative for mycoplasma.

Table 2.3. STR Result

Gene Locus	HMGUi001-A-1 (original cell line)	*ΔINK4 T2D Risk Region* iPSC line
D8S1179	13	13
D21S11	28, 29	28, 29
D7S820	9, 10	9, 10
CSF1PO	10.2, 11	10.2, 11
D3S1358	17.2	17.2
TH01	6, 9.3	6, 9.3
D13S317	13, 15	13, 15
D16S539	10, 12	10, 12
D2S1338	23, 26	23, 26
D19S433	13, 16	13, 16
vWA	17, 18.2	17, 18.2
TPOX	9	9
D18S51	18	18
AMEL	X	X
D5S818	11, 12	11, 12
FGA	24, 24.2	24, 24.2

2.2.3 Differentiating the Edited iPSC into Pancreatic Progenitors and Insulin Secreting Cells

After successfully characterizing the HMGUi001-A-5 iPSC line, we aimed to differentiate it into β-like cells. To this end, we used the Rezania protocol (Rezania et al., 2014) for the differentiation of iPSCs. In this protocol, the iPSCs were differentiated to anterior definitive endoderm (*ADE*), primitive gut tube (*PGT*), pancreatic progenitor (PP), endocrine progenitor (EP) and stem cell-derived *β* cells (*SC-β*). To test the efficiency of differentiation, the major TFs that have functions in the development of each stage were measured and quantified by immunostaining and FACS analyses, respectively. For example, the pancreatic progenitors were measured for the expression of PDX1 and NKX6.1 on day 12 of differentiation. The endocrine cells were tested for the expression of NKX2.2 and NKX6.1 on day 15 of differentiation. Eventually, β-like cells were evaluated for the expression of C-peptide, Glucagon and NKX6.1 at the final stage of differentiation on day 20 (Figure 2.24).

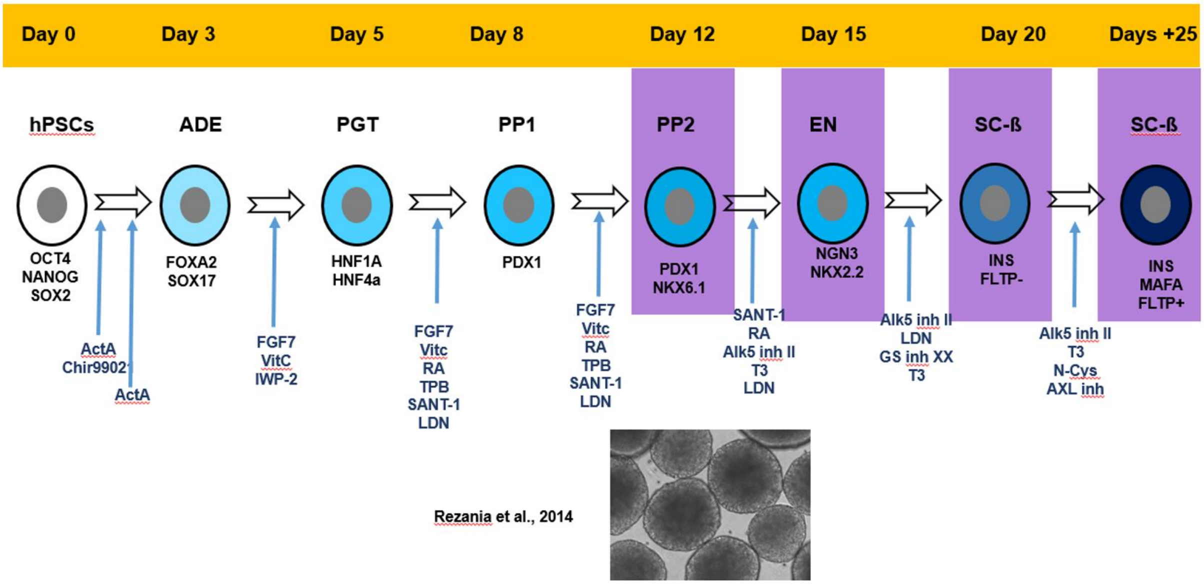


Figure 2-24: Protocol details for iPSC differentiation towards insulin secreting cells. The details of Rezania's protocol for iPSC differentiation to pancreatic progenitors and β*-like cells including timelines and chemical reagents*. *FGF7* (Fibroblast Growth Factor 7), VitC (Vitamin C), RA (Retinoic Acid), CHIR (CHIR99021), IWP-2 (Inhibitor of Wnt Production-2), *SANT-1* (Inhibitor of hedgehog signaling), Alk5 inh II (The TGFβ type I receptor kinase inhibitor II), T3 (3,3',5-Triiodo-L-thyronine), LDN (BMP inhibitor), TPB (Protein kinase C activator), N-Cys (N-acetyl cysteine), Trolox (vitamin E analogue), R428 (AXL receptor tyrosine kinase inhibitor), GSiXX (gamma secretase inhibitor XX).

2.2.3.1 Morphology and Size Changes of the Cell Clusters at Different Stages of Differentiation

The differentiation process was carried out in 3D culture. At the beginning of the process, we disassociated iPSCs and made single cells. Then, the cells preferred to form cell clusters or aggregates in a 3D culture medium. The morphology and size of aggregates were evaluated during the time of differentiation (stage 1 to stage 6). As illustrated in Figure 2.25, the aggregates derived from HMGUi001-A-5 have a very similar shape in comparison to the control aggregates. However, the HMGUi001-A-5 aggregates seem smaller compared to the control ones at stage 4, stage 5 and stage 6.

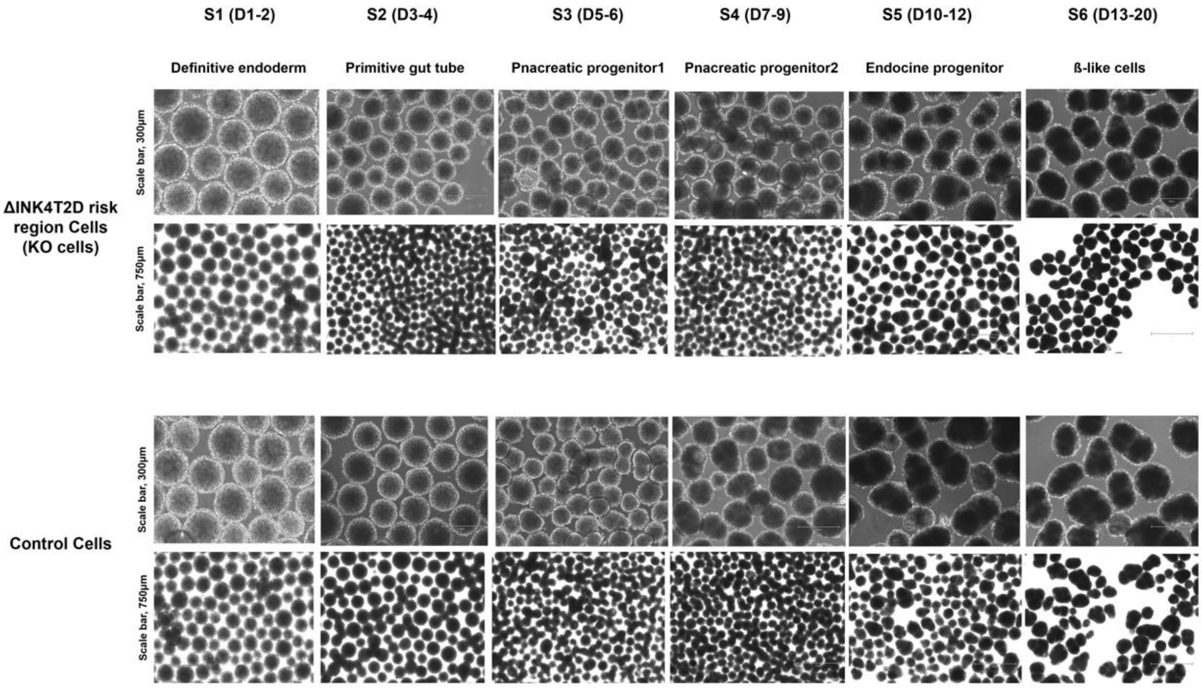


Figure 2-25: Morphology of the cell clusters or aggregates during differentiation. Comparison of the morphology and size of aggregates during the time course of differentiation (scale bars, 300 µm and 750 µm).

2.2.3.2 The INK4 Genes Show Distinct Expression Pattern During the *β*-cells Differentiation

As mentioned above, the T2D risk region at the INK4 locus may harbor an enhancer or silencer regulatory element. Deleting this element could alter the expression of the adjacent genes. To this end, we delineate the expression patterns of INK4 genes, *P15*, *P16*, *ANRIL* and the neighboring genes, *MTAP* and *DMRTA1*, upon the course of *β*-cells differentiation. Hence, we harvested the cell clusters or aggregates at the end of each differentiation stage. Following RNA isolation and cDNA synthesis, we performed qPCR using TaqMan probes. We could not report the expression of INK4 genes at iPSCs and early stages of differentiation. From stage 3 onwards, the expression of ANRIL and P15 gradually increased and showed the highest level of expression at the late stages of differentiation, stage 6 and stage 7.

However, the expression of ANRIL and P15 in the HMGUi001-A-5 cells was significantly lower than in the control cells at stage 5, stage 6 and stage 7 (n=3, P-value < 0.01) (Figure 2.26). The assay failed to detect any expression for P16 at the RNA level in iPSCs and during the following stages of differentiation; however, we could show that P16 was expressed at protein level at a later stage. Failing in qPCR could be due to the presence of several transcript variants for the P16 or the low quality of its TaqMan probe. Furthermore, the expression of two genes, MTAP and DMRTA1 in the vicinity of INK4 locus, were measured. MTAP and DMRTA1 are located > 100 kb upstream and downstream of the INK4 locus, respectively. As illustrated in Figure 2.26, both genes are expressed at high levels at stage 4 of differentiation. Deleting the INK4-T2D risk region did not affect the expression of DMRTA1 and MTAP upon the differentiation of HMGUi001-A-5 towards β-like cells.

Figure 2-26: Quantitative PCR of INK4 genes during β-like cells differentiation. The mRNA expression patterns of INK4 genes and the neighbor genes, MTAP and DMRTA1 during the HMGUi001-A-5 iPSCS differentiation to pancreatic progenitor, endocrine cells, and *β*-like cells. The mRNA expression levels were quantified by real time PCR (n = 3) and normalized to endogenous gene expression GAPDH. Student's t test with two-tailed distribution and three-sample equal variance were applied for statistics analysis.

2.2.3.3 Expression Analysis of PDX1 and NKX6.1 Markers in HMGUi001-A-5 Derived Pancreatic Progenitors

According to the published data (Annicotte et al., 2009), INK4 genes might have roles upstream of several regulatory pathways in β-cell proliferation and/or insulin secretion. To address this hypothesis, we looked for a phenotype in proliferating pancreatic progenitors or mature β-cells. We carried out immunostainings to verify the expression of PDX1, P15 and P16 at stage 4 or day12 of differentiation (pancreatic progenitors) using wild-type iPSCs. As illustrated in Figure 2.27, P16 is localized in the nucleus, whereas P15 is in both the cytoplasm and nucleus. As mentioned earlier, we could not detect the expression of P16 at the RNA level while it is detectable at the protein level as shown in Figure 2.27.

The sectioned clusters were stained for pancreatic progenitor markers, PDX1 and NKX6.1 at day 12 or stage 5, as shown in Figure 2.28. Apparently, clusters derived from HMGUi001-A-5 do not show any significant difference for PDX1 and NKX6.1 expression. FCAS was used to measure the protein levels of PDX1 and NKX6.1. Immunostaining and FACS results showed that 59% of HMGUi001-A-5 cells are expressed PDX1 versus 61% of the control cells. Co-staining for PDX1 and NKX6.1 showed that 35% of HMGUi001-A-5 cells and 36% of the control cells are double-positive for both markers (n=3). According to the immunostaining and FACS results, we could not report any significant difference between the groups in the percentage of the cells expressing PDX1 and NKX6.1 (Figure 2.28, and Figure 2.29).

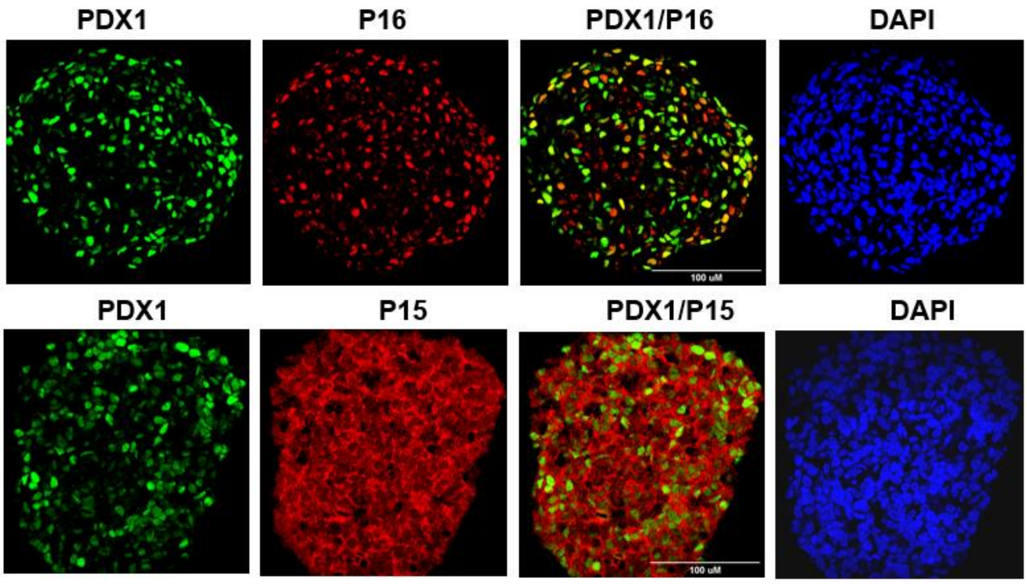


Figure 2-27: Immunostaining for the pancreatic progenitor's markers. Immunostaining of sectioned cell clusters derived from wild type iPSCs stained for PDX1, NKX6.1, and P15/P16 or with the nuclei marker DAPI at stage 4 or the pancreatic progenitor's stage. Scale bar, 100 mm.

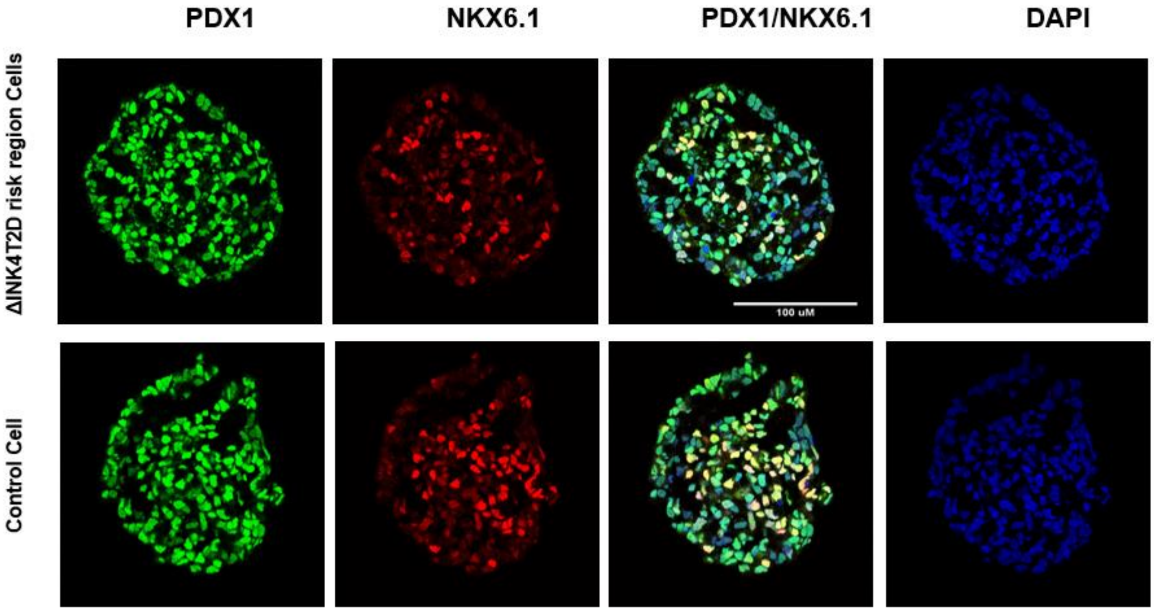


Figure 2-28: Immunostaining of HMGUi001-A-5 cells for the pancreatic progenitor's markers at stage 4. Immunostaining of sectioned aggregates or clusters stained for PDX1, and NKX6.1 or with the nuclei marker DAPI at stage 4 or the pancreatic progenitor's stage. Scale bar, 100 mm.

a

b

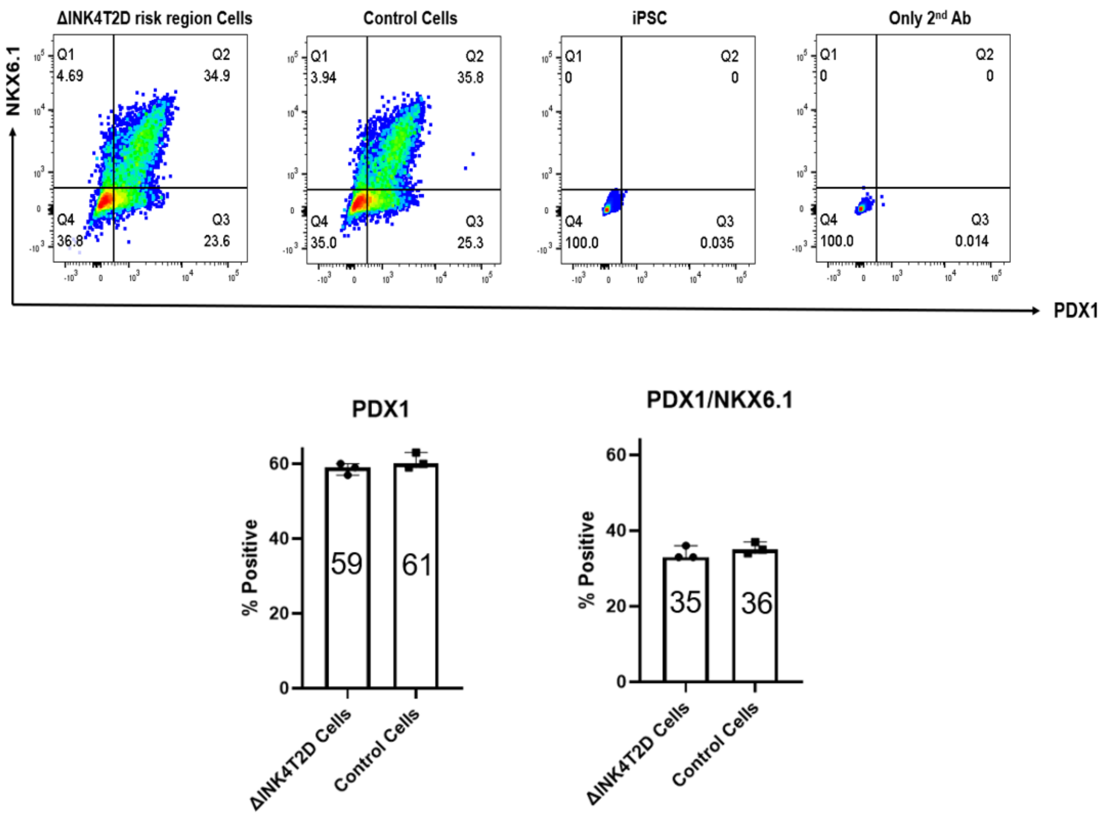


Figure 2-29: Quantification of PDX-1/NKX6.1 at pancreatic progenitor stage. A: Representative FACS dot plots for co-staining of PDX-1 and NKX6.1 at *pancreatic progenitor* cells stage. **B:** Quantification of FACS analysis for percentage of PDX-1+ and NKX6.1 + cells at *pancreatic progenitor* cells stage from three independent assays. iPSCs and Only-secondary-antibody were considered as negative control. The data were represented as mean + S.D. (n = 3).

2.2.3.4 Expression Analysis of PDX1 and NKX2.2 Markers in HMGUi001-A-5 Derived Endocrine Cells

To analyze how the deletion of the T2D risk region at the INK4 locus affects the endocrine induction, we collected cell aggregates in the middle of stage 4 (day 15) of differentiation in which the early endocrine progenitor cells express PDX1 and NKX2.2. Immunostaining and FACS analyses were carried out to measure the expression of PDX1 and NKX2.2. Furthermore, the cell clusters or aggregates were immunostained for P15 and P16 (Figure 2.30). Immunostaining and FACS results showed that 59% of HMGUi001-A-5 cells expressed PDX1 versus 66% of the control cells on day 15. Co-staining for PDX1 and NKX2.2 showed that 42% of HMGUi001-A-5 cells were double-positive versus 53% of the control cells. There is no significant difference in the expression of PDX1 between KO cells (HMGUi001-A-5 cells) and control. But KO cells express more PDX1/NKX2.2 (n=3, P-value < 0.01) (Figure 2.31). This is the first hint that deletion of the INK4-T2D risk region affects the expression of endocrine markers. This result may show a difference in the functionality of the KO-derived endocrine cells.

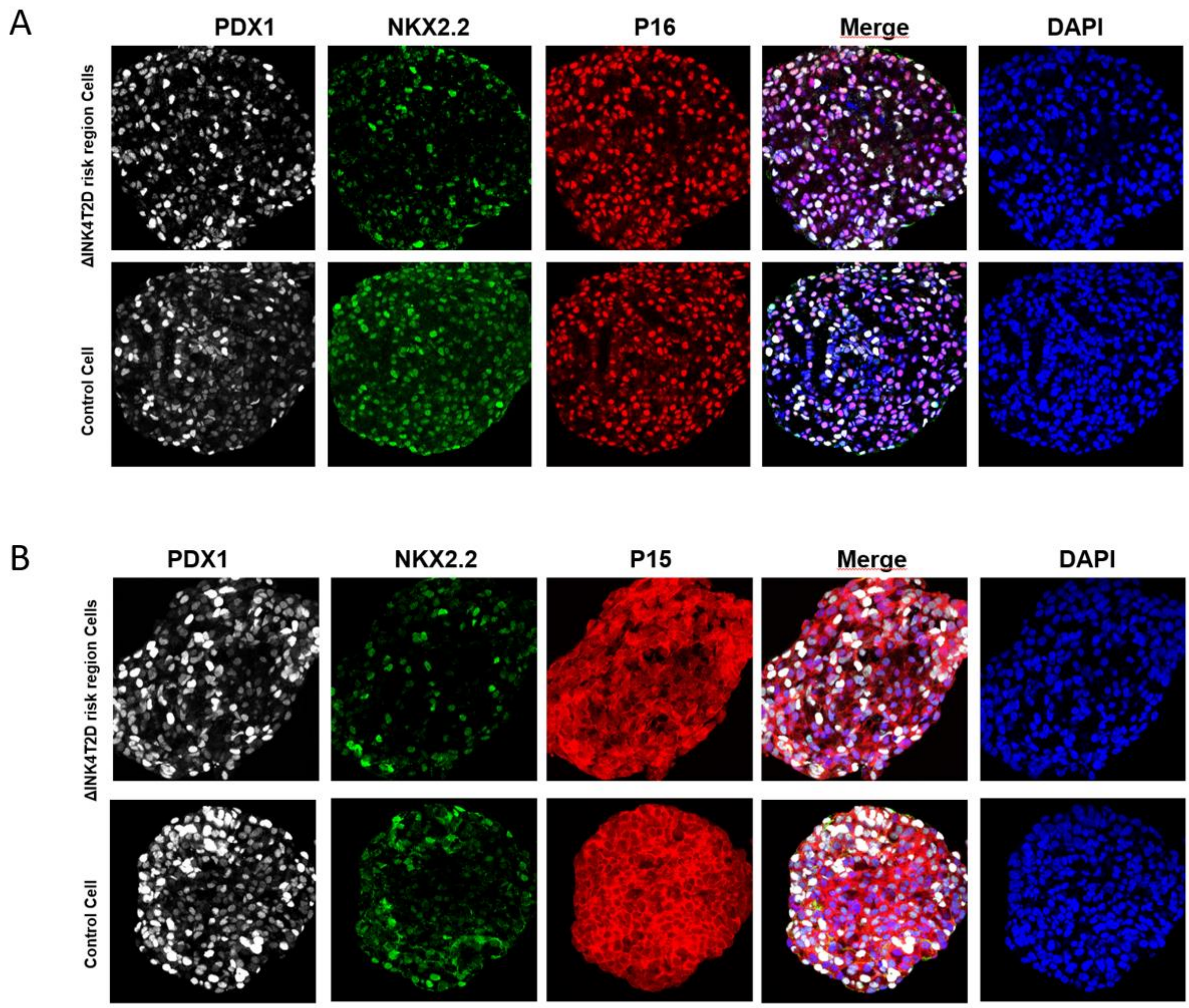


Figure 2-30: Immunostaining of HMGUi001-A-5 cells for the endocrine's markers and P15/P16. A and **B:** Immunostaining of sectioned aggregates stained for PDX1, NKX2.2 and P15/P16 or with the nuclei marker DAPI at stage 5 or the endocrine cells stage. Scale bar, 100 mm.

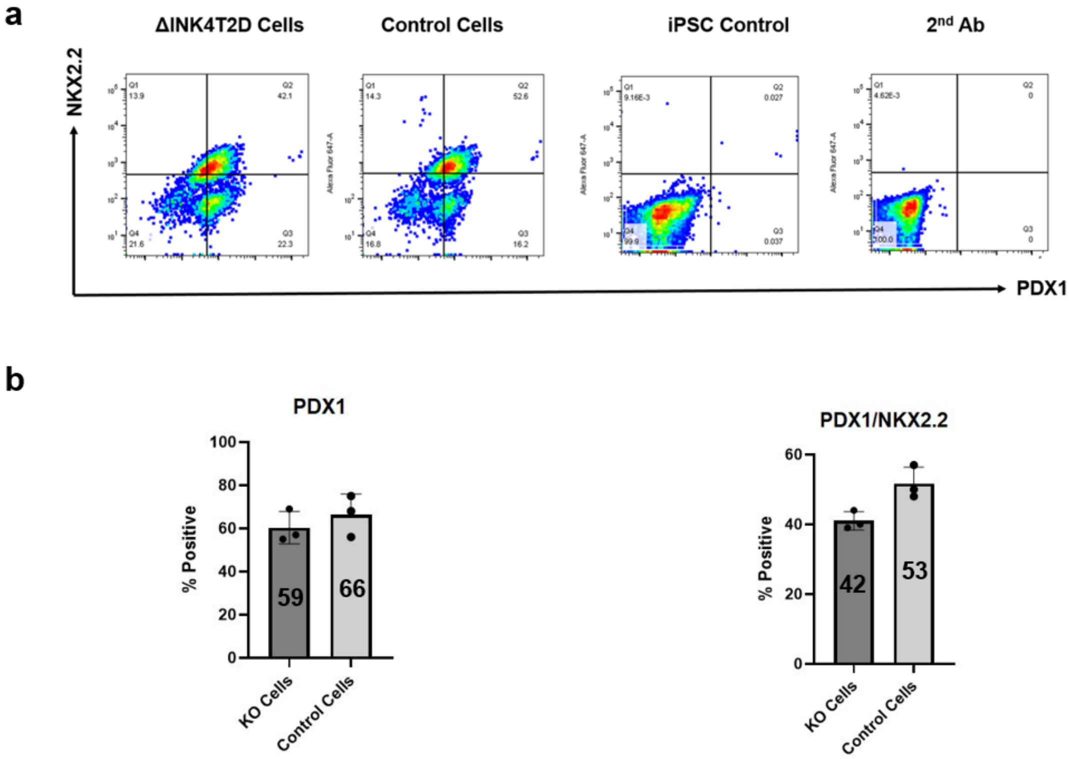


Figure 2-31: KO cells produce less PDX-1/NKX2.2 at endocrine cells stage. A: Representative FACS dot plots for co-staining of PDX-1 and NKX2.2 at endocrine cells stage. B: Quantification of FACS analysis for percentage of PDX-1+ and NKX2.2 + cells at endocrine cells stage from three independent assays. The data were shown as mean + S.D. (n = 3), $p < 0.01$.

2.2.3.5 Expression Analysis of C-peptide and Glucagon in HMGUi001-A-5 Derived β-like Cells

We further analyzed clusters derived from iPSC for the expression of C-peptide and Glucagon at the late stage of differentiation (Figure 2.32 and Figure 2.33). To achieve this, the aggregates were picked at the end of stage 6 (day 20) and prepared for immunostaining and FACS analyses. The result showed that 20% of cell clusters derived from HMGUi001-

A-5 expressed C-peptide versus 36% of the control cells at stage 6. 12% of KO cells expressed Glucagon versus 17% of the control cells. 3.1% of KO cells expressed both C-peptide and Glucagon versus 3.2% of the control cells (n=3, P-value < 0.001). Moreover, 6% of KO cells were double positive for C-peptide and NKX6.1 versus 18% of the control cells (n=3, P-value < 0.001). Therefore, the KO cells express less Insulin and NKX6.1 than the control cells (Figure 2.33). This result shows that the T2D risk region at the INK4 locus might affect β-cell functionality rather than β-cells specification.

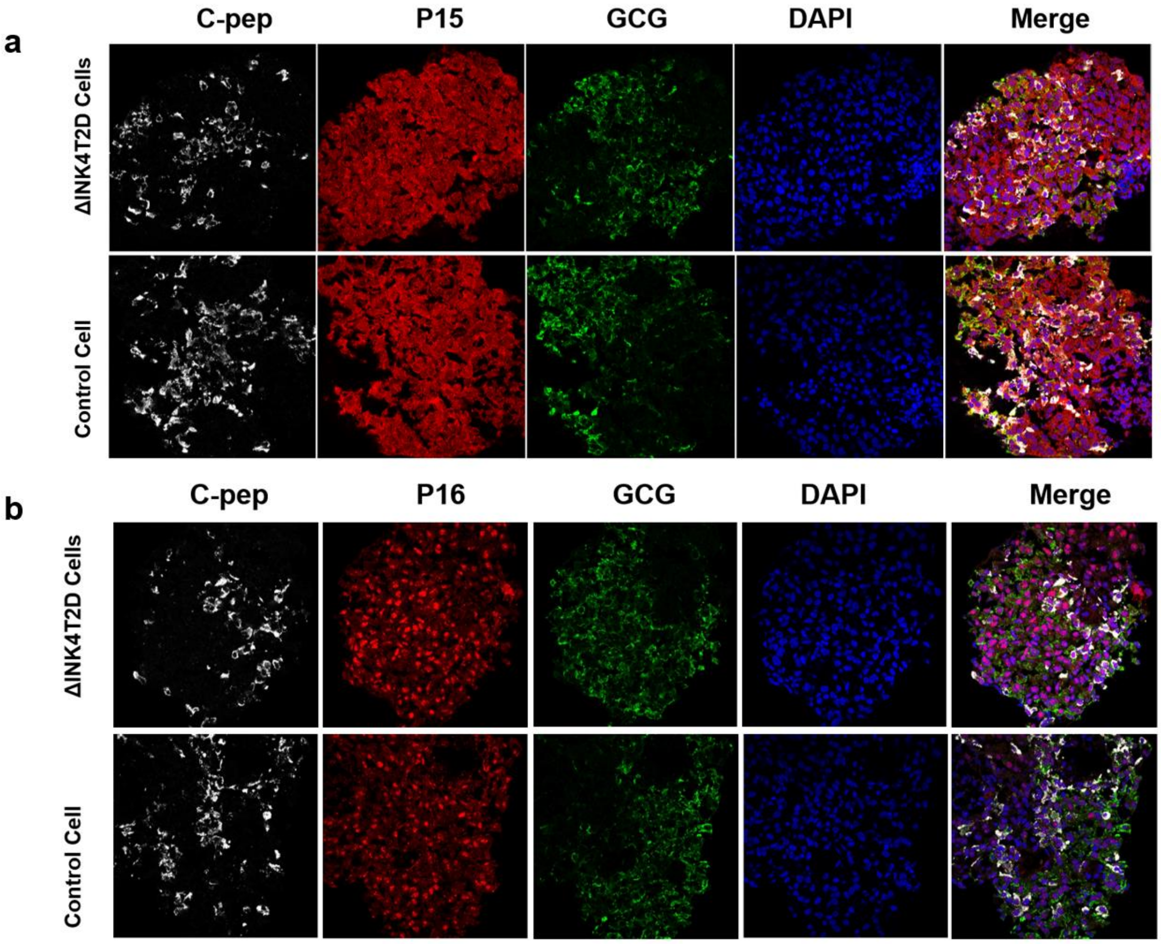


Figure 2-32: Immunostaining of β-like cells for hormones and P15/P16. A and B: Representative immunostaining images of sectioned clusters stained for C-pep and GCG (Glucagon), P15/P16 or with the nuclei marker DAPI at stage 6 or the β-like cells stage. Scale bar, 100 mm.

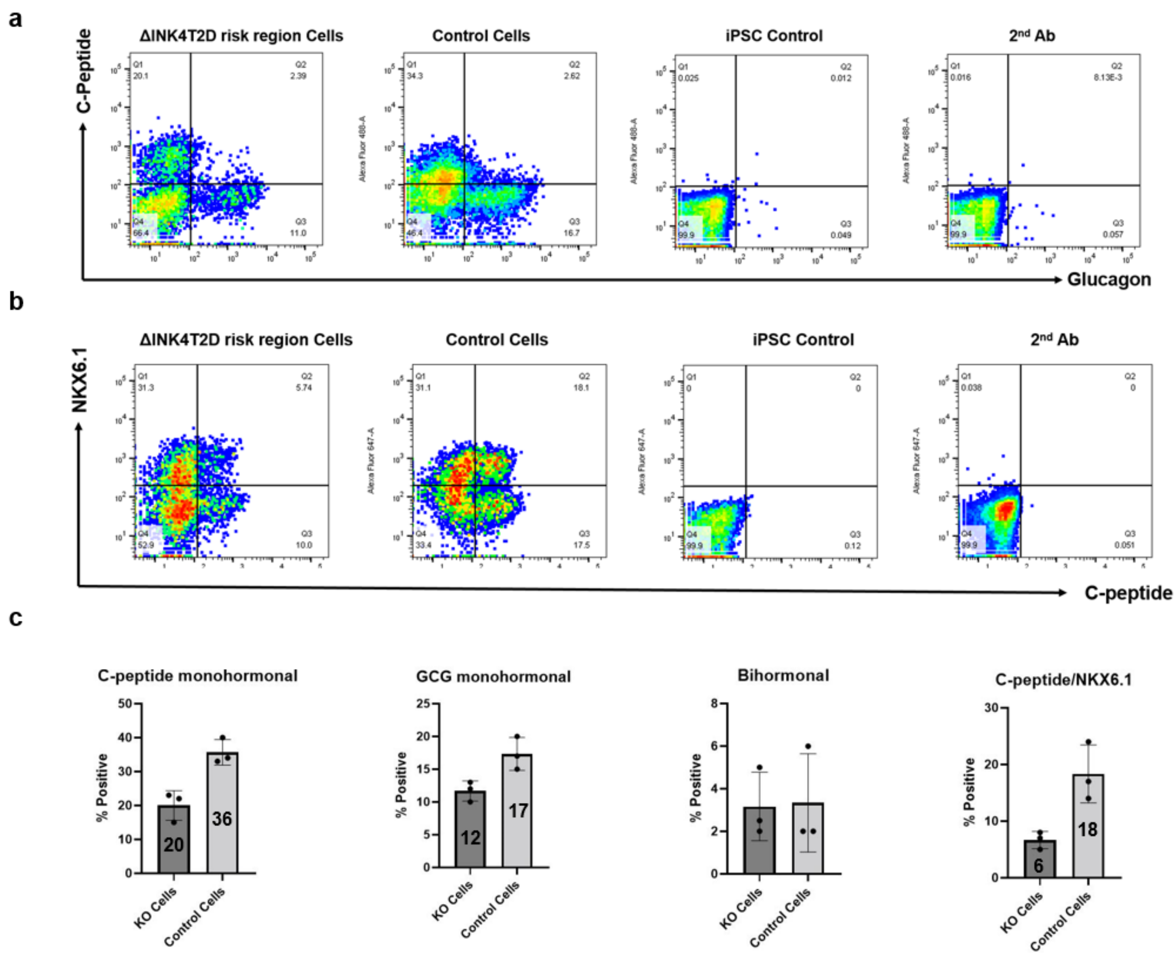


Figure 2-33: KO cells produce less insulin and NKX6.1 in *β*-like cells. A: Representative FACS dot plots for co-staining of C-pep and GCG at *β*-like cells stage. **B:** Representative FACS dot plots for co-staining of C-pep and NKX6.1 at *β*-like cells stage. **C:** Quantification of FACS analysis for percentage of C-pep+, GCG+ and C-pep+/NKX6.1+ cells at the stage of *β*-like cells from three independent assays. The data were represented as mean + S.D. (n = 3), p < 0.001.

2.2.3.6 Deletion of T2D Risk Region at the INK4 Locus Resulted in Diminished Proliferation Rate in Pancreatic Progenitors, Endocrine Cells and *β*-like Cells

As mentioned above, the INK4 genes negatively regulate the cell cycle and cell proliferation. By deletion of the T2D risk region at the INK4 locus, we expect to see an alteration in the expression of INK4 genes and possibly in cell proliferation. We performed EdU incorporation test to assess the effect of deleting the INK4-T2D risk region on cell proliferation during differentiation of iPSC towards β-like cells. Then, to evaluate the cell proliferation rate in the KO cells, we treated the iPSC-derived cell clusters at three time points of differentiation e.g. pancreatic progenitors (day 12), endocrine cells (day 15) and SC-β like cells (day 20). EdU was added to the cells, and waited for at least 6 hours, then the aggregates were collected for FACS analysis. Interestingly, the proliferation rate decreased in the KO-derived cell clusters at all three time points. The proliferation rates in the KO cells were12.5%, 3.4% and 1.35% at day 12, day 15 and day 20 of differentiation, respectively. While these ratios were 28%, 5.9% and 3.3% in the control-derived cell clusters (Figure 2.34). The KO cells approximately showed a two-fold decrease in proliferation rate (n=3, P-value < 0.001). This indicates that deletion of the T2D risk region at the INK4 locus might affect the expression of ANRIL and CDKN2B, leading to a reduction in cell proliferation.

We further analyzed the proliferation rate in the C-peptide and glucagon positive cells at stage 6 or day 20 of differentiation. To this end, we co-stained EdU and C-peptide or Glucagon and analyzed them with FACS. Among the C-peptide positive cells, 1.15% were proliferating in the KO cells versus 3.5% in the control cells. Moreover, 2.2% of Glucagon positive cell were EdU positive or proliferating versus 3.2% in the control cells (n=3, P-value < 0.001) (Figure 2.35).

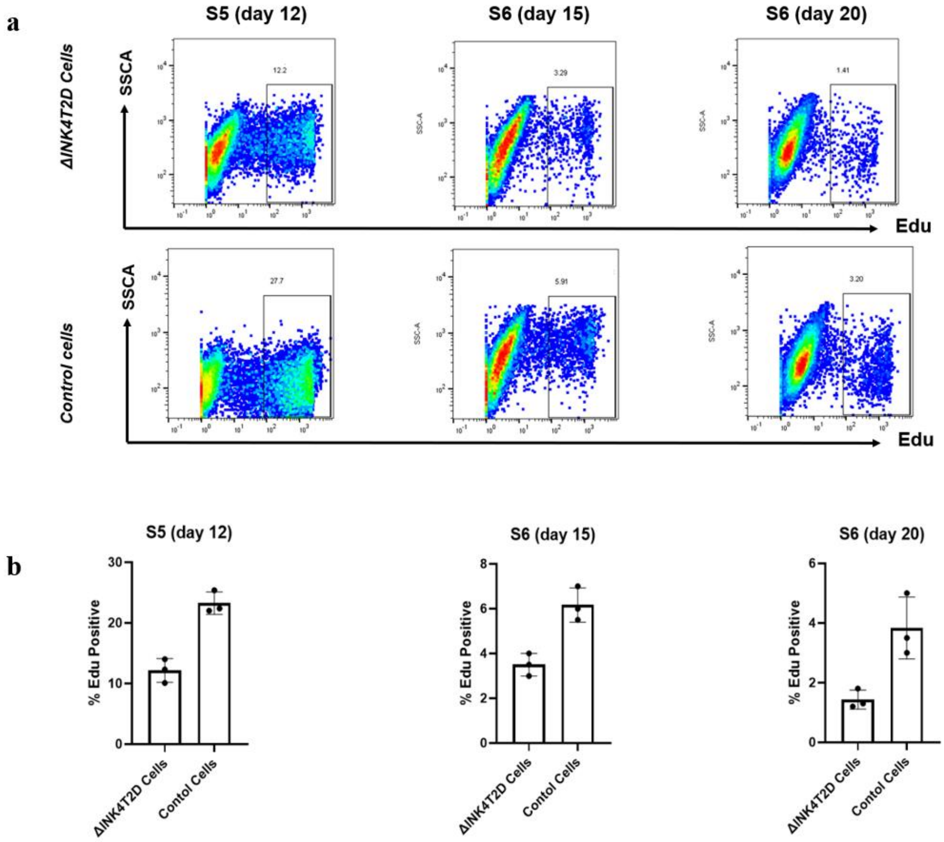


Figure 2-34: Analysis of cell proliferation at the stages of pancreatic progenitors, endocrine cells and *β*-like cells. A: Representative FACS dot plots of staining with EdU at pancreatic progenitors, endocrine cells and *β*-like cells stage. **B:** Quantification of FACS analysis for percentage of proliferating cells (EdU+) from three independent assays. The data were represented as mean + S.D. ($n = 3$), $p < 0.001$.

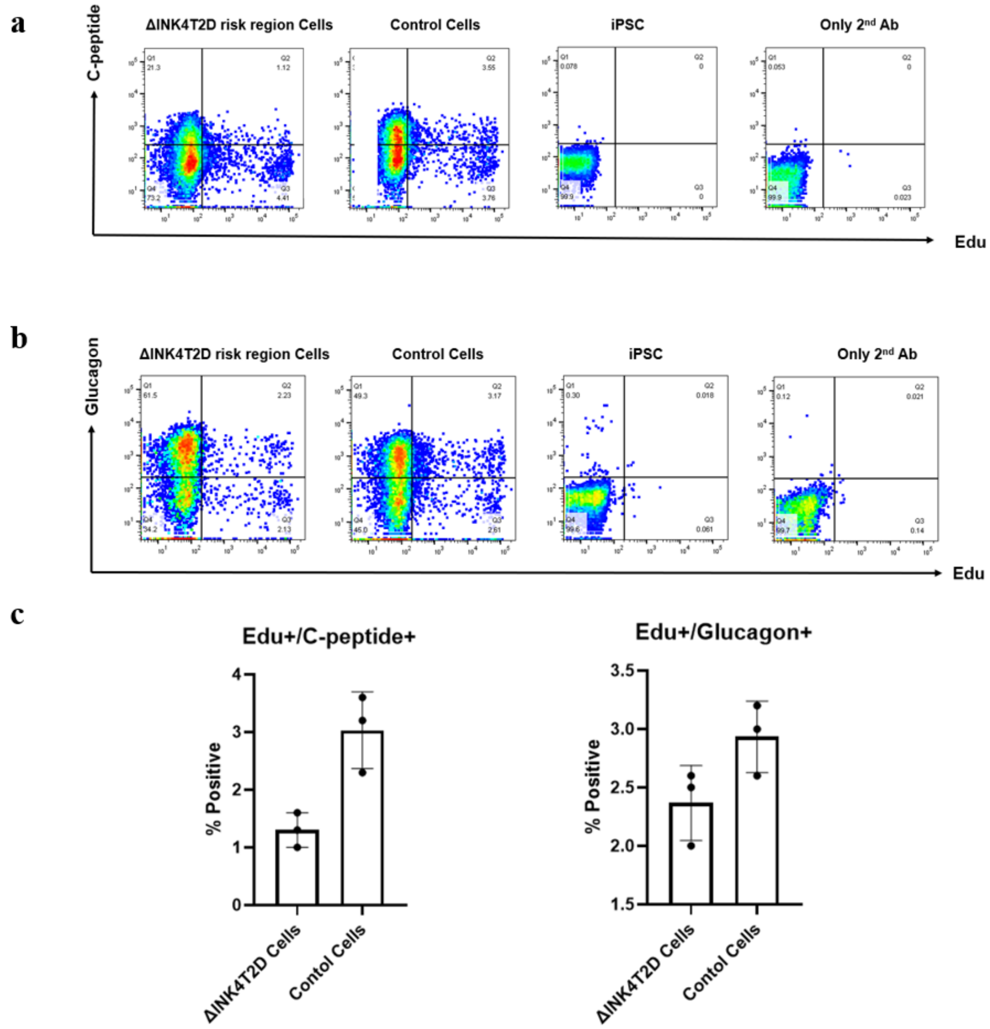


Figure 2-35: Analysis of cell proliferation for hormone+ cells at *β*-like cells stage. and B: Representative FACS dot plots of C-pep or GCG co-staining with proliferation staining marker EdU for *β*-like cells stage. (C) Quantification of FACS analysis for percentage of C-pep+ or GCG+ cells within EdU+ cells from three independent assays. The wild type iPSCs and Only-secondary-antibody samples were considered as negative controls. The data were represented as mean + S.D. (n = 3), $p < 0.001$.

2.2.3.7 Deletion of INK4-T2D Risk Region Affects Insulin Secretion in SC- β-like Cells

Finally, we performed *in vitro* functional assays on HMGUi001-A-5 -derived β-like cells at stage 6 for insulin secretion using both static and dynamic glucose-stimulated insulin secretion (GSIS). HMGUi001-A-5 cells exhibited a different pattern in insulin secretion compared to the control cells. According to the static GSIS, the ratio of C-peptide secreted in medium with high glucose (16.7 mM) to low glucose (2.8 mM) was 1.2-fold and 2.5-fold for HMGUi001-A-5 cells and control, respectively (Figure 2.36a). With dynamic GSIS, the KO cells displayed a slow first-phase insulin release following the high glucose exposure (16.7Mm). In the second phase, the ratio of secreted insulin/total insulin increased by 0.6 and 4 in the KO and control cells, respectively (Figure 2.36b). Then, the β-cells-derived KO cells showed impairment in insulin secretion. As mentioned above, the β-cells-derived KO cells showed a reduced proliferation rate and expressed less insulin and NKX6.1. Hence, reduced insulin secretion could be due to less β-cells mass or proliferation rate or less functionality in the KO cells.

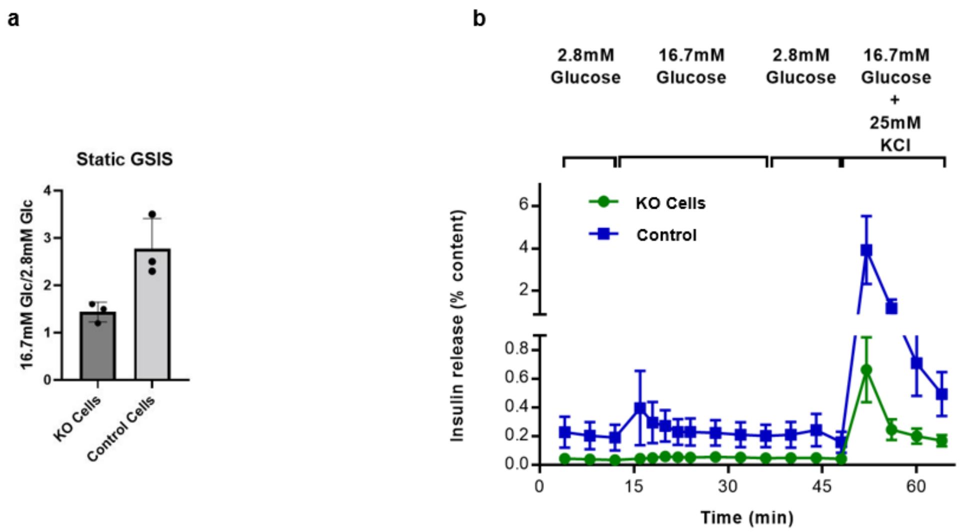


Figure 2-36: ***In vitro* analysis of static and dynamic GSIS at the *β*-like cells stage. A**: Static GSIS assay of stage 6 cells treated to either 2.8- or 18.7-mM glucose. **B**: Dynamic GSIS assay of stage 6 cells treated either 2.8 or 18.7 mM glucose.The data were shown as mean + S.D. (n = 3), p < 0.001.

3. Discussion

3.1 Increasing the Efficiency of CRISPR/Cas9

3.1.1 Improving Gene Editing Efficiency for CRISPR/Cas9 via Small RNAs in iPSCs

The CRISPR/Cas9 system is highly used because it is more efficient and easier to handle than other platforms. Despite huge progress in engineering CRISPR/Cas9 to improve its activity and efficiency, the system still needs to be more efficient, reliable and safe (Shahryari et al., 2021a). Here, we developed two methods to enhance gene editing efficiency for CRISPR/Cas9 platform. NHEJ and HDR pathways play major roles in repairing DNA upon a DNA double strand break. Since the NHEJ occurs faster than HDR, we targeted the main players of NHEJ, e.g. *XRCC4* and *DNAPKs,* by shRNA in order to boost the activity of HDR. To this end, we transfected our two established iPSC lines with a SOX2-T2A-tdTomato targeting plasmid and a triple expression cassette, sgRNA-shRNAI-shRNAII to downregulate *XRCC4* and *DNAPKs* genes. The gene-editing efficiency increased by ∼2-fold. Furthermore, the number of used expression vectors decreased to one and this could provide co-expression of shRNAs and sgRNAs at the same time. Our second approach improving the gene-editing efficiency aimed to minimize stress and thus apoptosis that results from either the transfection method or a high level of Cas9 enzyme. Small RNA miR21 shows anti-apoptotic activity, and it can improve cell viability. To increase cell survival and interfere with stress-stimulated cell death, we added miR21 expression cassette into the Cas9 expression plasmid. This approach improved the gene editing efficiency for adding the T2A-tdTomato reporter to the *SOX2* locus by ∼3-fold in GFP+ iPSCs. Eventually, we tested whether both methods together, miR21and shRNAs, have an accumulative effect on DNA editing efficiency for CRISPR/Cas9 in human iPSCs. The combination of both methods demonstrated a 3.5-fold increase in DNA editing efficiency for the CRISPR system, showing that a combined approach is beneficial over a single method.

We improved the efficiency of HDR-based gene manipulation such as gene KI or integration via temporary modulating the NHEJ pathway in human iPSCs. Partial inhibiting the major actors of NHEJ event with small chemicals improves the efficiency of precise HDR mediated DNA editing. SCR7, a known suppressor of ligase IV, boosts gene-editing efficiency for CRISPR/Cas9 system. The gene-editing efficiencies for small DNA alterations ranged from 3 to 19-fold improvement in various cell types (Maruyama et al., 2015) (Chu et al., 2015). However, in one study, the efficiency of HDR dependent gene editing did not improve by SCR7 (Pinder et al., 2015). These findings showed that the impact of SCR7 on HDR dependent DNA editing counts on the target DNA and cell types. Additionally, the efficiency of homologous recombination and precise DNA editing improved in the presence of other small chemicals such as RS-1 (activates Rad51), NU7441 and Ku-0060648 (both inhibit DNAPK) (Leahy et al., 2004) (Munck et al., 2012). We could replace small compounds (interfering with key players of NHEJ) with shRNAs that specifically targeted XRCC4 and DNAPK. The shRNA cassettes were inserted into the CRISPR plasmid, and this strategy abolished the use of chemical components. Cloning two shRNA expression segments into a single CRISPR/Cas9 vector guaranteed their co-expression and increased the SOX2 gene targeting efficiency by two-fold.

Cellular stress upon the transfection or toxicity due to high expression of Cas9 can negatively affect the efficiency of gene editing. DNA damage can activate TP53 target genes, which induces rapid cell death in human stem cells (Hug et al., 2016). Furthermore, stem cells are criticized by DSBs generated via CRISPR/Cas9, which led to the activation of TP53 pathway, resulting in diminished cell survival and induction of apoptosis (Liu et al., 2013) (Conti and Di Micco, 2018). In line with this, enormous cell death was detected upon plasmid electroporation, which reduced gene editing efficiency for CRISPR/Cas9. To counteract this, overexpression of BCL-XL (Li et al., 2018), a gene with anti-apoptotic role, increased iPSC survival and thereby enhanced the efficiency of gene targeting. Hence, it increased the efficiency of both HDR- and NHEJ-mediated gene manipulations.

To increase cell viability during the gene targeting, instead of BCL-XL, we applied miRNA-21, a miRNA that has a role in downregulating apoptosis. Apoptotic Peptidase Activating Factor 1 (APAF1) binds to free cytochromes, generating APAF1/pro-caspase9 complex. The expression of APAF1 is inhibited by miR21, and further caspase activation is downregulated. The active caspase9 stimulates caspase3 activity resulting in apoptosis. Therefore, miR21 downregulates the activation of caspase3 and thereby interferes with apoptosis. Moreover, overexpression of miR21 resulted in elevated cell survival (Buscaglia and Li, 2011) (Papagiannakopoulos et al., 2008). The small size of the encoding region for miR21 allows adding of this into the CRISPR/Cas9 vector, while the BCL-XL strategy requires an additional vector, and this reduces targeting efficiency. In contrast, incorporating sgRNA-miR21 cassette into the CRISPR/Cas9 plasmid guaranteed their co-expression in the iPSCs and therefore increased cell survival and the efficiency for *SOX2* locus targeting.

3.2 Functional Analysis of INK4 Locus in Endocrine Development and Function

3.2.1 INK4 Locus Polymorphisms and Risk of T2D

SNPs are highly regular forms of genetic diversity. Most SNPs are embedded within the non-coding genomic loci; hence they frequently are non-pathogenic. However, many SNPs can raise the risk of special disorders. For example, most SNPs associated with T2D locate in the non-coding loci and potentially rise the susceptibility to T2D. However, the mechanism by which these SNPs regulate local genome architecture remains secret for most genomic regions. Risk alleles might function in several ways, interconnecting with different genes at different loci and other polymorphisms in a tissue-specific manner. There are several potential mechanisms for how polymorphism can increase the risk of diabetes. The risk SNPs might locate within regulatory elements either enhancer or silencer and then can influence the binding of TFs. For example, T2D-associated SNPs might affect the binding of major transcription factors, such as MAFB, NKX6.1, NFAT, FOXA2, NFκB, HNF1 and PDX1 that regulate β-cell fate, development, and maturation. Additionally, the SNPs might also affect the regulation of microRNA for transcription and/or translation (Harismendy et al., 2011) (Pasquali et al., 2014). Most of these SNPs increase the risk T2D by affecting the development and/or function of the islet: The cost and limited accessibility of human islets, and the useless of non-human organisms for the human genome study, are obstacles uncovering the molecular mechanisms of the risk SNPs (Fuchsberger et al., 2016) (Gaulton, 2017) (Teumer et al., 2016) (Pasquali et al., 2014).

There are several SNPs at the INK4 locus that can increase risk of T2D and other related disorders including gestational diabetes mellitus, cystic fibrosis-related diabetes, and post-transplant diabetes, indicating a central diabetogenic mechanism. The INK4 locus encodes three genes: *CDKN2A*, *CDKN2B*, and the noncoding RNA *ANRIL*. *CDKN2A* and *CDKN2B* genes encode cell cycle inhibitors that have roles in the regulation of RB and P53 pathways. Their aberrant expressions are observed in aging, senescence, and tumorigenesis. These SNPs include rs2383208, rs10965250, rs10811661, rs10757283, rs1333051 and rs7018475 embedded within an 8 kb genomic block (termed INK4 T2D risk region) that is downstream of the ANRIL gene. Yet, the exact molecular mechanisms of how these SNPs can regulate the INK4 locus or raise the risk of T2D are still unknown (Pasmant et al., 2010) (Kong et al., 2016) (Kim and Sharpless, 2006) (Sharpless and Sherr, 2015a).
To analyze the potential functions of T2D-associated SNPs at the *INK4* locus, we aimed to create a model system by generating a human iPSC line lacking the 8 kb genomic block containing the T2D risk SNPs termed HMGUi001-A-5 line. Then, following successful characterization ofHMGUi001-A-5 line, we differentiated towards endocrine cells and/or β-like cells and examined if the deletion of the INK4-T2D risk region 1) affects the efficiency of differentiation or 2) changes the expression patterns of INK4 genes during this development or 3) affects proliferation of pancreatic progenitor or alters insulin secretion in β-like cells?

3.2.2 T2D-associated SNPs of INK4 Locus Might Affect Gene Expression

First, we wanted to address whether the deletion of T2D-associated SNPs at the *INK4* locus affects gene expression of the locus and/or neighboring genes. To this end, we tested the expression of *INK4* genes, and two neighboring genes, *MTAP* and *DMRTA1*C, at several time points during the differentiation towards β-like cells of wild type and HMGUi001-A-5 iPSCs. The *INK4* genes did not show any expression at the initial steps of differentiation. The first detectable expression of ANRIL and CDKN2B were observed during the pancreatic progenitor stage (S4) and reached its maximum after the development of endocrine and β-like cells (S6/S7) in the wild type cells. Interestingly, the expression of ANRIL and CDKN2B in the HMGUi001-A-5 cells was significantly lower than in the wild-type cells in the endocrine and β-like cells. In other words, deletion of the INK4 related T2D risk region resulted in diminished expression of ANRIL and P15 (CDKN2B) at stage 6 of differentiation or β-like cells. Both *MTAP* and *DMRTA1* genes are expressed at the maximum level at the pancreatic progenitor stage or stage 4. However, the 8 kb deletion did not affect their expression during differentiation. The data indicates that the 8 kb genomic block apparently affects the gene expression inside the INK4 locus rather than the outside.

Current data of the expression quantitative trait locus (eQTL) do not show a precise mechanism by which SNPs at the *INK4* locus increase T2D risk along local gene expression. However, those studies may not have been tested on the appropriate cell types, developmental stages, and environmental or food states to look for the exact phenotypes (Morris, 2014). These SNPs were not correlated with the expression of *INK4* genes, *CDKN2A* and *CDKN2B* in the pancreas, colon, and liver (Fadista et al., 2014). ANRIL expression is correlated with the expression of *INK4* genes, *CDKN2A*, and *CDKN2B* in several tissues, indicating coordinated regulation of the locus; however, independent regulation is also observed in several reports. Cardiovascular risk SNPs at the *INK4* locus were associated with ANRIL expression. Both T2D and non-T2D SNPs located at the *INK4* locus apparently reveal a more substantial effect on the expression of *ANRIL* in compared with other *INK4* genes expression (Cunnington et al., 2010b) (Gil and Peters, 2006) (Folkersen et al., 2009).

Kong et al. studied whether T2D risk SNPs located at the *INK4* locus impact gene expression of the locus, insulin release and β-cell proliferation in human islets. Totally, 95 islets from healthy donors without diabetes were evaluated for SNPs genotype of *rs10811661*, *rs2383208*, *rs564398*, and *rs10757283*, and gene expression of CDKN2A, CDKN2B, ARF, *MTAP*, and *ANRIL*. The expression analysis showed the INK4 genes are coordinately expressed in human islets. The expression patterns of *ARF*, *CDKN2A*, and *ANRIL* (but not *CDKN2B*) were highly correlated with each other, which increased with age. On the other hand, CDKN2B expression was significantly correlated with the expression of *MTAP*. Hence, the *INK4* locus genes are co-regulated in the human islets in two physically overlapped partners: *ARF-CDKN2A-ANRIL* and *MTAP-CDKN2B*. Risk alleles at *rs10811661* and *rs2383208* were associated with high expression of *ANRIL*, but not *ARF*, *CDKN2A/B* and *MTAP*, in an age-dependent manner e.g., younger donors with homozygous risk genotype revealed higher *ANRIL* expression. Interestingly, they detected combinations of several risk SNPs that might influence the expression of the locus indicating the presence of potential mechanisms by which SNPs can control islet's function and biology (Kong et al., 2018).

In mouse, deletion of the noncoding cardiovascular risk region (part of ANRIL and its downstream) on chromosome 4 resulted in diminished aortic expression of Arf and Cdkn2b (but not Cdkn2b), elevated CDK-dependent Smad2 linker phosphorylation, and decreased canonical TGF-β-dependent Smad2 phosphorylation. These results are correlated with raised susceptibility to aneurysm development. The KO mice are partly saved via a single therapy with a CDK inhibitor (Loinard et al., 2014).

As mentioned earlier, other INK4 polymorphisms significantly increase the risk of CAD. These SNPs are located in 58 kb interval, a part of ANRIL gene and its downstream sequence. Visel et al. showed that deletion of its orthologous sequence at the mouse Ink4 locus on chromosome 4 affects the expression of the neighboring genes in several vascular-relevant cell types. Cardiac expression of Cdkn2a/b, is significantly reduced in KO mice, demonstrating that regulatory roles are embedded within the CAD risk region at the mice Ink4 locus. Interestingly, allele-specific expression of Cdkn2b in heterozygous mice demonstrated that the deletion has an impact on expression via a cis acting manner. The results yielded strong proof that the CAD risk interval regulates cardiac expression of Cdkn2a/b in mice (Visel et al., 2010).

There is a regulatory element termed RDINK4/ARF upstream of the CDKN2B gene and shows a positive effect on the expression of *CDKN2A/B* and *ARF* genes. Evidently, the genetic changes of RDINK4/ARF, either heterozygous or homozygous deletions, can transcriptionally downregulate all of the *INK4* genes e.g. *CDKN2A/B* and *ARF*, thus contributing to cancer initiation and/or progression. The RDINK4/ARF fragment is frequently deleted (monoallelic or biallelic) in various human cancer cell lines, demonstrating its potential role in carcinogenesis. INK4 genes are highly deleted in gastrinomas and pancreatic neuroendocrine tumors (rare endocrine tumors emerging from the islet cells). Poi et al. reported deletion of RDINK4/ARF in pancreatic neuroendocrine tumors, demonstrating a functional role in islet cells fate and development (Poi et al., 2014) (Gonzalez et al., 2017) (Li et al., 2014) (Evers et al., 1994) (Muscarella et al., 1998).

Unbalanced metabolic inputs such as overfeeding or food limitation can affect CpG methylation at the human *INK4* locus. Additionally, DNA methylation can be changed by risk alleles of polymorphisms such as SNPs in the human genome. These changes could affect chromatin state and/or local gene expression. We have not examined DNA methylation patterns for the SNPs that are in the INK4 T2D-risk genomic block. However, there is a T2D-associated SNP, *rs564398,* outside of the INK4-T2D risk region, located inside the *ANRIL* gene. The risk allele of this SNP disrupts a DNA methylation CpG site, resulting in diminished methylation of neighboring CpG sites. Although it does not affect local gene expression, it decreases insulin content in human islets. However, its molecular mechanism still remains unclear (Dayeh et al., 2013) (Popov and Gil, 2010) (Jacobsen et al., 2012) (Daniel and Tollefsbol, 2015).

3.2.3 The Functions of INK4 Genes in Regulation of β-cell Proliferation

Many reports demonstrate fundamental roles for *INK4* genes in the regulation of the cell cycle and/or cell proliferation. Overexpression of *CDKN2A* gene yielded a decrease in β-cell proliferation in young mice. Accordingly, the knockout mice model for CDKN2A restored the age-related loss of proliferation (Krishnamurthy et al., 2006). In line with this, the ectopic expression of EZH2, a component of PRC2, downregulated CDKN2A, consequently increasing β-cell proliferation in young mice. The similar phenotype was also reported in older mice upon the downregulation of CDKN2A (Zhou et al., 2013). Downregulating CDKN2A via siRNA technology reduced the loss of β-cell proliferation in an *ex vivo* system. Furthermore, targeting WIP1/p38MAPK/BMI1 or PTEN/E2F/EZH2 pathways downregulated CDKN2A expression resulting in elevated β-cell proliferation in aging mice (Pascoe et al., 2012) (Wong et al., 2009) (Zeng et al., 2013). P18INK4c as an INK family inhibitor could synergistically increase β-cell proliferation rate in the CDKN2A knockout mice in a CDK4-dependent manner. (Ramsey et al., 2007). Increased expression level of CDKN2A reduced islet regenerative capacity with age. An antagonist of HNF4α induced β-cell proliferation by downregulating the CDK inhibitor, CDKN2A. Furthermore, suppressing HNF4α increased the cell proliferation rate in α-cells, β-cells and δ-cells (Chen et al., 2009)

(Kim et al., 2019). All these studies support the idea that CDKN2A is associated with decreased β-cell mass with age.

Second, we wanted to address "does deletion of the T2D risk region at the *INK4* locus affect β-cell proliferation or not? To this end, we treated the cell at three time points of differentiation e.g., pancreatic progenitors, endocrine cells and β-like cells with Edu. The rate of proliferation was significantly reduced (approximately two-fold decrease) in pancreatic progenitors, endocrine cells, and hormone-producing cells in the KO cells compared to the wild type cells. Furthermore, both C-peptide positive cells and Glucagon positive cells showed less proliferation rate in the KO cells compared with their counterparts in the control cells.

In mice, CDKN2A and CDKN2B, also known as INK family inhibitors, physically inhibit CDK4/6. The inhibitor of INK family, and CDK4 play fundamental roles in islet biology and function. Knockout mice models for CDK4 represent severe insulin-deficient diabetes due to hypoplastic islets. CDK4 knockout also affects other endocrine systems, male and female infertility, impaired growth and proliferation, and pituitary defects. In this model, the morphology of the islets appears normal at the early stages of life, indicating CDK4 is not mandatory for pancreatic formation; however, islets show no proliferation during postnatal growth. Hence, CDK4 is crucial for regulating postnatal mouse β-cell mass (Rane et al., 1999).

Visel et al. also showed that deletion of mouse orthologous for INK4-CAD risk region affects the proliferation rate of vascular cells. KO mice showed a lower survival rate both during development and adulthood. In KO mice, primary cultures of aortic smooth muscle cells revealed more proliferation rate and reduced senescence; both were consistent with the pathogenesis of accelerated CAD. Altogether, their data suggested that the CAD risk interval at the INK4 locus has a significant role in regulating the locus. It can affect CAD progression by changing the proliferation dynamics of cardiovascular cells (Visel et al., 2010).

A challenge is that human β-cell mass and proliferation could not be calculated in living human. Nevertheless, few studies examined whether INK4 polymorphism influences human β-cell proliferation or mass. Kong et al. measured cell cycle entry in 47 islets from healthy donors via the BrdU incorporation. To examine if any of the INK4 T2D SNPs influence β-cell proliferation, the proliferation index was stratified by SNP genotypes. Risk alleles of *rs2383208*, *rs10811661*, and *rs10757283* did not differ in the proliferation index. However, the risk allele of rs564398 was highly correlated with a reduced β-cell proliferation index. The homozygous risk alleles demonstrate approximately half induction of proliferation rate compared with islets carrying healthy alleles for this SNP. This suggests that *rs564398* or perhaps ANRIL have a functional role in maintaining human β-cell mass or proliferation (Kong et al., 2018).

3.2.4 The Functions of INK4 Genes in Regulation of Insulin Secretion

Several studies have reported that misregulation of *INK4* genes and their downstream target genes impair insulin signaling pathways. For example, the CDKN2A Knockdown increased insulin secretion capacity in the EndoC-bH1 human β-cell line (Pal et al., 2016). In mice, telomerase haploinsufficiency induced CDKN2A expression resulting in decreased in insulin secretion via impaired regulation of exocytosis. Glucose intolerance was also observed; however, β-cell mass was normal, indicating the main impact on β-cell function rather than β-cell proliferation (Pulizzi et al., 2009). The downstream target of INK4, CDK4 can also regulate insulin secretion. This is triggered through Rb-dependent transcriptional regulation activation mediated by E2F1. The E2F1 transcription factor induces transcription of *Kir6.2* gene coding potassium inward rectifying channel involved in insulin secretion. In humans, patients suffering from familial melanoma with heterozygous loss of function in the *CDKN2A* gene revealed elevated insulin secretion, reduced insulin sensitivity and decreased hepatic insulin clearance (Annicotte et al., 2009). These data support the *INK4* genes, mainly *CDKN2A* could impact on insulin secretory role, insulin clearance and insulin sensitivity.

Third, we aimed to address whether the deletion of the T2D risk region at the INK4 locus affects insulin secretion in the β-cell. To this end, we evaluated insulin secretion via the GSIS method in iPSC-derived β-cells. Deletion of the 8 kb genomic block reduced rate of insulin secretion. The β-cells derived from HMGUi001-A-5 expressed less insulin/NKX6.1 than the control cells. The decreased insulin secretion could be due to less proliferation rate or less β-cells mass, or less functionality.

The T2D risk SNPs located in the human *INK4* locus might affect diabetes risk via insulin secretory capacity of β-cells in pancreas and insulin sensitivity of other organs. Therefore, *CDKN2A/B* locus SNPs could influence the biology of pancreatic islets and other metabolic tissues. It has been reported that the link between *CDKN2A/B* polymorphism *rs10811661* and T2D is influenced by age. This agrees with the identified interaction between age and CDKN2A role in islets (Perry and Frayling, 2008) (Peng et al., 2013). T allele of *rs10811661* as a risk allele is linked to a diminished insulin secretion capacity after both oral and intravenous glucose challenges. The reduced insulin secretory capacity due to *INK4* polymorphisms could be due to failures in β-cell functions such as glucose sensing, insulin production, stimulus-secretion coupling or reduced β-cell proliferation and mass (Grarup et al., 2007) (Hribal et al., 2011).

Kong et al. tested 61 islets from healthy donors for insulin secretion stimulation index. This index was associated with BMI but indicated no correlation with donor age and sex. However, when the insulin release index was stratified by the *INK4* SNP genotype, the T2D risk alleles did not demonstrate any proof for failure in ex vivo glucose sensing, insulin production and release in this small group (Kong et al., 2018).

3.3 Conclusion

Our methods revealed 1) temporary co-expression of shRNAs downregulating the major actors of NHEJ, XRCC4 and DNA-PK, and 2) temporary co-expression of miR21 with sgRNA and Cas9 improve the efficiency of HDR dependent DNA editing for CRISPR/Cas9. Our novel CRISPR/Cas9 targeting method abolishes the use of more vectors and/or compounds and still interferes with the NHEJ pathway or stress-induced apoptosis. The improved DNA editing efficiency was consumable due to the diminished downstream workload necessary for screening the cells having the desired gene editing.

Initial attempts to use GWAS data to diagnose human diseases and develop therapies failed. Despite progress and valuable genetic data, T2D-associated SNPs in the *CDKN2A/B* locus do not yet have significant clinical implications and have low prognosis or disease risk value (van Hoek et al., 2008) (Majithia and Florez, 2009). There are many reports on the molecular functions of the *INK4* locus genes in regulating the biology of islets and other metabolic organs, mainly in rodents and humans. *INK4* genes regulate functions of pancreatic islet, adipose, muscle, liver and immune cells at various stages ranging from embryonic development to aging. The *INK4* genes, especially *CDKN2A* have critical roles in the regulation of rodent β-cell mass, but the exact molecular mechanism of how *INK4* SNPs increase the risk of T2D is not fully understood. Genomic deletion of the 8 kb risk region harboring all T2D-related SNPs at the *INK4* locus let to diminished β-cell proliferation and reduced insulin content and secretion in iPSC-derived β-cell. Gene regulation analyses do not still demonstrate a clear relationship between these SNPs and local or distal gene expression. To examine the exact molecular functions polymorphism at the *INK4* locus, there is an urgent need to focus on each single SNPs. Another point is that analyses should perform in the proper developmental stage, proper tissue, metabolic context and/or subpopulation. Less availability of human samples in terms of tissues and developmental stages is a significant challenge. Hopefully, *INK4* polymorphisms associated with T2D will improve our understanding of GWAS and diabetes and provide clinical applications in the future.

4. Materials and Methods

4.1 Solution and Buffers

The compounds and chemical that were used for preparation of solutions and buffers are listed in Table 4.1

Table 4.1. Solutions and buffers

Solutions and buffers	Composition
Solutions and buffers for immunostainings	10x PBS: 1.37 M NaCl, 26.8 mM KCl, 0,101 M Na_2HPO_4, 13.8 mM KH_2PO_4
	PBST: 1x PBS + 0.1% Tween20 (adjust to pH 7.4)
	4% PFA: 1.3 M PFA in 1x PBS (adjust to pH7.2-7.4)
	Permeabilisation (sections): 0.2% TritonX-100, 100 mM Glycin in dH_2O
	DAPI: 5 mg DAPI in 25 ml PBS
	Blocking solution: 5% FCS, 1% serum (goat or donkey) in PBST
	Elvanol (embedding): 0.015 mM Polyvinyl-alcohol, 24 mM Tris pH 6.0, 2 g DABCO in 90 ml H_2O and 37.8 ml Glycerol
	Antigen Retrieval: 10N HCl in H_2O

	10x Tris-Borat-Buffer: 10 mM $Na_2B_2O_7$ in dH_2O
Glucose stimulated insulin secretion (GSIS)	10x Krebs buffer: 1.2 M NaCl, 48 mM KCl, 25 mM $CaCl_2*2H_2O$, 12 mM $MgCl_2$ in dH_2O
	1x Modified Krebs buffer: 1x Krebs buffer, 5 mM HEPES, 0.025 mM $NaHCO_3$, 0.1% BSA in H_2O (adjust to pH7.4)
	FACS buffer: 1x PBS (-Ca/Mg), 3% FCS, 5 mM EDTA
	DNA lysis buffer: 100 mM Tris pH 8.0, 5 mM EDTA pH 8.0, 200 mM NaCl, 0.2% SDS in H_2O
Solutions and buffers for cell culture	DPBS (-Ca/-Mg), Gibco Trypsin-EDTA, 0.05% or 0.25% Trypsin, 0.53 mM EDTA•4Na, Gibco Penicillin/Streptomycin (100x) Gibco iPS Brew medium, Gibco MCBD131, Gibco Accutase, Gibco

4.2 sgRNA design and CRISPR-Cas9 plasmid construction

To add the tdTomato reporter sequence to the end of the *SOX2* gene, we designed sgRNA to target 3' end of the *SOX2* ORF. The CRISPOR web tool (*crispor.tefor.net)* was used to design the sgRNA. For better expression and compatibility with the BbsI site of the vector, CACCGGG oligo was added to the 5' end of the sense sgRNA and AAAC and CCC oligos to 5' and 3' end of the antisense (the sgRNA sequence is listed in Table 4.2). To clone the sgRNA oligos, PU6-(BbsI) sgRNA_CAG-GFP-bpA plasmid (Addgene ID86985) was digested with BbsI. Following the cloning, the expression plasmid was subjected to Sanger sequencing.

To generate a large deletion at the INK4 locus, we designed sgRNAs via the web tool CRISPETa (crispeta.crg.eu) (the sequences of the oligos are listed in Table 4.3). Initially, sgRNAs DNA binding sites were sequenced due to genome variation and low conservation. The oligos CACCGGG, AAAC and CCC were added to the sgRNAs sequences as described above. PU6-(BbsI) sgRNA_CAG-Cas9-GFP-bpA plasmid, Addgene ID86985, containing BbsI site was used for cloning single gRNAs.

Table 4.2. List of oligonucleotides and primers

sgRNA, shRNA and primer Oligos	**Sequence (5´--3´)**
sgRNA sense / antisense	caccgggCGGCCCTCACATGTGTGAGA / aaacTCTCACACATGTGAGGGCCGccc
shRNA_XRCC4_F	CCGGGCATGGACTGGGACAGTTTCTCTCGAGAGAAACTGTCCCAGTCCATGCTTTTTG
shRNA_XRCC4_R	AATTCAAAAAGCATGGACTGGGACAGTTTCTCTCGAGAGAAACTGTCCCAGTCCATGC
shRNA_DNAPK_F	CCGGGCATCCAGAGTAGCGAATACTCTCGAGAGTATTCGCTACTCTGGATGCTTTTTG
shRNA_DNAPK_R	AATTCAAAAAGCATCCAGAGTAGCGAATACTCTCGAGAGTATTCGCTACTCTGGATGC
EP 1738-AF	CAGGAAACAGCTATGACCATGAGGGCCCCCTTCACCGAGGGCCTATTTC
EP 1739-AR	CCGATGGCCAGGCCGATGCTGTGATCAAAAAAAGCACCGACTCGG
EP 1740-BF	ACAGCATCGGCCTGGCCATCGGGCCCCCTTCACCGAGGGCCTATTTC
EP 1808-BR	CTTGGCCATCTCGTTGCTGAAGATCTCTGCTGTCCCTGTAATAAACCC
EP 1742-CF	TTCAGCAACGAGATGGCCAAGGCCCCCTTCACCGAGGGCCTATTTC
EP 1809-CR	GTCAATAATCAATGTCGAATCCGGGATCTCTGCTGTCCCTGTAATAAACCC
XRCC4_Fwd / XRCC4_Rev	AGGAGACAGCGAATGCAAAG / CTTCTGGGCTGCTGTTTCTC (PCR product: 178 bp)
DNAPK_Fwd / DNAPK_Rev	GGAACAGCAGCATGTCATGG / CTGGCGTGTGAAACTTAGGC (PCR product: 148 bp)
GAPDH_Fwd / GAPDH_Rev	GGCCAAGGTCATCCATGA / TCAGTGTAGCCCAGGATG (PCR product: 354 bp)
EP 034/ EP 035	GGAAACAGCTATGACCATG / CTCCGAGGACAACAACATGG (PCR product: 1888 bp)

Table 4.3. SgRNA Oligos and screening primers

Name	Sequence
sgRNA1-sense	caccgggATATGGCTAAATAGTCCGTA
sgRNA1-antisense	aaacTACGGACTATTTAGCCATATccc
sgRNA2-sense	caccgggCCACCATGATCTAGCACTAA
sgRNA2-antisense	aaacTTAGTGCTAGATCATGGTGGccc
Forward-A (FA)	CCAAATTGCCTCAGCCAATG
Reverse-A (RA)	CAAATGGCCTTAGCCAGAGC
Forward-B (FB)	AAGCCACTTAGCTAGAGTAAGG
Reverse-B (RB)	CACCAGTCGTGTTGGATAAATG
Forward-C (FC)	GGAGCCATTCTATCGTGAACAG
Reverse-C (RC)	AAGCATAGGTGGGTTTCACTTC

4.3 shRNA and miRNA-21 expressing plasmids construction

ShRNAs were designed using Invitrogen's Block-It RNAi Designer (*https://rnaidesigner.invitrogen.com/rnaiexpress*). shRNAs supposed to target the leading players of NHEJ repair pathway; DNAPK, and XRCC4 (Table 4.2). Two shRNAs per gene were designed. AgeI and EcoRI sites were added to the 5´ and 3´sites of shRNA oligos, respectively. Furthermore, *CTCGAG* oligo as a palindromic loop was incorporated between the complementary regions of the oligos. Then complement oligos were annealed and ligated into pLKO.1 plasmid (Addgene: ID8453) that was already digested with AgeI and EcoRI. The sequences for shRNAs and scramble are listed in Table 4.9. Furthermore, to generate miRNA-21 expressing vector, a 72-bp sequence of pre-miRNA-21 (NC_000017.11) with 50 nucleotides harboring sequences at both sides was amplified by PCR using human genomic

DNA as a template. AgeI and EcoR1 sites were also added to the 5´ sites of the forward and reverse primers, respectively. The PCR product was added to the pLKO.1. vector. Following the cloning, all the plasmids were subjected to Sanger sequencing.

4.4 Gibson assembly reaction for generating sgRNA-shRNA and sgRNA-miRNA cassettes

As Yumlu et al. (Yumlu et al., 2019) described we used Gibson assembly technique to generate a dual small RNA expression cassette in the vector either sgRNA or shRNA. Both single and dual gRNAs expression vectors were subjected to Sanger sequencing. Following incorporation of the cassettes into the PU6-(BbsI) sgRNA_CAG-Cas9-GFP-bpA plasmid, the sgRNAI-sgRNAII, sgRNA-shRNAI-shRNAII or sgRNA-miRNA-21 were generated. For amplification, Yumlu et al protocol was changed by adding two more primers EP1808-BR and EP1809-CR for amplification of the miRNA (Table 4.9). Following the cloning, the resulting vectors were subjected to Sanger sequencing.

4.5 Human iPSC culture and plasmids transfection

We used two iPSCs lines, HMGUi001-A and HMGUi002-A. The cell lines were cultured on Geltrex-coated 6-well plates in the medium of StemMACS iPS-Brew. The cells were disassociated using 5 mM EDTA in PBS at 37 °C for 3 min or incubated with StemPro Accutase Cell Dissociation Reagent. To avoid cell death or cellular stress during the splitting time, 10 µM Y-27632 was added to the medium. The iPSCs were maintained at 37 °C, 20% O2 and 5% CO2. For the lipofectamine-based transfection, CRISPR-Cas9 plasmid containing the cassettes of sgRNA-shRNAI-shRNAII or the cassettes of sgRNA-miRNA21SOX2 were mixed with targeting plasmid of pUC19-SOX2-T2A-2xNLS-tdTomato-F2A-Puro (Addgene ID89991). 24 hours before the transfection, the cells were seeded at the density of $2x10^5$ cells per well of the 6-well plate. 1 µg CRISPR-Cas9 plasmid and 2 µg SOX2 targeting plasmid were transfected into the iPSCs using 5µl Lipofectamine™ Stem Transfection Reagent and 200 µl OptiMEM medium, per each well according to the manufacturing instruction. Three days following the transfection, the cells were dissociated and harvested for flowcytometry analysis as well as RNA extractions.

4.6 Fluorescence activated cell sorting (FACS) enrichment of transfected iPSCs

72 hours (SOX2 gene targeting) after the transfection, iPSCs were disassociated as described above and collected for the FACS. The cells were mixed in 0.5 ml iPS-Brew medium containing 10 µM Rock inhibitor and DAPI. Then, the cells suspension was filtered by a cell strainer and used for sorting. The efficiency of gene targeting in the transfected cell population was calculated by analyzing GFP- and Tomato- signal using FACSAria III. For establishing a stable fluorescent SOX2 reporter iPSC line double-positive GFP/tdTomato cells were FACS sorted and two thousand GFP+ cells were seeded on Geltrex coated 10 cm dishes. After 7-9 days single cell-derived iPSCs clones were picked, transferred to new plates, expanded, and subjected to PCR screening on isolated genomic DNA.

For the INK4 gene targeting, 48 hours after the transfection, the cells were collected for the FACS analysis. Using the FACS Aria III, the cells with high GFP signal levels were sorted. Around 4000 cells were seeded in 10 cm plates. After one week, single cell-derived clones were picked, expanded, and screened by PCR. To quantify the clones' pluripotency expression, flow cytometry was used for OCT4 and SOX2 markers. The result data was analyzed using FlowJo software.

4.7 Clone screening for CRIPR/Cas9-mediated genome editing

The appropriate PCR primers for sequencing and clones screening for both SOX2 and INK4 gene targeting were designed using Clone Manager Molecule software (Table 4.9 and Table 4.10). For SOX2 gene KI, three-primer PCR using EP1890, EP671 and EP1891 primers for the left arm yields a 1270 or1409 bp representing KI or WT alleles, respectively. Furthermore, PCR for the right arm using EP1892, EP1893 and EP036 primers yields 852 or 1265 bp representing KI or WT alleles, respectively. PCR reactions were performed using Taq DNA Polymerase enzyme using the following thermal conditions: initial denaturation at 95°C for 3 min, amplification for 40 cycles with denaturation at 95°C for 30 sec, annealing at 60°C for 30 sec and extension at 72°C for 90 sec. Direct DNA sequencing was performed to confirm the authenticity of the PCR products.

In INK4 gene targeting, as listed in Table 4.10, FA-RB and FC-RC primer pairs were used to detect the deletion amplicons. Long Amp Taq DNA Polymerase enzyme (NEB) was used for PCR reactions. The 750 bp (the PCR product of FA and RB primers) and 1910 bp (the PCR product of FC-RC primers) PCR products were indicators for biallelic deletion and wild type, respectively. To confirm the deletion's authenticity at both alleles, the PCR products were cloned into a TA vector (NEB) and then were sequenced.

4.8 Characterization of the hiPSCs for karyotyping and STR analysis

To carry out karyotyping, the positive clones (hiPSCs P29) were collected at logarithmic phase point during the cellular growth. For this, the cells were incubated with the Colcemid chemical for 120 min, disassociated with Accutase, and then treated with hypotonic solution (0.075 M KCl) for 20 min. Then, the cells were fixed by methanol/acetic acid solution at a ratio of 3:1. Metaphase chromosomes were determined using the standard G banding approach. Around 50 counts were performed at the metaphase stage, and the average of 85% was defined as the final karyotype. Finally, AmpFℓSTR™Identifiler™ PCR Amplification Kit (appliedbiosystems, Cat# 4322288) was applied for STR analysis according to the manufacturer's instructions.

4.9 Three germ layers differentiation

The HMGUi001-A-5 hiPSC line was directly differentiated towards three germ layers using StemMACS™ Trilineage Differentiation Kit (Miltenyi Biotec, Cat# 130-115-660) according to manufacturer's instructions. Next, the differentiated cells were immunostained for expression of endoderm (SOX17 and FOXA2), mesoderm (CD144 and SM22a) and ectoderm (Nestin and SOX2) markers.

4.10 Cell viability assay

In order to evaluate cell viability during the transfection, the cells were examined at two time points, 8 and 24 hours after lipofectamine transfection by 0.4% trypan blue dye. Cells supernatant containing dead cells were collected. Adherent cells were dissociated as described above. Floating dead cells and dissociated cells were centrifuged and, resuspended in a 1 ml medium. 20 µl resuspended cells and 20 µl of 0.4% trypan blue were mixed and waited for 3 min. Then, the number of dead and viable cells was counted on a hemocytometer.

4.11 Human iPSCs differentiation towards insulin producing ß-like cells

4.11.1 Rezania Protocol

Our iPSCs lines (HMGUi001-A-5 and SOX2-T2A-tdTomato reporter iPSCs) generated from HMGUi001-A were cultured on Geltrex coated 10 cm plates in the StemMACS iPS-Brew XF. The confluent iPSCs at confluency 80% were dissociated into single cells using *Accutase* and cultured at ~75,000 cells/cm^2 in the iPS-Brew medium supplemented with 10 µM Rock inhibitor in *Ultra*-Low *Attachment* 6-well plates. The medium was changed one day later and the differentiation process was started. The iPSCs were differentiated towards PDX1+NKX6.1+ pancreatic progenitor and ß-like cells according to the Rezania protocol (Rezania et al., 2014). The details of protocol, materials and concentrations are listed in Table 4.4, Table 4.5, Table 4.6, and Table 4.7. The aggregates were collected and fixed at the end of each time point of differentiation.

Table 4.4. Rezania protocol

Stage	Differentiation Protocol
S1 (3 d): **definitive endoderm**	2.0 x10^6 cells per each well of 6-well low-binding plate were seeded in 4 ml of MCDB 131 medium further supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax 10 mM final glucose concentration, 0.5% BSA, 100 ng/ml Activin-A, and 1.5 µM of CHIR-99021 for day 1 only. For day 2, cells were cultured in MCDB with same medium and component but 0.1 µM of CHIR-99021. The third day, CHIR was completely deleted.
S2 (2 d): **primitive gut tube**	Cells were 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 and 50 ng/ml of FGF7 for 2 days.
S3 (2 d): **posterior foregut**	Cultures were continued for 2 d in MCDB 131 medium further supplemented with 2.5 g/l sodium bicarbonate, 1× Glutamax, 10 mM final glucose concentration, 2% BSA, 0.25 mM ascorbic acid, 50 ng/ml of FGF7, 0.25 µM SANT-1, 1 µM

	retinoic acid (RA) 100 nM LDN193189 (LDN; BMP receptor inhibitor) 1:200 ITS-X, and 200 nM TPB (PKC activator)
S4 (3 d): **pancreatic endoderm, PDX1$^+$/NKX6.1$^+$ cells**	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.
S5 (3 d): **pancreatic endocrine precursors, PDX1$^+$/NKX6.1$^+$/NEUROD1$^+$**	The cells were exposed to MCDB 131 medium supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 20 mM final glucose concentration, 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), 10 μM ALK5 inhibitor II, μM zinc sulfate and 10 μg/ml of heparin for 3 d.
S6 (7–15 d): **NKX6.1$^+$/insulin$^+$ cells**	MCDB 131 medium further supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 20 mM final glucose concentration, 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 for the first 7 d only and 10 μg/ml of heparin for 7–15 d.
S7 (7–15 d): **NKX6.1$^+$/insulin$^+$/MAFA$^+$ cells**	MCDB 131 medium supplemented with 1.5 g/l sodium bicarbonate, 1× Glutamax, 20 mM 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, 10 μM Trolox (Vitamin E analogue), 2 μM R428 (AXL inhibitor) and 10 μg/ml of heparin for 7–15 d. Unless otherwise specified, for all stages, the cultures were fed daily.

Table 4.5. Components and concentrations in Rezania Protocol

Stage	day	media	supplement

S0 (2d)	d1	Seeded 75,000 hPSCs/cm2 on vitronectin-coated plates in E8 media with 5 uM Y-27632.	
	d2	Changed media.	
S1 (3d)	d0	S1-2	Activin A 100 ng/ml CHIR-99021 5 uM
	d1	S1-2	Activin A 100 ng/ml CHIR-99021 0.3 uM
	d2	S1-2	Activin A 100 ng/ml
	d3	S1-2	>80% DE cells expressing endoderm markers including CXCR4, SOX17 and FOXA
S2 (2d)	d3-4	S1-2	Vitamin C 0.25 mM FGF7 50 ng/ml, IWP-2 1.25 uM
S3 (2d)	d5-6	S3-4	Vitamin C 0.25 mM, FGF7 50 ng/ml, SANT-1 0.25 uM, RA 1 uM, LDN 100 nM, ITS-X 1:200, TPB 200 nM
S4 (3d)	d7-9	S3-4	Vitamin C 0.25 mM, FGF7 2 ng/ml, SANT-1 0.25 uM, RA 0.1 uM, LDN 200 nM, ITS-X 1:200, TPB 100 nM
S5 (3d)	d10-12	S5-7	T3 1 uM, Alk5i II 10 uM, SANT-1 0.25 uM, RA 0.05 uM, LDN 100 nM, ITS-X 1:200, ZnSO4 10 uM, Heparin 10 ug/ml
S6 (7d)	d13-19	S5-7	T3 1 uM, Alk5i II 10 uM, GSiXX 100 nM, LDN 100 nM, ITS-X 1:200, ZnSO4 10 uM, Heparin 10 ug/ml
S7 (2w)	d20-33	S5-7	T3 1 uM, Alk5i II 10 uM, N-Cys 1 mM, Trolox 10 uM, R428 2 uMITS-X 1:200, ZnSO4 10 uM, Heparin 10 ug/ml

Table 4.6. Medium components for Rezania Protocol

Media Stage	Media component (add corresponding supplements listed above)
S1-2	MCDB 131 GlutaMAX 1X BSA 0.5% NaHCO3 1.5 g/L Glucose 10 mM
S3-4	MCDB 131 GlutaMAX 1X BSA 2% NaHCO3 2.5 g/L Glucose 10 mM
S5-7	MCDB 131 GlutaMAX 1X BSA 2% NaHCO3 1.5 g/L Glucose 20 mM

Table 4.7. Commercial data of components at Rezania Protocol

Components	**Vendor**	**Cat. No.**
MCDB 131	GIBCO	10372-019
GlutaMAX	GIBCO	35050-061
NaHCO3	Fisher Scientific	144-55-8
D-Glucose	Sigma-Aldrich	G8769
BSA (bovine serum albumin)	LAMPIRE	7500855
Activin A	PeproTech	120-14E
CHIR-99021, GSK-3 inhibitor	Stemgent	04-0004
L-Ascorbic acid (vitamin C)	Sigma-Aldrich	A4544
FGF7 (KGF)	R&D	251-KG
SANT-1, Hedgehog inhibitor	Sigma	S4572
RA (retinoic acid)	Sigma	R2625
LDN, BMP inhibitor	Stemgent	04-0019
ITS-X (insulin-transferrin-selenium-ethanolamine)	GIBCO	51500-056
TPB, Protein kinase C (PKC) activator	EMD Millipore	565740
T3 (3,3',5-Triiodo-L-thyronine)	Sigma-Aldrich	T6397
ALK5i II (ALK5 inhibitor II)	Enzo Life Sciences	ALX-270-445
ZnSO4	Sigma-Aldrich	Z0251
Heparin	Sigma-Aldrich	H3149
GSiXX (gamma secretase inhibitor XX)	EMD Millipore	565789
N-Cys (N-acetyl cysteine)	Sigma-Aldrich	A9165
Trolox, vitamin E analogue	EMD Millipore	648471

R428, AXL receptor tyrosine kinase inhibitor	Selleck Chemicals	S2841
IWP-2, Wnt antagonist	Tocris Bioscience	3533

4.11.2 Millman Protocol

To initiate differentiation with Millman protocol (Millman et al., 2016), undifferentiated iPSCs were single-cell dispersed using Accutase. Then the cells were seeded at a density of $6x10^5$ cells/mL in StemMACS iPS-Brew XF containing 10 μM Y27632 in a 30-ml spinner flask. Cells were then passaged at least three times and cultured in the differentiation medium. The details of protocol, materials and concentrations are listed in Table 4.8, and Table 4.9.

Table 4.8. Components and concentrations in Millman Protocol

stage	Media + Supplement
Stage 1 (3 days)	S1 media + 100 ng/ml Activin A + 3 μM Chir99021 for 1 day. S1 media + 100 ng/ml Activin A for 2 days
Stage 2 (3 days)	S2 media + 50 ng/ml KGF
Stage 3 (1 day)	S3 media + 50 ng/ml KGF + 200 nM LDN193189 + 500 nM PdBU + 2 μM RA+ 0.25 μM Sant1 + 10 μM Y27632
Stage 4 (5 days)	S4 media + 5 ng/mL Activin A + 50 ng/mL KGF + 0.1 μM Retinoic Acid + 0.25 μM SANT1 + 10 μM Y27632
Stage 5 (7 days)	S5 media + 10 μM ALK5i II + 20 ng/mL Betacellulin + 0.1 μM Retinoic Acid + 0.25 μM SANT1 + 1 μM T3 + 1 μM XXI
Stage 6 (7-35 days)	ESFM differentiation medium

Table 4.9. The medium formulation in Millman Protocol

Medium	Formulations
S1 media	500mL MCDB 131 supplemented with 0.22 g glucose, 1.23 g sodium bicarbonate, 10 g BSA, 10 μL ITS-X, 5 mL GlutaMAX, 22 mg vitamin C, and 5 mL P/S solution

S2 media	500mL MCDB 131 supplemented with 0.22 g glucose, 0.615 g sodium bicarbonate, 10 g BSA, 10 µL ITS-X, 5 mL GlutaMAX, 22 mg vitamin C, and 5 mL P/S.
S3 media	500mL MCDB 131 supplemented with 0.22 g glucose, 0.615 g sodium bicarbonate, 10 g BSA, 2.5 mL ITS-X, 5 mL GlutaMAX, 22 mg vitamin C, and 5 mL P/S.
S5 media	500mL MCDB 131 supplemented with 1.8 g glucose, 0.877 g sodium bicarbonate, 10 g BSA, 2.5 mL ITS-X, 5 mL GlutaMAX, 22 mg vitamin C, 5 mL P/S, and 5 mg heparin
ESFM	500mL MCDB 131 supplemented with 0.23 g glucose, 10.5 g BSA, 5.2 mL GlutaMAX, 5.2 mL P/S, 5 mg heparin, 5.2 mL MEM nonessential amino acids, 84 µg ZnSO4, 523 µL Trace Elements A, and 523 µL Trace Elements B

4.12 **Dynamic Glucose-Stimulated Insulin Secretion (dGSIS)**

70 stem cell-derived islet-like aggregates (100000-150000 cells) were initially resuspended and incubated in KRBH buffer containing 2.8 mM glucose for 30 min. Then, they were loaded on a nylon filter in a plastic perfusion chamber containing Bio-Gel P-4 acrylamide microbeads (solutions, and aggregates were maintained in a water bath at 37 °C). The aggregates were then sequentially perfused with 2.8 mM glucose (low glucose) for 12 min, followed by 20 mM glucose (high glucose) for 24 min, again low glucose (2.8 mM) for 12 min and finally with 25 mM KCl for 12 min at a constant flow rate of 100 µl per 180 s using the BioRep perfusion system (Model No. PERI-4.2) maintained at 37 °C in a temperature-controlled chamber. At the same time, flow-through fractions were collected on a 96-well plate which was maintained at 4 °C. According to the manufacturer's instruction, the plate was quantified for insulin content using Human Insulin ELISA (Mercodia, catalog no. 10-1113-01).

4.13 Insulin Content

The islet-like aggregates from the S6 stage of Millman protocol differentiation were washed twice with PBS and dissociated using Accutase reagent. Single cells were counted, and one thousand cells were considered to calculate insulin content. The cells were resuspended in

Acid-EtOH solution (1.5% HCl and 70% EtOH) and stored on a shaker at 4 °C overnight. The solution was centrifuged at 4000 rpm for 15 min. The supernatant was transferred into a new tube and an equal volume of neutralization buffer, 1 M Tris (pH 7.5), was added. Next, human insulin was calculated using the Mercodia Human Insulin ELISA kit according to the manufacturer's instruction.

4.14 Cell proliferation analysis using Edu staining

At three time points, day 12, day 15 and day 20, iPSC-derived aggregates were treated with 10 μM EdU for 8 h and so that EdU could bind to the DNA during cell proliferation. Next, the aggregates were washed with PBS and treated with Accutase to produce single cells. Following centrifuge, the cells were washed three times with PBS containing 3% BSA and fixed with 4% formaldehyde for 20 min, then washed again with 3% BSA in PBS and permeabilized with 0.5% Triton X-100 in PBS for 20 min. Finally, the cells were stained with Click-iT reaction cocktail prepared according to the manufacturer's instructions (Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor™ 488 dye, C10337, Invitrogen™, Thermo Fisher Scientific) for 30 min protected from light. The nuclei were stained with 5 µg/ml Hoechst 33342 solution. Quantification of the EdU positive cells was performed by the FACS machine.

4.15 Cryopreservation

The iPSCs clones were thawed fast in a 37°C warm water bath and transferred into a culture dish containing rock inhibitor and culture medium. The next day, the medium was renewed, and the cells were passaged for at least three times prior to an experiment. To cryopreserve iPSCs, the cells were treated with Accutane as described above and re-suspended in a freezing medium (2 ml) containing medium, DMSO and FBS with the ratio of 5:1:4. After transferring the cells into cryovials, the cells were stored in freezing boxes at -80°C for three days and then transferred into liquid N_2 for long-term storage.

4.16 Cryosections

The islet-like aggregates were collected from the plates, washed twice with PBS, and fixed in 2% paraformaldehyde (PFA) for 20 min at RT. Next, the aggregates were cryoprotected in a sequential gradient of 15% and 30% sucrose in PBS for 2 h. The samples were inoculated overnight in 30% sucrose in PBS and tissue embedding medium (Leica) with the ratio of 1:1. Then, the aggregates were placed in a 100% tissue embedding medium, in an embedding mold, frozen using dry ice and stored at -80°C. In order to prepare cryosections, the embedded and frozen aggregates were cut in 20 µm sections using a cryostat (Leica), mounted on a glass slide, and dried for 30 min at RT before use or storage at -20°C.

4.17 Immunofluorescence imaging

The sections from cryosections step were permeabilized with 0.1 M Glycine and 0.2% Triton in PBS for 30 min. Next, they were blocked with the blocking solution (3% serum donkey, 0.1% BSA and Tween20). The primary antibodies were added and incubated overnight at 4 °C in blocking buffer. The next day, the cells were washed three times with PBS containing 0.1% Tween-20 (PBS-T). Then, the secondary antibodies were diluted in blocking buffer, added, and incubated for 4 h at RT (the primary and secondary antibodies and their dilution are listed in Table 4.10 and Table 4.11). Following washing three times with PBS-T, nuclei were stained with DAPI diluted in PBS for 30 min. The slides were washed 3 times in PBS-T and mounted with Evanol on glass slides. Images were taken by *Zeiss* confocal microscope. 10 aggregates were analyzed per conditions in z-stacks of 10 µm distance. Finally, the images were analyzed using Fiji software (Fiji).

Table 4.10. Primary antibody

ID	Protein Name	Generated in	Dilution	Company
817	CDKN2B / p15 INK4b	rabbit	IF 1:300	Life sicence
814	CDKN2A/p16INK4a	rabbit	IF 1:300	Abcam
815	CDKN2A/p14ARF	rabbit	IF 1:300	Abcam
48	Glucagon	guinea pig	IF 1:500	Millipore
82	Ki67	rabbit	IF 1:300	Novocastra
121	Insulin	guinea pig	IF 1:300	Thermo Fisher Scientific
123	Pdx1	rabbit	IF 1:300	NEB

192	Nkx6.1	goat	IF 1:200	R&D systems
197	Nkx6.1	rabbit	IF 1:300	Acris/Novus
199	Ki67	rabbit	IF 1:300	Abcam
215	Insulin	rabbit	IF 1:300	Thermo Fisher Scientific
216	Glucagon	guinea pig	IF 1:500	TAKARA
221	SOX2	goat	IF 1:500	Santa Cruz
227	OCT4	goat	IF 1:500	Santa Cruz
277	FOXA2	rabbit	IF 1:500	Cell Signaling
302	SOX17	goat	IF 1:500	Neuromics
315	Nestin	mouse	IF 1:500	Abcam

Table 4.11. Secondary antibody

ID	Name	Conjugated	Dilution	Company
11	Alexa Fluor phalloidin	546	IC 1:40	Invitrogen
18	Donkey anti-goat IgG	633	IC 1:500	Invitrogen
23	Donkey anti-mouse IgG	488	IC 1:500	Invitrogen
24	Donkey anti-rabbit IgG	555	IC 1:500	Invitrogen
28	Donkey anti-chicken IgY	488	IC 1:500	Dianova
45	donkey anti-rat IgG	649	IC 1:500	Dianova
46	donkey anti-guineapig	649	IC 1:500	Dianova
56	Donkey anti-mouse IgG	594	IC 1:500	Invitrogen

62	Donkey anti-rat IgG	647	IC 1:500	Dianova
63	Donkey anti-goat IgG	594	IC 1:500	Invitrogen
64	Donkey anti-rabbit IgG	594	IC 1:500	Invitrogen

4.18 RNA isolation/cDNA synthesis/quantitative real-time PCR

RNA isolation from the iPSCs was carried out using Trizol Reagent (Invitrogen, Carlsbad, CA). The RNA was eluted in 50 µl of nuclease-free water. Then it was stored at -80°C. The DNA or RNA concentration was measured by a NanoDrop. The purity of the DNA and RNA was assessed by the quotient of E_{260nm}/E_{280nm} and E_{260nm}/E_{230nm} which had to be between 1.8 and 2.0. cDNA was synthesized using the SuperScript VILO cDNA synthesis kit (Invitrogen, Carlsbad, CA). To this end, the RNA solution (100-500 ng RNA), 5x VILO™ reaction mix, and 10x SuperScript™ enzyme were incubated at 25°C for 10 min before 120 min at 85°C. Afterward, the cDNA was stored at -20°C or -80°C.

For the SOX2 gene targeting project, qPCR was carried out using SsoAdvanced Universal SYBR Green Supermix. qPCR reactions were carried out for XRCC4, DNAPK, and GAPDH genes (GenBank accession numbers: NM_003401.5, NM_006904.7, and NM_002046.3, respectively). For the INK4 project, the qPCR was performed using TaqMan™ probes (Table 4.12) and 100 ng cDNA per reaction. Each reaction has a total of 20 µl, consisting 2 µl cDNA in nuclease-free water, 10 µl TaqMan™ Advanced master mix (Life Technologies), and 1 µl TaqMan probe™ (Life Technologies) and 7 µl nuclease-free water. After sealing the 96 well plates (Life Technologies) and its centrifugation for 1 min at 1000 rpm, the qPCR was performed using Viia7 (Thermo Fisher Scientific) using the following thermal conditions: initiation at 95°C for 30 sec, amplification for 40 cycles with denaturation at 95°C for 10 sec, annealing/extending at 60°C for 1 min.

Table 4.12. Taqman primers and probes

Gene	**Order Information**	**Gene**	**Order Information**

ANRIL	Hs03300540_m1	*P14*	Hs99999189_m1
P15	Hs00793225_m1	*MTAP*	Hs00559618_m1
P16	Hs02902543_Mh	*DMRTA1*	Hs00403012_m1
GAPDH	Hs04420632_g1	*P16gama*	Hs07290632_m1
P12	Hs04189686_m1		

TaqMan primers and Probes were purchased from Life Technologies

4.19 Statistical Analysis

To analyze qPCR data, the Ct-values, a point of the linear slope of fluorescence, were normalized among samples, transformed to linear expression values, and normalized on reference genes and the control samples as are shown by the following formula:

Relative expression (gene) = $(2^{\text{Ct (mean genes)} - \text{Ct (gene)}}) / (2^{\text{Ct (mean references)} - \text{Ct (reference)}})$

Normalized expression (gene) = Relative expression (gene) / Relative expression $_{\text{control}}$ (gene)

Significance was determined using a two-tailed unpaired and Welch corrected t-test. The expression of each gene transcript was calculated by normalizing the respective housekeeping gene. The expression of target genes was normalized to the expression value of the housekeeping gene of *GAPDH*. The P-values were calculated using a two-tailed Student’s t-test.

We applied GraphPad Prism software (GraphPad, San Diego, CA, USA) to carry out statistical analyses. All experimental tests were repeated at least two or three times. Data represented are the mean ± S.D. using two-tailed Student's t-test or one-way analysis of variance (ANOVA). The criterion for statistical significance were * indicated *P*-values smaller than 0.05, ** > 0.01, *** > 0.001 and **** > 0.0001.

References

AHLGREN, U., JONSSON, J. & EDLUND, H. 1996. The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development,* 122**,** 1409-1416.

AHNFELT-RONNE, J., JORGENSEN, M. C., KLINCK, R., JENSEN, J. N., FUCHTBAUER, E. M., DEERING, T., MACDONALD, R. J., WRIGHT, C. V., MADSEN, O. D. & SERUP, P. 2012. Ptf1a-mediated control of Dll1 reveals an alternative to the lateral inhibition mechanism. *Development,* 139**,** 33-45.

AL-KAABI, A., VAN BOCKEL, L. W., POTHEN, A. J. & WILLEMS, S. M. 2014. p16INK4A and p14ARF gene promoter hypermethylation as prognostic biomarker in oral and oropharyngeal squamous cell carcinoma: a review. *Dis Markers,* 2014**,** 260549.

ALI, O. 2013. Genetics of type 2 diabetes. *World J Diabetes,* 4**,** 114-23.

ANDERS, C., NIEWOEHNER, O., DUERST, A. & JINEK, M. 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature,* 513**,** 569-+.

ANNICOTTE, J. S., BLANCHET, E., CHAVEY, C., IANKOVA, I., COSTES, S., ASSOU, S., TEYSSIER, J., DALLE, S., SARDET, C. & FAJAS, L. 2009. The CDK4-pRB-E2F1 pathway controls insulin secretion. *Nature Cell Biology,* 11**,** 1017-U247.

ARDEHALI, R., INLAY, M. A., ALI, S. R., TANG, C., DRUKKER, M. & WEISSMAN, I. L. 2011. Overexpression of BCL2 enhances survival of human embryonic stem cells during stress and obviates the requirement for serum factors. *Proc Natl Acad Sci U S A,* 108**,** 3282-7.

ATTALI, M., STETSYUK, V., BASMACIOGULLARI, A., AIELLO, V., ZANTA-BOUSSIF, M. A., DUVILLIE, B. & SCHARFMANN, R. 2007. Control of beta-cell differentiation by the pancreatic mesenchyme. *Diabetes,* 56**,** 1248-58.

AUGSORNWORAWAT, P., MAXWELL, K. G., VELAZCO-CRUZ, L. & MILLMAN, J. R. 2021. Single-cell transcriptome profiling reveals beta cell maturation in stem cell-derived islets after transplantation. *Cell Rep,* 34**,** 108850.

AVRAHAMI, D., LI, C. H., ZHANG, J., SCHUG, J., AVRAHAMI, R., RAO, S., STADLER, M. B., BURGER, L., SCHUBELER, D., GLASER, B. & KAESTNER, K. H. 2015. Aging-Dependent Demethylation of Regulatory Elements Correlates with Chromatin State and Improved beta Cell Function. *Cell Metabolism,* 22**,** 619-632.

BAE, S., KWEON, J., KIM, H. S. & KIM, J. S. 2014. Microhomology-based choice of Cas9 nuclease target sites. *Nature Methods,* 11**,** 705-706.

BAKHTI, M., BOTTCHER, A. & LICKERT, H. 2019. Modelling the endocrine pancreas in health and disease. *Nature Reviews Endocrinology,* 15**,** 155-171.

BALBOA, D., WELTNER, J., NOVIK, Y., EUROLA, S., WARTIOVAARA, K. & OTONKOSKI, T. 2017. Generation of a SOX2 reporter human induced pluripotent stem cell line using CRISPR/SaCas9. *Stem Cell Research,* 22**,** 16-19.

BANKAITIS, E. D., BECHARD, M. E. & WRIGHT, C. V. E. 2015. Feedback control of growth, differentiation, and morphogenesis of pancreatic endocrine progenitors in an epithelial plexus niche. *Genes & Development,* 29**,** 2203-2216.

BARRANGOU, R. 2013. CRISPR-Cas systems and RNA-guided interference. *Wiley Interdisciplinary Reviews-Rna,* 4**,** 267-278.

BARRANGOU, R. 2015. The roles of CRISPR-Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology,* 32**,** 36-41.

BASTIDAS-PONCE, A., SCHEIBNER, K., LICKERT, H. & BAKHTI, M. 2017. Cellular and molecular mechanisms coordinating pancreas development. *Development,* 144**,** 2873-2888.

BAUMGARTNER, B. K., CASH, G., HANSEN, H., OSTLER, S. & MURTAUGH, L. C. 2014. Distinct requirements for beta-catenin in pancreatic epithelial growth and patterning. *Dev Biol,* 391**,** 89-98.

BERES, T. M., MASUI, T., SWIFT, G. H., SHI, L., HENKE, R. M. & MACDONALD, R. J. 2006. PTF1 is an organ-specific and notch-independent basic helix-loop-helix complex containing the mammalian suppressor of hairless (RBP-J) or its paralogue, RBP-L. *Molecular and Cellular Biology,* 26**,** 117-130.

BHUSHAN, A., ITOH, N., KATO, S., THIERY, J. P., CZERNICHOW, P., BELLUSCI, S. & SCHARFMANN, R. 2001. Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. *Development,* 128**,** 5109-5117.

BRANZEI, D. & FOIANI, M. 2008. Regulation of DNA repair throughout the cell cycle. *Nature Reviews Molecular Cell Biology,* 9**,** 297-308.

BUNTING, S. F., CALLEN, E., WONG, N., CHEN, H. T., POLATO, F., GUNN, A., BOTHMER, A., FELDHAHN, N., FERNANDEZ-CAPETILLO, O., CAO, L., XU, X. L., DENG, C. X., FINKEL, T., NUSSENZWEIG, M., STARK, J. M. & NUSSENZWEIG, A. 2010. 53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks. *Cell,* 141**,** 243-254.

BURLISON, J. S., LONG, Q., FUJITANI, Y., WRIGHT, C. V. & MAGNUSON, M. A. 2008. Pdx-1 and Ptf1a concurrently determine fate specification of pancreatic multipotent progenitor cells. *Dev Biol,* 316**,** 74-86.

BUSCAGLIA, L. E. & LI, Y. 2011. Apoptosis and the target genes of microRNA-21. *Chin J Cancer,* 30**,** 371-80.

CANNAVO, E. & CEJKA, P. 2014. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature,* 514**,** 122-+.

CAPITO, C., SIMON, M. T., AIELLO, V., CLARK, A., AIGRAIN, Y., RAVASSARD, P. & SCHARFMANN, R. 2013. Mouse Muscle As an Ectopic Permissive Site for Human Pancreatic Development. *Diabetes,* 62**,** 3479-3487.

CATALDI, S., COSTA, V., CICCODICOLA, A. & APRILE, M. 2021. PPARgamma and Diabetes: Beyond the Genome and Towards Personalized Medicine. *Curr Diab Rep,* 21**,** 18.

CHEN, H., GU, X., SU, I. H., BOTTINO, R., CONTRERAS, J. L., TARAKHOVSKY, A. & KIM, S. K. 2009. Polycomb protein Ezh2 regulates pancreatic beta-cell Ink4a/Arf expression and regeneration in diabetes mellitus. *Genes Dev,* 23**,** 975-85.

CHENG, X. A., SHI, L. S., NIE, S. F., WANG, F., LI, X. C., XU, C. Q., WANG, P. Y., YANG, B. F., LI, Q. X., PAN, Z. W., LI, Y., XIA, H., ZHENG, C. H., KE, Y. H., WU, Y. X., TANG, T. T., YAN, X. X., YANG, Y., XIA, N., YAO, R., WANG, B. B., MA, X., ZENG, Q. T., TU, X., LIAO, Y. H. & WANG, Q. K. 2011. The Same Chromosome 9p21.3 Locus Is Associated With Type 2 Diabetes and Coronary Artery Disease in a Chinese Han Population. *Diabetes,* 60**,** 680-684.

CHU, V. T., WEBER, T., WEFERS, B., WURST, W., SANDER, S., RAJEWSKY, K. & KUHN, R. 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology,* 33**,** 543-U160.

CISSELL, M. A., ZHAO, L., SUSSEL, L., HENDERSON, E. & STEIN, R. 2003. Transcription factor occupancy of the insulin gene in vivo - Evidence for direct regulation by Nkx2.2. *Journal of Biological Chemistry,* 278**,** 751-756.

CLEAVER, O. & DOR, Y. 2012a. Vascular instruction of pancreas development. *Development,* 139**,** 2833-2843.

CLEAVER, O. & DOR, Y. 2012b. Vascular instruction of pancreas development. *Development,* 139**,** 2833-43.

COLLOMBAT, P., HECKSHER-SORENSEN, J., BROCCOLI, V., KRULL, J., PONTE, I., MUNDIGER, T., SMITH, J., GRUSS, P., SERUP, P. & MANSOURI, A. 2005. The simultaneous loss of Arx and Pax4 genes promotes a somatostatin-producing cell fate specification at the expense of the alpha- and beta-cell lineages in the mouse endocrine pancreas. *Development,* 132**,** 2969-2980.

COLLOMBAT, P., MANSOURI, A., HECKSHER-SORENSEN, J., SERUP, P., KRULL, J., GRADWOHL, G. & GRUSS, P. 2003. Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes & Development,* 17**,** 2591-2603.

CONG, L., RAN, F. A., COX, D., LIN, S. L., BARRETTO, R., HABIB, N., HSU, P. D., WU, X. B., JIANG, W. Y., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science,* 339**,** 819-823.

CONTI, A. & DI MICCO, R. 2018. p53 activation: a checkpoint for precision genome editing? *Genome Med,* 10**,** 66.

COSENTINO, C., TOIVONEN, S., DIAZ VILLAMIL, E., ATTA, M., RAVANAT, J. L., DEMINE, S., SCHIAVO, A. A., PACHERA, N., DEGLASSE, J. P., JONAS, J. C., BALBOA, D., OTONKOSKI, T., PEARSON, E. R., MARCHETTI, P., EIZIRIK, D. L., CNOP, M. & IGOILLO-ESTEVE, M. 2018. Pancreatic beta-cell tRNA hypomethylation and fragmentation link TRMT10A deficiency with diabetes. *Nucleic Acids Res,* 46**,** 10302-10318.

CUGINO, D., GIANFAGNA, F., SANTIMONE, I., DE GAETANO, G., DONATI, M. B., IACOVIELLO, L. & DI CASTELNUOVO, A. 2012. Type 2 diabetes and polymorphisms on chromosome 9p21: a meta-analysis. *Nutr Metab Cardiovasc Dis,* 22**,** 619-25.

CUNNINGTON, M. S., KOREF, M. S., MAYOSI, B. M., BURN, J. & KEAVNEY, B. 2010a. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. *Plos Genetics,* 6.

CUNNINGTON, M. S., SANTIBANEZ KOREF, M., MAYOSI, B. M., BURN, J. & KEAVNEY, B. 2010b. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. *PLoS Genet,* 6**,** e1000899.

D'AMOUR, K. A., AGULNICK, A. D., ELIAZER, S., KELLY, O. G., KROON, E. & BAETGE, E. E. 2005. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol,* 23**,** 1534-41.

DALEY, J. M., VANDER LAAN, R. L., SURESH, A. & WILSON, T. E. 2005. DNA joint dependence of Pol X family polymerase action in nonhomologous end joining. *Journal of Biological Chemistry,* 280**,** 29030-29037.

DANIEL, M. & TOLLEFSBOL, T. O. 2015. Epigenetic linkage of aging, cancer and nutrition. *J Exp Biol,* 218**,** 59-70.

DAVIS, A. J. & CHEN, D. J. 2013. DNA double strand break repair via non-homologous end-joining. *Translational Cancer Research,* 2**,** 130-143.

DAYEH, T. A., OLSSON, A. H., VOLKOV, P., ALMGREN, P., RONN, T. & LING, C. 2013. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia,* 56**,** 1036-46.

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**,** 871-882.

DELOUS, M., YIN, C. Y., SHIN, D. H., NINOV, N., CARTEN, J. D., PAN, L. Y., MA, T. P., FARBER, S. A., MOENS, C. B. & STAINIER, D. Y. R. 2012. sox9b Is a Key Regulator of Pancreaticobiliary Ductal System Development. *Plos Genetics,* 8.

DOW, L. E., FISHER, J., O'ROURKE, K. P., MULEY, A., KASTENHUBER, E. R., LIVSHITS, G., TSCHAHARGANEH, D. F., SOCCI, N. D. & LOWE, S. W. 2015. Inducible in vivo genome editing with CRISPR-Cas9. *Nature Biotechnology,* 33**,** 390-U98.

DRAK ALSIBAI, K., VACHER, S., MESEURE, D., NICOLAS, A., LAE, M., SCHNITZLER, A., CHEMLALI, W., CROS, J., LONGCHAMPT, E., CACHEUX, W., PIGNOT, G., CALLENS, C., PASMANT, E., ALLORY, Y. & BIECHE, I. 2019. High Positive Correlations between ANRIL and p16-CDKN2A/p15-CDKN2B/p14-ARF Gene Cluster Overexpression in Multi-Tumor Types Suggest Deregulated Activation of an ANRIL-ARF Bidirectional Promoter. *Noncoding RNA,* 5.

DUGGIRALA, R., BLANGERO, J., ALMASY, L., DYER, T. D., WILLIAMS, K. L., LEACH, R. J., O'CONNELL, P. & STERN, M. P. 1999. Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. *American Journal of Human Genetics,* 64**,** 1127-1140.

DUPRE, A., BOYER-CHATENET, L. & GAUTIER, J. 2006. Two-step activation of ATM by DNA and the Mre11-Rad50-Nbs1 complex. *Nat Struct Mol Biol,* 13**,** 451-7.

EVERS, B. M., RADY, P. L., SANDOVAL, K., ARANY, I., TYRING, S. K., SANCHEZ, R. L., NEALON, W. H., TOWNSEND, C. M. & THOMPSON, J. C. 1994. Gastrinomas Demonstrate Amplification of the Her-2/Neu Protooncogene. *Annals of Surgery,* 219**,** 596-604.

FADISTA, J., VIKMAN, P., LAAKSO, E. O., MOLLET, I. G., ESGUERRA, J. L., TANEERA, J., STORM, P., OSMARK, P., LADENVALL, C., PRASAD, R. B., HANSSON, K. B., FINOTELLO, F., UVEBRANT, K., OFORI, J. K., DI CAMILLO, B., KRUS, U., CILIO, C. M., HANSSON, O., ELIASSON, L., ROSENGREN, A. H., RENSTROM, E., WOLLHEIM, C. B. & GROOP, L. 2014. Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proc Natl Acad Sci U S A,* 111**,** 13924-9.

FAJAS, L., BLANCHET, E. & ANNICOTTE, J. S. 2010. CDK4, pRB and E2F1: connected to insulin. *Cell Division,* 5.

FOLKERSEN, L., KYRIAKOU, T., GOEL, A., PEDEN, J., MALARSTIG, A., PAULSSON-BERNE, G., HAMSTEN, A., FRANCO-CERECEDA, A., GABRIELSEN, A., ERIKSSON, P., WATKINS, H. & CONSORTIA, H. W. P. 2009. Relationship between CAD Risk Genotype in the Chromosome 9p21 Locus and Gene Expression. Identification of Eight New ANRIL Splice Variants. *Plos One,* 4.

FOWDEN, A. L. & HILL, D. J. 2001. Intra-uterine programming of the endocrine pancreas. *British Medical Bulletin,* 60**,** 123-142.

FRIDLYAND, L. E., TAMARINA, N. & PHILIPSON, L. H. 2003. Modeling of Ca2+ flux in pancreatic beta-cells: role of the plasma membrane and intracellular stores. *Am J Physiol Endocrinol Metab,* 285**,** E138-54.

FUCHSBERGER, C., FLANNICK, J., TESLOVICH, T. M., MAHAJAN, A., AGARWALA, V., GAULTON, K. J., MA, C., FONTANILLAS, P., MOUTSIANAS, L., MCCARTHY, D. J., RIVAS, M. A., PERRY, J. R. B., SIM, X., BLACKWELL, T. W., ROBERTSON, N. R., RAYNER, N. W., CINGOLANI, P., LOCKE, A. E., TAJES, J. F., HIGHLAND, H. M., DUPUIS, J., CHINES, P. S., LINDGREN, C. M., HARTL, C., JACKSON, A. U., CHEN, H., HUYGHE, J. R., VAN DE BUNT, M., PEARSON, R. D., KUMAR, A., MULLER-NURASYID, M., GRARUP, N., STRINGHAM, H. M., GAMAZON, E. R., LEE, J., CHEN, Y. H., SCOTT, R. A., BELOW, J. E., CHEN, P., HUANG, J., GO, M. J., STITZEL, M. L., PASKO, D., PARKER, S. C. J., VARGA, T. V., GREEN, T., BEER, N. L., DAY-WILLIAMS, A. G., FERREIRA, T., FINGERLIN, T., HORIKOSHI, M., HU, C., HUH, I., IKRAM, M. K., KIM, B. J., KIM, Y., KIM, Y. J., KWON, M. S., LEE, J., LEE, S., LIN, K. H., MAXWELL, T. J., NAGAI, Y., WANG, X., WELCH, R. P., YOON, J., ZHANG, W., BARZILAI, N., VOIGHT, B. F., HAN, B. G., JENKINSON, C. P., KUULASMAA, T., KUUSISTO, J., MANNING, A., NG, M. C. Y., PALMER, N. D., BALKAU, B., STANCAKOVA, A., ABBOUD, H. E., BOEING, H., GIEDRAITIS, V., PRABHAKARAN, D.,

GOTTESMAN, O., SCOTT, J., CAREY, J., KWAN, P., GRANT, G., SMITH, J. D., NEALE, B. M., PURCELL, S., BUTTERWORTH, A. S., HOWSON, J. M. M., LEE, H. M., LU, Y. C., KWAK, S. H., ZHAO, W., DANESH, J., LAM, V. K. L., PARK, K. S., SALEHEEN, D., et al. 2016. The genetic architecture of type 2 diabetes. *Nature,* 536**,** 41-+.

GAO, N., LELAY, J., VATAMANIUK, M. Z., RIECK, S., FRIEDMAN, J. R. & KAESTNER, K. H. 2008. Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes & Development,* 22**,** 3435-3448.

GAULTON, K. J. 2017. Mechanisms of Type 2 Diabetes Risk Loci. *Current Diabetes Reports,* 17.

GAULTON, K. J., WILLER, C. J., LI, Y., SCOTT, L. J., CONNEELY, K. N., JACKSON, A. U., DUREN, W. L., CHINES, P. S., NARISU, N., BONNYCASTLE, L. L., LUO, J. C., TONG, M., SPRAU, A. G., PUGH, E. W., DOHENY, K. F., VALLE, T. T., ABECASIS, G. R., TUOMILEHTO, J., BERGMAN, R. N., COLLINS, F. S., BOEHNKE, M. & MOHLKEL, K. L. 2008. Comprehensive Association Study of Type 2 Diabetes and Related Quantitative Traits With 222 Candidate Genes. *Diabetes,* 57**,** 3136-3144.

GHASEMI, A. & NOROUZIRAD, R. 2019. Type 2 Diabetes: An Updated Overview. *Crit Rev Oncog,* 24**,** 213-222.

GIL, J. & PETERS, G. 2006. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. *Nature Reviews Molecular Cell Biology,* 7**,** 667-677.

GITTES, G. K. 2009. Developmental biology of the pancreas: a comprehensive review. *Dev Biol,* 326**,** 4-35.

GONZALEZ, S., KLATT, P., DELGADO, S., CONDE, E., LOPEZ-RIOS, F., SANCHEZ-CESPEDES, M., MENDEZ, J., ANTEQUERA, F. & SERRANO, M. 2017. Oncogenic activity of Cdc6 through repression of the INK4/ARF locus (Retraction of Vol 440, Pg 702, 2006). *Nature,* 547**,** 246-246.

GRADWOHL, G., DIERICH, A., LEMEUR, M. & GUILLEMOT, F. 2000. neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 1607-1611.

GRANT, S. F. A., THORLEIFSSON, G., REYNISDOTTIR, I., BENEDIKTSSON, R., MANOLESCU, A., SAINZ, J., HELGASON, A., STEFANSSON, H., EMILSSON, V., HELGADOTTIR, A., STYRKARSDOTTIR, U., MAGNUSSON, K. P., WALTERS, G. B., PALSDOTTIR, E., JONSDOTTIR, T., GUDMUNDSDOTTIR, T., GYLFASON, A., SAEMUNDSDOTTIR, J., WILENSKY, R. L., REILLY, M. P., RADER, D. J., BAGGER, Y., CHRISTIANSEN, C., GUDNASON, V., SIGURDSSON, G., THORSTEINSDOTTIR, U., GULCHER, J. R., KONG, A. & STEFANSSON, K. 2006. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nature Genetics,* 38**,** 320-323.

GRARUP, N., ROSE, C. S., ANDERSSON, E. A., ANDERSEN, G., NIELSEN, A. L., ALBRECHTSEN, A., CLAUSEN, J. O., RASMUSSEN, S. S., JORGENSEN, T., SANDBAEK, A., LAURITZEN, T., SCHMITZ, O., HANSEN, T. & PEDERSEN, O. 2007. Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects - Validation and extension of genome-wide association studies. *Diabetes,* 56**,** 3105-3111.

GU, G. Q., DUBAUSKAITE, J. & 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.

GUSCHIN, D. Y., WAITE, A. J., KATIBAH, G. E., MILLER, J. C., HOLMES, M. C. & REBAR, E. J. 2010. A rapid and general assay for monitoring endogenous gene modification. *Methods Mol Biol,* 649**,** 247-56.

GUZ, Y., MONTMINY, M. R., STEIN, R., LEONARD, J., GAMER, L. W., WRIGHT, C. V. E. & 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.
HALD, J., SPRINKEL, A. E., RAY, M., SERUP, P., WRIGHT, C. & MADSEN, O. D. 2008. Generation and characterization of Ptf1a antiserum and localization of Ptf1a in relation to Nkx6.1 and Pdx1 during the earliest stages of mouse pancreas development. *J Histochem Cytochem,* 56**,** 587-95.
HANI, E. H., BOUTIN, P., DURAND, E., INOUE, H., PERMUTT, M. A., VELHO, G. & FROGUEL, P. 1998. Missense mutations in the pancreatic islet beta cell inwardly rectifying K+ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of Type II diabetes mellitus in Caucasians. *Diabetologia,* 41**,** 1511-5.
HANIS, C. L., BOERWINKLE, E., CHAKRABORTY, R., ELLSWORTH, D. L., CONCANNON, P., STIRLING, B., MORRISON, V. A., WAPELHORST, B., SPIELMAN, R. S., GOGOLIN-EWENS, K. J., SHEPARD, J. M., WILLIAMS, S. R., RISCH, N., HINDS, D., IWASAKI, N., OGATA, M., OMORI, Y., PETZOLD, C., RIETZCH, H., SCHRODER, H. E., SCHULZE, J., COX, N. J., MENZEL, S., BORIRAJ, V. V., CHEN, X., LIM, L. R., LINDNER, T., MEREU, L. E., WANG, Y. Q., XIANG, K., YAMAGATA, K., YANG, Y. & BELL, G. I. 1996. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet,* 13**,** 161-6.
HANLEY, K. P., HEARN, T., BERRY, A., CARVELL, M. J., PATCH, A. M., WILLIAMS, L. J., SUGDEN, S. A., WILSON, D. I., ELLARD, S. & HANLEY, N. A. 2010. In vitro expression of NGN3 identifies RAB3B as the predominant Ras-associated GTP-binding protein 3 family member in human islets. *Journal of Endocrinology,* 207**,** 151-161.
HARA, K., FUJITA, H., JOHNSON, T. A., YAMAUCHI, T., YASUDA, K., HORIKOSHI, M., PENG, C., HU, C., MA, R. C., IMAMURA, M., IWATA, M., TSUNODA, T., MORIZONO, T., SHOJIMA, N., SO, W. Y., LEUNG, T. F., KWAN, P., ZHANG, R., WANG, J., YU, W., MAEGAWA, H., HIROSE, H., CONSORTIUM, D., KAKU, K., ITO, C., WATADA, H., TANAKA, Y., TOBE, K., KASHIWAGI, A., KAWAMORI, R., JIA, W., CHAN, J. C., TEO, Y. Y., SHYONG, T. E., KAMATANI, N., KUBO, M., MAEDA, S. & KADOWAKI, T. 2014a. Genome-wide association study identifies three novel loci for type 2 diabetes. *Hum Mol Genet,* 23**,** 239-46.
HARA, K., FUJITA, H., JOHNSON, T. A., YAMAUCHI, T., YASUDA, K., HORIKOSHI, M., PENG, C., HU, C., MA, R. C. W., IMAMURA, M., IWATA, M., TSUNODA, T., MORIZONO, T., SHOJIMA, N., SO, W. Y., LEUNG, T. F., KWAN, P., ZHANG, R., WANG, J., YU, W. H., MAEGAWA, H., HIROSE, H., KAKU, K., ITO, C., WATADA, H., TANAKA, Y., TOBE, K., KASHIWAGI, A., KAWAMORI, R., JIA, W. P., CHAN, J. C. N., TEO, Y. Y., SHYONG, T. E., KAMATANI, N., KUBO, M., MAEDA, S., KADOWAKI, T. & CONSORTIUM, D. 2014b. Genome-wide association study identifies three novel loci for type 2 diabetes. *Human Molecular Genetics,* 23**,** 239-246.
HARISMENDY, O., NOTANI, D., SONG, X., RAHIM, N. G., TANASA, B., HEINTZMAN, N., REN, B., FU, X. D., TOPOL, E. J., ROSENFELD, M. G. & FRAZER, K. A. 2011. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature,* 470**,** 264-8.
HER, J. & BUNTING, S. F. 2018. How cells ensure correct repair of DNA double-strand breaks. *Journal of Biological Chemistry,* 293**,** 10502-10511.
HEUVEL-BORSBOOM, H., DE VALK, H. W., LOSEKOOT, M. & WESTERINK, J. 2016. Maturity onset diabetes of the young: Seek and you will find. *Neth J Med,* 74**,** 193-200.

HOGREBE, N. J., AUGSORNWORAWAT, P., MAXWELL, K. G., VELAZCO-CRUZ, L. & MILLMAN, J. R. 2020. Targeting the cytoskeleton to direct pancreatic differentiation of human pluripotent stem cells. *Nature Biotechnology,* 38**,** 460-+.

HOHL, M., KWON, Y., GALVAN, S. M., XUE, X. Y., TOUS, C., AGUILERA, A., SUNG, P. & PETRINI, J. H. J. 2011. The Rad50 coiled-coil domain is indispensable for Mre11 complex functions. *Nature Structural & Molecular Biology,* 18**,** 1124-U58.

HOLMSTROM, S. R., DEERING, T., SWIFT, G. H., POELWIJK, F. J., MANGELSDORF, D. J., KLIEWER, S. A. & MACDONALD, R. J. 2011. LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. *Genes & Development,* 25**,** 1674-1679.

HORIE, I., ABIRU, N., ETO, M., SAKO, A., AKESHIMA, J., NAKAO, T., NAKASHIMA, Y., NIRI, T., ITO, A., NOZAKI, A., HARAGUCHI, A., AKAZAWA, S., MORI, Y., ANDO, T. & KAWAKAMI, A. 2018. Sex differences in insulin and glucagon responses for glucose homeostasis in young healthy Japanese adults. *Journal of Diabetes Investigation,* 9**,** 1283-1287.

HRIBAL, M. L., PRESTA, I., PROCOPIO, T., MARINI, M. A., STANCAKOVA, A., KUUSISTO, J., ANDREOZZI, F., HAMMARSTEDT, A., JANSSON, P. A., GRARUP, N., HANSEN, T., WALKER, M., STEFAN, N., FRITSCHE, A., HARING, H. U., PEDERSEN, O., SMITH, U., LAAKSO, M., SESTI, G. & CONSORTIUM, E. 2011. Glucose tolerance, insulin sensitivity and insulin release in European non-diabetic carriers of a polymorphism upstream of CDKN2A and CDKN2B. *Diabetologia,* 54**,** 795-802.

HUANG, J., ELLINGHAUS, D., FRANKE, A., HOWIE, B. & LI, Y. 2012. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. *European Journal of Human Genetics,* 20**,** 801-805.

HUG, N., LONGMAN, D. & CACERES, J. F. 2016. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Research,* 44**,** 1483-1495.

HWANG, J. Y., SIM, X., WU, Y., LIANG, J., TABARA, Y., HU, C., HARA, K., TAM, C. H. T., CAI, Q. Y., ZHAO, Q., JEE, S., TAKEUCHI, F., GO, M. J., ONG, R. T. H., OHKUBO, T., KIM, Y. J., ZHANG, R., YAMAUCHI, T., SO, W. Y., LONG, J. R., GU, D. F., LEE, N. R., KIM, S., KATSUYA, T., OH, J. H., LIU, J. J., UMEMURA, S., KIM, Y. J., JIANG, F., MAEDA, S., CHAN, J. C. N., LU, W., HIXSON, J. E., ADAIR, L. S., JUNG, K. J., NABIKA, T., BAE, J. B., LEE, M. H., SEIELSTAD, M., YOUNG, T. L., TEO, Y. Y., KITA, Y., TAKASHIMA, N., OSAWA, H., LEE, S. H., SHIN, M. H., SHIN, D. H., CHOI, B. Y., SHI, J. J., GAO, Y. T., XIANG, Y. B., ZHENG, W., KATO, N., YOON, M., HE, J., SHU, X. O., MA, R. C. W., KADOWAKI, T., JIA, W. P., MIKI, T., QI, L., TAI, E. S., MOHLKE, K. L., HAN, B. G., CHO, Y. S. & KIM, B. J. 2015. Genome-Wide Association Meta-analysis Identifies Novel Variants Associated With Fasting Plasma Glucose in East Asians. *Diabetes,* 64**,** 291-298.

IHRY, R. J., WORRINGER, K. A., SALICK, M. R., FRIAS, E., HO, D., THERIAULT, K., KOMMINENI, S., CHEN, J., SONDEY, M., YE, C., RANDHAWA, R., KULKARNI, T., YANG, Z., MCALLISTER, G., RUSS, C., REECE-HOYES, J., FORRESTER, W., HOFFMAN, G. R., DOLMETSCH, R. & KAYKAS, A. 2018. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med,* 24**,** 939-946.

JACOBSEN, S. C., BRONS, C., BORK-JENSEN, J., RIBEL-MADSEN, R., YANG, B., LARA, E., HALL, E., CALVANESE, V., NILSSON, E., JORGENSEN, S. W., MANDRUP, S., LING, C., FERNANDEZ, A. F., FRAGA, M. F., POULSEN, P. & VAAG, A. 2012. Effects of short-term high-fat overfeeding on genome-wide DNA methylation in the skeletal muscle of healthy young men. *Diabetologia,* 55**,** 3341-9.

JASIN, M. 1996. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics,* 12**,** 224-228.

JASIN, M. & ROTHSTEIN, R. 2013. Repair of Strand Breaks by Homologous Recombination. *Cold Spring Harbor Perspectives in Biology,* 5.

JENNINGS, R. E., BERRY, A. A., KIRKWOOD-WILSON, R., ROBERTS, N. A., HEARN, T., SALISBURY, R. J., BLAYLOCK, J., PIPER HANLEY, K. & HANLEY, N. A. 2013. Development of the human pancreas from foregut to endocrine commitment. *Diabetes,* 62**,** 3514-22.
JENNINGS, R. E., BERRY, A. A., STRUTT, J. P., GERRARD, D. T. & HANLEY, N. A. 2015. Human pancreas development. *Development,* 142**,** 3126-37.
JEON, J., CORREA-MEDINA, M., RICORDI, C., EDLUND, H. & DIEZ, J. A. 2009. Endocrine Cell Clustering During Human Pancreas Development. *Journal of Histochemistry & Cytochemistry,* 57**,** 811-824.
JIN, T. R. 2016. Current Understanding on Role of the Wnt Signaling Pathway Effector TCF7L2 in Glucose Homeostasis. *Endocrine Reviews,* 37**,** 254-277.
JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science,* 337**,** 816-821.
JONCKHEERE, N., MAYES, E., SHIH, H. P., LI, B., LIOUBINSKI, O., DAI, X. & SANDER, M. 2008. Analysis of mPygo2 mutant mice suggests a requirement for mesenchymal Wnt signaling in pancreatic growth and differentiation. *Developmental Biology,* 318**,** 224-235.
JORGENSEN, M. C., AHNFELT-RONNE, J., HALD, J., MADSEN, O. D., SERUP, P. & HECKSHER-SORENSEN, J. 2007. An illustrated review of early pancreas development in the mouse. *Endocr Rev,* 28**,** 685-705.
KAHN, S. E., COOPER, M. E. & DEL PRATO, S. 2014. Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. *Lancet,* 383**,** 1068-83.
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., FOLEY, J. & JETTEN, A. M. 2009. Transcription factor Glis3, a novel critical player in the regulation of pancreatic beta-cell development and insulin gene expression. *Mol Cell Biol,* 29**,** 6366-79.
KAWAGUCHI, Y., COOPER, B., GANNON, M., RAY, M., MACDONALD, R. J. & WRIGHT, C. V. 2002. The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet,* 32**,** 128-34.
KEENAN, D. M., VELDHUIS, J. D., BASU, A. & BASU, R. 2019. A novel measure of glucose homeostasis (or loss thereof) comprising the joint dynamics of glucose, insulin, glucagon, and cortisol. *American Journal of Physiology-Endocrinology and Metabolism,* 316**,** E998-E1011.
KESAVAN, G., SAND, F. W., GREINER, T. U., JOHANSSON, J. K., KOBBERUP, S., WU, X., BRAKEBUSCH, C. & SEMB, H. 2009. Cdc42-mediated tubulogenesis controls cell specification. *Cell,* 139**,** 791-801.
KETTUNEN, J. L. T. & TUOMI, T. 2020. Human Physiology of Genetic Defects Causing Beta-cell Dysfunction. *J Mol Biol,* 432**,** 1579-1598.
KIM, J. H., EOM, H. J., LIM, G., PARK, S., LEE, J., NAM, S., KIM, Y. H. & JEONG, J. H. 2019. Differential effects, on oncogenic pathway signalling, by derivatives of the HNF4 alpha inhibitor BI6015. *British Journal of Cancer,* 120**,** 488-498.
KIM, K. P. & MIRKIN, E. V. 2018. So similar yet so different: The two ends of a double strand break. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis,* 809**,** 70-80.
KIM, S. K. & HEBROK, M. 2001. Intercellular signals regulating pancreas development and function. *Genes & Development,* 15**,** 111-127.
KIM, W. Y. & SHARPLESS, N. E. 2006. The regulation of INK4/ARF in cancer and aging. *Cell,* 127**,** 265-275.

KONG, Y., SHARMA, R. B., LY, S., STAMATERIS, R. E., JESDALE, W. M. & ALONSO, L. C. 2018. CDKN2A/B T2D Genome-Wide Association Study Risk SNPs Impact Locus Gene Expression and Proliferation in Human Islets. *Diabetes,* 67**,** 872-884.

KONG, Y., SHARMA, R. B., NWOSU, B. U. & ALONSO, L. C. 2016. Islet biology, the CDKN2A/B locus and type 2 diabetes risk. *Diabetologia,* 59**,** 1579-93.

KRAPP, A., KNOFLER, M., LEDERMANN, B., BURKI, K., BERNEY, C., ZOERKLER, N., HAGENBUCHLE, O. & 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 & Development,* 12**,** 3752-3763.

KRISHNAMURTHY, J., RAMSEY, M. R., LIGON, K. L., TORRICE, C., KOH, A., BONNER-WEIR, S. & SHARPLESS, N. E. 2006. p16INK4a induces an age-dependent decline in islet regenerative potential. *Nature,* 443**,** 453-7.

LAMMERT, E., CLEAVER, O. & MELTON, D. 2003. Role of endothelial cells in early pancreas and liver development. *Mech Dev,* 120**,** 59-64.

LANGENBERG, C. & LOTTA, L. A. 2018. Genomic insights into the causes of type 2 diabetes. *Lancet,* 391**,** 2463-2474.

LARSEN, H. L. & GRAPIN-BOTTON, A. 2017. The molecular and morphogenetic basis of pancreas organogenesis. *Semin Cell Dev Biol,* 66**,** 51-68.

LAU, H. H., NG, N. H. J., LOO, L. S. W., JASMEN, J. B. & TEO, A. K. K. 2018. The molecular functions of hepatocyte nuclear factors - In and beyond the liver. *J Hepatol,* 68**,** 1033-1048.

LAVIN, M. F. 2008. Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer (vol 9, pg 759, 2008). *Nature Reviews Molecular Cell Biology,* 9**,** 927-927.

LEAHY, J. J. J., GOLDING, B. T., GRIFFIN, R. J., HARDCASTLE, I. R., RICHARDSON, C., RIGOREAU, L. & SMITH, G. C. M. 2004. Identification of a highly potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor (NU7441) by screening of chromenone libraries. *Bioorganic & Medicinal Chemistry Letters,* 14**,** 6083-6087.

LEVINE, A. J., MOMAND, J. & FINLAY, C. A. 1991. The p53 tumour suppressor gene. *Nature,* 351**,** 453-6.

LI, H. X., GAN, W., LU, L., DONG, X., HAN, X. Y., HU, C., YANG, Z., SUN, L., BAO, W., LI, P. T., HE, M. A., SUN, L. D., WANG, Y. Q., ZHU, J. W., NING, Q. Q., TANG, Y., ZHANG, R., WEN, J., WANG, D., ZHU, X. L., GUO, K. Q., ZUO, X. B., GUO, X. H., YANG, H. D., ZHOU, X. H., ZHANG, X. J., QI, L., LOOS, R. J. F., HU, F. B., WU, T. C., LIU, Y., LIU, L. G., YANG, Z., HU, R. M., JIA, W. P., JI, L. N., LI, Y. X., LIN, X., CONSORTIUM, D. & CONSORTIUM, A.-T. D. 2013. A Genome-Wide Association Study Identifies GRK5 and RASGRP1 as Type 2 Diabetes Loci in Chinese Hans. *Diabetes,* 62**,** 291-298.

LI, J. A., KNOBLOCH, T. J., POI, M. J., ZHANG, Z. X., DAVIS, A. T., MUSCARELLA, P. & WEGHORST, C. M. 2014. Genetic alterations of RDINK4/ARF enhancer in human cancer cells. *Molecular Carcinogenesis,* 53**,** 211-218.

LI, X. L., LI, G. H., FU, J., FU, Y. W., ZHANG, L., CHEN, W. Q., ARAKAKI, C., ZHANG, J. P., WEN, W., ZHAO, M., CHEN, W. V., BOTIMER, G. D., BAYLINK, D., ARANDA, L., CHOI, H., BECHAR, R., TALBOT, P., SUN, C. K., CHENG, T. & ZHANG, X. B. 2018. Highly efficient genome editing via CRISPR-Cas9 in human pluripotent stem cells is achieved by transient BCL-XL overexpression. *Nucleic Acids Research,* 46**,** 10195-10215.

LIN, C. L. V. & VUGUIN, P. M. 2012. Determinants of Pancreatic Islet Development in Mice and Men: A Focus on the Role of Transcription Factors. *Hormone Research in Paediatrics,* 77**,** 205-213.

LIU, J. C., GUAN, X., RYAN, J. A., RIVERA, A. G., MOCK, C., AGRAWAL, V., LETAI, A., LEROU, P. H. & LAHAV, G. 2013. High Mitochondrial Priming Sensitizes hESCs to DNA-Damage-Induced Apoptosis (vol 13, pg 483, 2013). *Cell Stem Cell,* 13**,** 634-634.

LIU, M., REHMAN, S., TANG, X., GU, K., FAN, Q., CHEN, D. & MA, W. 2018. Methodologies for Improving HDR Efficiency. *Front Genet,* 9**,** 691.

LOBO, V. J. S. A., FERNANDEZ, L. C., CARRILLO-DE-SANTA-PAU, E., RICHART, L., COBO, I., CENDROWSKI, J., MORENO, U., DEL POZO, N., MEGIAS, D., BREANT, B., WRIGHT, C. V., MAGNUSON, M. & REAL, F. X. 2018. c-Myc downregulation is required for preacinar to acinar maturation and pancreatic homeostasis. *Gut,* 67**,** 707-718.

LOINARD, C., BASATEMUR, G., MASTERS, L., BAKER, L., HARRISON, J., FIGG, N., VILAR, J., SAGE, A. P. & MALLAT, Z. 2014. Deletion of chromosome 9p21 noncoding cardiovascular risk interval in mice alters Smad2 signaling and promotes vascular aneurysm. *Circ Cardiovasc Genet,* 7**,** 799-805.

LYTTLE, B. M., KRISHNAMURTHY, J. L. M., FELLOWS, F., WHEELER, M. B., GOODYER, C. G. & WANG, R. 2008. Transcription factor expression in the developing human fetal endocrine pancreas. *Diabetologia,* 51**,** 1169-1180.

MA, C. J., GIBB, B., KWON, Y., SUNG, P. & GREENE, E. C. 2017. Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Research,* 45**,** 749-761.

MA, R., YANG, H., LI, J., YANG, X., CHEN, X., HU, Y., WANG, Z., XUE, L. & ZHOU, W. 2016. Association of HNF4alpha gene polymorphisms with susceptibility to type 2 diabetes. *Mol Med Rep,* 13**,** 2241-6.

MAECHLER, P., CAROBBIO, S. & RUBI, B. 2006. In beta-cells, mitochondria integrate and generate metabolic signals controlling insulin secretion. *Int J Biochem Cell Biol,* 38**,** 696-709.

MAHADDALKAR, P. U., SCHEIBNER, K., PFLUGER, S., ANSARULLAH, STERR, M., BECKENBAUER, J., IRMLER, M., BECKERS, J., KNOBEL, S. & LICKERT, H. 2020. Generation of pancreatic beta cells from CD177(+) anterior definitive endoderm. *Nat Biotechnol,* 38**,** 1061-1072.

MAHAJAN, A., GO, M. J., ZHANG, W. H., BELOW, J. E., GAULTON, K. J., FERREIRA, T., HORIKOSHI, M., JOHNSON, A. D., NG, M. C. Y., PROKOPENKO, I., SALEHEEN, D., WANG, X., ZEGGINI, E., ABECASIS, G. R., ADAIR, L. S., ALMGREN, P., ATALAY, M., AUNG, T., BALDASSARRE, D., BALKAU, B., BAO, Y. Q., BARNETT, A. H., BARROSO, I., BASIT, A., BEEN, L. F., BEILBY, J., BELL, G. I., BENEDIKTSSON, R., BERGMAN, R. N., BOEHM, B. O., BOERWINKLE, E., BONNYCASTLE, L. L., BURTT, N., CAI, Q. Y., CAMPBELL, H., CAREY, J., CAUCHI, S., CAULFIELD, M., CHAN, J. C. N., CHANG, L. C., CHANG, T. J., CHANG, Y. C., CHARPENTIER, G., CHEN, C. H., CHEN, H., CHEN, Y. T., CHIA, K. S., CHIDAMBARAM, M., CHINES, P. S., CHO, N. H., CHO, Y. M., CHUANG, L. M., COLLINS, F. S., CORNELIS, M. C., COUPER, D. J., CRENSHAW, A. T., VAN DAM, R. M., DANESH, J., DAS, D., DE FAIRE, U., DEDOUSSIS, G., DELOUKAS, P., DIMAS, A. S., DINA, C., DONEY, A. S. F., DONNELLY, P. J., DORKHAN, M., VAN DUIJN, C., DUPUIS, J., EDKINS, S., ELLIOTT, P., EMILSSON, V., ERBEL, R., ERIKSSON, J. G., ESCOBEDO, J., ESKO, T., EURY, E., FLOREZ, J. C., FONTANILLAS, P., FOROUHI, N. G., FORSEN, T., FOX, C., FRASER, R. M., FRAYLING, T. M., FROGUEL, P., FROSSARD, P., GAO, Y. T., GERTOW, K., GIEGER, C., GIGANTE, B., GRALLERT, H., GRANT, G. B., GROOP, L. C., GROVES, C. J., GRUNDBERG, E., GUIDUCCI, C., HAMSTEN, A., HAN, B. G., HARA, K., HASSANALI, N., et al. 2014. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature Genetics,* 46**,** 234-+.

MAJITHIA, A. R. & FLOREZ, J. C. 2009. Clinical translation of genetic predictors for type 2 diabetes. *Current Opinion in Endocrinology Diabetes and Obesity,* 16**,** 100-106.

MAKAROVA, K. S., WOLF, Y. I., ALKHNBASHI, O. S., COSTA, F., SHAH, S. A., SAUNDERS, S. J., BARRANGOU, R., BROUNS, S. J. J., CHARPENTIER, E., HAFT, D. H., HORVATH, P., MOINEAU, S., MOJICA, F. J. M., TERNS, R. M., TERNS, M. P., WHITE, M. F., YAKUNIN, A. F., GARRETT, R. A., VAN DER OOST, J., BACKOFEN, R. & KOONIN, E. V. 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews Microbiology,* 13**,** 722-736.

MAKHARASHVILI, N. & PAULL, T. T. 2015. CtIP: A DNA damage response protein at the intersection of DNA metabolism. *DNA Repair,* 32**,** 75-81.

MALI, P., YANG, L. H., ESVELT, K. M., AACH, J., GUELL, M., DICARLO, J. E., NORVILLE, J. E. & CHURCH, G. M. 2013. RNA-Guided Human Genome Engineering via Cas9. *Science,* 339**,** 823-826.

MANNING, A. K., HIVERT, M. F., SCOTT, R. A., GRIMSBY, J. L., BOUATIA-NAJI, N., CHEN, H., RYBIN, D., LIU, C. T., BIELAK, L. F., PROKOPENKO, I., AMIN, N., BARNES, D., CADBY, G., HOTTENGA, J. J., INGELSSON, E., JACKSON, A. U., JOHNSON, T., KANONI, S., LADENVALL, C., LAGOU, V., LAHTI, J., LECOEUR, C., LIU, Y. M., MARTINEZ-LARRAD, M. T., MONTASSER, M. E., NAVARRO, P., PERRY, J. R. B., RASMUSSEN-TORVIK, L. J., SALO, P., SATTAR, N., SHUNGIN, D., STRAWBRIDGE, R. J., TANAKA, T., VAN DUIJN, C. M., AN, P., DE ANDRADE, M., ANDREWS, J. S., ASPELUND, T., ATALAY, M., AULCHENKO, Y., BALKAU, B., BANDINELLI, S., BECKMANN, J. S., BEILBY, J. P., BELLIS, C., BERGMAN, R. N., BLANGERO, J., BOBAN, M., BOEHNKE, M., BOERWINKLE, E., BONNYCASTLE, L. L., BOOMSMA, D. I., BORECKI, I. B., BOETTCHER, Y., BOUCHARD, C., BRUNNER, E., BUDIMIR, D., CAMPBELL, H., CARLSON, O., CHINES, P. S., CLARKE, R., COLLINS, F. S., CORBATON-ANCHUELO, A., COUPER, D., DE FAIRE, U., DEDOUSSIS, G. V., DELOUKAS, P., DIMITRIOU, M., EGAN, J. M., EIRIKSDOTTIR, G., ERDOS, M. R., ERIKSSON, J. G., EURY, E., FERRUCCI, L., FORD, I., FOROUHI, N. G., FOX, C. S., FRANZOSI, M. G., FRANKS, P. W., FRAYLING, T. M., FROGUEL, P., GALAN, P., DE GEUS, E., GIGANTE, B., GLAZER, N. L., GOEL, A., GROOP, L., GUDNASON, V., HALLMANS, G., HAMSTEN, A., HANSSON, O., HARRIS, T. B., HAYWARD, C., HEATH, S., HERCBERG, S., HICKS, A. A., HINGORANI, A., HOFMAN, A., HUI, J., HUNG, J., et al. 2012. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature Genetics,* 44**,** 659-U81.

MAO, Z. Y., BOZZELLA, M., SELUANOV, A. & GORBUNOVA, V. 2008. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair,* 7**,** 1765-1771.

MARRAFFINI, L. A. & SONTHEIMER, E. J. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet,* 11**,** 181-90.

MARTY-SANTOS, L. & CLEAVER, O. 2015. Progenitor Epithelium: Sorting Out Pancreatic Lineages. *J Histochem Cytochem,* 63**,** 559-74.

MARUYAMA, T., DOUGAN, S. K., TRUTTMANN, M. C., BILATE, A. M., INGRAM, J. R. & PLOEGH, H. L. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol,* 33**,** 538-42.

MASUI, T., LONG, Q., BERES, T. M., MAGNUSON, M. A. & MACDONALD, R. J. 2007. Early pancreatic development requires the vertebrate Suppressor of Hairless (RBPJ) in the PTF1 bHLH complex. *Genes & Development,* 21**,** 2629-2643.

MASUI, T., SWIFT, G. H., DEERING, T., SHEN, C. C., COATS, W. S., LONG, Q. M., ELSASSER, H. P., MAGNUSON, M. A. & 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.

MATHIASEN, D. P. & LISBY, M. 2014. Cell cycle regulation of homologous recombination in Saccharomyces cerevisiae. *Fems Microbiology Reviews,* 38**,** 172-184.

MATSCHINSKY, F. M. 1996. Banting Lecture 1995. A lesson in metabolic regulation inspired by the glucokinase glucose sensor paradigm. *Diabetes,* 45**,** 223-41.
MAYHEW, C. N. & WELLS, J. M. 2010. Converting human pluripotent stem cells into beta-cells: recent advances and future challenges. *Current Opinion in Organ Transplantation,* 15**,** 54-60.
MAYLE, R., CAMPBELL, I. M., BECK, C. R., YU, Y., WILSON, M., SHAW, C. A., BJERGBAEK, L., LUPSKI, J. R. & IRA, G. 2015. DNA REPAIR. Mus81 and converging forks limit the mutagenicity of replication fork breakage. *Science,* 349**,** 742-7.
MCGRATH, P. S., WATSON, C. L., INGRAM, C., HELMRATH, M. A. & WELLS, J. M. 2015. The Basic Helix-Loop-Helix Transcription Factor NEUROG3 Is Required for Development of the Human Endocrine Pancreas. *Diabetes,* 64**,** 2497-2505.
MEIER, J. J., KOHLER, C. U., ALKHATIB, B., SERGI, C., JUNKER, T., KLEIN, H. H., SCHMIDT, W. E. & FRITSCH, H. 2010. beta-cell development and turnover during prenatal life in humans. *European Journal of Endocrinology,* 162**,** 559-568.
MEIGS, J. B. 2019. The Genetic Epidemiology of Type 2 Diabetes: Opportunities for Health Translation. *Curr Diab Rep,* 19**,** 62.
MILLMAN, J. R., XIE, C., VAN DERVORT, A., GURTLER, M., PAGLIUCA, F. W. & MELTON, D. A. 2016. Generation of stem cell-derived beta-cells from patients with type 1 diabetes. *Nat Commun,* 7**,** 11463.
MIRALLES, F., CZERNICHOW, P. & SCHARFMANN, R. 1998. Follistatin regulates the relative proportions of endocrine versus exocrine tissue during pancreatic development. *Development,* 125**,** 1017-24.
MISHRA, V., NAYAK, P., SHARMA, M., ALBUTTI, A., ALWASHMI, A. S. S., ALJASIR, M. A., ALSOWAYEH, N. & TAMBUWALA, M. M. 2021. Emerging Treatment Strategies for Diabetes Mellitus and Associated Complications: An Update. *Pharmaceutics,* 13.
MIZUKAMI, H., TAKAHASHI, K., INABA, W., OSONOI, S., KAMATA, K., TSUBOI, K. & YAGIHASHI, S. 2014. Age-associated changes of islet endocrine cells and the effects of body mass index in Japanese. *Journal of Diabetes Investigation,* 5**,** 38-47.
MLADENOV, E., MAGIN, S., SONI, A. & ILIAKIS, G. 2016. DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. *Seminars in Cancer Biology,* 37-38**,** 51-64.
MORADI, S., MAHDIZADEH, H., SARIC, T., KIM, J., HARATI, J., SHAHSAVARANI, H., GREBER, B. & MOORE, J. B. T. 2019. Research and therapy with induced pluripotent stem cells (iPSCs): social, legal, and ethical considerations. *Stem Cell Res Ther,* 10**,** 341.
MORRIS, A. P. 2014. Fine mapping of type 2 diabetes susceptibility loci. *Curr Diab Rep,* 14**,** 549.
MOSHOUS, D., CALLEBAUT, I., DE CHASSEVAL, R., CORNEO, B., CAVAZZANA-CALVO, M., LE DEIST, F., TEZCAN, I., SANAL, O., BERTRAND, Y., PHILIPPE, N., FISCHER, A. & DE VILLARTAY, J. P. 2001. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell,* 105**,** 177-186.
MUNCK, J. M., BATEY, M. A., ZHAO, Y., JENKINS, H., RICHARDSON, C. J., CANO, C., TAVECCHIO, M., BARBEAU, J., BARDOS, J., CORNELL, L., GRIFFIN, R. J., MENEAR, K., SLADE, A., THOMMES, P., MARTIN, N. M. B., NEWELL, D. R., SMITH, G. C. M. & CURTIN, N. J. 2012. Chemosensitization of Cancer Cells by KU-0060648, a Dual Inhibitor of DNA-PK and PI-3K. *Molecular Cancer Therapeutics,* 11**,** 1789-1798.
MURARO, M. J., DHARMADHIKARI, G., GRUN, D., GROEN, N., DIELEN, T., JANSEN, E., VAN GURP, L., ENGELSE, M. A., CARLOTTI, F., DE KONING, E. J. & VAN OUDENAARDEN, A. 2016. A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst,* 3**,** 385-394 e3.

MURPHY, R., ELLARD, S. & HATTERSLEY, A. T. 2008. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab,* 4**,** 200-13.

MUSCARELLA, P., MELVIN, W. S., FISHER, W. E., FOOR, J., ELLISON, E. C., HERMAN, J. G., SCHIRMER, W. J., HITCHCOCK, C. L., DEYOUNG, B. R. & WEGHORST, C. M. 1998. Genetic alterations in gastrinomas and nonfunctioning pancreatic neuroendocrine tumors: An analysis of p16/MTS1 tumor suppressor gene inactivation. *Cancer Research,* 58**,** 237-240.

NANDA, V., DOWNING, K. P., YE, J., XIAO, S., KOJIMA, Y., SPIN, J. M., DIRENZO, D., NEAD, K. T., CONNOLLY, A. J., DANDONA, S., PERISIC, L., HEDIN, U., MAEGDEFESSEL, L., DALMAN, J., GUO, L., ZHAO, X., KOLODGIE, F. D., VIRMANI, R., DAVIS, H. R., JR. & LEEPER, N. J. 2016. CDKN2B Regulates TGFbeta Signaling and Smooth Muscle Cell Investment of Hypoxic Neovessels. *Circ Res,* 118**,** 230-40.

NIHAD, M., SHENOY, P. S. & BOSE, B. 2021. Cell therapy research for Diabetes: Pancreatic beta cell differentiation from pluripotent stem cells. *Diabetes Res Clin Pract,* 181**,** 109084.

NISHIMASU, H., RAN, F. A., HSU, P. D., KONERMANN, S., SHEHATA, S. I., DOHMAE, N., ISHITANI, R., ZHANG, F. & NUREKI, O. 2014. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell,* 156**,** 935-949.

NOSTRO, M. C. & KELLER, G. 2012. Generation of beta cells from human pluripotent stem cells: Potential for regenerative medicine. *Seminars in Cell & Developmental Biology,* 23**,** 701-710.

OFFIELD, M. F., JETTON, T. L., LABOSKY, P. A., RAY, M., STEIN, R. W., MAGNUSON, M. A., HOGAN, B. L. M. & WRIGHT, C. V. E. 1996. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development,* 122**,** 983-995.

PAGLIUCA, F. W., MILLMAN, J. R., GURTLER, M., SEGEL, M., VAN DERVORT, A., RYU, J. H., PETERSON, Q. P., GREINER, D. & MELTON, D. A. 2014. Generation of functional human pancreatic beta cells in vitro. *Cell,* 159**,** 428-39.

PAL, A. & MCCARTHY, M. I. 2013. The genetics of type 2 diabetes and its clinical relevance. *Clin Genet,* 83**,** 297-306.

PAL, A., POTJER, T. P., THOMSEN, S. K., NG, H. J., BARRETT, A., SCHARFMANN, R., JAMES, T. J., BISHOP, D. T., KARPE, F., GODSLAND, I. F., VASEN, H. F., NEWTON-BISHOP, J., PIJL, H., MCCARTHY, M. I. & GLOYN, A. L. 2016. Loss-of-Function Mutations in the Cell-Cycle Control Gene CDKN2A Impact on Glucose Homeostasis in Humans. *Diabetes,* 65**,** 527-33.

PALMER, N. D., GOODARZI, M. O., LANGEFELD, C. D., WANG, N., GUO, X. Q., TAYLOR, K. D., FINGERLIN, T. E., NORRIS, J. M., BUCHANAN, T. A., XIANG, A. H., HARITUNIANS, T., ZIEGLER, J. T., WILLIAMS, A. H., STEFANOVSKI, D., CUI, J. R., MACKAY, A. W., HENKIN, L. F., BERGMAN, R. N., GAO, X. Y., GAUDERMAN, J., VARMA, R., HANIS, C. L., COX, N. J., HIGHLAND, H. M., BELOW, J. E., WILLIAMS, A. L., BURTT, N. P., AGUILAR-SALINAS, C. A., HUERTA-CHAGOYA, A., GONZALEZ-VILLALPANDO, C., OROZCO, L., HAIMAN, C. A., TSAI, M. Y., JOHNSON, W. C., YAO, J., RASMUSSEN-TORVIK, L., PANKOW, J., SNIVELY, B., JACKSON, R. D., LIU, S. M., NADLER, J. L., KANDEEL, F., CHEN, Y. D. I., BOWDEN, D. W., RICH, S. S., RAFFEL, L. J., ROTTER, J. I., WATANABE, R. M. & WAGENKNECHT, L. E. 2015. Genetic Variants Associated With Quantitative Glucose Homeostasis Traits Translate to Type 2 Diabetes in Mexican Americans: The GUARDIAN (Genetics Underlying Diabetes in Hispanics) Consortium. *Diabetes,* 64**,** 1853-1866.

PAN, F. C. & BRISSOVA, M. 2014. Pancreas development in humans. *Curr Opin Endocrinol Diabetes Obes,* 21**,** 77-82.

PANICO, P., SALAZAR, A. M., BURNS, A. L. & OSTROSKY-WEGMAN, P. 2014. Role of calpain-10 in the development of diabetes mellitus and its complications. *Arch Med Res,* 45**,** 103-15.

PAPAGIANNAKOPOULOS, T., SHAPIRO, A. & KOSIK, K. S. 2008. MicroRNA-21 targets a network of key tumor-suppressive pathways in glioblastoma cells. *Cancer Res,* 68**,** 8164-72.
PARRA, E. J., BELOW, J. E., KRITHIKA, S., VALLADARES, A., BARTA, J. L., COX, N. J., HANIS, C. L., WACHER, N., GARCIA-MENA, J., HU, P., SHRIVER, M. D., KUMATE, J., MCKEIGUE, P. M., ESCOBEDO, J., CRUZ, M. & META-A, D. G. R. 2011. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. *Diabetologia,* 54**,** 2038-2046.
PASCOE, J., HOLLERN, D., STAMATERIS, R., ABBASI, M., ROMANO, L. C., ZOU, B., O'DONNELL, C. P., GARCIA-OCANA, A. & ALONSO, L. C. 2012. Free fatty acids block glucose-induced beta-cell proliferation in mice by inducing cell cycle inhibitors p16 and p18. *Diabetes,* 61**,** 632-41.
PASMANT, E., LAURENDEAU, I., HERON, D., VIDAUD, M., VIDAUD, D. & BIECHE, I. 2007. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res,* 67**,** 3963-9.
PASMANT, E., LAURENDEAU, I., SABBAGH, A., PARFAIT, B., VIDAUD, M., VIDAUD, D. & BIECHE, I. 2010. [The amazing story of ANRIL, a long non-coding RNA]. *Med Sci (Paris),* 26**,** 564-6.
PASMANT, E., SABBAGH, A., VIDAUD, M. & BIECHE, I. 2011. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J,* 25**,** 444-8.
PASQUALI, L., GAULTON, K. J., RODRIGUEZ-SEGUI, S. A., MULARONI, L., MIGUEL-ESCALADA, I., AKERMAN, I., TENA, J. J., MORAN, I., GOMEZ-MARIN, C., VAN DE BUNT, M., PONSA-COBAS, J., CASTRO, N., NAMMO, T., CEBOLA, I., GARCIA-HURTADO, J., MAESTRO, M. A., PATTOU, F., PIEMONTI, L., BERNEY, T., GLOYN, A. L., RAVASSARD, P., SKARMETA, J. L. G., MULLER, F., MCCARTHY, M. I. & FERRER, J. 2014. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet,* 46**,** 136-143.
PATTERSON, C. C., DAHLQUIST, G. G., GYURUS, E., GREEN, A., SOLTESZ, G. & GROUP, E. S. 2009. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study. *Lancet,* 373**,** 2027-33.
PEARSON, E. R., STARKEY, B. J., POWELL, R. J., GRIBBLE, F. M., CLARK, P. M. & HATTERSLEY, A. T. 2003. Genetic cause of hyperglycaemia and response to treatment in diabetes. *Lancet,* 362**,** 1275-81.
PENG, F., HU, D., GU, C. H., LI, X. B., LI, Y. Q., JIA, N., CHU, S. L., LIN, J. X. & NIU, W. Q. 2013. The relationship between five widely-evaluated variants in CDKN2A/B and CDKAL1 genes and the risk of type 2 diabetes: A meta-analysis. *Gene,* 531**,** 435-443.
PERRY, J. R. B. & FRAYLING, T. M. 2008. New gene variants alter type 2 diabetes risk predominantly through reduced beta-cell function. *Current Opinion in Clinical Nutrition and Metabolic Care,* 11**,** 371-377.
PETERSEN, M. C. & SHULMAN, G. I. 2018. Mechanisms of Insulin Action and Insulin Resistance. *Physiol Rev,* 98**,** 2133-2223.
PFANNKUCHE, K., HANNES, T., KHALIL, M., NOGHABI, M. S., MORSHEDI, A., HESCHELER, J. & DROGE, P. 2010. Induced pluripotent stem cells: a new approach for physiological research. *Cell Physiol Biochem,* 26**,** 105-24.
PINDER, J., SALSMAN, J. & DELLAIRE, G. 2015. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Research,* 43**,** 9379-9392.
PIPATPOLKAI, T., USHER, S., STANSFELD, P. J. & ASHCROFT, F. M. 2020. New insights into K-ATP channel gene mutations and neonatal diabetes mellitus. *Nature Reviews Endocrinology,* 16**,** 378-393.

POI, M. J., DROSDECK, J., FRANKEL, W. L., MUSCARELLA, P. & LI, J. A. 2014. Deletions of RDINK4/ARF Enhancer in Gastrinomas and Nonfunctioning Pancreatic Neuroendocrine Tumors. *Pancreas,* 43**,** 1009-1013.

POI, M. J., KNOBLOCH, T. J., YUAN, C. H., TSAI, M. D., WEGHORST, C. M. & LI, J. N. 2013. Evidence that P12, a specific variant of P16(INK4A), plays a suppressive role in human pancreatic carcinogenesis. *Biochemical and Biophysical Research Communications,* 436**,** 217-222.

POPOV, N. & GIL, J. 2010. Epigenetic regulation of the INK4b-ARF-INK4a locus: in sickness and in health. *Epigenetics,* 5**,** 685-90.

PULIZZI, N., LYSSENKO, V., JONSSON, A., OSMOND, C., LAAKSO, M., KAJANTIE, E., BARKER, D. J., GROOP, L. C. & ERIKSSON, J. G. 2009. Interaction between prenatal growth and high-risk genotypes in the development of type 2 diabetes (vol 52, pg 825, 2009). *Diabetologia,* 52**,** 2671-2672.

QUELLE, D. E., ASHMUN, R. A., HANNON, G. J., REHBERGER, P. A., TRONO, D., RICHTER, K. H., WALKER, C., BEACH, D., SHERR, C. J. & SERRANO, M. 1995. Cloning and characterization of murine p16INK4a and p15INK4b genes. *Oncogene,* 11**,** 635-45.

QUEREDA, V., MARTINALBO, J., DUBUS, P., CARNERO, A. & MALUMBRES, M. 2007. Genetic cooperation between p21(Cip1) and INK4 inhibitors in cellular senescence and tumor suppression. *Oncogene,* 26**,** 7665-7674.

RAMSEY, M. R., KRISHNAMURTHY, J., PEI, X. H., TORRICE, C., LIN, W., CARRASCO, D. R., LIGON, K. L., XIONG, Y. & SHARPLESS, N. E. 2007. Expression of p16Ink4a compensates for p18Ink4c loss in cyclin-dependent kinase 4/6-dependent tumors and tissues. *Cancer Res,* 67**,** 4732-41.

RANE, S. G., DUBUS, P., METTUS, R. V., GALBREATH, E. J., BODEN, G., REDDY, E. P. & BARBACID, M. 1999. Loss of Cdk4 expression causes insulin-deficient diabetes and Cdk4 activation results in beta-islet cell hyperplasia. *Nat Genet,* 22**,** 44-52.

REZANIA, A., BRUIN, J. E., ARORA, P., RUBIN, A., BATUSHANSKY, I., ASADI, A., O'DWYER, S., QUISKAMP, N., MOJIBIAN, M., ALBRECHT, T., YANG, Y. H., JOHNSON, J. D. & KIEFFER, T. J. 2014. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol,* 32**,** 1121-33.

RIEDEL, M. J., ASADI, A., WANG, R., AO, Z., WARNOCK, G. L. & KIEFFER, T. J. 2012. Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia,* 55**,** 372-381.

RIOPEL, M., LI, J. M., FELLOWS, G. F., GOODYER, C. G. & WANG, R. N. 2014. Ultrastructural and immunohistochemical analysis of the 8-20 week human fetal pancreas. *Islets,* 6.

ROBERTSON, K. D. & JONES, P. A. 1999. Tissue-specific alternative splicing in the human INK4a/ARF cell cycle regulatory locus. *Oncogene,* 18**,** 3810-3820.

RUBIO-CABEZAS, O., CODNER, E., FLANAGAN, S. E., GOMEZ, J. L., ELLARD, S. & HATTERSLEY, A. T. 2014. Neurogenin 3 is important but not essential for pancreatic islet development in humans. *Diabetologia,* 57**,** 2421-4.

RUCHAT, S. M., WEISNAGEL, S. J., VOHL, M. C., RANKINEN, T., BOUCHARD, C. & PERUSSE, L. 2009. Evidence for Interaction between PPARG Pro12Ala and PPARGC1A Gly482Ser Polymorphisms in Determining Type 2 Diabetes Intermediate Phenotypes in Overweight Subjects. *Experimental and Clinical Endocrinology & Diabetes,* 117**,** 455-459.

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., BLELLOCH, R., SZOT, G. L., ARVAN, P. & HEBROK, M. 2015. Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. *EMBO J,* 34**,** 1759-72.

RUTTER, G. A. 2001. Nutrient-secretion coupling in the pancreatic islet beta-cell: recent advances. *Mol Aspects Med,* 22**,** 247-84.
SAEEDI, P., PETERSOHN, I., SALPEA, P., MALANDA, B., KARURANGA, S., UNWIN, N., COLAGIURI, S., GUARIGUATA, L., MOTALA, A. A., OGURTSOVA, K., SHAW, J. E., BRIGHT, D., WILLIAMS, R. & COMMITTEE, I. D. F. D. A. 2019. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. *Diabetes Res Clin Pract,* 157**,** 107843.
SALIS, E. R., REITH, D. M., WHEELER, B. J., BROADBENT, R. S. & MEDLICOTT, N. J. 2017. Insulin resistance, glucagon-like peptide-1 and factors influencing glucose homeostasis in neonates. *Archives of Disease in Childhood-Fetal and Neonatal Edition,* 102**,** F162-F166.
SALISBURY, R. J., BLAYLOCK, J., BERRY, A. A., JENNINGS, R. E., DE KRIJGER, R., HANLEY, K. P. & HANLEY, N. A. 2014. The window period of NEUROGENIN3 during human gestation. *Islets,* 6.
SALSMAN, J. & DELLAIRE, G. 2017. Precision genome editing in the CRISPR era. *Biochemistry and Cell Biology,* 95**,** 187-201.
SALTIEL, A. R. 2016. Insulin Signaling in the Control of Glucose and Lipid Homeostasis. *Handb Exp Pharmacol,* 233**,** 51-71.
SANDER, M., SUSSEL, L., CONNERS, J., SCHEEL, D., KALAMARAS, J., DELA CRUZ, F., SCHWITZGEBEL, V., HAYES-JORDAN, A. & GERMAN, M. 2000. Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas. *Development,* 127**,** 5533-5540.
SANDHU, M. S., WEEDON, M. N., FAWCETT, K. A., WASSON, J., DEBENHAM, S. L., DALY, A., LANGO, H., FRAYLING, T. M., NEUMANN, R. J., SHERVA, R., BLECH, I., PHAROAH, P. D., PALMER, C. N. A., KIMBER, C., TAVENDALE, R., MORRIS, A. D., MCCARTHY, M. I., WALKER, M., HITMAN, G., GLASER, B., PERMUTT, M. A., HATTERSLEY, A. T., WAREHAM, N. J. & BARROSO, I. 2007. Common variants in WFS1 confer risk of type 2 diabetes. *Nature Genetics,* 39**,** 951-953.
SAPRANAUSKAS, R., GASIUNAS, G., FREMAUX, C., BARRANGOU, R., HORVATH, P. & SIKSNYS, V. 2011. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res,* 39**,** 9275-82.
SARTORI, A. A., LUKAS, C., COATES, J., MISTRIK, M., FU, S., BARTEK, J., BAER, R., LUKAS, J. & JACKSON, S. P. 2007. Human CtIP promotes DNA end resection. *Nature,* 450**,** 509-U6.
SAXENA, R., VOIGHT, B. F., LYSSENKO, V., BURTT, N. P., DE BAKKER, P. I. W., CHEN, H., ROIX, J. J., KATHIRESAN, S., HIRSCHHORN, J. N., DALY, M. J., HUGHES, T. E., GROOP, L., ALTSHULER, D., ALMGREN, P., FLOREZ, J. C., MEYER, J., ARDLIE, K., BOSTROM, K. B., ISOMAA, B., LETTRE, G., LINDBLAD, U., LYON, H. N., MELANDER, O., NEWTON-CHEH, C., NILSSON, P., ORHO-MELANDER, M., RASTAM, L., SPELIOTES, E. K., TASKINEN, M. R., TUOMI, T., GUIDUCCI, C., BERGLUND, A., CARLSON, J., GIANNINY, L., HACKETT, R., HALL, L., HOLMKVIST, J., LAURILA, E., SJOGREN, M., STERNER, M., SURTI, A., SVENSSON, M., SVENSSON, M., TEWHEY, R., BLUMENSTIEL, B., PARKIN, M., DEFELICE, M., BARRY, R., BRODEUR, W., CAMARATA, J., CHIA, N., FAVA, M., GIBBONS, J., HANDSAKER, B., HEALY, C., NGUYEN, K., GATES, C., SOUGNEZ, C., GAGE, D., NIZZARI, M., GABRIEL, S. B., CHIRN, G. W., MA, Q. C., PARIKH, H., RICHARDSON, D., RICKE, D., PURCELL, S., IN, D. G. I. B. & RES, N. I. B. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science,* 316**,** 1331-1336.
SCHAFFER, A. E., FREUDE, K. K., NELSON, S. B. & SANDER, M. 2010. Nkx6 Transcription Factors and Ptf1a Function as Antagonistic Lineage Determinants in Multipotent Pancreatic Progenitors. *Developmental Cell,* 18**,** 1022-1029.

SCHISLER, J. C., JENSEN, P. B., TAYLOR, D. G., BECKER, T. C., KNOP, F. K., TAKEKAWA, S., GERMAN, M., WEIR, G. C., LU, D. H., MIRMIRA, R. G. & NEWGARD, C. B. 2005. The Nkx6.1 homeodomain transcription factor suppresses glucagon expression and regulates glucose-stimulated insulin secretion in islet beta cells. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 7297-7302.

SCHWITZGEBEL, V. M., SCHEEL, D. W., CONNERS, J. R., KALAMARAS, J., LEE, J. E., ANDERSON, D. J., SUSSEL, L., JOHNSON, J. D. & GERMAN, M. S. 2000. Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development,* 127**,** 3533-3542.

SCOTT, L. J., MOHLKE, K. L., BONNYCASTLE, L. L., WILLER, C. J., LI, Y., DUREN, W. L., ERDOS, M. R., STRINGHAM, H. M., CHINES, P. S., JACKSON, A. U., PROKUNINA-OLSSON, L., DING, C. J., SWIFT, A. J., NARISU, N., HU, T., PRUIM, R., XIAO, R., LI, X. Y., CONNEELY, K. N., RIEBOW, N. L., SPRAU, A. G., TONG, M., WHITE, P. P., HETRICK, K. N., BARNHART, M. W., BARK, C. W., GOLDSTEIN, J. L., WATKINS, L., XIANG, F., SARAMIES, J., BUCHANAN, T. A., WATANABE, R. M., VALLE, T. T., KINNUNEN, L., ABECASIS, G. R., PUGH, E. W., DOHENY, K. F., BERGMAN, R. N., TUOMILEHTO, J., COLLINS, F. S. & BOEHNKE, M. 2007. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science,* 316**,** 1341-1345.

SEELY, E. W. 2006. Does treatment of gestational diabetes mellitus affect pregnancy outcome? *Nat Clin Pract Endocrinol Metab,* 2**,** 72-3.

SERRANO, M., HANNON, G. J. & BEACH, D. 1993. A New Regulatory Motif in Cell-Cycle Control Causing Specific-Inhibition of Cyclin-D/Cdk4. *Nature,* 366**,** 704-707.

SEYMOUR, P. A., FREUDE, K. K., TRAN, M. N., MAYES, E. E., JENSEN, J., KIST, R., SCHERER, G. & SANDER, M. 2007. SOX9 is required for maintenance of the pancreatic progenitor cell pool. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 1865-1870.

SEYMOUR, P. A., SHIH, H. P., PATEL, N. A., FREUDE, K. K., XIE, R. Y., LIM, C. J. & SANDER, M. 2012. A Sox9/Fgf feed-forward loop maintains pancreatic organ identity. *Development,* 139**,** 3363-3372.

SHAHRYARI, A., BURTSCHER, I., NAZARI, Z. & LICKERT, H. 2021a. Engineering Gene Therapy: Advances and Barriers. *Advanced Therapeutics,* 4.

SHAHRYARI, A., MOYA, N., SIEHLER, J., WANG, X. M., BLOCHINGER, A. K., BURTSCHER, I., BAKHTI, M., MOWLA, S. J. & LICKERT, H. 2020. Generation of a human iPSC line harboring a biallelic large deletion at the INK4 locus (HMGUi001-A-5). *Stem Cell Research,* 47.

SHAHRYARI, A., MOYA, N., SIEHLER, J., WANG, X. M., BURTSCHER, I. & LICKERT, H. 2021b. Increasing Gene Editing Efficiency for CRISPR-Cas9 by Small RNAs in Pluripotent Stem Cells. *Crispr Journal,* 4**,** 491-501.

SHARMA, S., LEONARD, J., LEE, S., CHAPMAN, H. D., LEITER, E. H. & MONTMINY, M. R. 1996. Pancreatic islet expression of the homeobox factor STF-1 relies on an E-box motif that binds USF. *Journal of Biological Chemistry,* 271**,** 2294-2299.

SHARPLESS, N. E. & SHERR, C. J. 2015a. Forging a signature of in vivo senescence. *Nat Rev Cancer,* 15**,** 397-408.

SHARPLESS, N. E. & SHERR, C. J. 2015b. Forging a signature of in vivo senescence (vol 15, pg 397, 2015). *Nature Reviews Cancer,* 15.

SHIH, H. P., KOPP, J. L., SANDHU, M., DUBOIS, C. L., SEYMOUR, P. A., GRAPIN-BOTTON, A. & SANDER, M. 2012. A Notch-dependent molecular circuitry initiates pancreatic endocrine and ductal cell differentiation. *Development,* 139**,** 2488-99.

SINGH, V. K., KALSAN, M., KUMAR, N., SAINI, A. & CHANDRA, R. 2015. Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery. *Front Cell Dev Biol,* 3**,** 2.
SOLAR, M., CARDALDA, C., HOUBRACKEN, I., MARTIN, M., MAESTRO, M. A., DE MEDTS, N., XU, X., GRAU, V., HEIMBERG, H., BOUWENS, L. & FERRER, J. 2009. Pancreatic exocrine duct cells give rise to insulin-producing beta cells during embryogenesis but not after birth. *Dev Cell,* 17**,** 849-60.
SONG, E. K., LEE, Y. R., KIM, Y. R., YEOM, J. H., YOO, C. H., KIM, H. K., PARK, H. M., KANG, H. S., KIM, J. S., KIM, U. H. & HAN, M. K. 2012. NAADP mediates insulin-stimulated glucose uptake and insulin sensitization by PPARgamma in adipocytes. *Cell Rep,* 2**,** 1607-19.
SONG, Z. W., WU, W. C., CHEN, M. Y., CHENG, W. Z., YU, J. T., FANG, J. Y., XU, L. L., YASUNAGA, J., MATSUOKA, M. & ZHAO, T. J. 2018. Long Noncoding RNA ANRIL Supports Proliferation of Adult T-Cell Leukemia Cells through Cooperation with EZH2. *Journal of Virology,* 92.
STANFEL, M. N., MOSES, K. A., SCHWARTZ, R. J. & ZIMMER, W. E. 2005. Regulation of organ development by the Nkx-homeodomain factors: An Nkx code. *Cellular and Molecular Biology,* 51**,** Ol785-Ol799.
STEINBUSCH, L. K., LUIKEN, J. J., VLASBLOM, R., CHABOWSKI, A., HOEBERS, N. T., COUMANS, W. A., VROEGRIJK, I. O., VOSHOL, P. J., OUWENS, D. M., GLATZ, J. F. & DIAMANT, M. 2011. Absence of fatty acid transporter CD36 protects against Western-type diet-related cardiac dysfunction following pressure overload in mice. *Am J Physiol Endocrinol Metab,* 301**,** E618-27.
STERNBERG, S. H., REDDING, S., JINEK, M., GREENE, E. C. & DOUDNA, J. A. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature,* 507**,** 62-+.
SUSSEL, L., KALAMARAS, J., HARTIGAN-O'CONNOR, D. J., MENESES, J. J., PEDERSEN, R. A., RUBENSTEIN, J. L. R. & GERMAN, M. S. 1998. Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. *Development,* 125**,** 2213-2221.
SYMINGTON, L. S. 2014. End Resection at Double-Strand Breaks: Mechanism and Regulation. *Cold Spring Harbor Perspectives in Biology,* 6.
SZCZELKUN, M. D., TIKHOMIROVA, M. S., SINKUNAS, T., GASIUNAS, G., KARVELIS, T., PSCHERA, P., SIKSNYS, V. & SEIDEL, R. 2014. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 9798-9803.
TABASSUM, R., CHAUHAN, G., DWIVEDI, O. P., MAHAJAN, A., JAISWAL, A., KAUR, I., BANDESH, K., SINGH, T., MATHAI, B. J., PANDEY, Y., CHIDAMBARAM, M., SHARMA, A., CHAVALI, S., SENGUPTA, S., RAMAKRISHNAN, L., VENKATESH, P., AGGARWAL, S. K., GHOSH, S., PRABHAKARAN, D., SRINATH, R. K., SAXENA, M., BANERJEE, M., MATHUR, S., BHANSALI, A., SHAH, V. N., MADHU, S. V., MARWAHA, R. K., BASU, A., SCARIA, V., MCCARTHY, M. I., VENKATESAN, R., MOHAN, V., TANDON, N., BHARADWAJ, D., DIAGRAM & INDICO 2013. Genome-Wide Association Study for Type 2 Diabetes in Indians Identifies a New Susceptibility Locus at 2q21. *Diabetes,* 62**,** 977-986.
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**,** 861-72.
TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* 126**,** 663-76.
TAKEUCHI, F., SERIZAWA, M., YAMAMOTO, K., FUJISAWA, T., NAKASHIMA, E., OHNAKA, K., IKEGAMI, H., SUGIYAMA, T., KATSUYA, T., MIYAGISHI, M., NAKASHIMA, N., NAWATA, H.,

NAKAMURA, J., KONO, S., TAKAYANAGI, R. & KATO, N. 2009. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes,* 58**,** 1690-9.

TANEERA, J., FADISTA, J., AHLQVIST, E., ZHANG, M. Z., WIERUP, N., RENSTROM, E. & GROOP, L. 2013. Expression profiling of cell cycle genes in human pancreatic islets with and without type 2 diabetes. *Molecular and Cellular Endocrinology,* 375**,** 35-42.

TATEISHI, K., HE, J., TARANOVA, O., LIANG, G. Y., D'ALESSIO, A. C. & ZHANG, Y. 2008. Generation of Insulin-secreting Islet-like Clusters from Human Skin Fibroblasts. *Journal of Biological Chemistry,* 283**,** 31601-31607.

TEUMER, A., TIN, A., SORICE, R., GORSKI, M., YEO, N. C., CHU, A. Y., LI, M., LI, Y., MIJATOVIC, V., KO, Y. A., TALIUN, D., LUCIANI, A., CHEN, M. H., YANG, Q., FOSTER, M. C., OLDEN, M., HIRAKI, L. T., TAYO, B. O., FUCHSBERGER, C., DIEFFENBACH, A. K., SHULDINER, A. R., SMITH, A. V., ZAPPA, A. M., LUPO, A., KOLLERITS, B., PONTE, B., STENGEL, B., KRAMER, B. K., PAULWEBER, B., MITCHELL, B. D., HAYWARD, C., HELMER, C., MEISINGER, C., GIEGER, C., SHAFFER, C. M., MULLER, C., LANGENBERG, C., ACKERMANN, D., SISCOVICK, D., BOERWINKLE, E., KRONENBERG, F., EHRET, G. B., HOMUTH, G., WAEBER, G., NAVIS, G., GAMBARO, G., MALERBA, G., EIRIKSDOTTIR, G., LI, G., WICHMANN, H. E., GRALLERT, H., WALLASCHOFSKI, H., VOLZKE, H., BRENNER, H., KRAMER, H., LEACH, I. M., RUDAN, I., HILLEGE, H. L., BECKMANN, J. S., LAMBERT, J. C., LUAN, J. A., ZHAO, J. H., CHALMERS, J., CORESH, J., DENNY, J. C., BUTTERBACH, K., LAUNER, L. J., FERRUCCI, L., KEDENKO, L., HAUN, M., METZGER, M., WOODWARD, M., HOFFMAN, M. J., NAUCK, M., WALDENBERGER, M., PRUIJM, M., BOCHUD, M., RHEINBERGER, M., VERWEIJ, N., WAREHAM, N. J., ENDLICH, N., SORANZO, N., POLASEK, O., VAN DER HARST, P., PRAMSTALLER, P. P., VOLLENWEIDER, P., WILD, P. S., GANSEVOORT, R. T., RETTIG, R., BIFFAR, R., CARROLL, R. J., KATZ, R., LOOS, R. J. F., HWANG, S. J., COASSIN, S., BERGMANN, S., ROSAS, S. E., STRACKE, S., HARRIS, T. B., CORRE, T., et al. 2016. Genome-wide Association Studies Identify Genetic Loci Associated With Albuminuria in Diabetes. *Diabetes,* 65**,** 803-817.

THEIS, F. J. & LICKERT, H. 2019. A map of beta-cell differentiation pathways supports cell therapies for diabetes. *Nature,* 569**,** 342-343.

THORENS, B. 2014. Neural regulation of pancreatic islet cell mass and function. *Diabetes Obesity & Metabolism,* 16**,** 87-95.

TIMPSON, N. J., LINDGREN, C. M., WEEDON, M. N., RANDALL, J., OUWEHAND, W. H., STRACHAN, D. P., RAYNER, N. W., WALKER, M., HITMAN, G. A., DONEY, A. S. F., PALMER, C. N. A., MORRIS, A. D., HATTERSLEY, A. T., ZEGGINI, E., FRAYLING, T. M. & MCCARTHY, M. I. 2009. Adiposity-Related Heterogeneity in Patterns of Type 2 Diabetes Susceptibility Observed in Genome-Wide Association Data. *Diabetes,* 58**,** 505-510.

TULACHAN, S. S., DOI, R., HIRAI, Y., KAWAGUCHI, Y., KOIZUMI, M., HEMBREE, M., TEI, E., CROWLEY, A., YEW, H., MCFALL, C., PRASADAN, K., PREUETT, B., IMAMURA, M. & GITTES, G. K. 2006. Mesenchymal epimorphin is important for pancreatic duct morphogenesis. *Development Growth & Differentiation,* 48**,** 65-72.

UEDA, K., KAWANO, J., TAKEDA, K., YUJIRJ, T., TANABE, K., ANNO, T., AKIYAMA, M., NOZAKJ, J., YOSHINAGA, T., KOIZUMI, A., SHINODA, K., OKA, Y. & TANIZAWA, Y. 2005. Endoplasmic reticulum stress induces Wfs1 gene expression in pancreatic beta-cells via transcriptional activation. *European Journal of Endocrinology,* 153**,** 167-176.

VAN HOEK, M., DEHGHAN, A., WITTENTAN, J. C. M., VAN DUIIN, C. M., UITTERLINDEN, A. G., OOSTRA, B. A., HOFMAN, A., SIJBRANDS, E. J. G. & JANSSENS, A. C. J. W. 2008. Predicting

Type 2 Diabetes Based on Polymorphisms From Genome-Wide Association Studies A Population-Based Study. *Diabetes,* 57**,** 3122-3128.
VANHOOSE, A. M., SAMARAS, S., ARTNER, I., HENDERSON, E., HANG, Y. & STEIN, R. 2008. MafA and MafB regulate Pdx1 transcription through the Area II control region in pancreatic beta cells. *Journal of Biological Chemistry,* 283**,** 22612-22619.
VELAZCO-CRUZ, L., GOEDEGEBUURE, M. M. & MILLMAN, J. R. 2020. Advances Toward Engineering Functionally Mature Human Pluripotent Stem Cell-Derived beta Cells. *Frontiers in Bioengineering and Biotechnology,* 8.
VELAZCO-CRUZ, L., SONG, J., MAXWELL, K. G., GOEDEGEBUURE, M. M., AUGSORNWORAWAT, P., HOGREBE, N. J. & MILLMAN, J. R. 2019. Acquisition of Dynamic Function in Human Stem Cell-Derived beta Cells. *Stem Cell Reports,* 12**,** 351-365.
VILLASENOR, A., CHONG, D. C. & CLEAVER, O. 2008. Biphasic Ngn3 expression in the developing pancreas. *Dev Dyn,* 237**,** 3270-9.
VILLASENOR, A., CHONG, D. C., HENKEMEYER, M. & CLEAVER, O. 2010. Epithelial dynamics of pancreatic branching morphogenesis. *Development,* 137**,** 4295-4305.
VISEL, A., ZHU, Y., MAY, D., AFZAL, V., GONG, E., ATTANASIO, C., BLOW, M. J., COHEN, J. C., RUBIN, E. M. & PENNACCHIO, L. A. 2010. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature,* 464**,** 409-12.
VOIGHT, B. F., SCOTT, L. J., STEINTHORSDOTTIR, V., MORRIS, A. P., DINA, C., WELCH, R. P., ZEGGINI, E., HUTH, C., AULCHENKO, Y. S., THORLEIFSSON, G., MCCULLOCH, L. J., FERREIRA, T., GRALLERT, H., AMIN, N., WU, G. M., WILLER, C. J., RAYCHAUDHURI, S., MCCARROLL, S. A., LANGENBERG, C., HOFMANN, O. M., DUPUIS, J., QI, L., SEGRE, A. V., VAN HOEK, M., NAVARRO, P., ARDLIE, K., BALKAU, B., BENEDIKTSSON, R., BENNETT, A. J., BLAGIEVA, R., BOERWINKLE, E., BONNYCASTLE, L. L., BOSTROM, K. B., BRAVENBOER, B., BUMPSTEAD, S., BURTT, N. P., CHARPENTIER, G., CHINES, P. S., CORNELIS, M., COUPER, D. J., CRAWFORD, G., DONEY, A. S. F., ELLIOTT, K. S., ELLIOTT, A. L., ERDOS, M. R., FOX, C. S., FRANKLIN, C. S., GANSER, M., GIEGER, C., GRARUP, N., GREEN, T., GRIFFIN, S., GROVES, C. J., GUIDUCCI, C., HADJADJ, S., HASSANALI, N., HERDER, C., ISOMAA, B., JACKSON, A. U., JOHNSON, P. R. V., JORGENSEN, T., KAO, W. H. L., KLOPP, N., KONG, A., KRAFT, P., KUUSISTO, J., LAURITZEN, T., LI, M., LIEVERSE, A., LINDGREN, C. M., LYSSENKO, V., MARRE, M., MEITINGER, T., MIDTHJELL, K., MORKEN, M. A., NARISU, N., NILSSON, P., OWEN, K. R., PAYNE, F., PERRY, J. R. B., PETERSEN, A. K., PLATOU, C., PROENCA, C., PROKOPENKO, I., RATHMANN, W., RAYNER, N. W., ROBERTSON, N. R., ROCHELEAU, G., RODEN, M., SAMPSON, M. J., SAXENA, R., SHIELDS, B. M., SHRADER, P., SIGURDSSON, G., SPARSO, T., STRASSBURGER, K., STRINGHAM, H. M., SUN, Q., SWIFT, A. J., THORAND, B., et al. 2011. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis (vol 42, pg 579, 2010). *Nature Genetics,* 43**,** 388-388.
WANDZIOCH, E. & ZARET, K. S. 2009. Dynamic Signaling Network for the Specification of Embryonic Pancreas and Liver Progenitors. *Science,* 324**,** 1707-1710.
WANG, J. F., ELGHAZI, L., PARKER, S. E., KIZILOCAK, H., ASANO, M., SUSSEL, L. & SOSA-PINEDA, B. 2004. The concerted activities of Pax4 and Nkx2.2 are essential to initiate pancreatic beta-cell differentiation. *Developmental Biology,* 266**,** 178-189.
WANG, X., CHEN, S., BURTSCHER, I., STERR, M., HIERONIMUS, A., MACHICAO, F., STAIGER, H., HARING, H. U., LEDERER, G., MEITINGER, T. & LICKERT, H. 2016a. 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**,** 273-276.
WANG, X., STERR, M., ANSARULLAH, BURTSCHER, I., BOTTCHER, A., BECKENBAUER, J., SIEHLER, J., MEITINGER, T., HARING, H. U., STAIGER, H., CERNILOGAR, F. M., SCHOTTA, G., IRMLER, M.,

BECKERS, J., WRIGHT, C. V. E., BAKHTI, M. & LICKERT, H. 2019. Point mutations in the PDX1 transactivation domain impair human beta-cell development and function. *Mol Metab,* 24**,** 80-97.

WANG, X., STERR, M., BURTSCHER, I., CHEN, S., HIERONIMUS, A., MACHICAO, F., STAIGER, H., HARING, H. U., LEDERER, G., MEITINGER, T., CERNILOGAR, F. M., SCHOTTA, G., IRMLER, M., BECKERS, J., HRABE DE ANGELIS, M., RAY, M., WRIGHT, C. V. E., BAKHTI, M. & LICKERT, H. 2018. Genome-wide analysis of PDX1 target genes in human pancreatic progenitors. *Mol Metab,* 9**,** 57-68.

WANG, X. M., CHEN, S., BURTSCHER, I., STERR, M., HIERONIMUS, A., MACHICAO, F., STAIGER, H., HAERING, H. U., LEDERER, G., MEITINGER, T. & LICKERT, H. 2016b. 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**,** 292-295.

WELLS, J. M., ESNI, F., BOIVIN, G. P., ARONOW, B. J., STUART, W., COMBS, C., SKLENKA, A., LEACH, S. D. & LOWY, A. M. 2007. Wnt/beta-catenin signaling is required for development of the exocrine pancreas. *Bmc Developmental Biology,* 7.

WELLS, J. M. & MELTON, D. A. 1999. Vertebrate endoderm development. *Annu Rev Cell Dev Biol,* 15**,** 393-410.

WETERINGS, E., VERKAIK, N. S., BRUGGENWIRTH, H. T., HOEIJMAKERS, J. H. J. & VAN GENT, D. C. 2010. The role of DNA dependent protein kinase in synapsis of DNA ends (vol 31, pg 7238, 2003). *Nucleic Acids Research,* 38**,** 3856-3856.

WIEDENHEFT, B., STERNBERG, S. H. & DOUDNA, J. A. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature,* 482**,** 331-8.

WILLIAMS, R. S., DODSON, G. E., LIMBO, O., YAMADA, Y., WILLIAMS, J. S., GUENTHER, G., CLASSEN, S., GLOVER, J. N. M., IWASAKI, H., RUSSELL, P. & TAINER, J. A. 2009. Nbs1 Flexibly Tethers Ctp1 and Mre11-Rad50 to Coordinate DNA Double-Strand Break Processing and Repair. *Cell,* 139**,** 87-99.

WONG, E. S., LE GUEZENNEC, X., DEMIDOV, O. N., MARSHALL, N. T., WANG, S. T., KRISHNAMURTHY, J., SHARPLESS, N. E., DUNN, N. R. & BULAVIN, D. V. 2009. p38MAPK controls expression of multiple cell cycle inhibitors and islet proliferation with advancing age. *Dev Cell,* 17**,** 142-9.

YANG, Y. S. & CHAN, L. 2016. Monogenic Diabetes: What It Teaches Us on the Common Forms of Type 1 and Type 2 Diabetes. *Endocrine Reviews,* 37**,** 190-222.

YAP, K. L., LI, S. D., MUNOZ-CABELLO, A. M., RAGUZ, S., ZENG, L., MUJTABA, S., GIL, J., WALSH, M. J. & ZHOU, M. M. 2010. Molecular Interplay of the Noncoding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a. *Molecular Cell,* 38**,** 662-674.

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., SLUKVIN, II & THOMSON, J. A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science,* 318**,** 1917-20.

YU, W. Q., GIUS, D., ONYANGO, P., MULDOON-JACOBS, K., KARP, J., FEINBERG, A. P. & CUI, H. M. 2008. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature,* 451**,** 202-U10.

YUMLU, S., BASHIR, S., STUMM, J. & KUHN, R. 2019. Efficient Gene Editing of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9. *Methods Mol Biol,* 1961**,** 137-151.

ZEGGINI, E. 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes (June, pg 1336, 2007). *Science,* 317**,** 1036-1036.

ZELENSKY, A., KANAAR, R. & WYMAN, C. 2014. Mediators of Homologous DNA Pairing. *Cold Spring Harbor Perspectives in Biology,* 6.

ZENG, N., YANG, K. T., BAYAN, J. A., HE, L., AGGARWAL, R., STILES, J. W., HOU, X., MEDINA, V., ABAD, D., PALIAN, B. M., AL-ABDULLAH, I., KANDEEL, F., JOHNSON, D. L. & STILES, B. L. 2013. PTEN controls beta-cell regeneration in aged mice by regulating cell cycle inhibitor p16ink4a. *Aging Cell,* 12**,** 1000-11.

ZHANG, D. H., JIANG, W., LIU, M., SUI, X., YIN, X. L., CHEN, S., SHI, Y. & DENG, H. K. 2009. Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells. *Cell Research,* 19**,** 429-438.

ZHOU, J. X., DHAWAN, S., FU, H., SNYDER, E., BOTTINO, R., KUNDU, S., KIM, S. K. & BHUSHAN, A. 2013. Combined modulation of polycomb and trithorax genes rejuvenates beta cell replication. *J Clin Invest,* 123**,** 4849-58.

ZHOU, Q., LAW, A. C., RAJAGOPAL, J., ANDERSON, W. J., GRAY, P. A. & MELTON, D. A. 2007. A multipotent progenitor domain guides pancreatic organogenesis. *Dev Cell,* 13**,** 103-14.

ZHOU, Q. & MELTON, D. A. 2018. Pancreas regeneration. *Nature,* 557**,** 351-358.

ZORN, A. M. & WELLS, J. M. 2009. Vertebrate endoderm development and organ formation. *Annu Rev Cell Dev Biol,* 25**,** 221-51.

Contributions

Alireza Shahryari performed CRISPR/Cas9 genome editing experiments (increasing gene editing efficiency of CRISPR/Cas9, large DNA deletion at the INK4 locus), shRNA experiments, generation of iPSC line and its characterization, differentiation of iPSC towards beta like cells, immunostaining, FACS experiment and analysis, and imaging and data analysis. Ingo Burtscher, Michael Sterr, Xianming Wang, Johanna Siehler and Weiwei Xu performed some bioinformatic analyses, some FACS experiments, and GSIS analysis.

Acknowledgements

First, I would like to thank all the people who were involved in the success of this dissertation! Without your help, I could not have accomplished this work! I would like to thank my PhD advisor, Prof. Heiko Lickert, for giving me the opportunity to prepare my dissertation in his lab. Thank you for providing an excellent scientific environment, for being a great supervisor and for all your support, ideas and enthusiasm! I have learned a lot during the past years and I'm very grateful for that! Moreover, I would like to thank Dr. Ingo Burtscher for teaching, supervising my project. Thank for all your great help and critical input. For the discussions, all the answers to countless questions and a great time in the stem cell research group! It was great to work with you and I have learned a lot from you! Special thanks go to Dr. Xianming Wang and Johanna Siehler, Dr. Michael Sterr, Melis Akguen, Dr. Weiwei Xu and Noel Moya for great help and support, and teamwork. I would like to thank the beta cell development and regeneration group, Dr Mostafa Bakhti and his team for the great teamwork and support, and a great time in the lab and office! The beta cell replacement group, Dr. Katharina Scheibner, Nicole Rogers, and Eunike Setyono, for the great friendship and support, and the good times we had in the last years! I want to thank all the IDR members, and in particular Donna Thomson, Kerstin Diemer, Emily Violette Baumgart and Anita Hofmann for a great time and support! Moreover, I want to thank Dr. Mara Catani and Dr. Zahra Nazari who helped me to correct this thesis! At last, I want to thank my family for their support! I want to thank my parents for all their help and support during my entire studies.

Publications

1. **Shahryari A,** Moya N, Siehler J, Burtscher I, Lickert H. Increasing Gene Editing Efficiency for CRISPR/Cas9 by small RNAs in Pluripotent Stem Cells. *CRISPR Journal*, 2021 Aug;4(4):491-501.

2. **Shahryari A***, Burtscher I, Nazari Z, Lickert H*. Engineering Gene Therapy: Advances and Barriers. *Advanced Therapeutics*, 2021 Sep;4(9):2100040. * Corresponding author

3. **Shahryari A***, Saghaeian Jazi M, Mohammadi S, Razavi Nikoo H, Nazari Z, Hosseini ES, Burtscher I, Mowla SJ*, Lickert H*. Development and Clinical Translation of Approved Gene Therapy Products for Genetic Disorders. *Front Genet.* 2019 Sep 25;10:868. * Corresponding author

4. **Shahryari A**, Moya N, Siehler J, Wang X, Karolina Blöchinger A, Burtscher I, Bakhti M, Mowla SJ, Lickert H. Generation of a human iPSC line harboring a biallelic large deletion at the INK4 locus (HMGUi001-A-5). *Stem Cell Res.* 2020 Jul 25;47:101927.

5. **Shahryari A***, Nazari Z, Saghaeian Jazi M, Hashemi-Shahraki F, Wißmiller K, Xu W, Burtscher I, Lickert H*, Pharmacological Aspects of Clinically Approved Gene Therapy Drugs and Products, (Chapter) Reference Module in Biomedical Sciences, Book Title: Comprehensive Pharmacology, Elsevier, 2022 *Corresponding author

6. Moya N, **Shahryari A**, Burtscher I, Beckenbauer J, Bakhti M, Lickert H. Generation of a homozygous ARX nuclear CFP (ARXnCFP/nCFP) reporter human iPSC line (HMGUi001-A-4). *Stem Cell Res.* 2020 Jul;46:101874.

7. Siehler J, Blöchinger AK, Akgün M, Wang X, **Shahryari A**, Geerlof A, Lickert H, Burtscher I. Generation of a heterozygous C-peptide-mCherry reporter human iPSC line (HMGUi001-A-8). *Stem Cell Res*. 2020 Dec 16;50:102126.

8. Blöchinger AK, Siehler J, Wißmiller K, **Shahryari A**, Burtscher I, Lickert H. Generation of an INSULIN-H2B-Cherry reporter human iPSC line. *Stem Cell Res.* 2020 May;45:101797.