Generation and characterization of a double reporter cell line for following maturation of hESC-derived β-cells

Kufrin, Vida

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:277341

University of Zagreb
Faculty of Science
Department of Biology

Vida Kufrin

Generation and characterization of a double reporter cell line for following maturation of hESC-derived β-cells

Master thesis

Zagreb, 2021.

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Vida Kufrin

Uspostavljanje i karakterizacija stanične linije sa sustavom dva reportera kao modela za praćenje sazrijevanja β-stanica izvedenih iz embrionalnih matičnih stanica

Diplomski rad

Zagreb, 2021.

This thesis was conducted in the laboratory of Dr. Anthony Gavalas at the Paul Langerhans Institute Dresden of the Helmholtz Zentrum München, TU Dresden, under the supervision of Dr. Anthony Gavalas, Prof. and co-supervision of Dr. Inga Urlić, Assoc. Prof. The thesis was submitted for an evaluation to the Department of Biology, Faculty of Science, University of Zagreb in order to acquire the academic title of Master of Molecular Biology.

Acknowledgments

First and foremost, I would like to express my profound gratitude to my supervisor, Dr. Anthony Gavalas for giving me this amazing opportunity to conduct my thesis in his lab and for his guidance and continuous support.

A special thank you goes to my labmates Elisa Zanfrini and Luka Jarc. Thank you for treating me as an equal and patiently training me, helping me with experiments and writing, motivating me, and always having my back. You showed me what research is all about and I could not have imagined having better supervisors than you. For all of that, to both of you, I am forever grateful.

I would also like to thank Dr. Inga Urlić for her co-supervision, all the valuable advice during the writing of this thesis and for being a huge help with the bureaucracy.

I would like to extend my thanks to the rest of the Gavalas group: Ivan, Mandy, Manuj and Zeina for making me look forward to coming into the lab every morning. I really enjoyed every lunch and coffee we had together. Also, I am grateful to Katja for solving all administration required for prolonging my stay in Dresden.

To all of my friends at home and in Germany, thank you for always being a listening ear, believing in me when I did not and for pushing me to be my best self. I would especially like to point out my molecool girls, Katarina, Lucija and Marija, you made the last 6 years truly memorable and I could not have done it without you.

Lastly, I would sincerely like to thank my wonderful family for their unfailing support and encouragement during my studies. You have always been and will be my greatest motivation.

BASIC DOCUMENTATION CARD

University of Zagreb
Faculty of Science
Department of Biology

Master Thesis

Generation and characterization of a double reporter cell line for following maturation of hESC-derived β-cells

Vida Kufrin

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Diabetes mellitus is a group of metabolic disorders primarily affecting pancreatic β-cells. Due to the limitations of existing therapy options for insulin-deficient diabetes, great efforts are put into developing cell-based replacement therapies utilizing human pluripotent stem cells. However, existing differentiation protocols, based on developmental principles *in vivo*, often result in the generation of cells that are not fully mature and functional. To improve and optimize existing protocols for hPSC reprogramming, we developed a novel H1 *INS-eGFP/MAFA-mCherry* double reporter cell line using CRISPR-Cas9 technology. Herein, detailed characterization and assessment of differentiation potential of novel cell line is reported. H1 *INS-eGFP/MAFA-mCherry* reporter line can reproducibly and efficiently differentiate to form β-cells while preserving pancreatic identity likewise paternal unmodified H1 cell line. Furthermore, H1 *INS-eGFP/MAFA-mCherry* derived β-cells express functionality and maturation related genes and can respond to glucose challenges *in vitro*. Also, both GFP and mCherry integrated reporter genes are fully functional in novel cell line. Reported findings led to the validation of a novel reporter cell line that will serve as a valuable tool in unraveling novel strategies to achieve a high-efficiency differentiation that presents one more step required for the translation of this approach to the clinic setting.

(48 pages, 13 figures, 5 tables, 90 references, original in: English)

Thesis is deposited in Central Biological Library.

Keywords: pancreas, stem cells, hESC, beta cells, diabetes, differentiation, GFP

Supervisor: Dr. Anthony Gavalas, Prof.
Co-supervisor: Dr. Inga Urlić, Assoc. Prof.
Reviewers: Dr. Inga Urlić, Assoc. Prof., Dr. Romana Gračan, Asst. Prof., Dr. Domagoj Đikić, Prof.

Thesis accepted: September 15th, 2021

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Diplomski rad

Uspostavljanje i karakterizacija stanične linije sa sustavom dva reportera kao modela za praćenje sazrijevanja β-stanica izvedenih iz embrionalnih matičnih stanica

Vida Kufrin

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Dijabetes je metabolički poremećaj koji primarno pogađa β-stanice gušterače. Budući da su postojeće terapijske opcije za pacijente sa smanjenom razinom inzulina ograničene, velike nade polažu se u stanične terapije temeljene na humanim pluripotentnim matičnim stanicama. Međutim, postojeći protokoli diferencijacije, temeljeni na razvojnim principima *in vivo,* često rezultiraju generacijom β-stanica koje nisu u potpunosti zrele i funkcionalne. Kako bi se protokoli diferencijacije unaprijedili i optimizirali, razvili smo novu dvostruku reportersku liniju H1 *INS-eGFP/MAFA-mCherry* koristeći tehnologiju CRISPR-Cas9. U ovom radu prikazana je detaljna karakterizacija i procjena diferencijacijskog potencijala nove stanične linije. Linija H1 *INS-eGFP/MAFA-mCherry* ima sposobnost efikasne i reproducibilne diferencijacije u β-stanice pritom zadržavajući identitet usporediv s nemodificiranom roditeljskom staničnom linijom H1. Nadalje, H1 *INS-eGFP/MAFA-mCherry* izvedene β-stanice eksprimiraju gene važne za funkcionalnost i sazrijevanje te imaju sposobnost odgovora na promjene u koncentraciji glukoze *in vitro*. Također, ugrađeni reporterski geni GFP i mCherry su potpuno stabilni i funkcionalni u ovoj staničnoj liniji. Provedeno istraživanje potvrdilo je ispravnost nove stanične linije koja će biti korištena kao vrijedan alat u dizajnu novih strategija za efikasnu diferencijaciju te predstavljati jedan korak više ka kliničkim ispitivanjima novih terapija za dijabetes.

(48 stranica, 13 slika, 5 tablica, 90 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u Središnjoj biološkoj knjižnici.

Ključne riječi: gušterača, matične stanice, hESC, beta stanice, dijabetes, diferencijacija, GFP

Voditelj: Prof. dr. sc. Anthony Gavalas
Suvoditelj: Izv. prof. dr. sc. Inga Urlić
Ocjenitelji: Izv. prof. dr. sc. Inga Urlić, Doc. dr. sc. Romana Gračan, Prof. dr. sc. Domagoj Đikić

Rad prihvaćen: 15.9.2021.

Table of contents

Abbreviations

AFP: Alpha-fetoprotein

BSA: Bovine albumin

CDX2: Caudal type homeobox 2

DE: Definitive endoderm

DM: Diabetes mellitus

DMEM/F-12: Dulbecco's modified eagle medium/nutrient mixture F-12

DPBS: Dulbecco's phosphate-buffered saline

EDTA: Ethylenediaminetetraacetic acid

FOXA2: Forkhead box protein A2

GCG: Glucagon

GFP: Green fluorescent protein

GK: Glucokinase

HHEX: Hematopoietically expressed homeobox

hESC: Human embryonic stem cell

INS: Insulin

MAFA: V-maf musculoaponeurotic fibrosarcoma oncogene homolog A

mCherry: Monomeric cherry

NEUROD1: Neurogenic differentiation 1

NGN3: Neurogenin 3

NKX6.1: NK6 homeobox 1

OCT4: Octamer-binding transcription factor 4

PDX1: Pancreatic and duodenal homeobox 1

PEP: Pancreatic endocrine precursors

PFA: Paraformaldehyde

PF: Posterior foregut

PGT: Primitive gut tube

PP: Pancreatic progenitors

PTF1A: Pancreas associated transcription factor 1a

ROCKi: Rho kinase inhibitor

RT: Room temperature

SOX2: SRY (sex determining region Y)-box 2

SOX9: SRY-box transcription factor 9

SOX17: SRY-box transcription factor 17

SST: Somatostatin

SUR1: Sulfonylurea receptor 1

TBP: TATA-binding protein

TF: Transcription factor

T1DM: Type 1 diabetes mellitus

T2DM: Type 2 diabetes mellitus

wpw: Well plate well

ZnT8: Zinc transporter 8

1 Introduction

1.1 Pancreas development

The pancreas is an organ with both exocrine and endocrine function that has a central role in nutrient metabolism. The exocrine pancreas is composed of acinar and ductal cells responsible for the production and secretion of enzymes important for digestion into the duodenum. On the other hand, the endocrine pancreas is responsible for glucose homeostasis regulation and is comprised of five distinct cell types, found in cell clusters called islets of Langerhans (Shih et al., 2013). During the development, the pancreas originates from the definitive endoderm (DE) as it forms from the duodenal part of the foregut. Pancreatogenesis is a fairly complex process coordinated by multiple master regulators and well-synchronized signaling pathways and is divided into the stages referred to as primary, secondary and tertiary transition. The primary transition is characterized by the emergence of dorsal and ventral buds followed by proliferation of multipotent progenitors, while the secondary transition is characterized by ductal tubulogenesis as well as exocrine and endocrine differentiation. Finally, the tertiary transition is identified by migration of differentiated cells and formation of compact islets of Langerhans (Gittes, 2009, Pictet et al., 1972). Understanding embryonic development of the pancreas, with an emphasis on cell fate decisions, is inevitably important for further advancements in designing cell replacement therapies (Shih et al., 2013).

1.1.1 Primary transition

Pancreas development starts with the emergence of two opposing epithelial buds, ventral and dorsal bud, elongating on the opposite sides of the foregut endoderm. Bud-forming cells are defined by inhibition of the sonic hedgehog pathway and the presence of retinoic acid that acts as an essential endodermal patterning signal (Chen et al., 2004). During gut rotation, the definitive pancreas is formed as a result of bud fusion. In this phase of development, cells are still undifferentiated and extensive proliferation leads to a morphology change in a process known as branching morphogenesis. The ventral bud eventually gives rise primarily to a part of pancreas head, while the dorsal bud gives rise to the rest of the pancreas (Shih et al., 2013). Among the earliest transcription factors (TFs) involved in determining pancreatic fate are PDX1, PTF1A, and SOX9 (Ahlgren et al., 1996, Kawaguchi et al., 2002, Seymour et al., 2007). Furthermore, mesenchymal signals are of high importance in this stage of development, with the FGF signaling pathway taking the lead. FGF10 acts as a pro-proliferative signal and is a regulator of PTF1A and SOX9 expression (Seymour et al., 2012). At the end of the primary transition, the dominant cell population are non-committed multipotent progenitor cells or so-called pancreatic progenitors (PPs) (Shih et al., 2013).

1.1.2 Secondary transition

The differentiation of PP cells towards acinar, ductal and endocrine lineage, along with proliferation and further branching of the pancreatic epithelium into organized tubular networks, is called secondary transition. Firstly, the tip and trunk domains are separated due to the segregated expression of master regulators and cross-repression of one another. The tip domain is marked by the expression of PTF1A, while the trunk domain is identified by NKX6.1 and SOX9 expression. The trunk domain eventually gives rise to endocrine and ductal cells, while the tip domain is restricted towards acinar lineage (Shih et al., 2013). It has been shown that Notch signaling plays a pivotal role in promoting trunk identity. However, it is not known how Notch signaling is regulated during this process (Afelik et al., 2012).

Specification to acinar, ductal or endocrine lineage is a multi-step process still not known in detail. As already discussed, acinar cells arise from precursor cells in the tip domain under PTF1A regulation through an autoactivation loop with acinar-specific genes enabling induction and maintenance of acinar phenotype (Kawaguchi et al., 2002). In the trunk domain cells are bipotent and give rise to either ductal or endocrine cells. The master regulator shifting cells towards the endocrine fate is TF NGN3. Thus, trunk cells that express NGN3 are destined to be endocrine, while the ones that do not activate NGN3 eventually acquire ductal fate (Johansson et al., 2007). Ductal cell determination has not yet been investigated in detail, but it is known that expression of several TFs important for trunk determination, including SOX9, later becomes restricted only to ducts (Seymour et al., 2007). Endocrine differentiation studies mostly focus on the endocrine lineage marker NGN3 and its regulation. It is thought that Notch signaling plays a key role in NGN3 regulation, with high Notch signaling resulting in NGN3 repression. Thus, to initiate endocrine differentiation, bipotent progenitors have to escape Notch signaling. By the end of the secondary transition, the pancreas acquires the organization of a mature organ and endocrine precursors stop arising from the ductal epithelium (Shih et al., 2013).

1.1.3 Tertiary transition

In the process of tertiary transition, $NGN3^{+}$ endocrine precursors exit the cell cycle and delaminate from the trunk epithelium. Delamination results in their migration to pancreatic stromal tissue where they later form clusters (Gouzi et al., 2011). This process is temporally controlled by the duration of exposure to the Notch signaling pathway. Endocrine progenitors differentiate to 5 different endocrine cell types: glucagon-producing α-cells, insulin-producing β-cells, somatostatin-producing δ-cells, pancreatic polypeptide–producing PP-cells, and ghrelin producing ε-cells (Shih et al., 2013). Specifically, fate decision between α- and β-cell precursors relies on mutual repression between opposite lineage determinants, with PDX1, NKX6.1 and PAX4 expression being essential for β-cell development, while ARX expression specifies α-cell identity (Collombat et al., 2003).

1.1.4 Maturation and maintenance of β-cells

Postnatally, β-cell number, phenotype and function are maintained by establishing novel transcription networks maintaining β-cell specific gene expression. Although regulatory pathways required for β-cell development *in utero* are known, mechanisms that maintain the differentiated state of adult β-cells are yet to be unraveled (Szabat et al., 2012). However, some TFs important in development, like PDX1, play a significant role in adulthood as well (Boj et al., 2001). Besides PDX1, the TFs NEUROD1 and MAFA are highly important since they serve as induction signals for insulin expression (Nishimura et al., 2006). It is important to note that prenatal development occurs without MAFA, and thus it is characterized as a gene that regulates mature cellular functions (Murtaugh, 2007).

The β-cell development, illustrated in Figure 1.1, is remarkably complex with multiple fate choices and lineage commitments that have to be made. The whole process starts from cells of the foregut that assume a pancreatic identity giving rise to PPs, to the bipotent trunk cells and the endocrine fate adoption directed by NGN3 expression, finally leading to β-cell determination controlled by Notch signaling (Murtaugh, 2007). The existing knowledge enabled the development of *in vitro* protocols that direct the differentiation of human embryonic stem cells (hESC) towards β-cell fate. Further research in the field, with an emphasis on elucidating endocrine specification and maturation processes, would present a tremendous step forward in further optimization of existing protocols.

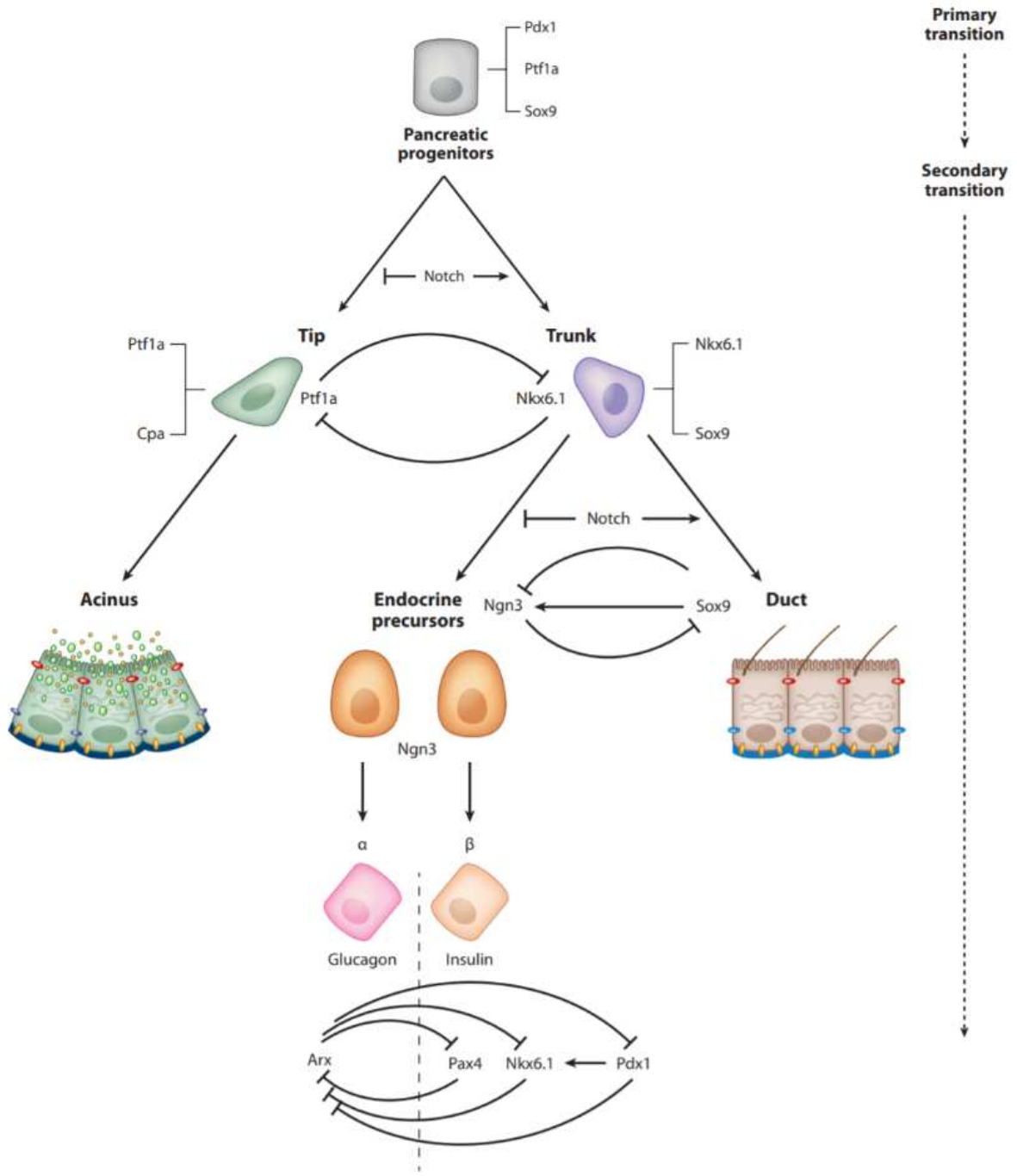

Figure 1.1 Lineage specification during pancreatogenesis. Starting from pancreatic progenitors maintained by expression of Pdx1, Sox9 and Ptf1a cells can adopt tip or trunk identity under Notch regulation. Tip cells later adopt acinar phenotype, while bipotent trunk cells are destined to shift to either endocrine or ductal lineage. Endocrine precursors later differentiate and form five distinct cell types, with α- or β-cell identity being determined by cross-repression of Arx and Pdx1 (Adapted from Shih et al., 2013).

1.2 Diabetes mellitus

1.2.1 Definition, classification and etiology

Diabetes mellitus (DM) is a group of metabolic disorders characterized by chronic hyperglycemia that is a result of a defect in insulin action, insulin secretion, or both (American Diabetes Association, 2014). Continuously elevated blood glucose has a severe impact on different organs and leads to long-term vascular complications including neuropathy, nephropathy and retinopathy or, in advanced cases, even stroke and myocardial infarction (Forbes and Cooper, 2013). Apart from vascular complications, DM has also been associated with an elevated risk for dementia (Gudala et al., 2013) and depression (Nouwen et al., 2011). According to American Diabetes Association, DM is classified into 3 categories: type 1 diabetes mellitus (T1DM), type 2 diabetes mellitus (T2DM) and a third class that includes various specific subtypes of diabetes such as monogenic diabetes, gestational diabetes and secondary diabetes (Canivelli and Gomis, 2014).

T1DM is characterized by autoimmune destruction of β-cells in the Langerhans islets resulting in abolished insulin secretion consequently leading to deregulation of rigorous glycemic control (Devendra et al., 2004). Autoimmunity is mainly T-cell mediated, even though humoral immunity also plays a significant role in pathogenesis. The presence of autoantibodies to β-cell specific antigens, which can be detected in serum before disease onset, is a hallmark of T1DM. However, the role of autoantibodies in pathogenesis is still not clear (Kharroubi and Darwish, 2015). In the past, T1DM has been referred to as juvenile diabetes, but several studies showed that incidence is comparable among adults (Mølbak et al., 1994). It has been shown that T1DM shows a strong HLA association, with linkage to gene variants of DR and DQ alleles resulting in either higher susceptibility or protection against developing T1DM. Apart from genetic background, several environmental factors including viral infection and early infant nutrition have been associated with T1DM (Devendra et al., 2004).

On the other hand, T2DM does not have an autoimmune origin but it is characterized by the unresponsiveness of tissues to insulin, called insulin resistance. Insulin resistance leads to an increased demand for insulin that cannot be met by β-cells which causes β-cell stress and gradually β-cell destruction (Kharroubi and Darwish, 2015). The prevalence of T2DM is higher in adults, but it has been increasing in the younger population mainly due to obesity (Rosenbloom et al., 2009). Even though more than 130 genetic variants have been associated with T2DM so far, which proves its complex genetic etiology, the specifics are still unclear. With an exception of obesity that is strongly linked with T2DM, there are no other environmental factors directly correlating with T2DM defined up to date (Skyler et al., 2017).

As of 2019, the International Diabetes Federation estimated that 463 million people currently live with some form of diabetes, with 90% of cases being T2D and less than 10% accounting for T1DM. Compared to 2009, when 285 million people had diabetes, it is clear that prevalence is rapidly growing.

Furthermore, by 2045, projections are that number will increase by more than 50%. (Saeedi et al., 2019). Apart from that, in 2017, health expenditure on diabetes was a striking 727 billion dollars, making it a huge economic burden (Bommer et al., 2018). Looking at diabetes in numbers, it is clear that it has emerged as a global epidemic that has to be addressed and requires immediate action and development of strategies to tackle it with new therapy and treatment options.

1.2.2 Existing treatment and therapy options

Due to differences in etiology and pathophysiology of T1DM and T2DM, therapies and treatments are also different. Since T1DM leads to complete loss of insulin, the current standard of care is life-long exogenous insulin administration. The discovery of insulin therapeutic potential was first described a century ago (Banting and Best, 1921), when animal pancreas extracts were used to establish normoglycemia in diabetic patients. Nowadays, there is a whole pallet of humanized insulin analogs, both rapid-acting and long-acting, with enhanced pharmaceutical properties available on the market (Evans et al., 2011). However, despite all advances made in the synthesis and administration of insulin replacement therapy, it is still limiting and cannot fully replicate the function of endogenous β-cells (Pathak et al., 2019).

An alternative approach that could ensure insulin independence in insulin-deficient patients is allogeneic transplantation of the whole pancreas or islets of Langerhans. However, due to the scarcity of donors, extensive cost and HLA compatibility requirements, for now, transplantation can be an option only for patients presenting extensive glycemic lability. The biggest risks of this approach are the possible rejection of the graft and poor vascularization post-transplantation. Another downfall is life-long immunosuppression which is more intensive compared to other transplants due to strong immunogenicity (Rickels and Robertson, 2019). Even though it has been shown that immunosuppression promotes long-term insulin independence, it is also tightly related to the increased risk of *de novo* malignancies (Tomimaru et al., 2015).

As mentioned above, T2DM differs from T1DM by the age of manifestation and factors triggering it. Based on that, useful treatment strategies are different and rely mostly on lifestyle change and pharmacotherapy in establishing glycemic control mainly using metformin as first-line therapy. However, after years of hypersecretion of insulin due to insulin resistance and inadequate glycemic control, β-cell mass can deplete for 65%, leading to demand for exogenous insulin or even transplantation (Butler et al., 2003). All stated calls for fast interventions in developing new strategies for reaching normoglycemia and insulin independence and puts a spotlight on cell replacement therapies using alternative cell sources with an emphasis on human pluripotent stem cells.

1.3 Human pluripotent stem cells (hPSC) and application in β-cell regeneration

Even though exogenous insulin administration and glucose monitoring are the golden standard in the present-day treatment of insulin-dependent diabetes, this approach represents a constant burden for patients and eventually leads to comorbidities and lower quality of life. Thus, one of the major challenges of regenerative medicine is to provide a solution for diabetic patients in the form of β-cell regeneration (Melton, 2020). By lineage tracing, it has been demonstrated that adult β-cells are mainly formed by the self-duplication of existing cells (Dor et al., 2004). Therefore, one of the initial attempts for β-cell regeneration was to simulate endogenous β-cell proliferation, but this approach was not efficient. Transdifferentiation of α or ductal cells emerged as an alternative option, but the requirement of ectopic expression of multiple TFs could not be met *in vivo* (Zhou et al., 2008). Because of that, directed differentiation of hPSC *in vitro* based on developmental principles serves as a promising approach.

1.3.1 hPSC

Human pluripotent stem cells include 2 cell types: human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs). They hold two main characteristics, being unlimited self-renewal capability *in vitro* along with pluripotency preservation. In other words, they retain unlimited differentiation potential, meaning that they can become cells from any germ layer (Zhu and Huangfu, 2013). hESC are extracted from the inner cell mass of the human blastocyst after *in vitro* fertilization. Multiple hESC cell lines, such as the H1 cell line, with normal karyotype, unlimited freezing/thawing ability, undifferentiated morphology, and a prospect to form all three germ layers are commercially available for more than 20 years. In the β-cell context, apart from cell-replacement therapies, β-cells derived by directed differentiation could be also very useful in basic developmental research and drug discovery (Thomson et al., 1998). Taking into consideration all ethical concerns arising from using hESC, along with rejection after transplantation due to alloimmunity, it may be better to generate pluripotent cells straight from patients' somatic cells. That was the idea behind the notable discovery of iPSCs, for which Yamanaka got the Nobel prize in 2012. In his study, Yamanaka identified and validated 4 distinct TFs Oct4, Sox2, c-Myc and Klf4 important for the maintenance of ES cell identity that were sufficient to induce pluripotency in adult fibroblasts cultures (Takashi and Yamanaka, 2006). Even though hiPSC share many properties with hESC, including pluripotency, self-renewal, morphology and genetic profile, their full characteristics and mechanistic details leading to reprogramming are far from being complete, making hESCs still a preferred option for regenerative purposes (Puri & Nagy, 2012).

1.3.2 Differentiation of hPSC to β-cells

The idea of using hESCs as a replacement for β-cells was first introduced in the 90s (Keller, 1995). The first breakthrough in developing directed differentiation protocols *in vitro* following developmental principles *in vivo,* along with the identification of main signaling pathways that drive differentiation, came 10 years later in the form of a 5-stage protocol resulting in hormone-expressing endocrine cells. However, these cells were highly polyhormonal and minimally responded to glucose, similarly to immature fetal β-cells. Also, expression of maturation marker MAFA was not detected. At the end of the differentiation, only 7% of cells accounted for insulin-positive ones (D'Amour et al., 2006). Two years later, the improved D'Amour protocol was tested *in vivo* by transplantation of endocrine precursors into rodents to assess the ability for further differentiation and maturation. Within 4 months, transplanted cells developed into endocrine cells that both morphologically and functionally resembled pancreatic islets and protected mice from β-cell toxin streptozotocin (Kroon et al., 2008). The success and reproducibility of this protocol served as a solid ground for further improvements and continuation of the research in finding the most optimal stepwise differentiation protocol.

However, the biggest challenge proved to be the production of cells that are functional, meaning that they can respond to multiple glucose challenges by regulated and synchronized insulin secretion without the need for further differentiation and maturation after transplantation (Melton, 2020). This achievement was independently demonstrated in 2014 by Melton and Kieffer labs. Melton's protocol contained 6 stages with fine-tuning of differentiation signals, with an emphasis on Notch inhibition and T3 stimulation in the final differentiation stages. The major technical difference compared to the D'Amour protocol is that this protocol was a suspension-based culture system. This protocol was able to reproducibly generate cells similar to adult β-cells both in size and in function without genetic modifications. Additionally, cell contact with other cell types (e.g., endothelial cells) was not needed since the signals were successfully replaced *in vitro* (Pagliuca et al., 2014). On the other hand, Kieffer's lab described 7-stage protocol combining planar 2D culture with 3D air-liquid interface culture, namely the Rezania protocol (Fig. 1.2). This was the first protocol able to generate around 50% of insulin-positive cells with a stable expression of TFs such as PDX1 and NKX6.1 as well as reduction of the percentage of polyhormonal cells. Furthermore, it was the first protocol that reported MAFA induction at levels similar to that of human islets. However, despite their ability to rapidly reverse diabetes 40 days upon transplantation, derived β-cells were functionally immature relative to human islets (Rezania et al., 2014).

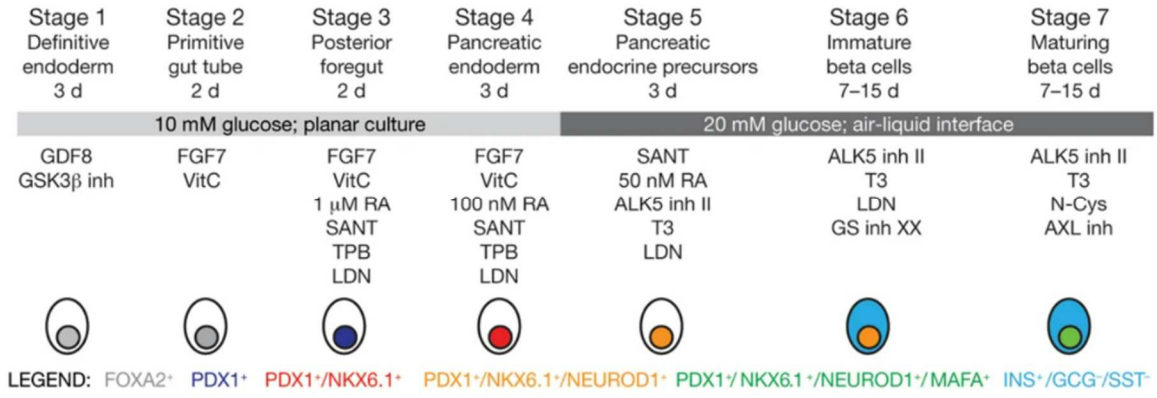

Figure 1.2 Rezania 7-stage differentiation protocol. Protocol combines planar culture and air-liquid interface and is directed by important growth factors and small molecules. Key molecular markers for each stage are illustrated below (Rezania et al., 2014).

The third most notable approach came from the Hebrok lab. The novelty introduced was endocrine cell clustering by isolating and reagreggating immature β-cells. This proved to be a good alternative that leads to a higher maturation rate and better similarity to endogenous counterparts. The protocol has mostly been done in suspension, with a novelty of Aggrewell plates being used for reaggregation. However, since immature β-cells are recognized using GFP reporter, this approach cannot be translated to a clinical setting (Nair et al., 2019).

It is important to note that each of the listed differentiation protocols requires strict monitoring for every respective stage to evaluate the efficiency and determine phenotype. Monitoring is usually done by immunofluorescent analysis, flow cytometry, and RT-qPCR (Brovkina and Dashin, 2020). All of the stated studies and differentiation protocols demonstrated the great possibility of using undifferentiated pluripotent stem cells to produce functional β-cells *in vitro.* Further improvements and optimizations have to be made to reduce the cost of the process and increase efficiency of differentiation. Also, the protocol has to be more reproducible and lead to the generation of more mature and functional cells (Melton, 2020), as well as removing unwanted undifferentiated cells that may be dangerous and form teratomas after transplantation (Hentze et al., 2009). Apart from cell-based therapies and possible clinical translation, hPSC-derived β-cells have a great potential to serve as a basic research tool to further unravel pathways important for function, maturation and longevity difficult to study *in vivo* (Melton, 2020).

This future cell-based replacement therapy utilizing hPSC is currently more restricted to T1DM. Several major challenges that will have to be addressed in the clinic are choosing the site for transplantation of islets and determining the most optimal cell number. Moreover, the longevity of transplanted islets will probably be one of the issues that will have to be solved due to both auto- and alloimmune rejection (Melton, 2020). However, possible solutions for both β-cell preservation and immune response evasion have already been proposed. Encapsulation using biocompatible membranes that permits diffusion of oxygen and nutrients while protecting against larger molecules e.g., antibodies

is the most promising approach (Deasi and Shea, 2017), as well as the approach of biological intervention reflected in modifying patients' immune system to evade the immune response by genetic modification of SC-islets prior transplantation (Cai et al., 2020)

To sum up, due to their unlimited differentiation capacity and self-renewal, hPSC are an attractive tool for the production of β-cells in vitro. Nowadays, the most common approach is using stepwise differentiation protocols directing cells to the DE stage, followed by foregut development, PPs, endocrine progenitors and finally, β-cells (Melton, 2020). However, even though these cells express key β-cell markers comparable to cadaveric β-cells (Rezania et al., 2014) and have the ability to reverse diabetes *in vivo* (Kroon et al., 2008, Rezania et al., 2012) there is still work left to do to improve maturation that would lead to increased functionality and hopefully translation of this approach to the clinic (Melton, 2020).

1.4 CRISPR-Cas9 technology and reporter cell lines

The common approach for the generation of reporter lines is based on 'knocking-in' of a reporter construct in the genome. That is done by creating a site-specific double-stranded break coupled with the introduction of the reporter sequence used for the repair. Strategies created for that purpose include TALEN, ZFNs and CRISPR-Cas9. Due to its simplicity and lower error rate, CRISPR-Cas9 emerged as the most popular approach for reporter line generation (Gaj et al., 2013).

Clustered regularly interspaced short palindromic repeat (CRISPR) DNA sequences are repeat elements naturally present in more than 40% of sequenced bacterial strains. In prokaryotes, CRISPR elements are crucial for providing acquired immunity as the sequence itself is derived from DNA fragments of bacteriophages that previously infected bacteria (Adli, 2018). Unlike other typical tandem repeat sequences in the genome, the CRISPR locus is special for several reasons. Firstly, repeating sequences can be easily distinguished due to separation with spacer sequences belonging to bacteriophages and mobile genetic elements (Mojica et al., 2000). Secondly, the CRISPR locus is tightly connected with well-conserved CRISPR-associated (Cas) enzymes (Jansen et al., 2002), with Cas9 being the one with catalytic activity in *S. thermophilus*. In particular, Cas proteins are nucleases that induce double-stranded brakes (DSB) in the DNA (Garneau et al., 2010). Thirdly, Cas enzyme activity is tightly regulated via guidance by short transcripts of spacer sequences called CRISPR RNAs (Brouns et al., 2008). Interestingly, this asset can be reprogrammed for targeting a specific DNA sequence using chimeric single-guide RNAs (sgRNAs) (Jinek et al., 2012). DSBs introduced in a Cas9 mediate fashion are repaired either by non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ is used for knock-out strategy, whereas HDR is the preferred option for knock-in strategy due to the lower error rate. Finally, it has been demonstrated that the CRISPR system is transferable between bacterial strains (Sapranauskas et al., 2011). All these independent discoveries lead the CRISPR system to evolve into a powerful genome-editing tool allowing us to modify genes in eukaryotic cells *in vivo* (Adli, 2018).

1.4.1 Reporter cell lines

The biggest advantages of CRISPR-Cas9 technology are its flexibility, simplicity and precision in a way that any location in the genome can be specifically targeted by designing complementary sgRNAs. Thus, this technology is extensively used in research, with the generation of reporter lines being one of its main applications by enabling targeted insertion of a reporter gene. Reporter genes are genes encoding proteins that are not naturally present in an organism of interest and therefore can be easily distinguished from endogenous proteins (Jurgielewicz et al., 2017). Due to their practicality, nontoxicity and relatively easy handling, many cell lines containing reporter systems have been generated for various purposes. Reporter lines are closely connected with various imaging technologies enabling visualization of

molecular processes and lineage tracing both *in vivo* and *in vitro.* In other words, reporter integration can be useful to identify when and in which cells a certain gene of interest is expressed or to understand the protein functionality (Li et al., 2018).

The usual approach for reporter line generation consists of proper sgRNA and vector design, transfection and targeting justification, followed by detailed characterization. Briefly, as shown in Figure 1.3, to generate reporter cell lines by using the CRISPR-Cas9 system, cells are transfected with the Cas9, sgRNAs and a donor vector carrying the 3´and the 5´ homology arms flanking the insertion cassette with the reporter sequence. sgRNA target site should be as close as to knock-in location to reduce off-targeting (Zhong et al., 2020). Cas9 nuclease, guided by sgRNAs, induces site-specific DSB on the target site in the genome. DSB is then repaired by the HDR mechanism, using donor vector as a repair template, leading to the integration of the reporter gene downstream of the gene of interest. The integrated reporter can be directly connected and translated with the protein of interest as a fusion protein or it can be linked by the T2A sequence. The T2A self-cleaving sequence integration allows the generation of a bi-cistronic reporter cassette. In other words, the reporter is co-transcribed with the gene of interest since it replaces STOP codon. Later, that asset enables co-translational cleavage of the gene of interest and the reporter, beneficial because both copies of the endogenous gene stay intact. Also, the gene of interest and reporter are present in equimolar concentrations (Blöchinger et al., 2020). After HDR, the success of the process is usually verified by sequencing and PCR. Finally, if the reporter is integrated an appropriate functional characterization of the novel cell line is needed (Zhong et al., 2020).

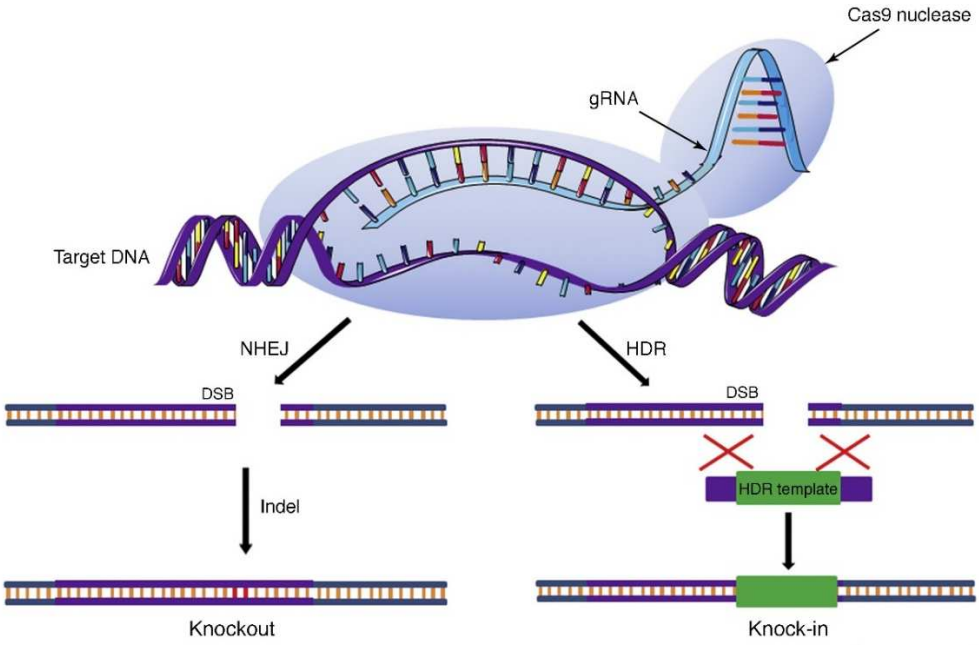

Figure 1.3 CRISPR-Cas9 mediated genome editing. sgRNA guides Cas9 nuclease to the target site to induce DSB. Subsequently, DSB is repaired via one of the two illustrated mechanisms. NHEJ repair leads to Indel mutations, while HDR pathway uses a template for repair and is used for knock-in strategies and reporter line generation (Liu et al., 2019)**.**

1.4.2 GFP and mCherry reporter genes

The green fluorescent protein (GFP) and the red fluorescent protein (RFP) are the most widespread and frequently used fluorescent reporter genes. GFP is a protein composed of 238 amino acids exhibiting green fluorescence when excited with light in the blue range. It was firstly isolated 60 years ago from the jellyfish *Aequorea victoria* (Shimomura et al., 1962) and a few years after the whole range of its variants, including enhanced green (eGFP), have been successfully developed and applied in research (Li et al., 2018). Nowadays, GFP and its variants are the best-established gene expression markers used in the whole range of model organisms and *in vitro* models. However, GFP variants have very narrow emission spectra and it is very challenging to differentiate them even with modern imaging systems when used together (Feng et al., 2000). For this reason, to obtain a double reporter system, it is necessary to choose a reporter that is spectrally distinct from GFP to enable easier identification while imaging. Thus, the RFP variants, consisting of several proteins emitting orange, red and far-red fluorescence, have been established by chromophore transformation (Subach & Verkhusha, 2012). RFPs can be further divided into groups of multimeric and monomeric proteins. Monomeric cherry (mCherry) is the most prominent example of monomeric RFPs. The advantages of mCherry over other RFPs are its high photostability and greater brightness, leading to its frequent usage (Li et al., 2018).

2 Objectives

The main objective of this study was to determine if the newly generated H1 *INS-eGFP/MAFA-mCherry* (*INS-MAFA*) double reporter line can effectively and reproducibly form functional β-cells *in vitro.* We hypothesized that this cell line will follow the same differentiation patterns as observed in the paternal H1 line used for its derivation. To test our hypothesis, we performed directed differentiation of hESC cells to β-cells following a defined 7-stage protocol for both lines. During differentiation, we completed a detailed characterization of H1 *INS-MAFA* cells at different stages side by side with H1 cells.

The specific goals of this research were:

- To assess the ability of H1 and H1 *INS-MAFA* hESC to form PPs using serum-free STEMdiff Pancreatic Progenitor Kit
- To further differentiate and characterize H1 and H1 *INS-MAFA* hESC-derived PP cells in micropatterned wells to generate uniform β-cell clusters
- To evaluate the potential of hESC-derived β-cells to respond to glucose challenges *in vitro*
- To confirm the functionality of integrated GFP and mCherry reporters both *in vitro* and *in vivo*

3 Materials and methods

3.1 Materials

This research was carried out with the hES H1 *INS-MAFA* reporter cell line previously generated in the lab. The targeting was performed by using the CRISPR-Cas9 system homologous recombination that allows a specific knock-in of the reporters GFP and mCherry downstream of the endogenous genes *INS* and *MAFA*, respectively. The targeting vectors were characterized by the 3' and the 5' homologous arms flanking the T2A-reporter sequence and the selection cassette between the FRT recombination sites. Briefly, wild-type H1 cells were first nucleofected with the Cas9 protein, the sgRNAs and the linearized donor vector carrying the GFP sequence to integrate the reporter downstream of the *INS* gene. The clones, selected by drug treatment, were validated by junctions PCRs and Sanger sequencing. The clones with the right *INS-GFP* targeting underwent the selection cassette removal and were confirmed again by junctions PCRs and sequencing. Stably targeted H1 *INS-GFP* cells were then used for the *MAFA-mCherry* targeting following the same procedure.

3.2 Methods

3.2.1 hESCs culture

During this research, all the cell culture work was done following the good cell culture practices guidelines.

3.2.1.1 Matrigel coating

A natural ECM-based hydrogel Matrigel hESC-Qualified Matrix (Corning) was first diluted 1:5 in cold DMEM/F-12 (Gibco) and stored at -20 ºC as a stock solution. Afterward, the stocked aliquots were further diluted 1:10 in cold DMEM/F-12 and 1 ml or 500 µl of the final 1:50 diluted Matrigel solution was added to a 6 well plate well (wpw) or a 12wpw, respectively. After 1 h incubation at room temperature (RT), the wells were washed once with 2 ml of DMEM/F-12 and kept until use.

3.2.1.2 Thawing of cryopreserved hESCs

The cells were stored in liquid nitrogen and one frozen cryovial was used to plate one 6wpw. Firstly, the cells were thawed as quickly as possible in a 37 °C water bath. Afterward, the content of the cryovial was topped with 500 µl of warm DMEM/F-12 and gently transferred in a 15 ml tube containing 5 ml of DMEM/F-12. Cells were then centrifuged at 300 G for 4 min at RT and cell pellet was gently resuspended in 1.5 ml of mTeSR medium (Stem Cell Technologies) with 10 µM ROCK inhibitor

(ROCKi, Miltenyi Biotec). The cell suspension was plated into a single precoated 6wpw and incubated at 37 °C, 5% CO_2. The cells were maintained by a daily wash with 2 ml of DPBS (Gibco) followed by fresh 1.5 ml mTeSR1 media change along with growth monitoring to determine optimal passaging time.

3.2.1.3 Passaging of hESCs

hES cells were passaged when they reached the optimal density of 70% with differentiated cells regions below 5%, usually 3-5 days after the seeding. Firstly, cells were washed once with 2 ml of DPBS. 1 ml of enzyme-free passaging reagent ReLeSR (Stem Cell Technologies) was then added and aspirated within 1 min so that the cells were left covered with a thin film of it. After 2 min at 37 °C, 5% CO_2, 1 ml of mTeSR was added to each well and the plate was firmly tapped to detach the colonies. Colonies were further broken up by pipetting with the p1000 to generate a uniform suspension of aggregates up to 10 cells each. Cells were split up to 1:10 ratio and seeded in precoated wells with 10 µM ROCKi for the first day after splitting. The cells were incubated at 37 °C, 5% CO_2 and maintained by daily media change as described in Section 3.2.1.2.

3.2.1.4 Cryopreserving of hESCs

The cells were occasionally frozen in order to maintain the stocks. When cells from one well were 70% confluent, they were frozen in one cryovial as colonies up to 10 cells each. Cells were detached and broken up as described in Section 3.2.1.3. The colonies suspension was then transferred in a 15 ml tube and centrifuged at 300 G for 4 min at RT. Afterward, the cell pellet was gently resuspended in 1 ml of cold mFreSR (Stem Cell Technologies) with 10 µM ROCKi and transferred to a labeled cryovial. Vials were frozen using a standard slow rate-controlled cooling protocol by reducing the temperature at 1 °C/min and storing at -80 °C overnight. The day after, cryovials were transferred and stored in liquid nitrogen.

3.2.2 2D differentiation: from hESCs to pancreatic progenitors

Before starting the differentiation, hES cells were passaged at least twice, but not more than 5 times. The cells were expanded to ensure enough cells for the differentiation. Requirements that had to be met for starting a successful differentiation were actively growing cells that were not too confluent and did not show any signs of differentiation in the hES culture. Prior to differentiation, cells had to be properly seeded as a single cell suspension.

3.2.2.1 hESCs seeding for differentiation

To seed cells for differentiation, mTeSR1 medium was firstly changed in the morning. After approximately 6 h, cells were washed once with 2 ml of DPBS and detached using 1 ml of TrypLE (Gibco) for 2 min at 37 °C, 5% Co_2. Colonies were then scraped using p1000 tip and subsequently gently pipetted a few times to obtain a single cell suspension. After pipetting, 1 ml of 2% bovine albumin (BSA, LSP)/DMEM/F-12 and 10 µM ROCKi was added to the well. The procedure was then repeated for up to 3 wells at the time. Afterward, the cells were pipetted a few more times, passed through a 40 µm strainer (pluriSelect) and centrifuged at 300 G for 4 min at RT. The pellet was then resuspended in 1 ml of mTeSR with 10uM ROCKi. To count the cells, 25 µl of cell solution was mixed with 25 µl of 0.4% Trypan Blue Solution (Thermo Fisher). 10 µl of the mix was loaded into the counting chamber of Countess Cell Counting Chamber Slides (Invitrogen) and the viability and cell number were determined using the Cell Countess II (Invitrogen). The counting was done in duplicate to ensure the correct number and the cells were seeded only if the viability was higher than 90%. The final volume of mTeSR1 containing 10 µM ROCKi was adjusted according to the cell number to a final concentration of 450 000 cells/ml. Subsequently, 900 000 or 250 000 cells were seeded in precoated 6wpw or 24wpw (Corning), respectively. The plate was then placed in an incubator at 37 °C, 5% CO_2. The next day, cells were washed with 2 ml of DPBS followed by mTeSR change. When the cells reached approximately 60% confluency (usually after 2 days), the differentiation was started.

3.2.2.2 Differentiation of hESC to PP stage

STEMdiff Pancreatic Progenitor Kit (STEMCELL Technologies) was used for the first four stages of the differentiation: DE, primitive gut tube (PGT), posterior foregut (PF) and PP stage. During the entire protocol lasting 13 days, the cells were gently washed with 2 ml of DPBS before media change to remove dead cells and debris. The medium for each respective stage was prepared following the instructions of the manufacturer (Fig. 3.1). Briefly, for Stage 1 of differentiation, Endoderm Basal Medium was used, while for Stage 2,3, and 4, Stage 2 – 4 Basal Medium was used. For each stage, specific supplements in 100x concentration were added to the relevant basal medium. The morphology of the cells was visually monitored throughout the differentiation and documented by photographs at the end of each stage.

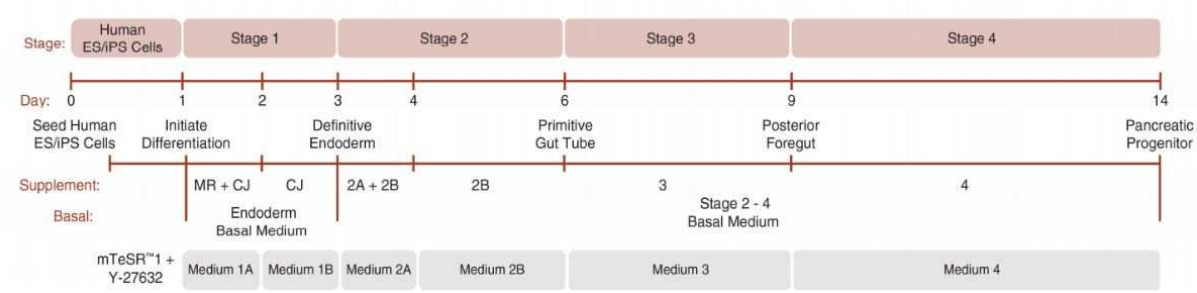

Figure 3.1 Differentiation protocol for the generation of pancreatic progenitors with STEMdiff Pancreatic Progenitor Kit. Protocol is defined by 4 stages and lasts 14 days (Product Information Sheet 2, STEMCELL Technologies).

3.2.2.3 Dissociation of PP cells

At the end of the 2D differentiation, PP cells were detached and clustered for further differentiation and maturation towards β-cells or collected for RT-qPCR and flow cytometry analysis. Maximum 3 wells at the time were first washed with 2 ml DPBS and then treated with 1 ml/well of Accumax (Sigma) for 4 min at 37 °C or until the cells were round. The cells were then pipetted using p1000 to generate single cell suspension and the dissociation was stopped by adding 1 ml of 2% BSA/DMEM/F-12 and 10 µM ROCKi. In addition, the cells were passed through a 40 µl strainer to ensure single-cell suspension. The suspension was thoroughly mixed and cells were counted as described in Section 3.2.2.1. The cells were then divided in appropriate numbers into separate 15 ml conical tubes for different purposes.

3.2.3 3D differentiation: from PPs to β-cells

Differentiation to β-cells lasts 25 days and it is characterized by 3 stages: Stage 5, Stage 6 and Stage 7 lasting 3,7 and 14 days, respectively. For these stages, AggreWell™800 24-well plates (STEMCELL Technologies) were used. These plates contain an array of microwells 800 µm in size per each well allowing the formation of uniform cell aggregates such as human islet-like clusters by enabling the seeding of a precise number of cells per microwell. The plates were prepared following the instructions of the manufacturer to ensure optimal performance during clustering. Firstly, the wells were pre-treated with 500 µl of Anti-Adherence Rinsing Solution (STEMCELL Technologies) and the plate was centrifuged in a swinging bucket rotor at 1300 G for 5 min at RT. After ensuring that there are no bubbles in the microwells the wells were rinsed with 2 ml of DMEM/F-12 and incubated with 1 ml of complete Stage 5 medium at 37 °C, 5% CO_2. Dissociated PP cells (Section 3.2.2.3) were counted as described in Section 3.2.2.1 and resuspended in Stage 5 medium at the concentration of 1.5 million cells/ml. This seeding number allowed to obtain 300 clusters consisted of 5000 cells each per microwell. Subsequently, 1 ml of the single cell suspension was added to one well of the AggreWell™800 plate and centrifuged at 100 G for 3 min at RT to capture the cells in the microwells. The plate was inspected to verify equal cell distribution among microwells. Finally, the plate was incubated at 37 °C, 5% CO_2 for 24 h to allow

clusters formation. After 24 h, half of the differentiation medium, composed of basal PEP-S7 medium with the addition of stage-specific factors and small molecules, was changed. The composition of the basal PEP-S7 medium and the supplementary factors for each stage are listed in Supplementary tables 8.1 and 8.2. During Stage 6 and 7, the plate was also placed on an orbital shaker at 100 rpm to ensure proper mixing of the media. The clusters morphology and size were checked daily throughout differentiation and photographs were taken at the end of each respective stage. Also, to follow the progress and direction of the differentiation, at the end of each stage, half well of clusters (~150 clusters) was collected for RT-qPCR analysis. In the end, the efficiency of the differentiation was evaluated by analyzing the derived β-cells using GSIS, flow cytometry and RT-qPCR.

3.2.4 Immunostaining of cell monolayers on coverslips

For immunostaining, hES, DE and PP cells were cultured on 12 mm diameter coverslips (Carl Roth). After a wash with DPBS, the cells were fixed with 400 µl of 4% paraformaldehyde in DPBS (PFA) for 20 min at RT and stored at 4 °C in 2 ml of DPBS if necessary. Furthermore, cells were washed three times with 1 ml of 0.3% triton/DPBS (DPBST) and blocked with 250 µl of 5% serum/DPBST for 1 h at RT. The serum used was from the same species as the species used for raising secondary antibodies. After blocking, cells were incubated with 250 µl of unconjugated primary antibodies diluted in 2.5% serum/DPBST under gentle rotation overnight at 4 °C. The list of primary antibodies can be found in Supplementary table 8.3. The next day, primary antibodies were washed three times fast and three times for 5 min with DPBST. Afterward, the cells were incubated with 250 µl of the secondary antibodies diluted in 2.5% serum/DPBST under gentle rotation for 1 h at RT in the dark. The list of secondary antibodies can be found in Supplementary table 8.3. The secondary antibodies were washed three times fast and three times for 5 min with DPBST. Subsequently, the nuclei were stained with 250 µl of Hoechst 33342 (Life Technologies) diluted 1:1500 in DPBST for 10 min in the dark. After several extensive washes with DPBS, coverslips were transferred from the wells to the slides and mounted using ProLongGold Antifade Mountant (without DAPI) (Life Technologies). Imaging was conducted using ApoTome.2 (ZEISS), while images were analyzed using ZEN 2.6. software (ZEISS) and Adobe Photoshop CS6.

3.2.5 Flow cytometry immunostaining

Firstly, DE and PP cells were dissociated as described in Section 3.2.3, washed with 3 ml of DPBS and fixed with 1 ml of 4% PFA for 10 min at 4 °C. Afterward, cells were centrifuged at 300 G for 4 min at RT, washed with 2 ml of DPBS and stored in 1 ml of DPBS for up to 10 days if necessary. For the staining, fixed cells were divided into 15 ml tubes (1-2 million cells per tube) for the staining and negative control (stained only with secondary antibodies). Immunofluorescence staining was done using the eBioscience Foxp3 / Transcription Factor Staining Buffer Set (Invitrogen). Firstly, cells were fixed and permeabilized with Fixation/Permeabilization solution for 1 h at 4 °C in the dark. After two washes with 2 ml of Permeabilization Buffer 1X (PB), cells were blocked with 100 µl of 5% serum/PB. After 15 min of incubation at RT, primary antibodies were directly added to the blocking solution according to the appropriate dilutions and incubated overnight at 4 °C at 1200 rpm. All the primary antibodies, along with the additional information, are listed in Supplementary table 8.4. The next day, primary antibodies were washed twice with 2 ml of PB. The cells were then incubated with the appropriate secondary antibodies in 5% serum/PB for 1 h at RT in the dark at 1200 rpm. The list of secondary antibodies can be found in Supplementary table 8.4. After the incubation, cells were washed twice with 3 ml of PB and resuspended in 300 µl of 5% serum/PB with the addition of 10 mM EDTA (Sigma). In the end, to ensure single cell suspension, cells were passed through a 40 µm strainer and then analyzed by BD FACSCanto II (BD Biosciences).

3.2.6 RNA extraction, cDNA synthesis and RT-qPCR

3.2.6.1 RNA extraction

For the RNA extraction, PP, PEP, S6 and S7 cells were collected in a 15 ml tube and pelleted at 300 G for 4 min. Cells were then lysed in 350 µl of RLT buffer and 100x β-mercaptoethanol (Sigma). Samples were stored at -20 °C until further processing. Extraction was performed using the Rneasy Micro Kit (Qiagen) with on-column digestion of genomic DNA following the instructions of the manufacturer. In the last step, RNA was eluted in 15 µl of Rnase-free water by centrifugation at full speed for 1 min. Elution step was repeated twice in order to increase RNA yield. Finally, the concentration and purity of isolated RNA were determined by NanoDrop (Thermo Fisher) and samples were stored at -20 °C.

3.2.6.2 cDNA synthesis

After RNA extraction, the first strand cDNA was synthesized using the PrimeScript RT Master 5X Mix (Takara). A maximum of 2 µg of RNA was used for 40 µl of reaction. For S7 samples, the final volume of the reverse transcription reaction was 60 µl. The cDNA synthesis was performed using the Mastercycler nexus gradient thermocycler (Eppendorf) and the program consisted of the first step at 37

°C for 15 min for the reverse transcription followed by the second step at 85 °C for 5 seconds for the reverse transcriptase inactivation. After the synthesis, cDNA was stored at -20 °C.

3.2.6.3 Quantitative reverse transcription polymerase chain reaction (RT-qPCR)

Relative gene expression for the genes of interest was determined by RT-qPCR using the Roche platform. Primers used for RT-qPCR were previously designed using the Primer3 software (SimGene). The parameters that were considered were: product size (90-120bp), primer size (18-22bp), primer Tm (58-62 °C) and GC content (40-60%). The primer pairs for each stage of the differentiation analyzed are reported in Supplementary table 8.5. For every stage of differentiation, biological triplicates were analyzed. Each amplification reaction consisted of 5 µl 2X FastStart Essential DNA Green Master (Roche), 1 µl of cDNA (10-50ng), 1 µl of each 10 µM primer and water up to the final volume of 10 µl. All the amplification reactions were carried out in technical triplicates using the LightCycler 480 II instrument (Roche) for 40 cycles. After amplification, the threshold cycle values (Ct) were used for the relative expression quantification. Briefly, Ct values were normalized with the Ct values of the housekeeping *TBP* gene giving the ΔCt, whereas the fold change was calculated from the ΔΔCt values obtained over the undifferentiated H1 hESC cells ΔCt.

3.2.7 Functionality assays for ß-cells

3.2.7.1 Glucose-stimulated insulin secretion (GSIS)

At the end of the differentiation (Day 14 of S7), the β-cell functionality was tested by GSIS. Clusters were incubated overnight in 2 ml of S7 medium with reduced glucose concentration (5.5 mM). The day after, clusters were collected from a full well and transferred to a 20 µm strainer (pluriSelect). The strainer with clusters was placed into a 12wpw and washed three times by transferring the strainer to a new well containing 2 ml of prewarmed KRBH buffer (129 mM NaCl, 4.8 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgSO_4$, 1 mM Na_2HPO_4, 1.2 mM KH_2PO_4, 5 mM $NaHCO_3$, 10 mM HEPES, 0.1% BSA in deionized water, filter sterilized, pH 7.4). Cells were equilibrated in the last wash in the hybridization oven for 1 h at 37 ºC. Afterward, cells were washed with 2 ml of low glucose solution (KRBH spiked with 2.8 mM glucose), transferred to a new well and 750 µl of low glucose solution was added on the top of the clusters. After 1 h incubation in the hybridization oven at 37 ºC, the supernatant was collected as a low glucose sample. Subsequently, strainer with clusters was transferred to the new well and 750 µl of high glucose solution (KRBH spiked with 16.7 mM glucose) was added on the top of the clusters. After 1 h incubation in the hybridization oven at 37 ºC, the supernatant was collected as a high glucose sample. Lastly, the strainer was transferred to the new well and 750 µl of high glucose/KCl solution (KRBH spiked with 16.7 mM glucose and 60 mM KCl) was added on the top of the clusters. After 1 h incubation in the hybridization oven at 37 ºC, the supernatant was collected as a KCl sample. Finally,

the stimulation samples were stored at -20 ºC, whereas the clusters on the strainer were collected for DNA extraction.

3.2.7.2 DNA extraction

DNA extraction was performed using Dneasy Blood & Tissue Kit (Qiagen) to normalize the GSIS results. After GSIS, the clusters were collected from the strainer in 500 µl DPBS and centrifuged at 300 G for 4 min. Cells were then resuspended in 200 µl of DPBS with the addition of 20 µl of proteinase K. The DNA extraction was done following the instructions of the manufacturer. In the last step, DNA was eluted in 20 µl of prewarmed Buffer AE. The elution step was repeated twice in order to increase DNA yield. Finally, the concentration and purity of isolated DNA was determined by NanoDrop and samples were stored at -20 °C.

3.2.7.3 Ultrasensitive C-peptide enzyme-linked immunosorbent assay (ELISA)

ELISA for human C-peptide was performed using Ultrasensitive C-peptide ELISA kit (Mercodia) following the manufacturer's protocol. Briefly, 50 µl of calibrators and GSIS samples were pipetted in duplicates in the appropriate wells and then 50 µl of Assay Buffer was added to each well. The plate was incubated for 1 h at RT at 700 rpm. Wells were then washed 6 times using 400 µl of wash buffer 1x solution. Subsequently, 200 µl of enzyme conjugate 1x solution was added to the wells followed by incubation for 1 h at RT at 700 rpm. After incubation, wells were washed as described above and 200 µl of the substrate TMB was added. After 30 min incubation at RT, 50 µl of Stop Solution was added to each well. Optical density (OD) was measured within 30 min at 450 nm using the Infinite 200 PRO plate reader (Tecan). To generate a calibration curve, the Curve expert software was used. In short, the mean of measured OD values for each calibrator with their known concentrations allowed the design of the standard curve. Generated rational curve was then used to extrapolate released C-peptide concentrations from the mean of the samples OD values. Finally, the respective concentrations were normalized to the DNA content allowing calculation of the stimulation index.

3.2.7.4 Live cell flow cytometry

Prior to flow cytometry analyses, S7 clusters had to be dissociated to single cells in order to properly quantify reporters' expression. For that, half well of clusters was collected from the microwells in a 15 ml tube. Cells were then centrifuged at 300 G for 2 min at RT and washed with 10 ml of DPBS. After spinning down, clusters were resuspended in 1 ml of dissociation solution containing 1x non-enzymatic cell dissociation solution (Sigma) with Dnase I (40 µg/ml) and 0.01% Trypsin-EDTA. Dissociation was performed for approximately 6 min in a 37 °C water bath. When all the clusters were dissociated into single cells, the process was stopped by adding 3 ml of 2% BSA/DMEM/F-12. Cells were then

centrifuged at 300 G for 4 min at RT and resuspended in 350 µl of DPBS containing 0.05% FCS and 0.02% EDTA. DRAQ7 (Miltenyi Biotec) was also added to the mixture to label dead cells. Before flow cytometry, cells were also counted as described in Section 3.2.2.1 to determine the total number of live cells. To ensure single-cell suspension, cells were passed through a 40 µm strainer and then analyzed by BD FACSCanto II (BD Biosciences). Cells were screened for the reporters GFP and mCherry fluorescent signals, green and red respectively. Dead cells labeled with DRAQ7 in far-red were excluded from the analysis.

3.2.8 Statistical analysis

In this study, flow cytometry data was analyzed using the FlowJo software (FLOWJO, LLC). For RT-qPCR, statistical analysis was performed using Graphpad Prism 8 software. Statistical significance was determined using the unpaired parametric Student's *t*-test for two-tailed distributions with Welch's correction for sets of data containing two groups. When comparing more than two independent groups, ordinary one-way ANOVA with Tukey's *post hoc* comparison was applied. The level of significance was set at $p<0.05$ and displayed visually as mean±standard deviation (SD), where applicable.

4 Results

As already mentioned, the main objective of this research was detailed characterization and assessment of differentiation potential of H1 *INS-MAFA* reporter cell line. After validation of a proper integration after CRISPR-Cas9 targeting (data not shown), line with modified *INS* and *MAFA* loci was used for further analysis and characterization (Fig. 4.1).

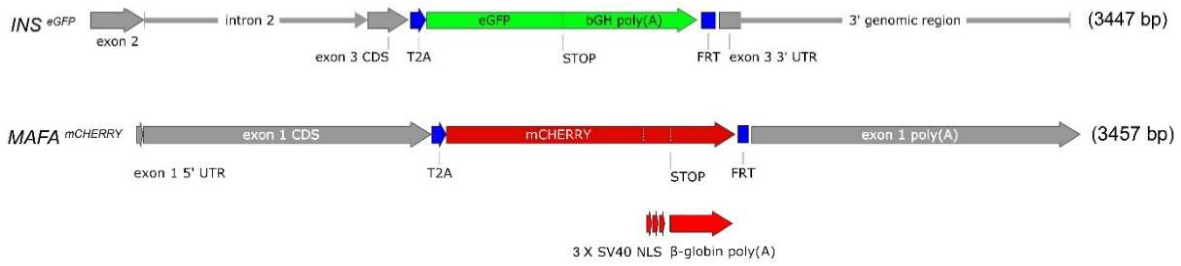

Figure 4.1 Reporter integration in target loci. eGFP and mCherry reporters are integrated downstream of the target gene and are transcribed under respective promoter regulation joint with target gene transcript by T2A sequence.

Further analysis was conducted by directed differentiation of H1 and H1 *INS-MAFA* hESC using differentiation protocol illustrated in Figure 4.2. The stepwise protocol directed hESC towards definitive endoderm (DE), primitive gut tube (PGT), posterior foregut (PF), pancreatic progenitors (PP), pancreatic endocrine precursors (PEP), immature β-cells and finally, maturing β-cells. Protocol design was based on developmental principles by modulation of signaling pathways important for lineage commitment.

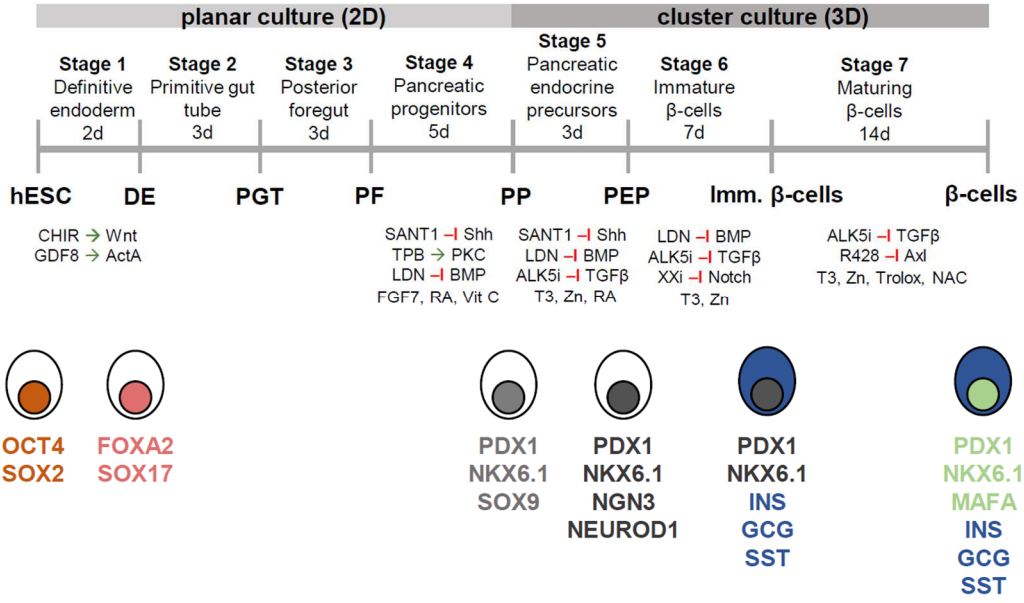

Figure 4.2 7-stage differentiation protocol. Respective differentiation stages, along with the duration in days, are depicted above. Modulated signaling pathways for a proper lineage commitment are shown below, along with main molecular markers followed for certain differentiation stages.

4.1 H1 *INS-MAFA* hESCs differentiate to pancreatic progenitors similarly to the parental H1 cells

In order to use the H1 *INS-MAFA* reporter line for improving the current differentiation protocol to obtain fully functional and mature hESC derived β-cells, the line had to be properly characterized to ensure that it retains a specific identity during differentiation to the PP stage. Characterization was performed by carrying out several analyses, namely immunofluorescence on coverslips, flow cytometry, and RT-qPCR at the hESC, DE and PP stage. Unmodified wild type H1 cell line was differentiated in the same manner and used as a reference for comparison.

The first step in the molecular characterization of H1 *INS-MAFA* cells was to verify the hESC identity by immunofluorescence on coverslips. The molecular machinery behind hESC heavily relies on the expression of SOX2 and OCT4. They regulate a large gene network supporting pluripotency and self-renewal as two main characteristics of hESC (Rizzino & Wuebben, 2016). Thus, the expression of these genes has to be maintained within narrow limits to enable proper differentiation as a response to developmental signals. Immunofluorescence stainings in Figure 4.3 show that H1 *INS-MAFA* hESCs retained the expression of these two pluripotency markers after the genomic modification similarly to H1 hESC. Confirming the hESC identity was an important requirement for further differentiation to PPs and later β-cells.

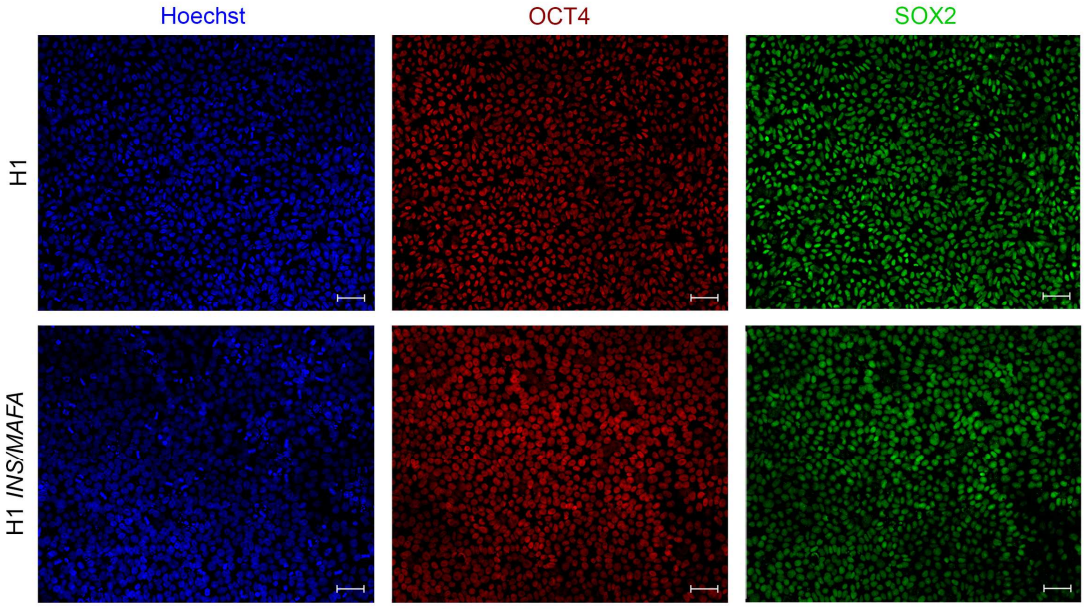

Figure 4.3 hESC. Immunofluorescence staining of H1 and H1 *INS-MAFA* cells for pluripotency markers OCT4 (red) and SOX2 (green). Nuclei are labeled with Hoechst 33342 in blue. Scale bar: 50 μm.

To efficiently reprogramme hESC towards DE fate, it is important to mimic developmental principles *in vivo*. Thus, two main contents of differentiation media for this stage include activator of Wnt signaling

and GSK8 as a regulator of Nodal signaling (Mahaddalkar et al., 2020; Shih et al., 2013). At the end of Stage 1 of planar differentiation, H1 and H1 *INS-MAFA* cells were analyzed for the expression of specific proteins by immunofluorescence on coverslips and flow cytometry to confirm the DE identity. The DE stage is characterized by co-expression of SOX17 and the anterior DE marker FOXA2 (D'Amour et al., 2006). As shown in Figure 4.4A, almost all of the H1 *INS-MAFA* cells stained positive for the two markers, likewise H1 cells. In addition, flow cytometry quantification confirmed the latter by demonstrating that 96.6% of the analyzed cells were $SOX17^{+}FOXA2^{+}$ double positive, comparable to the 95.7% of H1-derived DE cells (Fig. 4.4B). Thus, this confirmed the potential of the H1 *INS-MAFA* line to acquire DE fate with high efficiency like the original H1 line.

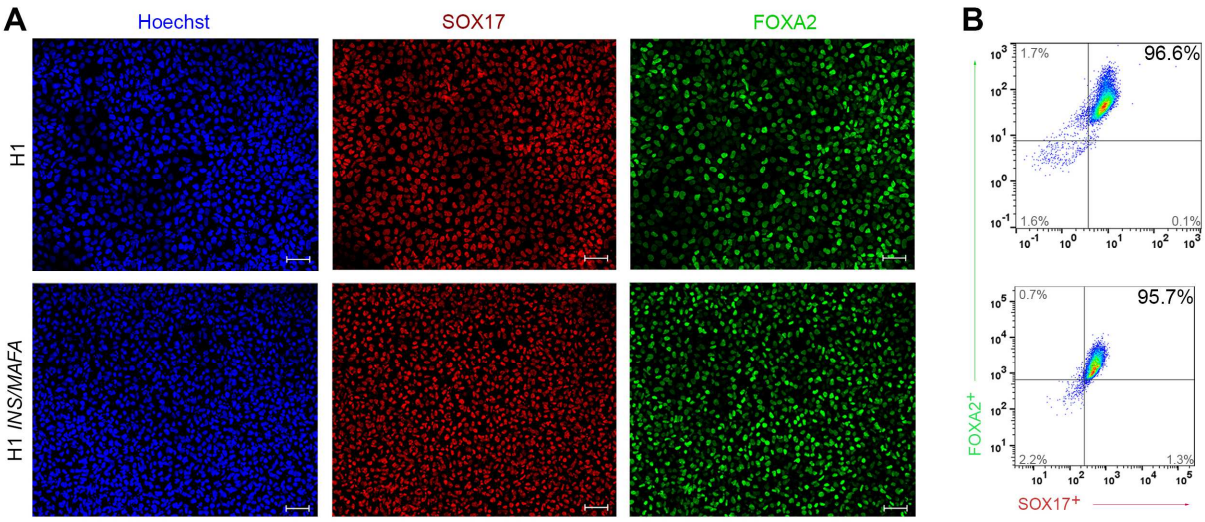

Figure 4.4 Definitive endoderm stage (Stage 1). (A) Immunofluorescence staining of H1 and H1 *INS-MAFA* cells for SOX17 (red) and FOXA2 (green). Nuclei are labeled with Hoechst 33342 in blue. Scale bar: 50 µm. **(B)** Flow cytometry plots of the H1 and H1 *INS-MAFA* cells showing the percentage of cells double-positive for SOX17 and FOXA2.

Due to the high reprogramming efficiency observed for the first differentiation stage, PGT and PF identity were not directly assessed and the analysis was continued at the PP stage. To direct cells towards PP fate, several signaling pathways had to be properly modulated. In particular, Shh and BMP signaling had to be inhibited for proper patterning, whereas PKC had to be activated as it induces PDX1 expression. Other molecules important for differentiation at this stage included FGF7, retinoic acid (RA) important for patterning of gut endoderm and vitamin C that promotes extracellular matrix production (Shih et al., 2013). At the end of the planar differentiation, the gene expression profile of the PP cells derived from both cell lines was first analyzed by RT-qPCR followed by immunofluorescence on coverslips and flow cytometry to confirm the results on protein level. Genes analyzed by RT-qPCR can be divided into 2 groups. The first one includes „desirable genes" (Fig. 4.5A), namely the TFs *PDX1*, *NKX6.1*, *SOX9*, *FOXA2*, and *PTF1A* that mark PP identity. Orchestrated expression of these TFs leads pancreas development and lineage commitment (Aigha & Abdelalim, 2020). More specifically, *PDX1*

is considered to be a master regulator of pancreatogenesis required for subsequent differentiation and specification of all pancreatic lineages and later maintenance of β-cell function (Zhu et al., 2017), whereas *NKX6.1* has a crucial role in guiding proper development towards functional β-cells (Aigha & Abdelalim, 2020). Thus, the expression of *PDX1* and *NKX6.1* is detrimental for proper reprogramming towards β-cells and was followed at every subsequent differentiation stage. On the other hand, *SOX9* expression is very important for the maintenance of progenitor phenotype, but later its expression becomes restricted exclusively to ductal cells (Seymour et al., 2007). Moreover, the expression of *FOXA2* and *PTF1A* was investigated since these genes are equally important for acquiring pancreatic fate (Kawaguchi et al., 2002). As expected, no significant difference in gene expression was observed for the main progenitor markers between H1 and H1 *INS-MAFA* lines. The only exception was *PTF1* that demonstrated significantly lower expression in H1 *INS-MAFA* cells.

The second group of genes (Fig. 4.5B) represents „undesirable genes“ whose expression *in vivo* marks either hepatic or intestinal linage. The main unwanted markers followed in this study include the hepatic markers *AFP* and *HHEX* (Jin et al., 2009; Watanabe et al., 2014) along with the intestinal marker *CDX2* (Silberg et al., 2000). Even though the expression of these markers is not expected throughout the differentiation, the hepatic marker *AFP* was highly expressed in both cell lines, especially in the H1 *INS-MAFA* line, whereas *HHEX* and *CDX2* were similarly expressed in both cell lines. To sum up, RT-qPCR showed a similar pattern of gene expression between cell lines, except for *PTF1A* and *AFP*. The significant difference in the expression observed for these two genes could be explained by an inter- and intra- differentiation variability that strongly depends on several factors such as the identification of the perfect time point for starting the differentiation.

Additionally, RT-qPCR results have been confirmed on protein level by immunofluorescence on coverslips and flow cytometry. Immunofluorescence staining (Fig. 4.5C) suggests that the vast majority of cells co-expressed the main PP markers PDX1, NKX6.1 and SOX9 similarly to H1 cells. Again, immunofluorescence was quantitatively confirmed by flow cytometry, showing that almost 100% of analyzed H1 and H1 *INS-MAFA* cells were double-positive for PDX1 and SOX9 (Fig. 4.5D). In addition, more than half of the double positive cells were also positive for NKX6.1 (Fig. 4.5E), 75% and 55% for H1 and H1 *INS-MAFA* cells, respectively.

All things considered, it has been shown that after the genome modification, H1 *INS-MAFA* cells retain their pluripotent identity and can successfully and reproducibly differentiate to the PP stage similarly to the H1 cells.

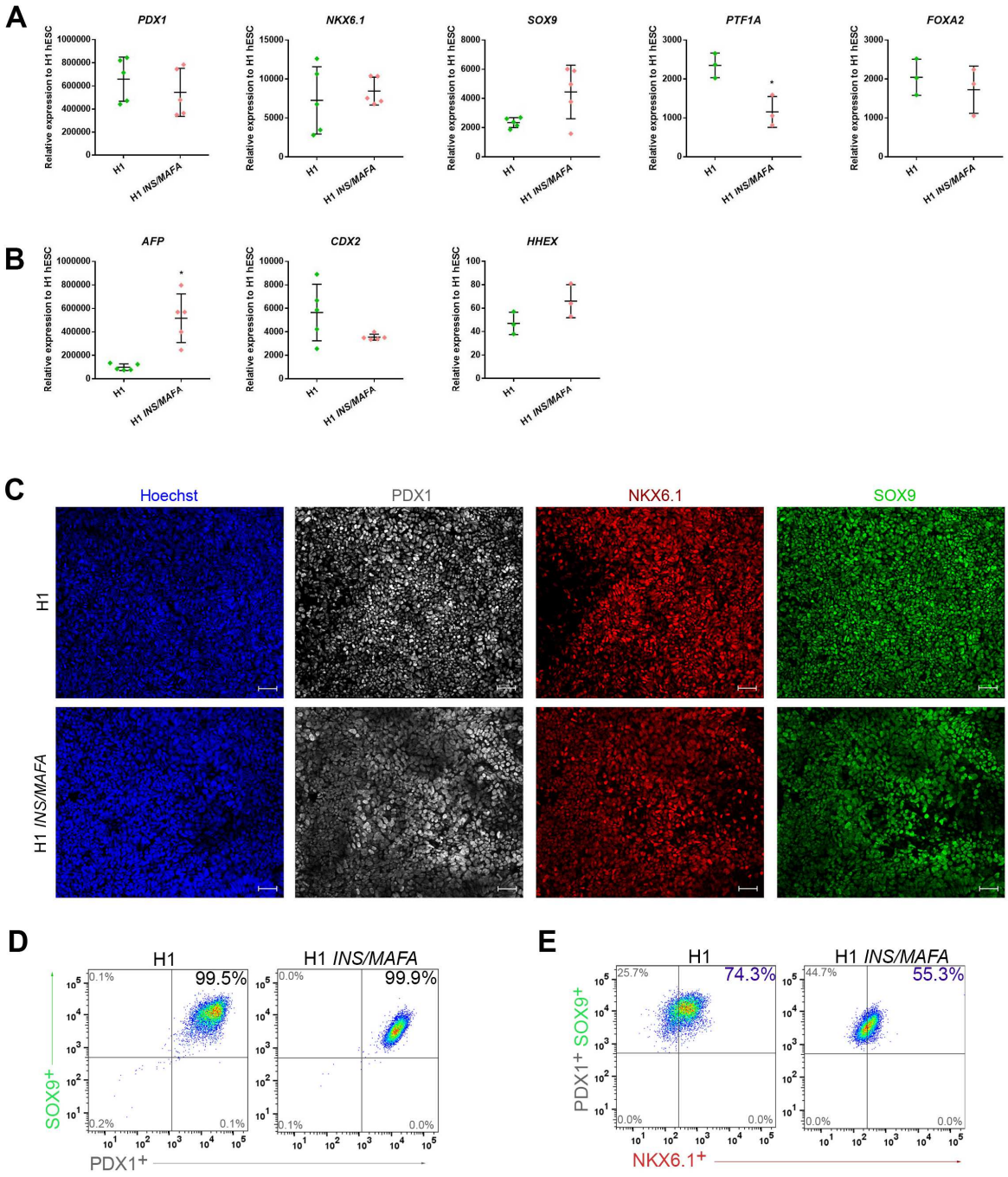

Figure 4.5 Pancreatic progenitor stage (Stage 4). (A, B) RT-qPCR gene expression analyses of the: **(A)** pancreatic progenitor markers *PDX1, NKX6.1, SOX9, PTF1A* and *FOXA2;* **(B)** hepatic markers *AFP* and *HHEX* and intestinal marker *CDX2*. Values are normalized to the expression of housekeeping h*TBP* and the fold change is relative to the hESC stage. The data shown are biological replicates (n = 5 for *PDX1*, *NKX6.1*, *SOX9*, *AFP*, *CDX2* and n = 3 for *PTF1A*, *FOXA2*, *HHEX*). Means were compared with parametric unpaired two-tailed t-test with Welch's correction. Significance is corresponding to p<0.0332 (*), 0.0021 (**), 0.0002 (***), <0.0001 (****). **(C)** Immunofluorescence staining of H1 and H1 *INS-MAFA* cells for PDX1 (gray), NKX6.1 (red) and SOX9 (green). Nuclei are labeled with Hoechst 33342 in blue. Scale bar = 50 um. (D, E) Flow cytometry plots of the H1 and H1 *INS-MAFA* cells showing the percentage of the cells double positive for PDX1 and SOX9 **(D)** and the PDX1$^+$/SOX9$^+$ cells triple positive for NKX6.1 **(E)**.

4.2 H1 and H1 *INS-MAFA* derived PP cells reproducibly differentiate to insulin-producing β-cells

At the end of planar differentiation, H1 and H1 *INS-MAFA* derived PPs were dissociated and clustered in micropatterned wells to acquire β-cell identity. In 3D culture, PP cells were firstly directed towards endocrine lineage until pancreatic endocrine progenitors (PEP, S5) guided by continuous Shh and BMP inhibition, as well as RA patterning. In this stage, triiodothyronine (T3) important for maturation and TGFβ inhibitor Alk5ii important for endocrine induction were introduced and used further in the protocol (Rezania et al., 2014). PEP specification was followed by differentiation towards immature β-cell cells (S6), where inhibition of Notch signaling pathway, important for maintenance of progenitor identity, proved to be essential (Afelik et al., 2012). To finally acquire a mature β-cell phenotype (S7), an Axl inbitor R428 and antioxidant N-acetyl cysteine were added to the media to induce the expression of maturation TF MAFA (Rezania et al., 2014). To evaluate differentiation efficiency, clusters were analyzed by RT-qPCR at the end of the S5, S6 and S7 to investigate the expression of the main molecular markers.

As shown in Figure 4.6A, key pancreatic markers *PDX1*, *NKX6.1* and *SOX9* expression was assessed for both PEP and S6 cells. Importantly, *PDX1* stayed stably expressed, whereas expression of *NKX6.1* was significantly upregulated in S6 cells for both cell lines. On the other hand, *SOX9* was significantly downregulated as its expression is restricted to the ductal lineage with further differentiation. Interestingly, the expression of the hepatic marker *AFP* was significantly downregulated in S6 cells, while intestinal marker *CDX2* stayed stably expressed (Fig. 4.6B). In addition, PEP cells were also analyzed for the expression of the endocrine markers *NGN3* and *NEUROD1*. In particular, *NGN3* is a proendocrine gene that is transiently expressed in endocrine precursors, while *NEUROD1* is its downstream effector that drives insulin expression (Zhu et al., 2017). H1 and *H1 INS-MAFA* derived PEP cells showed a high expression of *NGN3*, while *NEUROD1* expression was significantly lower in H1 *INS-MAFA* cells (Fig. 4.6C). Since *NEUROD1* expression is regulated by *NGN3*, significance can be explained by the fact that *NGN3* also showed a lower expression trend in H1 *INS-MAFA* cells, even though that difference was not significant. Nevertheless, high expression of both genes shows that cells successfully shifted towards endocrine lineage.

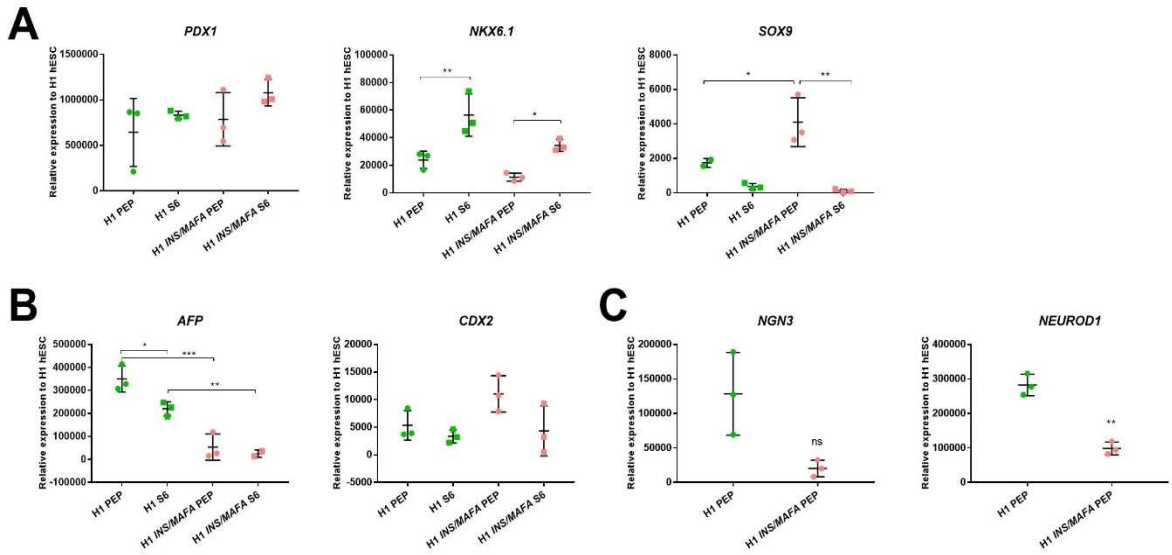

Figure 4.6 Pancreatic endocrine precursor (Stage 5) and immature β-cell stage (Stage 6). RT-qPCR gene expression analyses of the: **(A)** main pancreatic progenitor markers *PDX1*, *NKX6.1* and *SOX9*; **(B)** hepatic and liver markers *AFP* and *CDX2*; **(C)** endocrine lineage markers *NGN3* and *NEUROD1* at the end of Stage 5. Values are normalized to the expression of the housekeeping h*TBP* and the fold change is relative to the hESC stage. The data shown are biological replicates (n = 3). Means were compared with one-way ANOVA with Tukey test for multiple comparisons (A, B) or parametric unpaired two-tailed t-test with Welch's correction (C). Significance is corresponding to p<0.0332 (*), 0.0021 (**), 0.0002 (***), <0.0001 (****).

At the end of the differentiation, derived β-cells stably expressed *PDX1* and *NKX6.1* (Fig. 4.7A) and they also showed a downregulation trend for *AFP* expression, whereas *CDX2* expression remained stable (Fig. 4.7B). Importantly, both H1 and H1 *INS-MAFA* cells demonstrated high expression of *INS* already at the end of the S6, with an increasing trend during S7. Moreover, the cells also expressed the α-cell hormone *GCG* and δ-cell hormone *SST*. Surprisingly, both *GCG* and *SST* expression were significantly higher in H1 *INS-MAFA* cells compared to H1 cells. Additionally, *GCG* exhibited a strong upregulation trend between S6 and S7 H1 *INS-MAFA* cells, which was not the case for H1 cells (Fig. 4.7C).

In conclusion, even though some differences were observed for the expression of the endocrine markers *NGN3* and *NEUROD1* in S5, the cells of the two lines properly differentiated towards β-cells. In both cases, expression of the main pancreatic markers was preserved, whereas hepatic fate was suppressed. Consequently, derived β-cells were characterized by a high hormonal content, with an emphasis on *INS*.

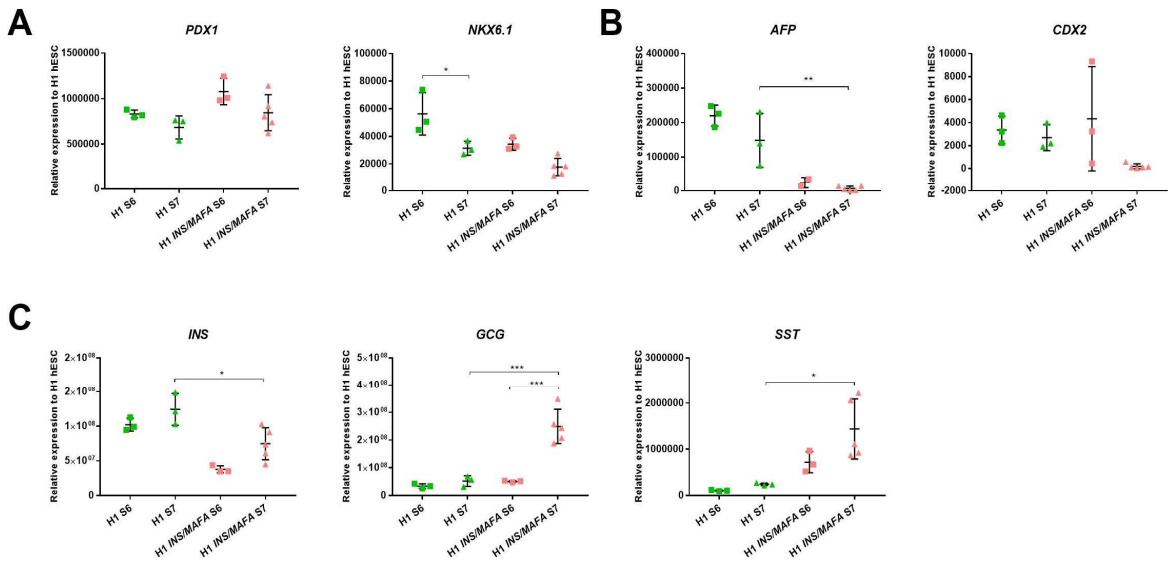

Figure 4.7 hESC derived β-cells (Stage 7). RT-qPCR gene expression analyses of the: **(A)** for the main pancreatic markers *PDX1* and *NKX6.1;* **(B)** hepatic and liver markers *AFP* and *CDX2;* **(C)** hormones *INS*, *GCG* and *SST*. Values are normalized to the expression of housekeeping h*TBP* and the fold change is relative to the hESC stage. The data shown are biological replicates (n = 3 and n = 5 only for H1 *INS/MAFA* S7). Means were compared with one-way ANOVA with Tukey test for multiple comparisons. Significance is corresponding to p<0.0332 (*), 0.0021 (**), 0.0002 (***), <0.0001 (****).

4.3 H1 and H1 *INS-MAFA* derived β-cells express main functionality related genes and can respond to glucose challenges in vitro

As discussed in Section 4.2, hESCs derived β-cells are characterized by high *INS* expression, but that is not the sole indicator of their functionality. Thus, to investigate the functionality, S7 cells were assessed for the expression of specific genes involved in the glucose response, insulin release and maturation. In addition, the real functionality was evaluated by glucose stimulation and insulin release was quantified by ELISA assay. In particular, the genes analyzed by RT-qPCR were *GK, SUR1* and *ZnT8*. Specifically, *GK* is a hexokinase isozyme serving as the main glucose sensor in β-cells (Shirakawa & Terauchi, 2020). *SUR1* is the main subunit of ATP-sensitive potassium channels important for stimulating insulin secretion (Gloyn et al., 2006), while *ZnT8* is a transporter that regulates zinc homeostasis by controlling its uptake in secretory granules (Davidson et al., 2014). As shown in Figure 4.8A, all of the functionality-related genes were strongly expressed compared to hESC cells for both H1 and H1 *INS-MAFA* cell lines, suggesting that cells acquired the desired phenotype.

Furthermore, an important characteristic of adult β-cells is the expression of the maturation TF *MAFA*. Through the MAFA targets, it is known that MAFA specifically binds to the conserved insulin enhancer element and regulates glucose-responsive insulin expression (Zhu et al., 2017). Finely tuned *MAFA* expression is detrimental for proper β-cell maturation both *in vivo* and *in vitro* and is shown to be highly beneficial in improving GSIS (Rezania et al., 2014). Even though *MAFA* was expressed in

both cell lines, it was not strongly upregulated (Fig. 4.8B), suggesting that further improvements in the differentiation protocol are still needed.

To test real functional capacity, derived β-cells were exposed to glucose stimulation in a form of sequential static challenges. Upon exposure to different stimuli (2.8 mM glucose, 16.7 mM glucose, 16.7 mM glucose/30 mM KCl), samples were collected for ELISA. By performing ELISA assay for each sample and stimulation event, the stimulation index was calculated as a ratio between normalized secreted C-peptide concentration for 16.7 mM high glucose stimuli and 2.8 mM low glucose stimuli, as well as a ratio between 16.7 mM high glucose/30 mM KCl stimuli and 2.8 mM low glucose stimuli. The final indexes give insight into the ability of the cells to respond to higher stimuli compared to basal secretion. As shown in Figure 4.8C, it is evident that both H1 and H1 *INS-MAFA* cells secreted 1.5 times more C-peptide when stimulated with 16.7 mM glucose. Subsequently, when stimulated with 16.7 mM glucose and 30 mM KCl, H1 and H1 *INS-MAFA* cells secreted 12 and 10 times more C-peptide compared to basal secretion rate, respectively.

Therefore, it was confirmed that the H1 *INS-MAFA* derived β-cells genetic profile fully resembles the one determined for the original H1 cells since no significant differences in the expression of functionality and maturation specific genes were assessed. Also, H1 and H1 *INS-MAFA* derived β-cells were capable of sensing low and high glucose impulses by regulating C-peptide secretion, indicating that cells are functional.

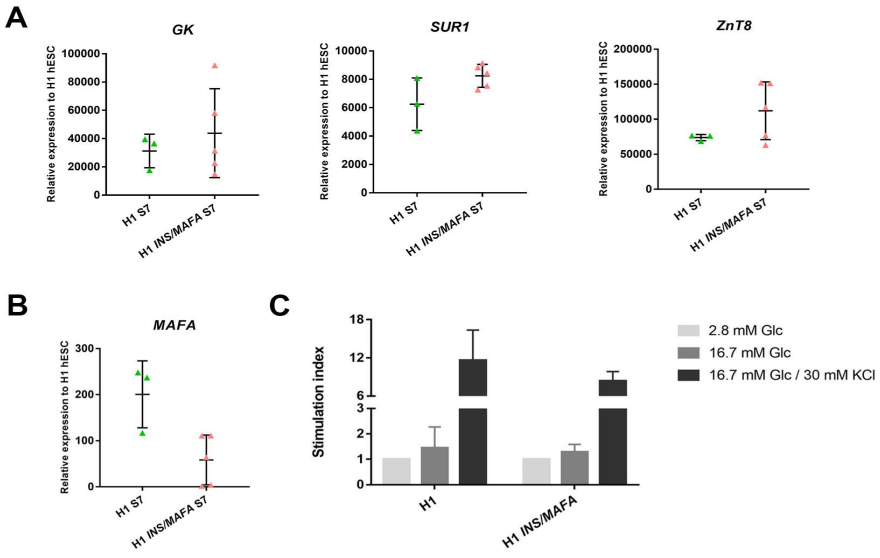

Figure 4.8 Functional analysis of hESC derived β-cells (S7). (A, B) RT-qPCR gene expression analyses of the: **(A)** functionality markers *GK*, *SUR1*, *ZnT8;* **(B)** maturation marker *MAFA*. Values are normalized to the expression of the housekeeping h*TBP* and the fold change is relative to the hESC stage. The data shown are biological replicates (n = 3 for H1 and n = 5 for H1 *INS-MAFA*). Means were compared with parametric unpaired two-tailed t-test with Welch's correction. Significance is corresponding to p<0.0332 (*), 0.0021 (**), 0.0002 (***), <0.0001 (****). **(C)** Static glucose-stimulated human C-peptide secretion. Data (n = 2) is shown using a bar chart for every independent challenge in a form of a stimulation index.

4.4 Newly integrated GFP and mCherry reporters are functional in H1 *INS-MAFA* derived β-cells

Apart from assessing the differentiation potential of H1 *INS-MAFA* cell line, one of the aims of this characterization was to determine the functionality of the reporters. Since GFP and mCherry were integrated downstream of the *INS* and *MAFA* genes, respectively, their expression is under the regulation of the endogenous genes' promoters. In this way, the reporters are transcribed as a single mRNA with the respective gene and translated separately due to the self-cleaving peptide T2A. To verify the reporters' functionality, flow cytometry analyses were carried out on single cells derived from dissociated clusters to quantify the GFP and mCherry positive cells, in addition to live cluster imaging.

Strikingly, in both independent differentiations, flow cytometry analyses showed that more than 70% of cells were GFP positive, whereas almost no cells were detected for mCherry (Fig. 4.9A). This could be explained by the low expression of *MAFA* in the derived β-cells as it was demonstrated by the RT-qPCR in Figure 4.8B. However, by performing high magnification imaging, a few cells positive for the nuclear mCherry were detected. Moreover, these cells co-expressed GFP, localized in the cytoplasm, confirming the functionality of both reporters (Fig. 4.9B). However, additional analysis by immunostaining on cryosections is necessary to confirm latter results and better visualize positive cells.

All in all, it was confirmed that the novel H1 *INS-MAFA* cell line retains the potential to differentiate to β-cells and that both GFP and mCherry reporters are properly functional, with a cytoplasmatic pattern for GFP and nuclear location for mCherry.

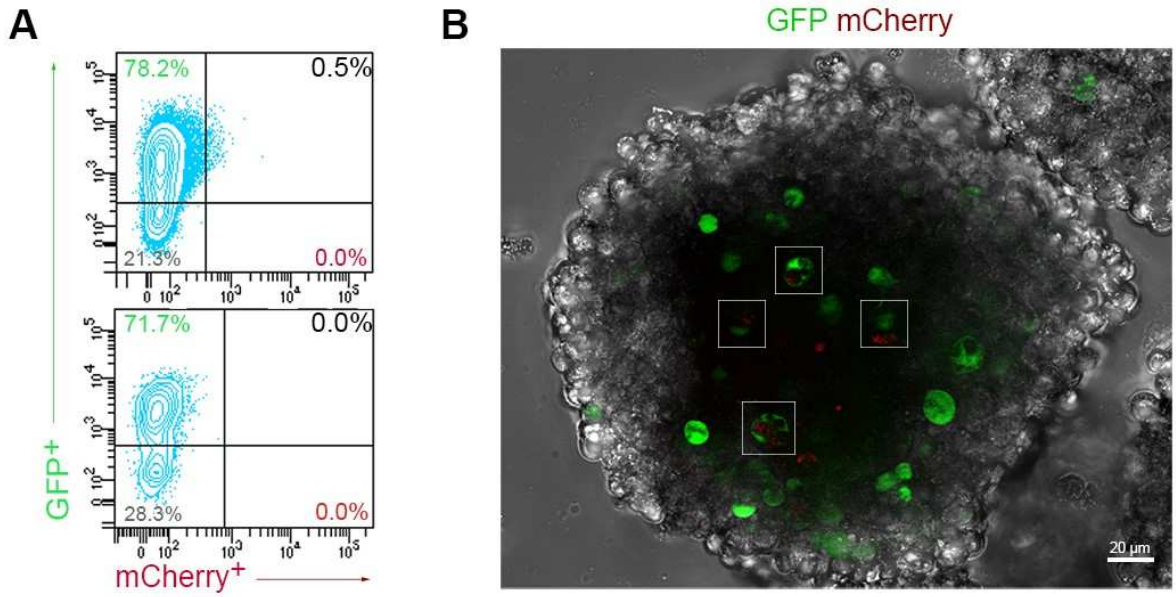

Figure 4.9 GFP and mCherry functionality in the β-cells derived from H1 *INS-MAFA* cell line. **(A)** Flow cytometry plots from two independent differentiations showing the GFP$^+$ (green), mCherry$^+$ (red) and the GFP$^+$/mCherry$^+$ cells (black). **(B)** Imaging of a live cluster at the end of the differentiation. The GFP$^+$/mCherry$^+$ cells are highlighted in the white boxes.

5 Discussion

Apart from exogenous insulin administration, pancreatic islet transplantation is a well-established treatment option for patients suffering from insulin-deficient diabetes (Rickels & Robertson, 2019). Even though successful transplantation can lead to acquiring insulin independence, this approach is still limited due to the scarcity of donor islets and life-long immunosuppression and there are multiple efforts put into making it more personalized and accessible. In order to overcome these pitfalls, cell based therapies utilizing hPSCs emerged as a promising alternative. *In vitro* directed differentiation of hPSCs has the potential to provide uniform and unlimited supply of human β-cells, as well as enable effective drug screening and disease modelling. However, protocols for the generation of β-cells *in vitro* are still not optimal as derived cells are not fully functional and mature. Furthermore, differentiation protocols are still not completely efficient and often lead to the generation of polyhormonal cells ($INS^+GCG^+SST^+$) (Pagliuca et al., 2014). To improve the efficiency in order to produce more functional β-cells from hESCs *in vitro* several options have been explored, such as developing a protocol for enrichment of $PDX1^+/NKX6.1^+$ PP population by replating at various densities (Memon et al., 2018), exploring a wide spectrum of factors required for maturation and β-cell commitment (Rezania et al., 2014) along with recapitulating endocrine cell clustering *in vitro* (Nair et al., 2019). However, even though a lot of progress has been made in this field, additional improvements are still needed to finally obtain fully functional β-cells that can respond to glucose challenges as β-cells *in vivo*.

To assess different conditions that could lead to a higher maturation rate and increase the efficiency of the existing protocols, several different reporter lines have been generated in the Gavalas lab. These lines will enable easier and quicker evaluation of differentiation efficiency as well as the assessment of derived β-cells functionality and maturation. In particular, H1 *INS-MAFA* reporter line described in this thesis allows us to follow the expression of insulin, the main endocrine hormone produced and secreted by β-cells, as well as MAFA that is a crucial TF for the acquisition and maintenance of mature β-cell phenotype (Nishimura et al., 2009, 2015). Thus, the biggest advantage of this double reporter line is to allow the evaluation of both the differentiation efficiency and maturation *in vitro* as well as *in vivo* by longitudinal live imaging after transplantation in the anterior chamber of the mouse eye (Cohrs et al., 2020).

H1 *INS-MAFA* line was generated using the CRISPR-Cas9 technology by precisely knocking in the GFP and mCherry reporter cassettes downstream of the endogenous *INS* and *MAFA* genes. So, the main aim of this research was the H1 *INS-MAFA* line characterization to assess its differentiation potential. First of all, hESCs were evaluated for their unlimited proliferative ability and differentiation capacity. In fact, it was shown that H1 *INS-MAFA* cells still expressed the known pluripotency markers OCT4 and SOX2 in a similar manner to the H1 cells, demonstrating that hESC identity was retained. With further differentiation, the cells were analyzed for the main stage-specific DE and PP markers.

Immunofluorescence stainings and flow cytometry analyses for both DE and PP cells, with additional RT-qPCRs for PP cells, confirmed that H1 *INS-MAFA* cells can successfully, reproducibly and efficiently differentiate into PPs that are stable and suitable for the subsequent differentiation to β-cells.

To further evaluate differentiation capacity and determine cell functionality, PP cells were clustered in micropatterned Aggrewells. In particular, the high expression of endocrine-specific markers, *NGN3* and *NEUROD1,* shows a proper differentiation to pancreatic endocrine precursors as their expression plays a pivotal role in driving and regulating insulin expression (Zhu et al., 2017). Apart from RT-qPCR, flow cytometry analyses for these two markers were also performed in the lab at the end of PEP stage and they confirmed the expression of NGN3 and NEUROD1 in the vast majority of the cells (data not shown).

In addition, expression of the main progenitor markers *PDX1*, *NKX6.1* and *SOX9* was verified throughout the differentiation. Specifically, the expression of *PDX1* was stable and comparable between PEP, S6 and S7 cells, whereas *NKX6.1* expression was slightly upregulated towards the end of the differentiation for both cell lines. High and stable co-expression of *PDX1* and *NKX6.1* is crucial for proper β-cell development since glucose-responsive monohormonal β-cell originate from progenitors co-expressing *PDX1* and *NKX6.1* (Aigha & Abdelalim, 2020). On the other hand, *SOX9* expression was drastically downregulated by the end of the S6 since its expression is then restricted only to the ductal lineage. Interestingly, the hepatic marker *AFP* expression was significantly reduced in every consecutive differentiation stage for both cell lines, while intestinal marker *CDX2* remained stably expressed. The impact of unwanted markers on differentiation efficiency and quality of derived β-cells still needs to be addressed, but it has been speculated that lower expression of *AFP* and *CDX2* in PP cells results in a better differentiation outcome. Hence, to reduce their expression in PP cells and to improve the differentiation efficiency as well as the quality of derived β-cells, new conditions are under test in the lab.

One of the main features of β-cells is insulin expression. Cells at the end of S6 already highly expressed *INS,* showing also an upregulation trend throughout S7 for both H1 and H1 *INS-MAFA* cells. However, *GCG* was highly expressed and even upregulated in the H1 *INS-MAFA* derived β-cells by the end of S7, suggesting the presence of α-cells or polyhormonal β-cells. In fact, previous research demonstrated that around 10% of derived β-cells are actually polyhormonal (Rezania et al., 2014), and this was also confirmed in the lab by flow cytometry analyses of live single cells derived from the H1 *INS-eGFP/GCG-mCherry,* another reporter line generated in the lab. On the other hand, human islets are composed of 50% β-cells, 40% α-cells and 10% δ-cells (Brissova et al., 2005), justifying the presence of glucagon producing cells. Thus, high glucagon content, especially in the H1 *INS-MAFA* derived cells, has to be clarified by immunostaining on cryosections at the end of S7.

Apart from the markers mentioned above, β-cells were analyzed for the expression of functionality related genes. RT-qPCR results indicate that *GK*, *SUR1* and *ZnT8* were highly expressed in both H1 and H1 *INS-MAFA* cells suggesting that components of signaling pathways important for

specific β-cell function are present. Moreover, the maturation marker *MAFA* was also expressed and comparable between the H1 and H1 *INS-MAFA* derived cells. However, the lower expression observed in the H1 *INS-MAFA* derived cells can be explained by the fact that RT-qPCR primers bind between the last codon and the polyA tail where, in the targeted allele, the mCherry reporter is introduced. Some researchers argue that low levels of *MAFA* are expected *in vitro* since in human islets *MAFA* starts to be expressed in puberty, whereas *MAFB* is considered as the TF indispensable for early functions (Arda et al., 2016; Nair et al., 2019). On the other hand, some researchers imply that *MAFA* is essential for early maturation *in vitro*. Hence, molecules that have the potential of inducing *MAFA* expression *in vitro* comparably to human islets have been identified. Those compounds include tyrosine kinase receptor AXL inhibitor R428 and the antioxidant N-acetyl cysteine that promotes the stability of the MAFA protein (Rezania et al., 2014). However, even though the differentiation media contains these molecules, the *MAFA* levels remained low relative to the levels reported for human islets, suggesting that there are other pathways involved in the *MAFA* induction. An inducible MAFA line has been developed in the lab to identify the downstream targets of MAFA, allowing then to introduce new small molecules in the media that can reproduce its effects and obtain more mature β-cells. Nevertheless, the clear function of *MAFA* and *MAFB* in cell differentiation *in vitro* and development *in vivo* is yet to be investigated.

Gene expression analyses are not necessarily indicative of the functionality of cells after differentiation. Therefore, the next step in the characterization of derived β-cells was performing GSIS followed by human C-peptide ELISA assay to evaluate their full functional potential reflected in rapid response to high glucose concentrations. In brief, insulin is synthesized and secreted primarily in response to glucose, with an increase in intracellular Ca^{2+} being the primary insulin secretory signal. Specifically, the biosynthesis of insulin starts with translation in preproinsulin form and further processing to proinsulin in the endoplasmic reticulum. Proinsulin is then transported to the Golgi apparatus where it enters into secretary vesicles. In the vesicles, it is additionally cleaved to yield insulin and C-peptide. Thus, C-peptide is stored and secreted in equimolar concentrations as insulin, but its higher stability makes it a better target to analyze insulin secretion (Fu et al., 2013). Within a one-hour time frame *in vitro* under 16.7 mM glucose static incubation stimulation, higher C-peptide release was observed compared to basal 2.8 mM glucose stimulation. The determined stimulation index was lower when compared to the one from Rezania et al. (2014), who observed stimulation similar in magnitude to that of adult human islets. On the other hand, direct KCl-induced depolarization event revealed the rapid response indicating the presence of a releasable insulin pool. Nevertheless, H1 and H1 *INS-MAFA* derived cells demonstrated slight differences in the magnitude of response. Due to the fact that only two independent differentiations were included in the analysis, more replicas are needed to determine if that difference is significant. Also, cadaveric human islets should be included in the analysis as a control. Overall, these results imply that H1 and H1 *INS-MAFA* derived β-cells can partially respond to glucose indicating their functionality, but further improvements are still necessary. Additionally, more sensitive assays such as perfusion analysis assessing the magnitude of response and single-cell imaging focusing

on Ca^{2+} signaling have to be performed in the future to determine possible differences that cannot be measured with static glucose stimulation assays.

Lastly, the functionality of the integrated GFP and mCherry reporters was evaluated along with differentiation efficiency analysis. Several studies assessing differentiation efficiency based on insulin expression are already published. For example, one study showed that around 50% of a total S7 population are insulin-positive cells, with 1/5 of those also expressing glucagon. That suggests that their protocol yields one insulin$^+$ cell from every two starting hESC (Rezania et al., 2014). Supporting that, Melton's lab demonstrated that around 50% of derived S7 cells are C-peptide positive (Pagliuca et al., 2014). Strikingly, flow cytometry analysis of H1 *INS-MAFA* cells showed that more than 70% of derived β-cells are GFP$^+$ suggesting that our differentiation protocol has higher efficiency. The major difference in our protocol is the usage of Aggrewell plates for the 3D culture, whereas other published protocols rely mostly on air-liquid interface culture system. The use of Aggrewell plates has firstly been described by Hebrok lab that used them for immature β-cells reaggreagtion to achieve endocrine enrichment (Nair et al., 2019), but their usage for differentiation from PP stage to β-cells has not been reported. Therefore, this data suggest that the Aggrewells may help in increasing the differentiation efficiency. On the downside, mCherry signal was not detected by flow cytometry analysis. However, there are no publications that assess MAFA expression in S7 cells by flow cytometry probably due to the low expression levels, as demonstrated by RT-qPCR. Nonetheless, by cluster imaging, a few GFP$^+$ / mCherry$^+$ cells were identified confirming the functionality of both reporters. However, additional immunostaining on cryosections for both GFP and mCherry will be performed to inspect if GFP$^+$ cells are in fact INS$^+$ and if mCherry$^+$ cells are also positive for MAFA.

As already discussed, several points have to be considered and addressed in the future. Firstly, some experiments, particularly GSIS assays, still need more replicas to assess if the observed differences are line-specific or due to variability of the differentiations. In fact, new conditions for the differentiation, particularly for the last three stages, are under testing in the lab to improve the efficiency as well as to increase the reproducibility. Furthermore, gene expression analysis and GSIS have to be done on several batches of cadaveric human islets in order to confirm that the differentiated clusters are comparable to human adult islets. Also, functionality has to be more precisely determined by using highly sensitive assays including dynamic GSIS and single-cell imaging. Finally, and most importantly, this line will be used for following and improving the maturation of β-cells *in vitro* along with several other lines generated in the lab. Small molecules and factors for the later stages of differentiation will be tested to increase *MAFA* expression, easily followed with the mCherry reporter. Apart from *in vitro* analyses, the clusters will be used for *in vivo* experiments to confirm if they retain their mature phenotype and function after transplantation.

6 Conclusion

In summary, the reported findings suggest that the H1 *INS-MAFA* cells maintain their pluripotency and can successfully differentiate to PP cells similarly to the parental H1 cells. Furthermore, H1 *INS-MAFA* cells form uniform clusters and can efficiently differentiate to endocrine progenitors and further to β-cells in a comparable manner to H1 cells. Moreover, H1 and H1 *INS-MAFA* derived β-cells express functionality and maturation related genes and can respond to glucose challenges *in vitro*. In addition, it was demonstrated that integrated GFP and mCherry reporters are fully functional. In conclusion, the novel H1 *INS-MAFA* reporter cell line can be used as a tool for testing new conditions to optimize the differentiation protocol that will potentially lead to increased maturation and generation of fully functional β-cells *in vitro*.

7 References

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nature Communications*, *9*(1), 1911.

Afelik, S., Qu, X., Hasrouni, E., Bukys, M. A., Deering, T., Nieuwoudt, S., Rogers, W., MacDonald, R. J., & Jensen, J. (2012). Notch-mediated patterning and cell fate allocation of pancreatic progenitor cells. *Development*, *139*(10), 1744–1753.

Ahlgren, U., Jonsson, J., & Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development, 122*(5), 1409–1416.

Aigha, I. I., & Abdelalim, E. M. (2020). NKX6.1 transcription factor: A crucial regulator of pancreatic β cell development, identity, and proliferation. *Stem Cell Research & Therapy*, *11*(1), 459.

American Diabetes Association. (2014). Diagnosis and classification of diabetes mellitus. *Diabetes Care*, *37*, S81–S90.

Arda, H. E., Li, L., Tsai, J., Torre, E. A., Rosli, Y., Peiris, H., Spitale, R. C., Dai, C., Gu, X., Qu, K., Wang, P., Wang, J., Grompe, M., Scharfmann, R., Snyder, M. S., Bottino, R., Powers, A. C., Chang, H. Y., & Kim, S. K. (2016). Age-dependent pancreatic gene regulation reveals mechanisms governing human β cell function. *Cell Metabolism*, *23*(5), 909–920.

Banan, M. (2020). Recent advances in CRISPR/Cas9-mediated knock-ins in mammalian cells. *Journal of Biotechnology*, *308*, 1–9.

Banting, F. G., & Best, C. H. (1990). Pancreatic extracts. 1922. *The Journal of Laboratory and Clinical Medicine*, *115*(2), 254–272.

Blöchinger, A. K., Siehler, J., Wißmiller, K., Shahryari, A., Burtscher, I., & Lickert, H. (2020). Generation of an INSULIN-H2B-Cherry reporter human iPSC line. *Stem Cell Research*, *45*, 101797.

Boj, S. F., Parrizas, M., Maestro, M. A., & Ferrer, J. (2001). A transcription factor regulatory circuit in differentiated pancreatic cells. *Proceedings of the National Academy of Sciences*, *98*(25), 14481–14486.

Bommer, C., Sagalova, V., Heesemann, E., Manne-Goehler, J., Atun, R., Bärnighausen, T., Davies, J., & Vollmer, S. (2018). Global economic burden of diabetes in adults: projections from 2015 to 2030. *Diabetes Care*, *41*(5), 963–970.

Bramswig, N. C., & Kaestner, K. H. (2011). Transcriptional regulation of α-cell differentiation. *Diabetes, Obesity and Metabolism*, *13*(s1), 13–20.

Brissova, M., Fowler, M. J., Nicholson, W. E., Chu, A., Hirshberg, B., Harlan, D. M., & Powers, A. C. (2005). Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *The Journal of Histochemistry and Cytochemistry*, *53*(9), 1087–1097.

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., & van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science, 321*(5891), 960–964.

Brovkina, O., & Dashinimaev, E. (2020). Advances and complications of regenerative medicine in diabetes therapy. *PeerJ*, *8*, e9746.

Butler, A. E., Janson, J., Bonner-Weir, S., Ritzel, R., Rizza, R. A., & Butler, P. C. (2003). Cell deficit and increased cell apoptosis in humans with type 2 diabetes. *Diabetes*, *52*(1), 102–110.

Cai, E. P., Ishikawa, Y., Zhang, W., Leite, N. C., Li, J., Hou, S., Kiaf, B., Hollister-Lock, J., Yilmaz, N. K., Schiffer, C. A., Melton, D. A., Kissler, S., & Yi, P. (2020). Genome-scale in vivo CRISPR screen identifies RNLS as a target for beta cell protection in type 1 diabetes. *Nature Metabolism*, *2*(9), 934–945.

Canivell, S., & Gomis, R. (2014). Diagnosis and classification of autoimmune diabetes mellitus. *Autoimmunity Reviews*, *13*(4–5), 403–407.

Chen, Y., Pan, F. C., Brandes, N., Afelik, S., Sölter, M., & Pieler, T. (2004). Retinoic acid signaling is essential for pancreas development and promotes endocrine at the expense of exocrine cell differentiation in Xenopus. *Developmental Biology*, *271*(1), 144–160.

Cohrs, C. M., Chen, C., & Speier, S. (2020). Transplantation of islets of Langerhans into the anterior chamber of the eye for longitudinal in vivo imaging. *Methods in Molecular Biology*, *2128*, 149–157.

Collombat, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes & Development*, *17*(20), 2591–2603.

D'Amour, K. A., Bang, A. G., Eliazer, S., Kelly, O. G., Agulnick, A. D., Smart, N. G., Moorman, M. A., Kroon, E., Carpenter, M. K., & Baetge, E. E. (2006). Production of pancreatic hormone–expressing endocrine cells from human embryonic stem cells. *Nature Biotechnology*, *24*(11), 1392–1401.

Davidson, H. W., Wenzlau, J. M., & O'Brien, R. M. (2014). Zinc transporter 8 (ZnT8) and β cell function. *Trends in Endocrinology and Metabolism: TEM*, *25*(8), 415–424.

Desai, T., & Shea, L. D. (2017). Advances in islet encapsulation technologies. *Nature Reviews Drug Discovery*, *16*(5), 338–350.

Devendra, D., Liu, E., & Eisenbarth, G. S. (2004). Type 1 diabetes: Recent developments. *BMJ*, *328*(7442), 750–754.

Dor, Y., Brown, J., Martinez, O. I., & Melton, D. A. (2004). Adult pancreatic β-cells are formed by self-duplication rather than stem-cell differentiation. *Nature*, *429*(6987), 41–46.

Evans, M., Schumm-Draeger, P. M., Vora, J., & King, A. B. (2011). A review of modern insulin analogue pharmacokinetic and pharmacodynamic profiles in type 2 diabetes: Improvements and limitations. *Diabetes, Obesity and Metabolism*, *13*(8), 677–684.

Feng, G., Mellor, R. H., Bernstein, M., Keller-Peck, C., Nguyen, Q. T., Wallace, M., Nerbonne, J. M., Lichtman, J. W., & Sanes, J. R. (2000). Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron*, *28*(1), 41–51.

Forbes, J. M., & Cooper, M. E. (2013). Mechanisms of diabetic complications. *Physiological Reviews*, *93*(1), 137–188.

Fu, Z., Gilbert, E. R., & Liu, D. (2013). Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes. *Current Diabetes Reviews*, *9*(1), 25–53.

Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, *31*(7), 397–405.

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. H., & Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67–71.

Gittes, G. K. (2009). Developmental biology of the pancreas: A comprehensive review. *Developmental Biology*, *326*(1), 4–35.

Gloyn, A. L., Siddiqui, J., & Ellard, S. (2006). Mutations in the genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) in diabetes mellitus and hyperinsulinism. *Human Mutation*, *27*(3), 220–231.

Gouzi, M., Kim, Y. H., Katsumoto, K., Johansson, K., & Grapin-Botton, A. (2011). Neurogenin3 initiates stepwise delamination of differentiating endocrine cells during pancreas development. *Developmental Dynamics*, *240*(3), 589–604.

Gudala, K., Bansal, D., Schifano, F., & Bhansali, A. (2013). Diabetes mellitus and risk of dementia: A meta-analysis of prospective observational studies. *Journal of Diabetes Investigation*, *4*(6), 640–650.

Hentze, H., Soong, P. L., Wang, S. T., Phillips, B. W., Putti, T. C., & Dunn, N. R. (2009). Teratoma formation by human embryonic stem cells: Evaluation of essential parameters for future safety studies. *Stem Cell Research*, *2*(3), 198–210.

Jaafar, R., Tran, S., Shah, A. N., Sun, G., Valdearcos, M., Marchetti, P., Masini, M., Swisa, A., Giacometti, S., Bernal-Mizrachi, E., Matveyenko, A., Hebrok, M., Dor, Y., Rutter, G. A., Koliwad, S. K., & Bhushan, A. (2019). MTORC1-to-AMPK switching underlies β cell metabolic plasticity during maturation and diabetes. *Journal of Clinical Investigation*, *129*(10), 4124–4137.

Jansen, R., Embden, J. D. A. van, Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), 1565–1575.

Jin, L., Long, L., Green, M. A., & Spear, B. T. (2009). The alpha-fetoprotein enhancer region activates the albumin and alpha-fetoprotein promoters during liver development. *Developmental Biology*, *336*(2), 294–300.

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*(6096), 816–821.

Johansson, K. A., Dursun, U., Jordan, N., Gu, G., Beermann, F., Gradwohl, G., & Grapin-Botton, A. (2007). Temporal control of Neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Developmental Cell*, *12*(3), 457–465.

Jurgielewicz, P., Harmsen, S., Wei, E., Bachmann, M. H., Ting, R., & Aras, O. (2017). New imaging probes to track cell fate: Reporter genes in stem cell research. *Cellular and Molecular Life Sciences: CMLS*, *74*(24), 4455–4469.

Kawaguchi, Y., Cooper, B., Gannon, M., Ray, M., MacDonald, R. J., & Wright, C. V. E. (2002). The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nature Genetics*, *32*(1), 128–134.

Keller, G. M. (1995). In vitro differentiation of embryonic stem cells. *Current Opinion in Cell Biology*, *7*(6), 862–869.

Kharroubi, A. T. (2015). Diabetes mellitus: The epidemic of the century. *World Journal of Diabetes*, *6*(6), 850.

Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J., Agulnick, A. D., D'Amour, K. A., Carpenter, M. K., & Baetge, E. E. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nature Biotechnology*, *26*(4), 443–452.

Li, S., Chen, L.-X., Peng, X.-H., Wang, C., Qin, B.-Y., Tan, D., Han, C.-X., Yang, H., Ren, X.-N., Liu, F., Xu, C.-H., & Zhou, X.-H. (2018). Overview of the reporter genes and reporter mouse models. *Animal Models and Experimental Medicine*, *1*(1), 29–35.

Liu, B., Saber, A., & Haisma, H. J. (2019). CRISPR/Cas9: A powerful tool for identification of new targets for cancer treatment. *Drug Discovery Today*, *24*(4), 955–970.

Mahaddalkar, P. U., Scheibner, K., Pfluger, S., Ansarullah, Sterr, M., Beckenbauer, J., Irmler, M., Beckers, J., Knöbel, S., & Lickert, H. (2020). Generation of pancreatic β cells from CD177+ anterior definitive endoderm. *Nature Biotechnology*, *38*(9), 1061–1072.

Melton, D. (2021). The promise of stem cell-derived islet replacement therapy. *Diabetologia*, *64*(5), 1030–1036.

Memon, B., Karam, M., Al-Khawaga, S., & Abdelalim, E. M. (2018). Enhanced differentiation of human pluripotent stem cells into pancreatic progenitors co-expressing PDX1 and NKX6.1. *Stem Cell Research & Therapy*, *9*(1), 15.

Mojica, F. J., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, *36*(1), 244–246.

Mølbak, A. G., Christau, B., Marner, B., Borch-Johnsen, K., & Nerup, J. (1994). Incidence of insulin-dependent diabetes mellitus in age groups over 30 years in Denmark. *Diabetic Medicine*, *11*(7), 650–655.

Murtaugh, L. C. (2007). Pancreas and beta-cell development: From the actual to the possible. *Development*, *134*(3), 427–438.

Nair, G. G., Liu, J. S., Russ, H. A., Tran, S., Saxton, M. S., Chen, R., Juang, C., Li, M., Nguyen, V. Q., Giacometti, S., Puri, S., Xing, Y., Wang, Y., Szot, G. L., Oberholzer, J., Bhushan, A., & Hebrok, M. (2019). Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells. *Nature Cell Biology*, *21*(2), 263–274.

Nishimura, W., Bonner-Weir, S., & Sharma, A. (2009). Expression of MafA in pancreatic progenitors is detrimental for pancreatic development. *Developmental Biology*, *333*(1), 108–120.

Nishimura, W., Kondo, T., Salameh, T., El Khattabi, I., Dodge, R., Bonner-Weir, S., & Sharma, A. (2006). A switch from MafB to MafA expression accompanies differentiation to pancreatic β-cells. *Developmental Biology*, *293*(2), 526–539.

Nishimura, W., Takahashi, S., & Yasuda, K. (2015). MafA is critical for maintenance of the mature beta cell phenotype in mice. *Diabetologia*, *58*(3), 566–574.

Nouwen, A., Nefs, G., Caramlau, I., Connock, M., Winkley, K., Lloyd, C. E., Peyrot, M., Pouwer, F., & for the European Depression in Diabetes (EDID) Research Consortium. (2011). Prevalence of depression in individuals with impaired glucose metabolism or undiagnosed diabetes: A systematic review and meta-analysis of the European Depression in Diabetes (EDID) Research Consortium. *Diabetes Care*, *34*(3), 752–762.

Pagliuca, F. W., Millman, J. R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J. H., Peterson, Q. P., Greiner, D., & Melton, D. A. (2014). Generation of functional human pancreatic β Cells in vitro. *Cell*, *159*(2), 428–439.

Pathak, V., Pathak, N. M., O'Neill, C. L., Guduric-Fuchs, J., & Medina, R. J. (2019). Therapies for type 1 diabetes: Current scenario and future perspectives. *Clinical Medicine Insights: Endocrinology and Diabetes*, *12*, 117955141984452.

Pictet, R. L., Clark, W. R., Williams, R. H., & Rutter, W. J. (1972). An ultrastructural analysis of the developing embryonic pancreas. *Developmental Biology*, *29*(4), 436–467.

Puri, M. C., & Nagy, A. (2012). Concise Review: embryonic stem cells versus induced pluripotent stem cells: the game is on. *Stem cells*, *30*(1), 10–14.

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. C., Johnson, J. D., & Kieffer, T. J. (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nature Biotechnology*, *32*(11), 1121–1133.

Rezania, A., Bruin, J. E., Riedel, M. J., Mojibian, M., Asadi, A., Xu, J., Gauvin, R., Narayan, K., Karanu, F., O'Neil, J. J., Ao, Z., Warnock, G. L., & Kieffer, T. J. (2012). Maturation of human embryonic stem cell–derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. *Diabetes*, *61*(8), 2016–2029.

Rickels, M. R., & Robertson, R. P. (2019). Pancreatic islet transplantation in humans: Recent progress and future directions. *Endocrine Reviews*, *40*(2), 631–668.

Rizzino, A., & Wuebben, E. L. (2016). Sox2/Oct4: A delicately balanced partnership in pluripotent stem cells and embryogenesis. *Biochimica Et Biophysica Acta*, *1859*(6), 780–791.

Rosenbloom, A. L., Silverstein, J. H., Amemiya, S., Zeitler, P., & Klingensmith, G. J. (2009). Type 2 diabetes in children and adolescents. *Pediatric Diabetes*, *10*, 17–32.

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. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice*, *157*, 107843.

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 Research*, *39*(21), 9275–9282.

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*, *104*(6), 1865–1870.

Seymour, P. A., Shih, H. P., Patel, N. A., Freude, K. K., Xie, R., Lim, C. J., & Sander, M. (2012). A Sox9/Fgf feed-forward loop maintains pancreatic organ identity. *Development*, *139*(18), 3363–3372.

Shih, H. P., Wang, A., & Sander, M. (2013). Pancreas organogenesis: From lineage determination to morphogenesis. *Annual Review of Cell and Developmental Biology*, *29*(1), 81–105.

Shimomura, O., Johnson, F. H., & Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology*, *59*, 223–239.

Shirakawa, J., & Terauchi, Y. (2020). Newer perspective on the coupling between glucose-mediated signaling and β-cell functionality. *Endocrine Journal*, *67*(1), 1–8.

Silberg, D. G., Swain, G. P., Suh, E. R., & Traber, P. G. (2000). Cdx1 and cdx2 expression during intestinal development. *Gastroenterology*, *119*(4), 961–971.

Skyler, J. S., Bakris, G. L., Bonifacio, E., Darsow, T., Eckel, R. H., Groop, L., Groop, P.-H., Handelsman, Y., Insel, R. A., Mathieu, C., McElvaine, A. T., Palmer, J. P., Pugliese, A., Schatz, D. A., Sosenko, J. M., Wilding, J. P. H., & Ratner, R. E. (2017). Differentiation of diabetes by pathophysiology, natural history, and prognosis. *Diabetes*, *66*(2), 241–255.

Subach, F. V., & Verkhusha, V. V. (2012). Chromophore transformations in red fluorescent proteins. *Chemical Reviews*, *112*(7), 4308–4327.

Szabat, M., Lynn, F. C., Hoffman, B. G., Kieffer, T. J., Allan, D. W., & Johnson, J. D. (2012). Maintenance of cell maturity and plasticity in the adult pancreas: Developmental biology concepts in adult physiology. *Diabetes*, *61*(6), 1365–1371.

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, *126*(4), 663–676.

Thomson, J. A. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, *282*(5391), 1145–1147.

Tomimaru, Y., Ito, T., Marubashi, S., Kawamoto, K., Tomokuni, A., Asaoka, T., Wada, H., Eguchi, H., Mori, M., Doki, Y., & Nagano, H. (2015). De novo malignancy after pancreas transplantation in Japan. *Transplantation Proceedings*, *47*(3), 742–745.

Veres, A., Faust, A. L., Bushnell, H. L., Engquist, E. N., Kenty, J. H.-R., Harb, G., Poh, Y.-C., Sintov, E., Gürtler, M., Pagliuca, F. W., Peterson, Q. P., & Melton, D. A. (2019). Charting cellular identity during human in vitro β-cell differentiation. *Nature*, *569*(7756), 368–373.

Watanabe, H., Takayama, K., Inamura, M., Tachibana, M., Mimura, N., Katayama, K., Tashiro, K., Nagamoto, Y., Sakurai, F., Kawabata, K., Furue, M. K., & Mizuguchi, H. (2014). HHEX promotes hepatic-lineage specification through the negative regulation of eomesodermin. *PloS One*, *9*(3), e90791.

Zhong, A., Li, M., & Zhou, T. (2020). Protocol for the generation of human pluripotent reporter cell lines using CRISPR/Cas9. *STAR Protocols*, *1*(2), 100052.

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., & Melton, D. A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to β-cells. *Nature*, *455*(7213), 627–632.

Zhu, Y., Liu, Q., Zhou, Z., & Ikeda, Y. (2017). PDX1, Neurogenin-3, and MAFA: Critical transcription regulators for beta cell development and regeneration. *Stem Cell Research & Therapy*, *8*(1), 240.

Zhu, Z., & Huangfu, D. (2013). Human pluripotent stem cells: An emerging model in developmental biology. *Development*, *140*(4), 705–717.

8 Supplementary data

Table 8.1 Formulation for basal PEP-S7 differentiation medium. All the compounds used, along with manufacturer and catalog number are listed.

Basal PEP-S7 differentiation medium		
Compound	**Manufacturer**	**Cat #**
MCDB medium w/o glutamine	Gibco	10372-019
14.5 mM Glucose	Sigma	G8769
1.5 g/L NaHCO3	Sigma	S6297
2% BSA	LSP	BSA-68700
1x Glutamax	Gibco	3505-038
1x Pen/Strep	Gibco	15140-122
0.5x ITS-X	Gibco	51500-056
1.8 U/mL Heparin	Sigma	2106-10VL

Table 8.2 Transcription factors and small molecules supplemented in basal PEP-S7 differentiation medium. All the compounds used for obtaining PEP, S6 and S7 media, along with their final concentration, manufacturer and catalog number are listed.

PEP medium			
Factor	**Final conc (µM)**	**Manufacturer**	**Cat #**
Sant1	0.25	Sigma	(S4572)
RA	0.05	Sigma	R2625
LDN	0.1	Sigma	SML0559
ALK5i	10	MACS Militenybiotec	130-117-340
T3	1	Sigma	T6397
Zn	10	Sigma	Z0251
S6 medium			
Factor	**Final conc (µM)**	**Manufacturer**	**Cat #**
XXi	0.1	Milipore	565790
LDN	0.1	Sigma	SML0559
ALK5i	10	MACS Militenybiotec	130-117-340
T3	1	Sigma	T6397
Zn	10	Sigma	Z0251
S7 medium			
Factor	**Final conc (µM)**	**Manufacturer**	**Cat #**
Trolox	10	Milipore	648471
R428	2	Selleckchem	S2841
NAC	1000	Sigma	A9165
ALK5i	10	MACS Militenybiotec	130-117-340
T3	1	Sigma	T6397
Zn	10	Sigma	Z0251

Table 8.3 Antibody information for immunofluorescent staining on coverslips. All the target antigens, along with primary antibody species, dilution, manufacturer, and secondary antibody are listed.

Stage	Antigen	pAb species	Dilution	Manufacturer	sAb	Fluorophore	Dilution	Manufacturer
hESC	OCT4	Mouse	1:100	Santa Cruz	goat anti-mouse	Alexa-568	1:500	Molecular probes
	SOX2	Rabbit	1:200	Abcam	goat anti-rabbit	Alexa-488	1:500	Molecular probes
DE	SOX17	Mouse	1:50	R&D Systems	goat anti-mouse	Alexa-568	1:500	Molecular probes
	FOXA2	Rabbit	1:200	Merck	goat anti-rabbit	Alexa-488	1:500	Molecular probes
PP	PDX1	Goat	1:40	R&D Systems	donkey anti-goat	Alexa-647	1:500	Molecular probes
	NKX6.1	Mouse	1:500	DSHB	donkey anti-mouse	Alexa-568	1:500	Life Technologies
	SOX9	Rabbit	1:1000	Millipore	donkey anti-rabbit	Alexa-488	1:500	Life Technologies

Table 8.4 Antibody information for immunostaining for flow cytometry. Unconjugated primary antibodies, along with dilution, manufacturer, and secondary antibodies are listed.

Stage	pAb	Dilution	Manufacturer	sAb	Fluorophore	Dilution	Manufacturer
DE	mouse anti-SOX17	1:50	R&D Systems	goat anti-mouse	Alexa-568	1:500	Molecular probes
	rabbit anti-FOXA2	1:200	Merck	goat anti-rabbit	Alexa-488	1:500	Molecular probes
PP	goat anti-PDX1	1:40	R&D Systems	donkey anti-goat	Alexa-647	1:500	Molecular probes
	mouse anti-NKX6.1	1:250	DSHB	donkey anti-mouse	Alexa-568	1:500	Life Technologies
	rabbit anti-SOX9	1:1000	Millipore	donkey anti-rabbit	Alexa-488	1:500	Life Technologies

Table 8.5 Primers used for RT-qPCR. Target genes for each differentiation stage analyzed, along with forward and reverse primer sequences, are listed.

Stage	Gene	Primer	Sequence (5'-3')
PP-S7	*hTBP*	Fw	TATCGTGGTCACACTGTTGAG
		Rev	ACCATCCTCTGAGACCGTTTT
	hPDX1	Fw	GGAGAAAGATGGACCCCTGG
		Rev	CAGCCTCTACCTCGGAACAG
	hNKX6.1	Fw	ATTCGTTGGGGATGACAGAG
		Rev	CGAGTCCTGCTTCTTCTTGG
	hAFP	Fw	GCAGCAGTCTGAATGTCCGTAC
		Rev	TGCCCAGTTTGTTCAAGAAGC
	hCDX2	Fw	GCTGGAGAAGGAGTTTCACTACAGT
		Rev	AACCAGATTTTAACCTGCCTCTCA
PP-S6	*hSOX9*	Fw	CGGAGGAAGTCGGTGAAG
		Rev	CCTTGAAGATGGCGTTGG
PP	*hFOXA2*	Fw	ACACCACTACGCCTTCAACC
		Rev	CCTTGAGGTCCATTTTGTGG
	hPTF1A	Fw	TTCACCGACCAGTCTTCACG
		Fw	GTGGCTAAGGAACTCCACCT
	hHHEX	Rev	ACGGTGAACGACTACACGC
		Fw	CTTCTCCAGCTCGATGGTCT
PEP	*hNGN3*	Rev	TGGGTGCTAAGGGTAAGGGA
		Fw	CAGCCAGGGAGAAGCAGAAG
	hNEUROD1	Rev	AGGCAGCCCTTTGGGTACTA
		Fw	TTGATCCCCTGTTTCTTCCA
S6-S7	*hINS*	Rev	AGATCACTGTCCTTCTGCCA
		Fw	CGCACAGGTGTTGGTTCA
	hGCG	Rev	GAGACATGCTGAAGGGACCT
		Fw	CTTCCTCGGCCTTTCACCAG
	hSST	Fw	AGCTGCTGTCTGAACCCAAC
		Rev	GCTCAAGCCTCATTTCATCC
S7	*hMAFA*	Fw	GCGGAGAACGGTGATTTCTA
		Rev	GAAGGTGGGAACGGAGAAC
	hGK	Fw	AGCGTGAAGACCAAACACCA
		Rev	ATGCTTGTCCAGGAAGTCGG
	hSUR1	Fw	TGAGAGCGAGGAGGATGACA
		Rev	AAGACCAGCAACGACAGGAG
	hZnT8	Fw	GGCCGTCATGGAGTTTCTT
		Rev	CACCGGTTTCTGTTGGAGTT

Curriculum vitae

I was born on 20th of December 1996 in Zagreb, Croatia. After finishing my high school education in the First Gymnasium in Zagreb in 2015, I enrolled in the Undergraduate program of Molecular biology at the Faculty of Science, University of Zagreb. In the academic year of 2016/2017, I was a laboratory demonstrator for an undergraduate class of Zoology. In 2017, I did laboratory practice at Ruđer Bošković Institute in the Laboratory for chemical biology under the supervision of dr. Branka Salopek Sondi as a part of PhytoBraCRo project. In 2018, I obtained the title of bacc. biol. mol. (*cum laude*) and enrolled in the Molecular biology master program at the same university. In the academic year 2018/2019 I completed laboratory practice at Ruđer Bošković Institute in the Laboratory for mitochondrial bioenergetics and diabetes under the supervision of dr. Iva Škrinjar as a part of SuMERA project. In the following year, I took part in the Erasmus+ exchange and spent 6 months studying at Regenerative Biology and Medicine master program at Centre for Regenerative Therapies, TU Dresden, Germany. In the academic year 2020/2021, I pursued Erasmus+ traineeship at Paul Langerhans Institute Dresden in the Laboratory for stem cells in pancreas development and disease where I also conducted my thesis work under the supervision of dr. Anthony Gavalas. Throughout my studies, I participated in science popularization events such as Night of biology. I was a winner of the national STEM scholarship (2017/2018) and a winner of the city of Zagreb scholarship for excellence (2019/2020).