UNIVERSITY OF CALIFORNIA, SAN DIEGO

Role of Pioneer Transcription Factors FOXA1 and FOXA2 during Pancreatic Specification of Human Embryonic Stem Cells.

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Biology

by

Dieter Ka Yeung Lam

Committee in charge:

Professor Maike Sander, Chair
Professor Kimberly Cooper, Co-Chair
Professor Gen-Sheng Feng

2016

The thesis of Dieter Ka Yeung Lam is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2016

DEDICATION

I dedicate this thesis to my parents, Andy and Diana, for their love and unwavering support.

I also dedicate this thesis to all of the close friends who have supported me in the recent years, especially:

Nicole and Daniel; for their wisdom and encouragement.

Calvin, Jay, Teresa, Nisha, and Tiffany; my time in graduate school was more meaningful and enjoyable because of these amazing people.

EPIGRAPH

“If at first you don’t succeed, redo your experiments. Yes, all of them.”

“You can’t fail if you don’t try.”
-Daniel Xu

“You also can’t succeed if you don’t try.”
-Nicole Choy

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to express my appreciation to Dr. Sander for giving me the opportunity to conduct research in her lab. I will always remember this lab as the beginning of my scientific career.

I would also like to thank Dr. Wang for his guidance and mentorship. He always encouraged me to challenge myself and played an integral role in the development of my research skills. His intellectual and professional influence will follow me wherever I go, and I am forever thankful.

Of course, I cannot forget Thomas C. Harper, M.S., for teaching me “the tools of the trade”, nor the other members of the Sander lab for their support and guidance. This is especially true of Fenfen, who first hired me into the lab has always been kind to me throughout the years.

Finally, I would like to thank the Juvenile Diabetes Research Foundation and the National Institutes of Health for funding our research.

ABSTRACT OF THE THESIS

Role of Pioneer Transcription Factors FOXA1 and FOXA2 during Pancreatic Specification of Human Embryonic Stem Cells.

by

Dieter Ka Yeung Lam

Master of Science in Biology

University of California, San Diego, 2016

Professor Maike Sander, Chair
Professor Kimberly Cooper, Co-chair

During pancreas development, undifferentiated cells must undergo stepwise progression toward particular cell fates. The emergence of the distinct cell types of the pancreas relies on the capacity of lineage intermediates to properly interpret and respond to environmental inductive cues, an ability termed developmental competence. Enhancers have been shown to be critical in spatiotemporal gene regulation during development. However, how enhancers acquire the ability to

translate signals from the extracellular environment into cell-type-specific transcriptional responses during development is poorly understood. Epigenetic priming and activation of enhancers by the pioneer transcription factors FOXA1 and FOXA2 has been proposed to confer developmental competence in various stages of pancreas lineage specification. We have previously shown that enhancers are first recognized by the pioneer transcription factors FOXA1 and FOXA2 when competence is acquired. To elucidate the roles and requirement of FOXA1 and FOXA2 for pancreas differentiation and enhancer priming, human embryonic stem cell (hESC) lines harboring shRNAs targeting *FOXA1* and *FOXA2* or were generated. Using differentiation of these hESC lines as a model for pancreas development, we employed qRT-PCR, ChIP-qPCR, and protein detection methods to address the effect, if any, the loss of these transcription factors have on pancreas differentiation. . As a follow up, CRIPSR/Cas9-mediated gene editing was also employed to generate *FOXA1*/2 knockout lines in order to serve as a reliable platform for future experiments.

I:

INTRODUCTION

A. Diabetes Mellitus and the Developing Pancreas

Type I Diabetes Mellitus is a debilitating endocrine disease that is characterized by the autoimmune loss of insulin-producing β-cells in the pancreas. The pancreas is a key player in metabolism that performs both exocrine and endocrine functions. While the exocrine pancreas is involved in digestion, the endocrine pancreas is responsible for the secretion of digestion-related hormones. The Islets of Langerhans within the pancreas is the epicenter of the endocrine pancreas, where endocrine cells such as α and β-cells reside and mediate glucose balance. Here, β-cells secrete insulin in response to glucose, whereby insulin invokes tissue to uptake glucose and store it as glycogen. In type I diabetes, autoimmune loss of these β-cells results in the impairment of proper insulin secretion, leading to chronic glucose imbalance and a host of complications that arise from it (Van Belle et al., 2011).

An effective and commonplace treatment of diabetes is the periodic replacement of insulin through injections using a syringe or insulin pump. While this treatment regimen is quite effective, it is very costly and an inconvenience for the patients who suffer from this disease (American Diabetes Association, 2002). This calls for a more permanent and reliable treatment solution that is less burdensome. One promising alternative is β-cell replacement therapy, whereby pancreatic β-cells or β-cell progenitors differentiated from human stem cells (hSCs) can be transplanted into type I diabetes patients.

Much work has been done to grow replacement β-cells *in vitro*. HSCs have recently been successfully differentiated towards functional pancreatic endocrine cells

capable of hormone secretion in response to normal signals. However, these efforts have not yet yielded homogenous populations of cells with the functional characteristics comparable to that of normal β-cells needed for stem cell therapy. For example, recent efforts have yielded functional β-cells that co-populate with other pancreatic cell types, which may not be desirable in a therapy approach that requires pure cell populations (Pagliuca et al., 2014; Rezania et al., 2014). Furthermore, the factors used for inducing β-cell differentiation in vitro are highly expensive, which may hinder the prospects of widespread stem cell therapy for the millions of patients who are affected by this disease. In order to improve the quality and efficiency of β-cell differentiation as well as the future outlook of stem cell therapy for the majority of diabetes patients, it is imperative that we elucidate the molecular mechanisms that underlie pancreas development.

B. Pancreas Differentiation of Human Embryonic Stem Cells

Current approaches to differentiate human embryonic stem cells towards pancreatic cells have been developed from and are based on the studies of pancreas development in model organisms such as mouse and zebrafish. In general, this process starts by the induction of embryonic stem cells to either one of the three major germ layers: endoderm, ectoderm, or mesoderm. For pancreas differentiation, embryonic stem cells are first differentiated to definitive endoderm (DE) germ layer cells. DE cells are then differentiated to primitive gut tube (GT) cells, where the primitive gut tube is an embryonic feature that eventually develops into the various endoderm-derived organs (Shih et al., 2013). The primitive gut tube can be

demarcated into three sections: the foregut, the midgut, and hindgut. The GT cells in the posterior portion of the foregut are precursors to cells of the hepatic and pancreatic lineages during development. Here, cells of the posterior foregut (PF) develop into pancreatic progenitors (PP) and eventually the pancreas, while cells of the other portions of the foregut develop into the liver. FOXA2 is expressed in the foregut, as well as HNF4A, which is considered a marker for this stage (D'Amour et al., 2006, Wang et al., 2012). A flowchart summarizing pancreas differentiation can be found in Appendix A.

Throughout these stages of differentiation, various cell types can be identified by the developmental markers they express. PDX1 is a hallmark marker of pancreas-lineage cells, having high levels of expression in both PF and PP cells. It is known to be a pancreas lineage determining transcription factor, in that it is crucial in cell fate specification toward pancreas (Yamamizu et al., 2013, Shih et al., 2013). Other critical regulators of the pancreatic fate are also expressed at the PF and PP stage such as PTF1A, NKX6.1 and SOX9. HNF4α, on the other hand, is a hallmark marker and lineage determining transcription factor for hepatic (liver) cells. (DeLaForest et al., 2011, Yamamizu et al., 2013).

C. Developmental Competence and FOXA Pioneer Transcription Factors

Certain transcription factors have been suggested to play key roles in establishing developmental competence. Developmental competence is defined as the ability of lineage intermediates to differentiate toward downstream cell types through proper interpretation and response to environmental inductive cues. One

proposed mechanism of developmental competence is through changes in chromatin modifications such as H3K4me1 and H3K27ac, allowing important genes to be activated during specific steps of development (Wang et al., 2015, Zaret, K., 1999, Friedman et al., 2006).

Multiple studies have implicated a class of transcription factors (TF) called “pioneer transcription factors” to play roles during this process. This class differs from traditional transcription factors by their ability to bind compacted chromatin without the need for cooperative binding with other factors. The interactions of these pioneer transcription factors result in either the “poising” or activation of condensed chromatin at developmentally relevant sites, such as at promoters or enhancers, opening them up or making them more accessible for subsequent factors to bind and activate gene expression during critical developmental steps (Cirillo et al., 2002, Zaret et al., 2011). Specifically, poised enhancers are marked by H3K4me1, while active enhancers are marked by both H3K4me1 and H3K27ac (Rada-Iglesias et al., 2011, Creyghton et al., 2010). Thus, it has been hypothesized that enhancers are poised prior to their activation by these pioneer transcription factors (Wang et al., 2015).

The FOXA family of transcription factors, which include FOXA1 and FOXA2, have been known to play important roles in the development of many endoderm-derived organs. In pancreas development, FOXA2 is highly induced at the DE stage at day 2 of pancreas differentiation followed by FOXA1 at the GT stage at day 5. Both factors continue to be expressed during the PF and PP stages immediately prior and during induction of the pancreatic lineage (refer to Fig. 1) (Friedman et al., 2006). FOXA1 and FOXA2 have been identified as prominent transcription factors in pancreas development, and they have been shown to exhibit high affinity binding to

chromatin (Cirillo et al., 2002, Zaret et al., 2011). Sequence motif analysis of poised pancreatic enhancers suggests that they have the potential to be regulated by FOXA1 and FOXA2. Furthermore, ChIP-sequencing of FOXA1 and FOXA2 in HSCs of early pancreas development indicated significant association with many types of these pancreatic enhancers (Wang et al., 2015*)*. This leads us to hypothesize that FOXA1 and FOXA2 act as pioneering factors to poise pancreatic enhancers for activation during pancreas development. Therefore, a major goal of this project is to identify the functional roles these two factors play during enhancer priming and pancreas development.

D. Short Hairpin RNA-mediated Knockdowns and CRISPR/Cas9-mediated Knockouts

A popular method to assess the functional requirement of transcription factors for biological processes, is to perform loss-of-function experiments via RNA interference (RNAi) mediated knockdown. In this study we employ short hairpin RNAs (shRNA). While shRNA delivery vectors are made externally and subsequently transduced, the shRNAs themselves are produced by the cellular machinery. shRNA primary transcripts are made in the nucleus, which are then processed as they make their into the cytoplasm. When associated with RISC and fully active, shRNAs bind and cleave their specific target gene transcript, marking them for further degradation by endonucleases. (Rao et al., 2009) Thus, shRNA-mediated RNAi knockdowns work at the post-transcriptional level. While this method is commonplace and effective, it is by no means a perfect system. shRNAs are easily degraded by RNases, and they

may sometimes be unable to keep up with transcripts that have high turnover rates. One better alternative is the deletion of target genes at the genomic level through gene editing.

An increasingly popular technique to test the functional requirement of genes is through the generation of homozygous null cell lines via the Clustered, Regularly Interspaced Palindromic Repeats/associated nuclease Cas9 (CRISPR/Cas9) gene editing system. The CRISPR system of gene editing was initially discovered as a defense mechanism against foreign genetic elements in prokaryotes (Bolotin et al., 2005, Mojica et al., 2005, Pourcel et al., 2005). This system has recently been utilized in eukaryotic cells to perform gene editing, including disruption or deletion of genes at the genomic level. The CRISPR/Cas9 gene editing system can be broken down to four major parts: the CRISPR associated Cas9 nuclease, a highly specific 20-nt guide CRISPR RNA (crRNA), an auxiliary trans-activating CRISPR RNA (tracrRNA), and the protospacer adjacent motif (PAM) sequence on the target genome. The crRNA and tracrRNA, as in the system used for this thesis, can be fused to generate a chimeric single guide RNA (sgRNA). When this sgRNA associates with the Cas9 nuclease, a single nucleotide cut can be made at the target site corresponding to the 17^{th} nucleotide in the guide sequence, provided it is upstream of a PAM sequence. (Hsu et al., 2013). In total, this method provides an invaluable tool for the performing true loss-of-function studies.

II:

RESULTS

Evaluation of shRNA mediated FOXA1 knockdown:

As mentioned previously, FOXA factors are highly expressed throughout pancreas development, and are proposed to play an integral role by facilitating the activation of pancreatic enhancers. To explore this, we employed shRNA-mediated RNAi of *FOXA1* in hopes of generating a knockdown *FOXA1* cell line. To do this, undifferentiated CyT49 hESCs were transduced with lentivirus containing either one of each of the eight shRNAs targeting *FOXA1* (FOXA1 sh1-8), a cocktail of four from sh1 to sh4, a cocktail of all eight (FOXA1 shAll), or a non-targeting control (shNTC). These CyT49 were then differentiated towards the pancreatic lineage, and aggregates were collected at the GT stage (day 5 of pancreatic differentiation). Expression levels of FOXA1 mRNA were quantified using qRT-PCR (Figure 1A), while FOXA1 protein levels were examined via western blot (Figure 1B). While the CyT49 cells infected with only one shRNA did not show any significant FOXA1 knockdown, the ones infected with cocktails of multiple shRNAs did. This was most notable in the shAll cells, where a four-fold knockdown of FOXA1 was observed. The shBot cells exhibited only a two-fold knockdown of FOXA1 (Figure 1A). Western blot of the shAll sample seems to confirm this, where an observable decrease of FOXA1 protein is observed relative to the shNTC control (Figure 1B).

Evaluation of shRNA mediated FOXA2 knockdown:

Similarly for FOXA2, undifferentiated CyT49 hESCs were transduced with lentivirus containing either one of each of the seven FOXA2 shRNAs (FOXA2 sh11-17), a cocktail of all seven (FOXA2 shStar), or a non-targeting control (shNTC). These CyT49 were then differentiated towards the pancreatic lineage, and

aggregates were collected at the DE stage, or day 2 of pancreatic differentiation. Expression levels of FOXA2 mRNA were quantified using qRT-PCR (Figure 2). Although there is an across-the-board decrease in FOXA2 mRNA expression for all samples, none of them were significant. The shStar cells displayed the greatest decrease of FOXA2 expression at around two-fold, but this was within the margin of error based on standard deviation of technical replicates.

Requirement of FOXA1 for Proper Pancreatic Expression:

To test the requirement of FOXA1 for proper expression of hESC-derived pancreatic cell types, another round of differentiation was performed with our newly produced knockdown FOXA1 cell line. (The shAll cells mentioned above were chosen as the cell line for all subsequent experiments requiring FOXA1 knockdown; "shFOXA1" shall henceforth refer to the shAll CyT49 hESC line.) These cells were differentiated to PP (day 10 of differentiation), and aggregates were collected from the GT, PF, and PP stages. Once again, qRT-PCR was used to quantify mRNA expression levels. *FOXA1* expression across these three stages exhibited only a two-fold decrease relative to shNTC control, as opposed to the four-fold decrease observed in the initial round of differentiation (Figures 1A and 3). At the GT stage, a decrease in *FOXA1* expression did not affect the expression of *FOXA2* and *HNF4a* (Figure 3A). However, the two-fold decrease in *FOXA1* expression resulted in a comparable decrease in expression of early pancreatic markers (Figure 3B). These results suggest that *FOXA1* is required for the proper pancreatic expression of hESC-derived pancreatic cell types.

Requirement of FOXA1 for Proper Hepatic Expression:

It is also known that *FOXA1* and *FOXA2* play a role in liver development. FOXA1 and FOXA2 have previously been shown to be required for liver specification in mouse (Lee et al., 2005). Additionally, FOXA1 and FOXA2 binding events have been observed occur during liver development in our own work (Wang et al., 2015). To confirm these previous studies, we sought to determine the requirement for these two transcription factors during liver differentiation in our hESC system. In a separate round of differentiation, two groups of shFOXA1 CyT49s were differentiated to day 8 of the hepatic lineage (liver) and day 7 (PF) of the pancreatic lineage in order to determine the role and requirement of FOXA1 for proper hepatic expression. Expression levels of *PDX1* at day 7 in the shFOXA1 cells were decreased two-fold, consistent with results from previous rounds of pancreatic differentiation (Figure 4). Expression levels of *AFP*, an early hepatic marker, did not change significantly between the shFOXA1 and shNTC control (Jones et al., 2000) (Figure 4). This suggests that *FOXA1* by itself is not required for proper hepatic expression. This is consistent with the observation in mouse that FOXA1 and 2 display functional redundancy and can play compensatory roles (Costa et al., 2003, Lee et al., 2005).

H3K4me1 Enrichment of FOXA1-bound pancreatic enhancers at during enhancer priming

Given that the temporal pattern of FOXA recruitment to pancreatic enhancers mirrors an increase of H3K4me1 levels at the GT stage (Lupien et al., 2008, Wang et al., 2015), we tested whether *FOXA1* is required for H3K4me1 deposition at pancreatic enhancers. To achieve this, ChIP-qPCR was employed to quantify

H3K4me1 enrichment at FOXA1-bound pancreatic enhancers in shFOXA1 and shNTC CyT49s differentiated to the GT stage (day 5). A noticeable difference in H3K4me1 enrichment pancreatic enhancers was not observed between the shFOXA1 and shNTC control cells (Figure 5). At face value, these results suggest that FOXA1 is not directly involved in the deposition of H3K4me1 to these enhancers at this stage of differentiation. However, the lack of an observable change may be attributed to the low knockdown efficiency of FOXA1 or compensatory effects by FOXA2

Generation and Evaluation of FOXA1 and FOXA2 CRISPR/Cas9-mediated knockout CyT49 hESC cell lines:

The two-fold knockdown of FOXA1 and the lack of any significant FOXA2 knockdown in the shRNA system called for an alternate strategy in downregulating these genes for our studies. Given its emerging popularity and prominence as a reliable system, CRISPR/Cas9-mediated gene editing for the knockout of FOXA1 and FOXA2 was chosen. The variant of CRISPR/Cas9 utilized for our studies is the Genome-scale CRISPR Knock-out system from the Zhang lab (Shalem et al., 2014), where components of the system are delivered through lentiviral vectors. To achieve this, undifferentiated CyT49 hESCs grown in multiple wells were transduced with either a single LgP-sgFOXA1 or sgFOXA2 lentivirus, a combination of LgP-sgFOXA1 and/or sgFOXA2 lentiviruses, lentivirus containing a scaffold control (unaltered LgP plasmid), or a combination of two non-targeting control LgP lentiviruses (GFP and AAVS1). The combinations of LgP-sgFOXA1 and/or LgP-sgFOXA2 lentiviruses

include the use both available sgRNAs targeting the upstream region of exon 1 and downstream region of exon 2 for each of the genes (sgFOXA1-d&f or sgFOXA2g&f) in order to induce an excision (Appendix E). CyT49 transduced with the combination of lentiviruses for the purpose of FOXA1 gene excision is hereby referred to as “sgFOXA1-dub”, while that for FOXA2 is hereby referred to as “sgFOXA2-dub”. All cell groups were also transduced with lentivirus containing the Lenti-Cas9-blast vector in order for the cells to express Cas9 nuclease.

Following transduction, the infected CyT49 cells were maintained in media containing 5ng/μL blasticidin and 1ng/μL puromycin for 7 days to select for the incorporation of the constructs. Given the nature of CyT49 expansion, colony generation and cell line purification was not performed. After recovery, a subset of the cells from each well was harvested for genomic DNA. The PCR primers used for PCR amplification, the expected band sizes, and the target location of the amplification can be found in Appendix F.

Gel electrophoresis images of the CyT49 cells infected with only one sgRNA construct suggest that PCR amplification was successful and yielded the expected band sizes for all samples (Figure 6A). Since gel electrophoresis cannot resolve fragment sizes resulting from a single nucleotide deletion, these PCR products were sequenced. Chromatograph snapshots of these PCR product sequences at the expected sgRNA/Cas9 nuclease cut sites show a transition from clean peaks to an erratic pattern of multiple competing peaks, suggesting successful nucleotide deletion in a subpopulation of cells for all four of these samples (Figure 6B). Excision of FOXA1 in the FOXA1-dub and FOXA2 in the FOXA2-dub samples was also detected through gel electrophoresis (Figure C). Clean bands of sizes 316nt and 365nt

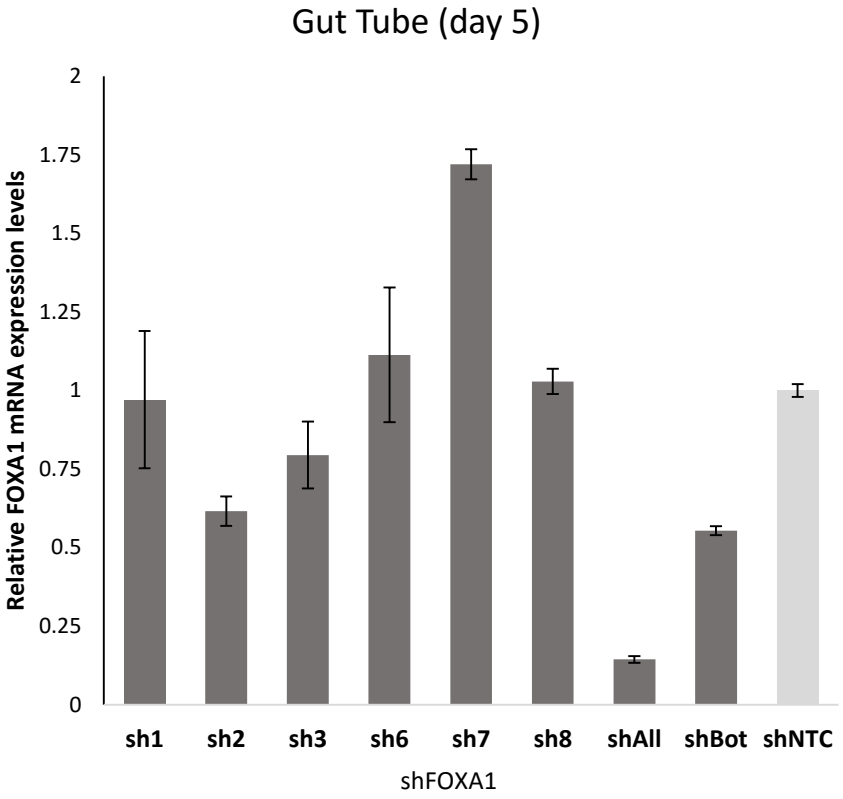
(FOXA1-dub and FOXA2-dub, respectively), as opposed to >3kb, are indicative of a double nucleotide deletion resulting in the excision of the FOXA1 and FOXA2 genes. Altogether, these results show that the GeCKO CRISPR/Cas9 gene editing system can be effective at making nucleotide deletions in FOXA1 and FOXA2.

Figure 1. Expression levels of FOXA1 in CyT49 hESCs differentiated to GT:

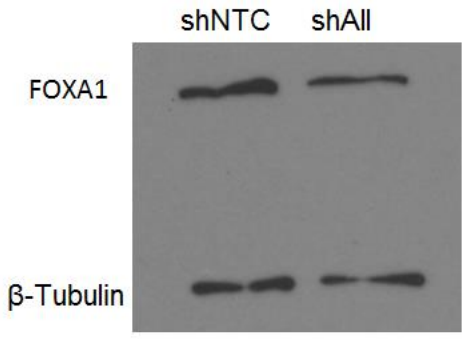
(A) Relative FOXA1 mRNA expression levels in CyT49 hESCs transduced with FOXA1 shRNAs differentiated to GT (day 5). These mRNA expression levels are relative to shNTC. shALL refers to cells transduced with all eight shRNAs, while shBot refers to cells transduced with sh1, sh2, sh3 and sh4. The relative mRNA levels were calculated using the ΔΔCt method as outlined in Livak & Schmittgen (2001). All bar graphs show the standard error of the mean.

(B) Western blot representing FOXA1 protein levels between shAll and shNTC CyT49 hESCs differentiated to gut tube (day 5 of differentiation). β-Tubulin was used as the loading control.

1A.


Gut Tube (day 5)
Relative FOXA1 mRNA expression levels
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0
sh1
sh2
sh3
sh6
sh7
sh8
shAll
shBot
shNTC
shFOXA1

1B.


shNTC
shAll
FOXA1
β-Tubulin

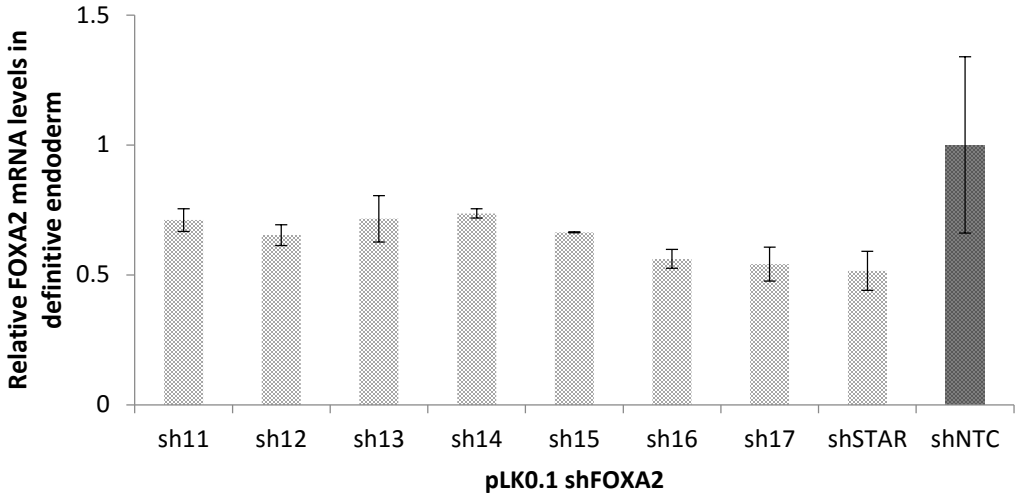
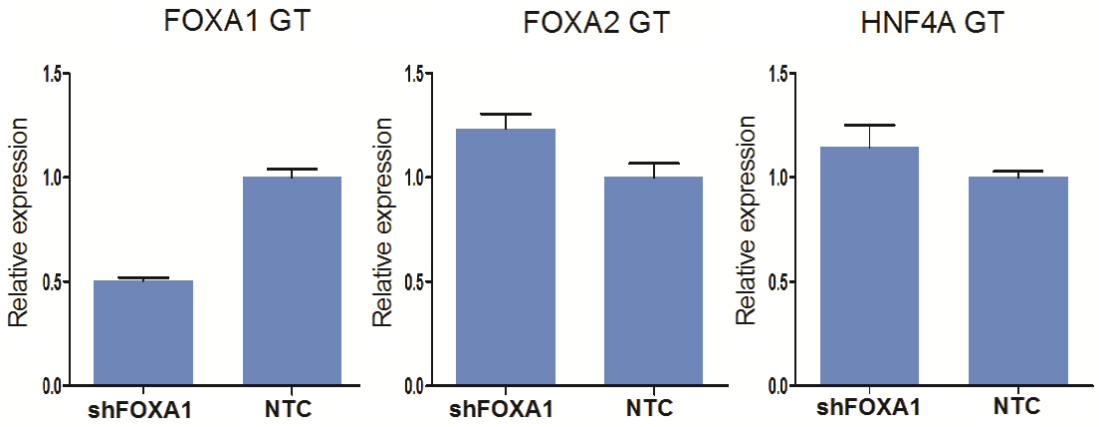


Figure 2. Expression levels of FOXA2 mRNA in CyT49 hESCs differentiated to DE:
Relative *FOXA2* mRNA expression levels of pLKO.1-derived *FOXA2* shRNAs in hESC cells differentiated to definitive endoderm (day 2). These mRNA expression levels are compared to a non-targeting control (shNTC). "shSTAR" refers to the sample transduced with a cocktail of 5µL each of sh11-sh17. The relative mRNA levels were calculated using the ΔΔCt method as outlined in Livak & Schmittgen (2001). All bar charts display the standard error of the mean.

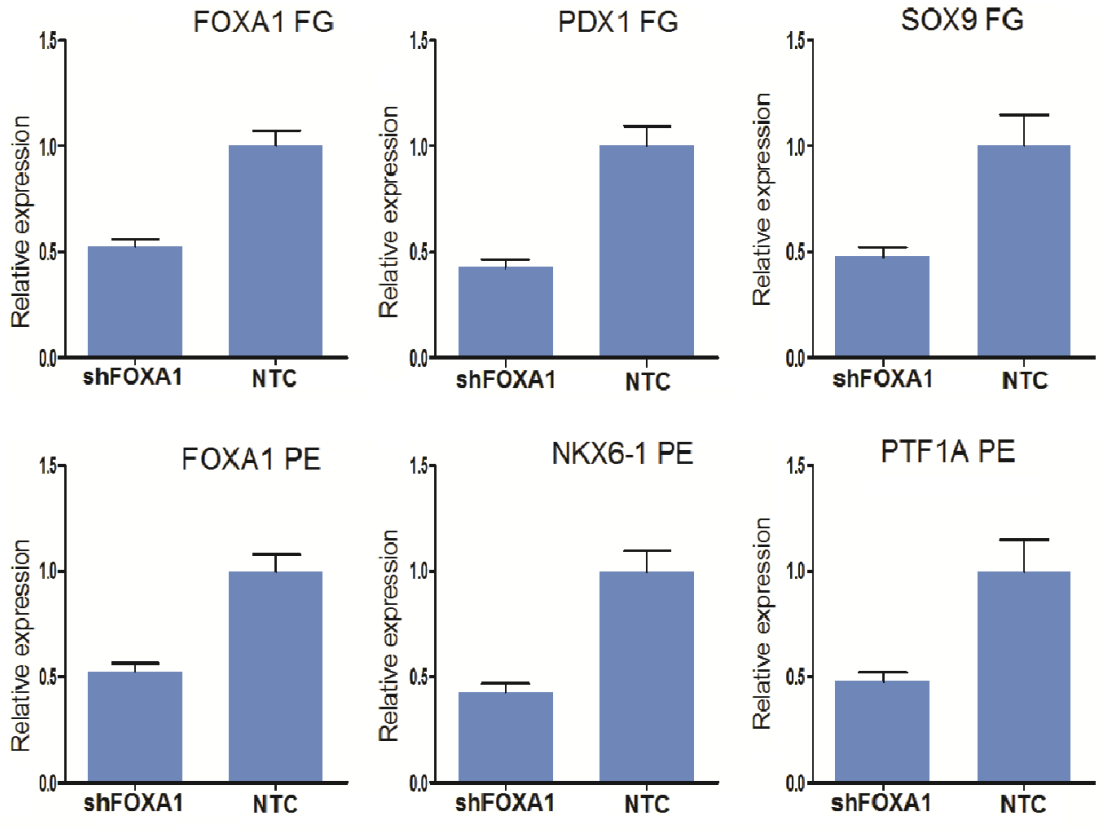
Figure 3: Role of FOXA1 in proper expression of hESC-derived pancreatic cell types and their progenitors:

(A) Relative FOXA1, FOXA2 and GT marker HNF4A mRNA expression levels in shFOXA1 CyT49 at the GT stage of pancreatic differentiation.
(B) Relative FOXA1, PF marker and PP marker mRNA expression levels in shFOXA1 CyT49 at the PF stage of pancreatic differentiation. The markers examined include PDX1, SOX9, NKX6.1 and PTF1A.

3A:


FOXA1 GT
FOXA2 GT
HNF4A GT
Relative expression
1.5
1.0
0.5
0.0
shFOXA1
NTC

3B:


FOXA1 FG
PDX1 FG
SOX9 FG
FOXA1 PE
NKX6-1 PE
PTF1A PE
Relative expression
1.5
1.0
0.5
0.0
shFOXA1
NTC

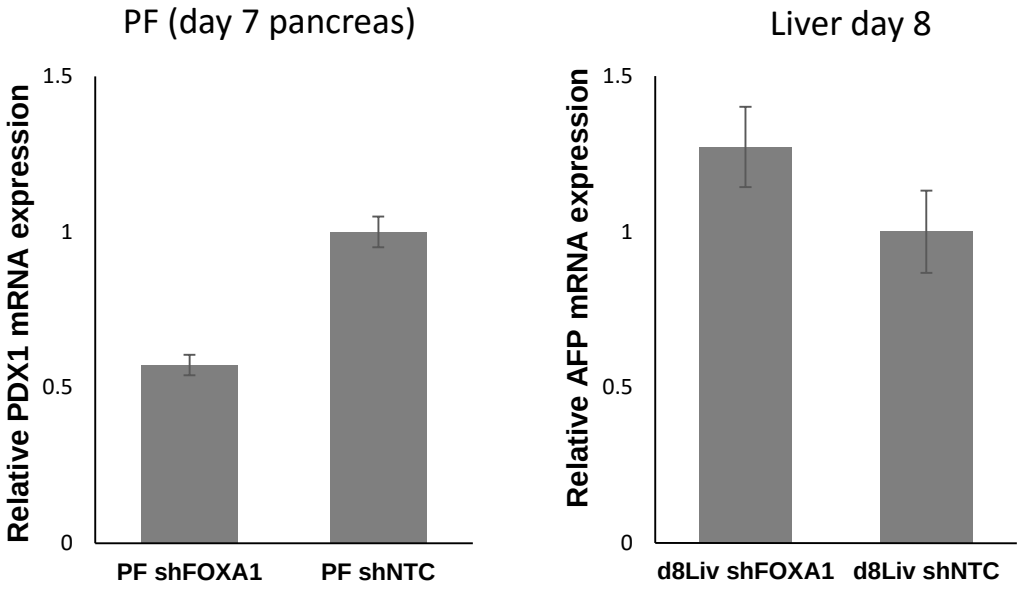


Figure 4: Role of FOXA1 in proper expression of hESC-derived hepatic cells:

(Right) Relative PDX1 mRNA expression levels of shFOXA1 and shNTC control CyT49 differentiated to the PF stage of pancreatic differentiation. (Left) Relative Alpha-Fetoprotein (AFP) mRNA expression levels of shFOXA1 and shNTC control CyT49 differentiated to day 8 of hepatic differentiation.

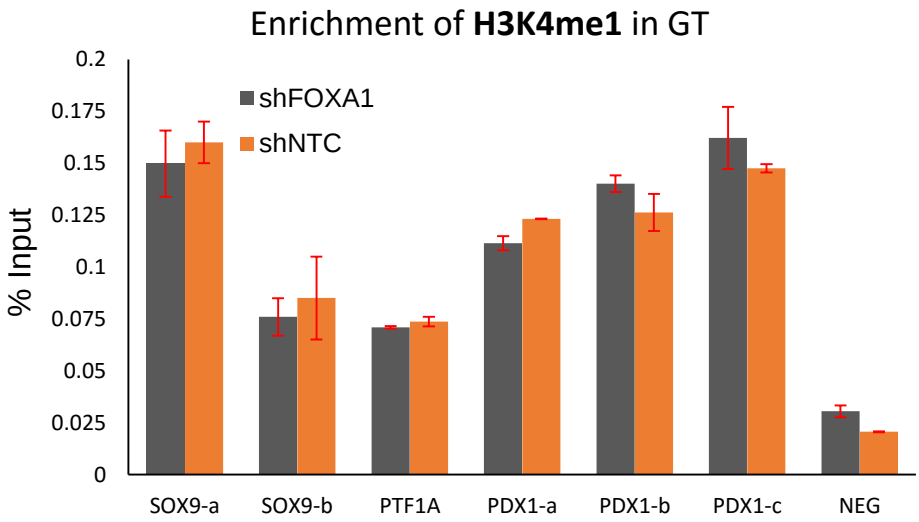


Figure 5: Enrichment of H3K4me1 at FOXA1-bound pancreatic enhancers: Quantification of H3K4me1 enrichment at FOXA1-bound pancreatic enhancers by ChIP-qPCR in shNTC control and *FOXA1* knockdown (shFOXA1) cells differentiated to GT (day 5). NEG is a negative control region of the genome.

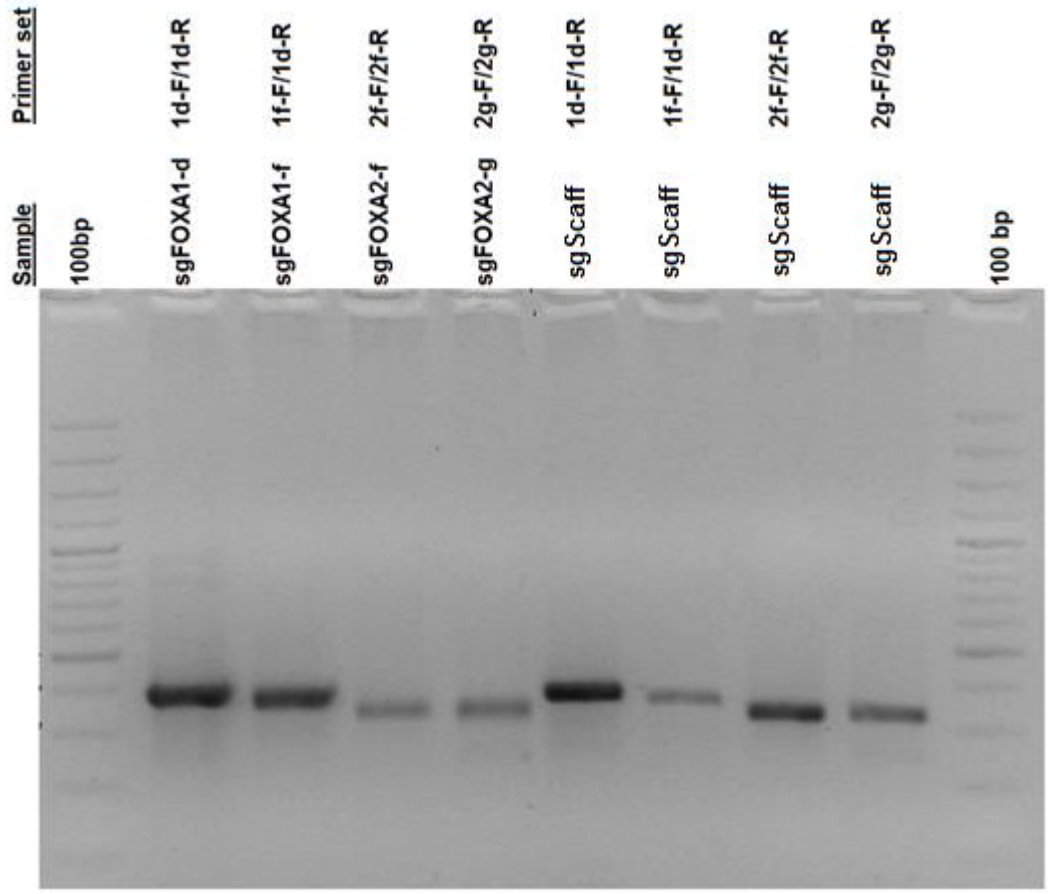
Figure 6: Evaluation of GeCKO CRISPR/Cas9 against FOXA1/FOXA2 in CyT49 hESC:

(A) Gel electrophoresis images of PCR DNA amplified from relevant regions of sgFOXA1-d (via primers 1d-F/R), sgFOXA1-f (via primers 1f-F/R), sgFOXA2-f (via primers 2f-F/R), and sgFOXA2-g(via primers 2gF/R) CyT49 hESCs. See Appendices E and F.

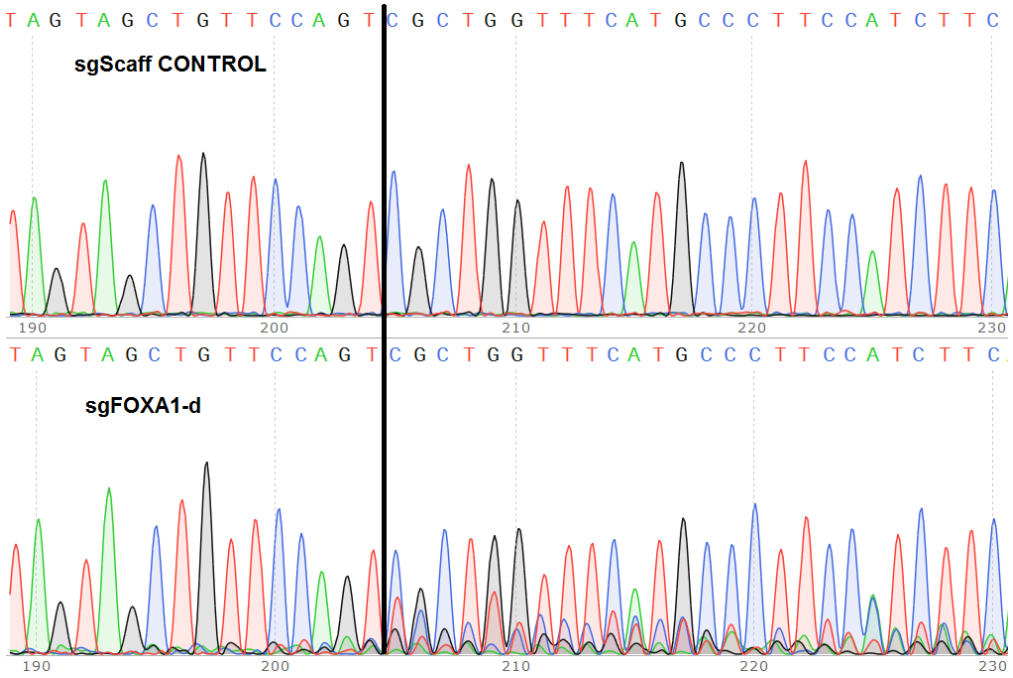
(B- E) Sequencing chromatograms of the PCR products described in (A) above. The region where the nucleotide deletion location is located for each sample is shown and emphasized by a solid black line.

(F) Gel electrophoresis images of PCR DNA amplified from FOXA1-dub (via primers 1d-R/1f-R), FOXA2-dub (via primers 2f-F/2g-R).

6A:


Sample
Primer set
100bp
sgFOXA1-d
1d-F/1d-R
sgFOXA1-f
1f-F/1d-R
sgFOXA2-f
2f-F/2f-R
sgFOXA2-g
2g-F/2g-R
sg Scaff
1d-F/1d-R
sg Scaff
1f-F/1d-R
sg Scaff
2f-F/2f-R
sg Scaff
2g-F/2g-R
100 bp

6B:



6C:

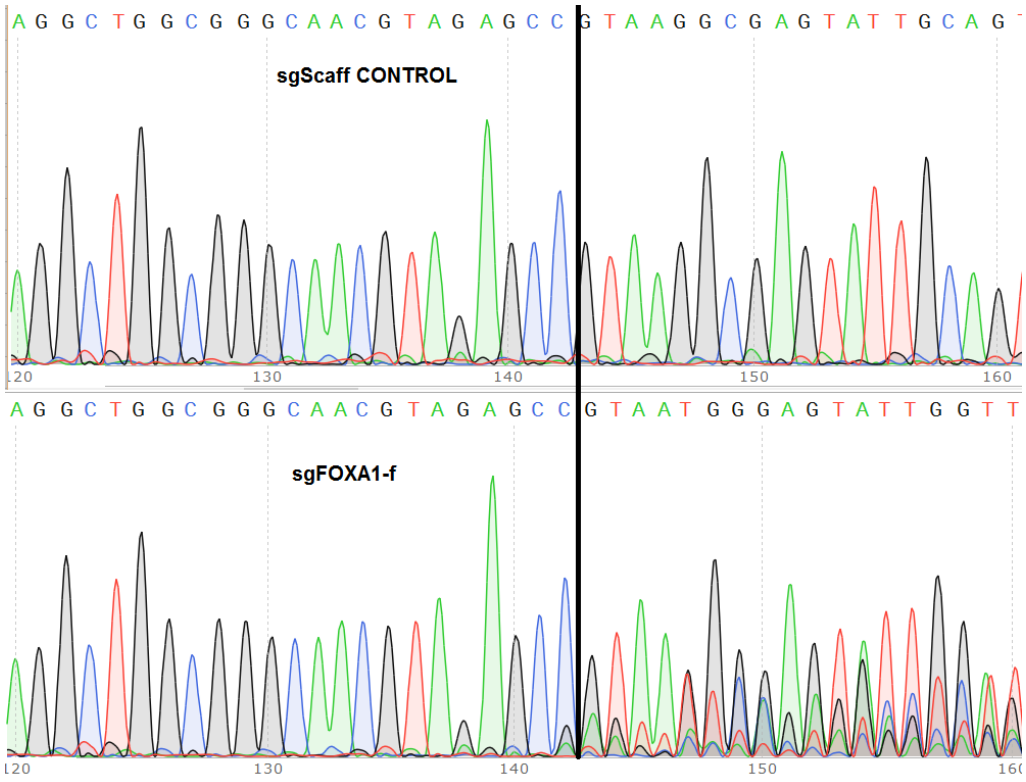
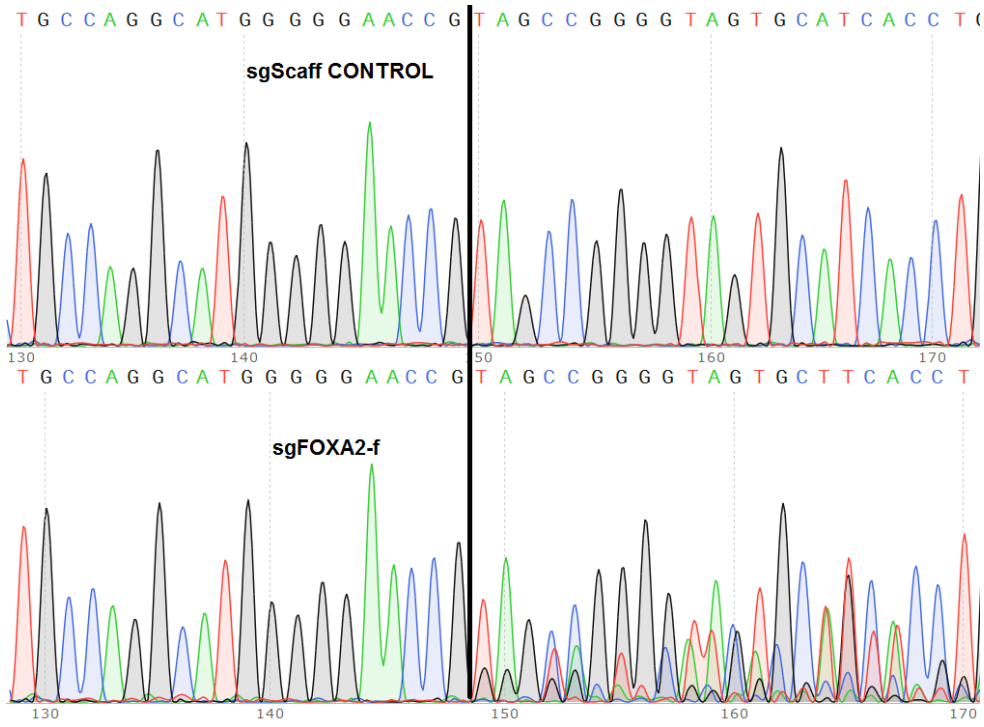


Figure 6, continued

6D:



6E:

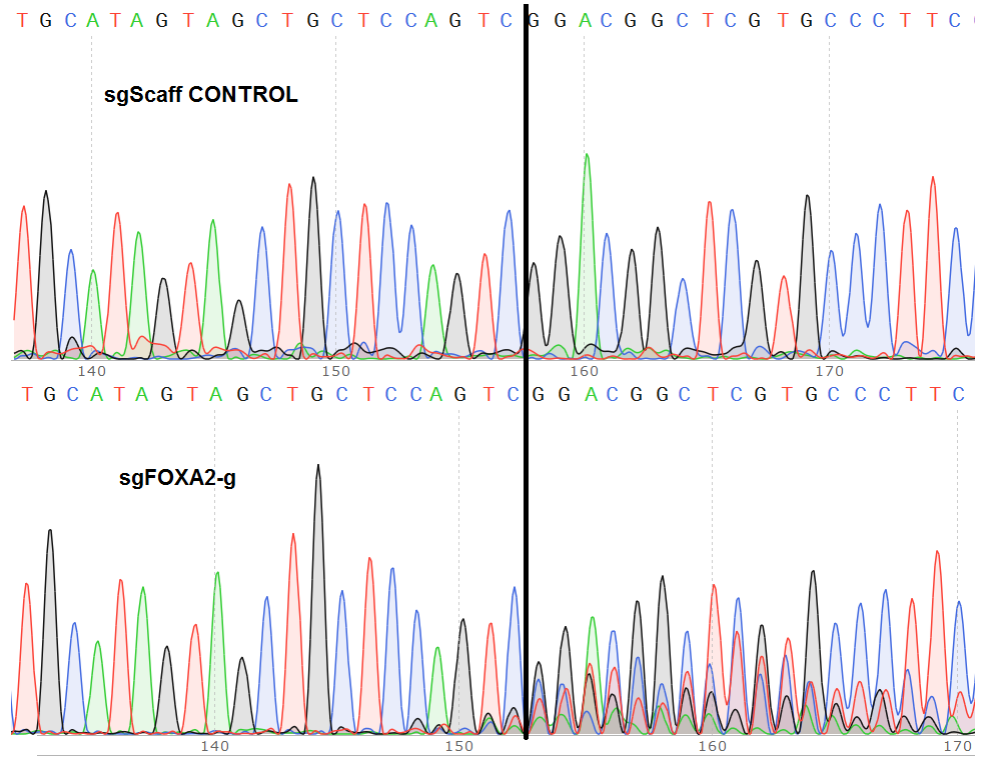


Figure 6, continued

6F:

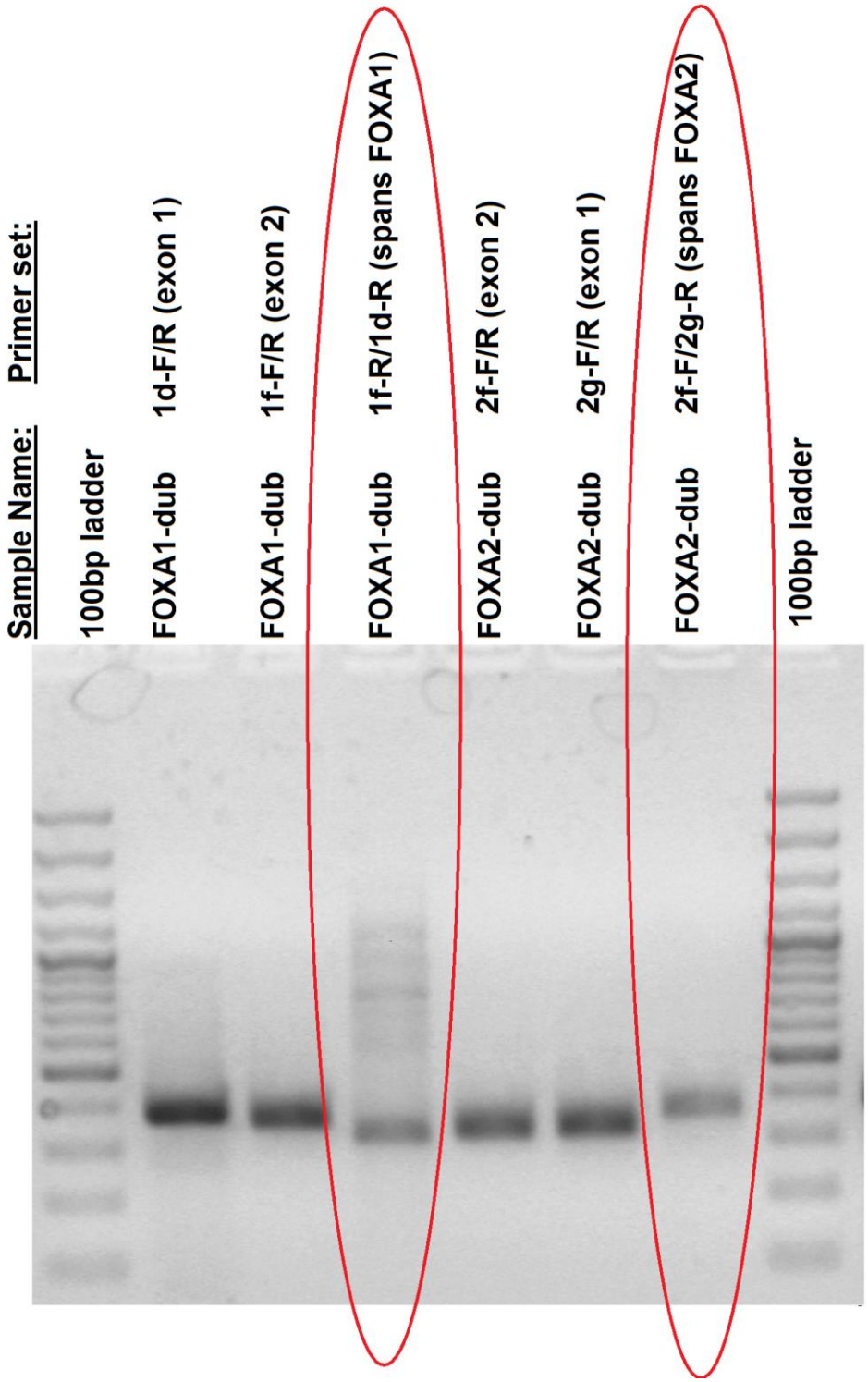


Figure 6, continued

III:

DISCUSSION

In this study, we sought to downregulate FOXA1 and FOXA2 to elucidate their role during pancreas differentiation of hESCs. To achieve this, we designed and utilized shRNAs in hopes of creating *FOXA1* and *FOXA2* knockdown CyT49 hESC lines. While we were able to generate a *FOXA1* cell line, we were unable to produce a knockdown *FOXA2* cell line with significant knockdown. The decrease in *FOXA1* expression ranged between two and four fold between rounds of experiments, which is modest at best. Nonetheless, we continued our studies with only the knockdown *FOXA1* cell lines on hand.

Expression of FOXA2 and HNF4A, which is normally expressed in GT, did not change in the knockdown FOXA1 cells. However, a decrease in expression of pancreas genes was observed in the PF and PP stages of differentiation in the knockdown FOXA1 cell lines. Furthermore, this decrease in expression was found to be about two-fold across all pancreatic genes tested, which mirrors the two-fold decrease of *FOXA1* expression in the knockdown cells. These results suggest that FOXA1 is required for the proper expression of hESC-derived pancreatic cell types.

As mentioned above, it has been previously found that the recruitment of FOXA1/2 to pancreatic enhancers coincides with an increase of H3K4me1 levels at the GT stage. To determine if FOXA1 is required for H3K4me1 deposition at pancreatic enhancers at this stage, ChIP-qPCR of these enhancers was performed. While we did not find any changes in H3K4me1 levels in the knockdown FOXA1 cells, it does mirror previous findings that suggest FOXA1 may merely play an indirect role by opening genomic regions marked by H3K4me1 (Lupien et al., 2008). Another possibly and likely explanation could be the low knockdown efficiency of our

shFOXA1 cells, where there is simply not enough of a decrease in FOXA1 expression to affect a change.

We also tested if FOXA1 is required for proper expression of hESC-derived hepatic cells. Previous studies in mouse show that FOXA proteins are critical for liver development (Lee et al., 2005). Expression levels of AFP however, did not significantly change in the knockdown FOXA1 cells. This can once again be attributed to the low knockdown efficiency of the cell line. However, there is also evidence in the literature that points to the redundancy of FOXA1 and FOXA2 in liver development (Costa et al., 2003). These results therefore suggest that proper hepatic differentiation is not dependent on FOXA1 alone.

Given the limitations of our FOXA1 knockdown cells and the shRNA system as a whole, we attempted to create FOXA1 and FOXA2 knockout CyT49s in hopes of increasing the efficiency of FOXA1 and FOXA2 downregulation. While shRNA has been a prime method of choice by many molecular biologists, it has its shortcomings. RNAi methods downregulate genes at the post-transcriptional level. While this is usually sufficient in suppressing gene expression, this may not be the case for robust genes such as the FOXA proteins. CRISPR/Cas9-mediated gene editing is an attractive alternative because it offers the possibility of silencing genes at the genomic level. In this study, we sought to disrupt *FOXA1* and *FOXA2* in two ways: frameshift disruption of the gene through a single nucleotide deletion, or excision of the genes via combinations of multiple guideRNAs. Both of these were performed on CyT49 hESCs in this study, and both resulted in successful nucleotide deletion or gene excision. Our next steps will be to differentiate these cell lines and characterize them.

While we are making progress in our FOXA studies, there are also questions to be raised as to the reliability of our CRISPR/Cas9 system variant. GeCKO was designed in a way such that CRISPR/Cas9 components could be delivered to cells using lentiviral vectors. While this allows us to quickly move forward with our experiments, lentiviral vectors add the parameter of off-target effects since there is random integration of the constructs into the host cell genome. This is further compounded by the difficulties of colony selection in our CyT49 hESCs. CyT49 cells in culture tend to spread out across the dish, and tend not to grow in colonies. To address this, we have recently acquired the H1 hESC cell line, which has a higher propensity to grow in colonies. Another issue that arises from our current lentiviral delivery system are the off-target effects of Cas9 nuclease and gRNA over time due to the constitutive expression of Cas9 nuclease and gRNA. To address this, we are in the process of testing an alternative method of delivering CRISPR/Cas9 components into hESCs whereby gRNAs and Cas9 nuclease are directly transfected into the cell. If successful, this method would eliminate concerns of off-target effects from random integration. We hope that these new tools can aid us in our studies of pancreas differentiation.

IV:

MATERIALS AND METHODS

Design and Cloning of FOXA1 and FOXA2 shRNAs for RNAi-mediated Knockdown:

A total of 8 FOXA1 and 7 FOXA2 shRNAs and were designed, synthesized, and subsequently cloned into the bacterial cloning vectors “pLKO.1” (provided by Addgene, Inc.) The shRNA constructs were generated using designs from the Broad Institute’s RNAi Consortium. (The sequences of these constructs are summarized in Appendices B and C.) All shRNAs were designed with the addition of endonuclease sites at both ends for ligation into cloning vectors. These sites are EcoRI & AgeI. Once designed, oligonucleotides of the constructs were purchased from Integrated DNA Technologies, Inc. These oligonucleotides were then annealed and ligated into their respective cloning vector, pLKO.1. Once ligation was completed, the plasmids containing the shRNA constructs were transformed into DH5α competent cells, which were then plated on Lennox Broth (LB) plates with 100μg/mL ampicillin and incubated overnight at 37ºC. Colonies were then picked for another overnight culture in LB broth with 100μg/mL ampicillin at 37ºC. Afterward, DNA was extracted from the overnight LB broth cultures using an Invitrogen PureLink™ Quick Miniprep Kit. A portion of the extracted DNA was then digested with restriction enzymes to verify the incorporation of shRNA construct into the vector backbone. The DNA was then sequenced by Eton BioSciences, Inc. for a final verification.

FOXA1/2 CRISPR/Cas9-mediated Knockout, GeCKO variant:

The “Genome-wide CRISPR/Cas9 KnockOut” (GeCKO) variant of CRISPR/Cas9-mediated knockout was employed. A total of two FOXA1 and two FOXA2 single guide RNAs (sgRNA) were designed, synthesized, and subsequently cloned into the bacterial cloning vector “Lenti-guide-puro” (originally obtained from

Addgene, Inc., and kindly provided by the Yeo lab, hereby referred to as LgP). LgP provides a scaffold for the incorporation the specific guideRNA into the tracrRNA, as well as the conferral of a puromycin selection marker. The sgRNAs were designed using CRISPR design tools from Benchling, Inc. An sgRNA for exon 1 and exon 2 was designed for each of these two genes. Each sgRNA was designed with the addition of forward 5'-CACCG-3' and reverse 5' AAAC 3' overhangs for proper insertion into the aforementioned cloning vector. (The sequences of these sgRNA designs are summarized in Appendix E.) The cloning of these sgRNAs into the LgP cloning vector was done in a manner similar to that of the FOXA1/FOXA2 shRNAs mentioned earlier. Lentivirus were packaged with LgP plasmids containing the FOXA1/FOXA2 sgRNAs. Lentivirus were also packaged with a separate "Lenti-Cas9-blast" (originally obtained from Addgene, Inc., and kindly provided by the Yeo lab) plasmid that is required for the expression of Cas9 protein in the target CyT49 hESCs.

Lentivirus production and transduction into hESCs:

Once the shRNA and sgRNA constructs were sequence-verified, lentiviruses were packaged using 293T cells. These 293T cells were grown to ~80% confluency in 1X DMEM/10% FBS media before transfection. The shRNA construct plasmids were then co-transfected with a lentivirus envelope plasmid and lentivirus packaging plasmid, all of which are required for successful lentivirus production. 293T cells begin to produce lentivirus around 24 hours after transfection, accumulating in the supernatant. This viral supernatant was collected over a course of two days, filtered to remove cell debris, and centrifuged at 19,500xG in an ultracentrifuge for 2 hours at

4ºC. The viral pellet was then collected and resuspended in 200µL of storage media (DMEM w/ High Glucose), and stored at -80ºC until use.

Transduction of hESCs with lentivirus:

Prior to transduction, undifferentiated hESCs were grown in 6-well tissue culture plates coated with Matrigel™ in monolayer form. The hESC cell line used for our studies is “Cythera 49” from ViaCyte, Inc., hereby abbreviated as “CyT49”. Once sufficient confluency was reached, the Cyt49 cells were transduced with pre-thawed lentivirus solution. In addition to transduction of FOXA1 and FOXA2 shRNA lentivirus, a control well of Cyt49 cells was transduced with a non-targeting control shRNA. Transduction was repeated the following day for the wells that had insufficient levels of cell death following the initial transduction. Selection of pLKO.1-derived shRNA transduced CyT49 hESCs was achieved by maintaining the cells in 1ng/µL puromycin for two weeks, followed by recovery. Selection of LgP-derived sgRNA and Lenti-Cas9-blast transduced CyT49 hESCs was achieved by maintaining cells in 1ng/µL puromycin and 5ng/µL blasticidin for one week, followed by recovery.

Preparation of CyT49 hESCs for Differentiation:

After drug selection, CyT49 hESCs were allowed to grow to >90% confluency and subsequently passaged onto ultra-low adhesion tissue culture plates (Corning, Inc., #3471) in preparation for differentiation. In accordance with the differentiation protocol outlined in Schulz et al. (2012), the Cyt49 cells initiate differentiation by aggregating into embryoid body form (“aggregates”). This was achieved by rotating the hESCs in liquid suspension on a rotating platform at 95rpm at 37°C. The day after

this initial step is designated as day 0 of the differentiation, by which time the Cyt49s have completely formed into aggregates. Differentiation toward downstream cell lineages required the daily replacement of cell media containing appropriate concentration of specific reagents and factors specific to each cell stage transition. Appendix A summarizes the different stages of differentiation, the reagents and factors needed to reach each stage, and the major genes that are active at each stage.

Collection of aggregates and cell processing for staining and qPCR analysis

Aggregates were collected at various stages of the differentiation for staining and qPCR analysis. For each day of collection, aggregates were fixed with 4% PFA and embedded into a O.C.T. block for cryosectioning. In general, these cryosections are stored in -80ºC for future staining. RNA was also extracted from collected aggregates using a Qiagen RNEasy™ Mini Kit. This RNA was subsequently reverse-transcribed to cDNA using the BioRad iSCRIPT reverse transcription kit, where it was later analyzed by qPCR.

Analysis of qPCR data:

qRT-PCR experiments were quantified using iQ SYBR green Supermix (BioRad) and the CFX96 real-time system (BioRad). Quantification of qPCR data into relative mRNA expression values was done via the ΔΔCt method, outlined in Livak & Schmittgen (2001). Specifically, mRNA expression levels of target genes (which are determined by the amplification kinetics of their respective cDNAs), were compared to

that of TATA-box binding protein (TBP) as a normalizer. The final relative mRNA levels were determined via the following expression:

$$\mathbf{2^{-(\Delta\Delta Ct)}, \text{ where: } \Delta\Delta Ct = (Ct_{GOI} - Ct_{norm})_{unknown} - (Ct_{GOI} - Ct_{norm})_{calibrator}}$$

Ct_{GOI} refers to the Ct values of the samples with FOXA2 as the target sequence, while Ct_{norm} refers to the Ct value of the corresponding samples where TBP is the target sequence. The expressions $(Ct_{GOI} - Ct_{norm})_{unknown}$ and $(Ct_{GOI} - Ct_{norm})_{calibrator}$ are also known as $\Delta Ct_{unknown}$ and $\Delta Ct_{calibrator}$, respectively, where $\Delta\Delta Ct$ is evaluated by taking the difference between $\Delta Ct_{unknown}$ and $\Delta Ct_{calibrator}$. The relative mRNA levels were evaluated from mean Ct, and the error displayed on all qRT-PCR data displays the standard error of the mean.

Western Blot:

Whole cell extracts were fractionated by SDS-PAGE and transferred to a polyvinylidene difluoride membrane using a transfer casette. After incubation with nonfat milk in PBS for 1hr, the membrane was washed once with PBST and incubated with antibodies against FOXA1(Abcam ab5089, 1:1000), and β-tubulin overnight at 4°C. Membranes were then washed and incubated with a 1:3000 dilution of horseradish peroxidase-conjugated antibody against the species corresponding to each primary antibody. Blots were washed with PBST and subsequently developed.

ChIP-qPCR:

CyT49 aggregates at the GT stage were collected for chromatin and sonicated in accordance to and in use of the Active Motif® ChIP-IT® High Sensitivity kit. Chromatin immunoprecipitation for H3K4me1 in Figure 5 was performed using 5µg of

H3K4me1 antibody (Abcam ab8895) on 100µg of chromatin collected at the GT stage. ChIP experiments were quantified using iQ SYBR green Supermix (BioRad) and the CFX96 real-time system (BioRad). Enrichment was calculated as percentage of input. Enrichment is calculated by comparing expression data from qPCR to that of input DNA. NEG control is a region of genomic DNA previously shown to show no enrichment002E

Genotyping Confirmation of CyT49 infected with FOXA1/FOXA2 sgRNAs

Genomic DNA was extracted using QuickExtract™ DNA Extraction Solution from Epicentre®. Standard PCR techniques were employed to amplify relevant fragments of the FOXA1 and FOXA2 genes. A table of all PCR primers used can be found in Appendix F. PCR products that were selected to be sequenced were done so by Genewiz, Inc, which provided the chromatograms featured in Figure 6.

APPENDICES

APPENDIX A: Schematic of Pancreatic Differentiation of CyT49 hESCs:
Differentiation of Cythera 49 hESCs toward pancreatic endocrine cells, as outlined in Shultz et al. 2012. Below each cell type are the major genes that are active for that stage, as well as the time for which the stage persists (for example, D2 = day 2 of differentiation). Shown above each stage transition are the factors or reagents needed by the cells to differentiate into the next desired cell stage.

Appendix A:


ES
DE
GT
PF
PP
PE
Actvin A, Wnt3A
KGF
RA, Noggin, Cyc
No Factors
No Factors
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Embryonic stem cell (OCT4, SOX2, NANOG)
Definitive endoderm (SOX17, FOXA2 CER, CXCR4)
Primitive gut tube (HNF1B, HNF4A)
Posterior foregut (PDX1, HNF6)
Pancreatic progenitor (SOX9, PDX1, NKX6.1, PTF1A)
Endocrine cells (INS,GLU)
D0
D2
D5
D7
D10
D13

APPENDIX B: FOXA1 shRNA oligonucleotide sequences (sense strands):

sh1	5'-CCGGGCGTACTACCAAGGTGTGTATCTCGAGATACACACCTTGGTAGTACGCTTTTTG -3'
sh2	5'-CCGGGCAGCATAAGCTGGACTTCAACTCGAGTTGAAGTCCAGCTTATGCTGCTTTTTG -3'
sh3	5'-CCGGGCGAAGTTTAATGATCCACAACTCGAGTTGTGGATCATTAAACTTCGCTTTTTG -3'
sh4	5'-CCGGATACGAACAGGCACTGCAATACTCGAGTATTGCAGTGCCTGTTCGTATTTTTTG -3'
sh5	5'-CCGGGTATTCCAGACCCGTCCTAAACTCGAGTTTAGGACGGGTCTGGAATACTTTTTG -3'
sh6	5'-CCGGCAAACCGTCAACAGCATAATACTCGAGTATTATGCTGTTGACGGTTTGTTTTTG -3'
sh7	5'-CCGGTCTAGTTTGTGGAGGGTTATTCTCGAGAATAACCCTCCACAAACTAGATTTTTG -3'
sh8	5'-CCGGGAACACCTACATGACCATGAACTCGAGTTCATGGTCATGTAGGTGTTCTTTTTG –3'

APPENDIX C: FOXA2 shRNA oligonucleotide sequences (sense strands):

sh11	5’-CCGGGAACGGCATGAACACGTACATCTCGAGATGTACGTGTTCATGCCGTTCTTTTTG -3’
sh12	5’-CCGGGCAAGGGAGAAGAAATCCATACTCGAGTATGGATTTCTTCTCCCTTGCTTTTTG -3’
sh13	5’-CCGGCTACGCCAACATGAACTCCATCTCGAGATGGAGTTCATGTTGGCGTAGTTTTTG -3’
sh14	5’-CCGGCTCCTCTTAAGAAGACGACCTCGAGGTCGTCTTCTTAAGAGGAGTTTTTG -3’
sh15	5’-CCGGACGGCATGAACACGTACATGACTCGAGTCATGTACGTGTTCATGCCGTTTTTTG -3’
sh16	5’-CCGGGGAACACCACTACGCCTTCAACTCGAGTTGAAGGCGTAGTGGTGTTCCTTTTTG -3’
sh17	5’-CCGGCCCATTATGAACTCCTCTTAACTCGAGTTAAGAGGAGTTCATAATGGGTTTTTG -3’

APPENDIX D: FOXA1-bound enhancers examined and their coordinates in the human chromosome, related to Figure 5:

Enhancer name	Coordinates
PDX1-a	chr13:27,381,114-27,382,132
PDX1-b	chr13:27,383,262-27,384,342
PDX1-c	chr13:27,910,499-27,910,518
PTF1a	chr10:23,220,296-23,220,315
SOX9-a	chr17:72,173,873-72,173,892
SOX9-b	chr17:71,680,847-71,680,866
NEG	chr13:65,364,902-65,364,921

APPENDIX E: FOXA1 and FOXA2 sgRNA oligonucleotide sequences:
Both the forward and reverse sequences are shown below. FOXA1 and FOXA2 both have two exons. The target region of each guide RNA pair is also shown below. Overhangs of the oligonucleotides for cloning purposes are shown as lower case nucleotides.

	Sequence	Target Region
sgFOXA1-d forward	caccgTAGTAGCTGTTCCAGTCGC	FOXA1 exon 1
sgFOXA1-d reverse	aaacGCGACTGGAACAGCTACTAC	
sgFOXA1-f forward	caccgCACTGCAATACTCGCCTTA	FOXA1 exon 2
sgFOXA1-f reverse	aaacTAAGGCGAGTATTGCAGTGC	
sgFOXA2-f forward	caccgCATGGGGGAACCGTAGCCG	FOXA2 exon 2
sgFOXA2-f reverse	aaacCGGCTACGGTTCCCCCATGC	
sgFOXA2-g forward	caccgTAGTAGCTGCTCCAGTCGGA	FOXA2 exon 1
sgFOXA2-g reverse	aaacTCCGACTGGAGCAGCTACTAc	

APPENDIX F: PCR primer combinations used to amplify relevant regions of CFOXA1 and FOXA2 for assay of CRISPR/Cas9-mediated deletions.

For primer pairs 5 and 6, a successful excision of their respective target genes by Cas9 nuclease is expected result in a band size of 316bp and 365bp, respectively.

Pair	Primer Pair 5'→ 3'	Expected size	Target Region
1	1d-F ACAGAGCAGGGCAGCAGGT 1d-R GTGACTGCAGCTGCTCAGC	372bp	FOXA1 exon 1
2	1f-F ACCAGCATGGCTATGCCAG 1f- R CTCTGTGCCCGCCTCTCA	357bp	FOXA1 exon 2
3	2f-F ATCCGGGGTGCCAGAGTTA 2f-R GAAGCCGGAACACCACTAC	326bp	FOXA2 exon 2
4	2g-F TGATTCCAAGGAGGGCGGAA 2g-R GGGACGGTGCTTTGGCTGA	330bp	FOXA2 exon 1
5	1d-R GTGACTGCAGCTGCTCAGC 1f- R CTCTGTGCCCGCCTCTCA	Depends on excision of FOXA1	FOXA1 in its entirety
6	2f-F ATCCGGGGTGCCAGAGTTA 2g-F TGATTCCAAGGAGGGCGGAA	Depends on excision of FOXA2	FOXA2 in its entirety

REFERENCES

American Diabetes Association (2002) Insulin Administration. *Diabetes Care* **25** S.112-S.115

Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S.D. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151** 2551-2561.

Cirillo L, Lin FR, Cuesta I, Jarnik M, Friedman D, Zaret K.S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FOXA) and GATA-4. *Mol. Cell* **9***,* 279–289.

Costa, R.H., Kalinichenko, V.V., Holterman, A.L., Wang, X. (2003) Transcription Factors in Liver Development, Differentiation, and Regeneration. *Hepatology* **38(6)** 1331-1347.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931–21936.

DeLaForest, A., Nagaoka, M., Si-Tayeb, K., Noto, F.K., Konopka, G., Battle, M.A., Duncan, S.A. (2011) HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. *Development*. **138**, 4143-4153.

D'Amour, K. A., Bang, A. G., Eliazer, S., Kelly, O. G., Agulnick, A. D., Smart, N. G., Baetge, E. E. (2006) Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat Biotechnol, **24*****(11)**, 1392-1401.

Friedman J.R., Kaestner K.H. (2006) The FoxA family of transcription factors in development and metabolism. *Cell Mol Life Sci **63**,* 2317-2328

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

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffinni, L.A., Bao, G., Zhang, F. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology. *31*(9), 827-834.

Jones, E.A., Clement-Jones, M., James, O.F.W., Wilson, D.I. (2000) Differences between human and mouse alpha-fetoprotein expression during early development. *J. Anat.* **198(5)**, 555-559.

Lee C.S., Friedman J.R., Fulmer J.T., Kaestner K.H. (2005) The initiation of liver development is dependent on Foxa transcription factors. *Nature* **435,** 944 –947.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., Soria, E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174-182

Pagliuca, F.W., Millman, J.R., Gürtler, M., Segel, M., Dervort, A.V., Ryu, J.H., Peterson, Q.P., Greiner, D., Melton, D.A. (2014). Generation of Functional Human Pancreatic ß Cells In Vitro. *Cell **159**,* 428-439.

Pourcel, C., Salvignol, G., Vergnaud, G. (2005) CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151** 653-663

Van Belle, T.L., Coppieters, K.T., Von Herrath, M.G. (2011) Type I Diabetes: Etiology, Immunology, and Therapeutic Strategies. *Physiological Rev* **91.1**, 79-118.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283.

Rao, D.D., Vorhies, J.S., Senzer, N., Nemunaitis, J. (2009) siRNA vs. shRNA: Similarities and differences. *Advanced Drug Delivery Reviews* **61**, 746-759.

Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, Al., 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**,* 1121-1133.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., Zhang, F. (2014) Genome-Scame CRISPR-Cas9 Knockout Screening in Human Cells. *Science **343**(**6166**)*, 84-87.

Shih, H.P., Wang, A., Sander, M. (2013) Pancreas Organogenesis: From Lineage Determination to Morphogenesis. *Annu. Rev. Cell Dev. Biol. **29**,* 19.1-19.25.

Wang, A., Sander, M., (2012) Generating cells of the gastrointestinal system: current approaches and applications for the differentiation of human pluripotent stem cells. *J Mol Med*. **90***,* 763-771.

Wang A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N.A., Muth, K., Palmer, J., Qiu, Y., Wang, J., **Lam, D.K.**, Raum, J.C., Stoffers, D.A., Ren, B., Sander, M. (2015) Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16***,* 386-399.

Yamamizu, K., Piao, Y., Sharov, A.A., Zsiros, V., Yu, H., Nakazawa, K., Schlessinger, D., Ko, M.S.H. (2013) Identification of transcription factors for lineage-specific ESC differentiation. *Stem Cell Reports* **1(6)** 545-559

Zaret, K. (1999) Developmental Competence of the Gut Endoderm: Genetic Potentiation by GATA and HNF3/Fork Head Proteins. *Dev. Bio.* **209-1***,* 1-10.

Zaret K.S., Carroll, J.S. (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes & Dev*.. **25:** 2227-2241