Loss Of Tbx3 Enhances Pancreatic Progenitor Generation From Human Pluripotent Stem Cells

Somdutta Mukherjee

University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Cell Biology Commons, and the Developmental Biology Commons

Recommended Citation


https://repository.upenn.edu/edissertations/4221

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/4221

For more information, please contact repository@pobox.upenn.edu.
Loss Of Tbx3 Enhances Pancreatic Progenitor Generation From Human Pluripotent Stem Cells

Abstract

Tbx3 has been identified as a regulator of fate decisions during liver development in the mouse, but whether it function similarly in humans remains unclear. We generated TBX3 knockout human pluripotent stem cell (PSCs) lines using CRISPR/Cas9 genome editing technology. We differentiated TBX3 knockout lines to hepatocytes and found a decrease in hepatic markers and in hepatocyte function, demonstrating that TBX3 is also important for liver differentiation in humans. Surprisingly, we detected expression of pancreatic markers, including PDX1 in our TBX3 knockout hepatocytes, suggesting the possibility that TBX3 may regulate liver development by suppressing a pancreatic fate. We next differentiated the TBX3 knockout lines to pancreatic progenitors to determine whether the loss of TBX3 impacts pancreatic differentiation. We found that TBX3 knockout PSCs generated more pancreatic progenitors, and that these progenitors had an enhanced pancreatic gene expression signature at the expense of hepatic gene expression. We also found that epithelial-to-mesenchymal transition was commonly dysregulated in TBX3 mutant pancreas and hepatocyte cells. This suggests that TBX3 may function during EMT in both liver and pancreas development. These data highlight a potential role of TBX3 in distinguishing between hepatic and pancreatic domains during foregut patterning, with implications for enhancing the generation of pancreatic progenitors from PSCs.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Cell & Molecular Biology

First Advisor

Paul J. Gadue

Keywords

Endoderm patterning, Pancreatic progenitors, Pluripotent Stem Cells, TBX3

Subject Categories

Cell Biology | Developmental Biology

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/4221
LOSS OF TBX3 ENHANCES PANCREATIC PROGENITOR GENERATION FROM HUMAN
PLURIPOTENT STEM CELLS

Somdutta Mukherjee
A DISSERTATION
in
Cell and Molecular Biology
Presented to the Faculties of the University of Pennsylvania
in
Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
2021

Supervisor of Dissertation

____________________
Paul J. Gadue, PhD
Associate Professor of Pathology and Laboratory Medicine

Graduate Group Chairperson

____________________
Daniel S. Kessler, PhD
Associate Professor of Cell and Developmental Biology

Dissertation Committee

Christopher J. Lengner, PhD; Associate Professor, Department of Biomedical Sciences
Kenneth S. Zaret, PhD; Joseph Leidy Professor, Department of Cell and Developmental Biology
Montserrat C. Anguera, PhD; Associate Professor, Department of Biomedical Sciences
Kathleen M. Loomes, MD; Professor of Pediatrics, Department of Pediatrics
LOSS OF TBX3 ENHANCES PANCREATIC PROGENITOR GENERATION FROM HUMAN PLURIPOTENT STEM CELLS

COPYRIGHT
2021
Somdutta Mukherjee
ACKNOWLEDGMENT

There are several people that I need to thank for helping and supporting me throughout my graduate career. First, and most importantly, I would like to thank my thesis advisor Paul Gadue. I have grown so much as a scientist over these last eight years, and I could not have done it without your guidance and support. When I changed my entire project at the end of my fifth year (a grad student’s worst nightmare), your enthusiasm and passion for science really helped encourage me to move forward and pursue this new direction. I have learned so much during my time in your lab (and discovered some cool things!), and I truly cannot thank you enough for giving me the opportunity to be a part of your lab.

I would also like to thank Debbie French for the feedback and support over the years. You’ve always asked insightful questions and given me great suggestions and really helped my project grow. Thank you for always encouraging me along the way.

I would also like to thank all the members, past and present, of the French-Gadue lab. This has been my home for the last eight years, and you all have made this such a fun place to work. Karla, thank you for helping me learn the pancreas differentiation (and commiserating with me when they don’t work). Jean Ann and Alyssa, I have learned so much from you both. Thank you for being great lab mates, but more importantly great friends, who I always have a good time with, both in and outside of lab. And to everyone else in the lab, thank you for sharing your knowledge and helping me out whenever I’m stuck on something. To all the past members of the lab, especially Chintan and Chia-min, thanks for going the extra mile to help me out whenever I’ve needed it. Finally, to all former grad students of the Gadue Lab, Amita, Xiuli, Leo, Sid, and Sara, I appreciate your encouragement, and always looked forward to our fun grad student dinners. Amita,
you always made late nights in lab a little less unbearable. Xiuli, thank you for all your helpful advice. Leo, thank you for teaching me how to grow stem cells all those years ago as a rotation student. Sid, thanks for helping with the pancreas and genomics experiments when I was completely lost. And Sara, our grad school “milestones” were often at the same time, and I’m glad I had someone to go through the ups and downs with. Thank you for all your support along the way, and for being a good friend.

I would like to thank my committee members for all their suggestions and guidance over the years. Their input has really shaped my project and helped it become a fully developed story. I’d also like to thank CAMB and the DSRB programs, in particular Steve DiNardo, Mary Mullins, and Meagan Schofer, for providing a supportive community of peers and faculty. I am so grateful to have been a part of DSRB, and to have had so many opportunities and learn and grow as a scientist.

To all of my CAMB friends, I am so glad to have met all of you. Whether we were having fun, or bonding over failed experiments, each and every one of you have helped me get through grad school. To Lisa, Lexy, Camille, Kelsey, and Becca, I’ve enjoyed all our wine and cheese nights, and celebrating life’s milestones with you. Steve and Terra, I’m so glad we were able to go on all our fun national park adventures and make some great memories along the way. I cannot imagine grad school without any of you, and I am so grateful to have made some lifelong friends.

Joe, I really could not ask for a better boyfriend. Thank you for being so supportive and for being there for me. Whether it was helping me troubleshoot my Western blots, making delicious cake, or simply just reminding me that everything will be okay, you’ve really made these last years of grad school so much easier for me.
Last, but not least, I have to thank my parents. Your love and support have meant so much to me. Whether it was sending me papers you thought might help with my project, or buying things in bulk for me from Costco so I don’t have to go shopping, I have always appreciated everything you have done for me. You have always encouraged me to keep going, especially when I was frustrated. Thank you for always pushing me to try harder and do my best, both as a scientist, and in life. I would not be the person I am today without you both. I don’t say it enough, but thank you for everything you do for me. Words cannot express how grateful I am for you.
ABSTRACT

LOSS OF TBX3 ENHANCES PANCREATIC PROGENITOR GENERATION FROM HUMAN PLURIPOTENT STEM CELLS

Somdutta Mukherjee
Paul J. Gadue

Tbx3 has been identified as a regulator of fate decisions during liver development in the mouse, but whether it functions similarly in humans remains unclear. We generated TBX3 knockout human pluripotent stem cell (PSCs) lines using CRISPR/Cas9 genome editing technology. We differentiated TBX3 knockout lines to hepatocytes and found a decrease in hepatic markers and in hepatocyte function, demonstrating that TBX3 is also important for liver differentiation in humans. Surprisingly, we detected expression of pancreatic markers, including PDX1 in our TBX3 knockout hepatocytes, suggesting the possibility that TBX3 may regulate liver development by suppressing a pancreatic fate. We next differentiated the TBX3 knockout lines to pancreatic progenitors to determine whether the loss of TBX3 impacts pancreatic differentiation. We found that TBX3 knockout PSCs generated more pancreatic progenitors, and that these progenitors had an enhanced pancreatic gene expression signature at the expense of hepatic gene expression. We also found that epithelial-to-mesenchymal transition was commonly dysregulated in TBX3 mutant pancreas and hepatocyte cells. This suggests that TBX3 may function during EMT in both liver and pancreas development. These data highlight a potential role of TBX3 in distinguishing between hepatic and pancreatic domains during foregut patterning, with implications for enhancing the generation of pancreatic progenitors from PSCs.
# TABLE OF CONTENTS

ACKNOWLEDGMENT ........................................................................................................III

ABSTRACT ..........................................................................................................................VI

LIST OF TABLES ...................................................................................................................IX

LIST OF ILLUSTRATIONS ..................................................................................................X

LIST OF ABBREVIATIONS ..................................................................................................XII

CHAPTER 1: INTRODUCTION AND OVERVIEW ...............................................................1

1.1 Characteristics of the liver ...........................................................................................1
  1.1.1 Liver structure ........................................................................................................1
  1.1.2 Liver function .........................................................................................................3

1.2 Characteristics of the Pancreas ...................................................................................4

1.3 Endoderm, liver, and pancreas development .............................................................6
  1.3.1 Endoderm development .........................................................................................6
  1.3.2 Liver development ................................................................................................11
  1.3.3 Pancreas development ..........................................................................................22

1.4 Liver versus pancreas development ..........................................................................33

1.5 Characteristics of TBX3 .............................................................................................36
  1.5.1 Structure of TBX3 ................................................................................................37
  1.5.2 Functions of TBX3 ................................................................................................39
  1.5.3 TBX3 in liver development ...................................................................................41

1.6 Pluripotent stem cells .................................................................................................43
  1.6.1 Discovery of embryonic stem cells .......................................................................43
  1.6.2 Induced pluripotent stem cells .............................................................................44
  1.6.3. Pluripotent stem cells as a model system ..........................................................45
  1.6.4 Differentiation of PSCs to hepatocytes and pancreatic β-cells ..............................47

1.7 Overview of research goals .......................................................................................50

CHAPTER 2: MATERIALS AND METHODS ....................................................................51

2.1 Pluripotent stem cell culture and maintenance .........................................................51

2.2 Generation of TBX3 mutant lines using CRISPR/Cas9 genome editing .................51
2.3 Hepatocyte differentiation ................................................................. 52
2.4 Pancreatic β-cell Differentiation ......................................................... 54
2.5 Genomic Analysis ............................................................................. 55
2.6 RNA isolation and quantitative real time polymerase chain reaction .... 56
2.7 Western Blot ..................................................................................... 57
2.8 Flow Cytometry ................................................................................ 58
2.9 Cytochrome P450 activity assay ......................................................... 60
2.10 Statistical Analysis .......................................................................... 60

CHAPTER 3: IMPACTS OF LOSS OF TBX3 ON HEPATOCYTE AND PANCREAS DIFFERENTIATION FROM HUMAN PLURIPOTENT STEM CELLS ........................................ 61

3.1 Introduction ....................................................................................... 61

3.2 Results ............................................................................................. 63
  3.2.1 Loss of TBX3 impairs hepatocyte differentiation in human PSCs .... 63
  3.2.2 Pancreas-specific genes are upregulated in TBX3 knockout cells ..... 67
  3.2.3 Loss of TBX3 enhances pancreatic progenitor generation from PSCs 69
  3.2.4 TBX2 does not compensate for the loss of TBX3 in hepatocyte or pancreas differentiation ......................................................... 71
  3.2.5 iPSC+ cells are enriched for a pancreatic gene signature .............. 75

3.3 Discussion ......................................................................................... 80

CHAPTER 4: SUMMARY AND SPECULATIONS ........................................... 88

4.1 Summary of Findings ......................................................................... 88

4.2 Enhancing cholangiocyte differentiation from PSCs .......................... 88

4.3 TBX3 and epithelial-mesenchymal transition in pancreas differentiation 92

4.4 Applications for in vitro differentiation of PSCs to pancreatic β-cells ..... 97

4.5 Summary and conclusions ................................................................. 99

BIBLIOGRAPHY ..................................................................................... 102
LIST OF TABLES

CHAPTER 2: MATERIALS AND METHODS

Table 2.1: Primers sequences utilized in generation of TBX3 mutant PSC lines........52
Table 2.2. Table of forward and reverse primers used for qRT-PCR..................57
Table 2.3: Primary and secondary antibodies....................................................59

CHAPTER 3: IMPACTS OF LOSS OF TBX3 ON HEPATOCYTE AND PANCREAS DIFFERENTIATION FROM HUMAN PLURIPOTENT STEM CELLS

Table 3.1: TBX3 mutant pluripotent stem cell lines........................................63
Table 3.2: Top pathways downregulated in iPSC+/primitive gut tube cells.........78
Table 3.3: Top pathways upregulated in iPSC+/PP1 cells.................................78
Table 3.4: Genes commonly up- and downregulated in iPSC+/primitive gut tube, PP1, and hepatoblast cells...............................................................80
LIST OF ILLUSTRATIONS

CHAPTER 1: INTRODUCTION AND OVERVIEW

Figure 1.1: Structure of hepatic sinusoid.................................................................2
Figure 1.2: Pancreas structure..................................................................................5
Figure 1.3: Anterior-posterior axis patterning in the endoderm during gut tube formation.........................................................................................................................10
Figure 1.4: Overview of liver development..............................................................15
Figure 1.5: Overview of pancreas development....................................................23
Figure 1.6: Liver versus ventral pancreas specification...........................................34
Figure 1.7: Structure and function of TBX3............................................................38
Figure 1.8 Differentiating PSCs to endodermal lineages.........................................48

CHAPTER 3: IMPACTS OF LOSS OF TBX3 ON HEPATOCYTE AND PANCREAS DIFFERENTIATION FROM HUMAN PLURIPOTENT STEM CELLS

Figure 3.1: Loss of TBX3 impairs hepatocyte differentiation in the iPSC+/− line....65
Figure 3.2: Pancreatic genes are expressed in the iPSC+/− line during hepatocyte differentiation.........................................................................................................................68
Figure 3.3: Loss of TBX3 enhances pancreatic progenitor generation in the iPSC+/− line...............................................................................................................................70
Figure 3.4 TBX2 does not compensate for the loss of TBX3 in hepatocyte or pancreas differentiation...........................................................................................................72
Figure 3.5: iPSC+/− PP1 cells are enriched for a pancreatic gene signature............76
Figure 3.6: Characterization of TBX3 mutant PSC lines.......................................82
Figure 3.7: Loss of TBX3 impairs hepatocyte differentiation in the ESC+/− line....83
Figure 3.8: Pancreatic genes are expressed in the ESC+/− line during hepatocyte differentiation....................................................................................................................85
Figure 3.9: Loss of TBX3 enhances pancreatic progenitor generation in the ESC+/− line...............................................................................................................................86
Figure 3.10: iPSC+/++ PGT and hepatoblast cells are enriched for hepatic gene signature.........................................................................................................................87
CHAPTER 4: SUMMARY AND SPECULATIONS

Figure 4.4: Cholangiocyte marker expression in TBX3 mutant PSCs…………….90
Figure 4.2: Pancreatic hormone production in TBX3 mutant PSCs……………...94
Figure 4.3: Proposed biphasic role of TBX3 during pancreas development……96
Figure 4.4: Summary of loss of TBX3 in β-cell differentiation from PSCs……..101
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>α-fetoprotein</td>
</tr>
<tr>
<td>ALK5i II</td>
<td>ALK 5 inhibitor II</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone morphogenetic protein</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>Cas9</td>
<td>CRISPR associated protein 9</td>
</tr>
<tr>
<td>C/EBPα</td>
<td>CCAAT/enhancer-binding protein alpha</td>
</tr>
<tr>
<td>CDX2</td>
<td>Caudal type homeobox 2</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>DE</td>
<td>Definitive endoderm</td>
</tr>
<tr>
<td>DEX</td>
<td>Dexamethasone</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s minimum essential medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>ESC</td>
<td>Embryonic stem cells</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>FGF</td>
<td>Fibroblast growth factor</td>
</tr>
<tr>
<td>gRNA</td>
<td>Guide RNA</td>
</tr>
<tr>
<td>HGF</td>
<td>Hepatocyte growth factor</td>
</tr>
<tr>
<td>hESC</td>
<td>Human embryonic stem cells</td>
</tr>
<tr>
<td>HHEX</td>
<td>Hematopoietically expressed homeobox</td>
</tr>
<tr>
<td>HNF4α</td>
<td>Hepatocyte nuclear factor 4 alpha</td>
</tr>
<tr>
<td>ICM</td>
<td>Inner cell mass</td>
</tr>
<tr>
<td>IHBD</td>
<td>Intrahepatic bile ducts</td>
</tr>
<tr>
<td>iPSC</td>
<td>Induced pluripotent stem cells</td>
</tr>
<tr>
<td>mESC</td>
<td>Mouse embryonic stem cells</td>
</tr>
<tr>
<td>MPC</td>
<td>Multipotent pancreatic progenitor cells</td>
</tr>
<tr>
<td>MTG</td>
<td>α-Monothioglycerol, α-Thioglycerol, 3-Mercapto-1,2-propanediol</td>
</tr>
<tr>
<td>SFD</td>
<td>Serum-free defined</td>
</tr>
<tr>
<td>NGN3</td>
<td>Neurogenin 3</td>
</tr>
<tr>
<td>OSM</td>
<td>Oncostatin M</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative-reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDX1</td>
<td>Pancreatic and duodenal homeobox 1</td>
</tr>
<tr>
<td>PROX1</td>
<td>Prospero related homeobox protein 1</td>
</tr>
<tr>
<td>PSC</td>
<td>Pluripotent stem cells</td>
</tr>
<tr>
<td>PTF1a</td>
<td>Pancreas transcription factor 1a</td>
</tr>
<tr>
<td>RA</td>
<td>Retinoic acid</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell Park Memorial Institute Medium</td>
</tr>
<tr>
<td>ROCK</td>
<td>Rho-associated kinase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SAA</td>
<td>Stabilized ascorbic acid</td>
</tr>
<tr>
<td>SHH</td>
<td>Sonic hedgehog</td>
</tr>
<tr>
<td>SOX9</td>
<td>Sry box factor 9</td>
</tr>
<tr>
<td>STM</td>
<td>Septum transversum mesenchyme</td>
</tr>
<tr>
<td>TALENs</td>
<td>Transcription-activator-like effector nucleases</td>
</tr>
<tr>
<td>TBX3</td>
<td>T-box transcription factor 3</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA binding protein</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffered saline</td>
</tr>
<tr>
<td>TGFβ</td>
<td>Transforming growth factor beta</td>
</tr>
<tr>
<td>TTR</td>
<td>Transthyretin</td>
</tr>
<tr>
<td>UMS</td>
<td>Ulnar-mammary syndrome</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION AND OVERVIEW

1.1 Characteristics of the liver

The liver is the largest internal organ in the body, and is a major center for metabolism, digestion, and detoxification. As part of this thesis focuses on the transcription factor TBX3 and its role in endoderm patterning and liver development, we will first review the basic structure and function of the liver.

1.1.1 Liver structure

The liver has four lobes that are made up of hepatic lobules, hexagonal structures that are constituted of several different types of cells arranged around a central vein (Figure 1.1). The portal triad, which is comprised of a hepatic artery, hepatic portal vein, and intrahepatic bile duct (IHBD), are located at each corner of the hepatic lobule. Hepatocytes are arranged in cords-like structures between the central vein and the portal triad, and are separated by sinusoidal vessels that carry blood through the lobule. Hepatocytes comprise about 80% of the liver, and carry out the majority of liver functions. The sinusoids are made up of specialized liver endothelial cells, and Kupffer cells and Stellate cells are located within the sinusoidal space. A small subset of cells adjacent to the portal vein differentiate into cholangiocytes, epithelial cells that line the bile ducts (Gordillo et al., 2015).

Hepatocytes are the main functional cell type in the liver. Although they all have the same structure, the specific function of a hepatocyte is determined by its location within the hepatic acinus, the functional unit of the liver. The different zones are determine by signaling gradients, and can change in response to environmental factors, demonstrating that zonation is a dynamic phenomenon (Kietzmann, 2017). Blood flow
Figure 1.1: Structure of hepatic sinusoid

Functionally mature hepatocytes (red) exist in cords between the central vein and the portal triad, made up of the hepatic artery, portal vein, and intrahepatic bile duct (dashed circle). The central vein and portal triad are connected by sinusoidal spaces which contain specialized endothelial cells (brown), Kupffer cells (blue), and stellate cells (yellow). Created with BioRender.com
creates an oxygen and nutrients gradients which contribute to the zonation process within the hepatic acinus, to determine the specific function of the hepatocytes (Trefts et al., 2017).

1.1.2 Liver function

The liver carries out several metabolic functions. It regulates glucose metabolism by taking up blood glucose and synthesizing glycogen in response to increased levels of insulin. Conversely, glycogen, which is stored in the liver, is broken down into glucose in response to increased levels of glucagon (Bechmann et al., 2012; Trefts et al., 2017). The liver also synthesizes lipids de novo and from fatty acids taken up from the blood, and also breaks them down. The liver synthesizes alternative sources of energy such as ketones and triglycerides from fat, and is a prominent site of cholesterol synthesis in the body (Nguyen et al., 2008; Trefts et al., 2017). Bile, which emulsifies and digests lipids, is also produced by hepatocytes and secreted into spaces called the canaliculi. As it travels through the canaliculi, cholangiocytes modify bile by secreting and absorbing solutes and water via transmembrane channels. These processes are regulated by hormones and signaling pathways including Ca^{2+} and cAMP (Banales et al., 2019). The modified bile then drains into the IHBDs, which connect to the gallbladder where bile is stored (de Buy Wenniger and Beuers, 2010; Pitt and Nakeeb, 2017). Finally, the liver makes and secretes many proteins into the bloodstream, including albumin, growth factors, clotting factors and other regulatory proteins. Amino acids metabolism also takes place in the liver (Trefts et al., 2017).

Another major function of the liver is detoxification. Drugs and toxins are metabolized in the liver through a three-step process. In phase I metabolism, compounds undergo oxidation/reduction or hydrolysis reactions by enzymes, such as cytochrome P450
enzymes, to convert them to a safer or more active form. In phase II metabolism, compounds from phase I undergo various modifications so that they are more stable and easier to export into the bile. During phase III metabolism, toxins and other compounds generated in phase II are excreted into the bile or transported to the kidneys where they are then removed from the body. Additionally, the liver is responsible for sequestering and excreting metals, such as iron or copper, which are essential to the body, but can be toxic at high levels (Blondet et al., 2018).

1.2 Characteristics of the Pancreas

The pancreas is an organ that functions in both metabolism and digestion. As part of this thesis focuses on the transcription factor TBX3 and its role in endoderm patterning and pancreatic progenitor generation from pluripotent stem cells, we will next review the basic structure and function of the pancreas.

The pancreas is an endodermal organ that originates from two sections of the posterior foregut endoderm that bud off and come together. The pancreas has exocrine function to aid in digestion, and endocrine function to regulate glucose homeostasis. The endocrine component of the pancreas is comprised of acinar cells, which secrete digestive enzymes and fluids, and ductal cells which form the pancreatic ducts through which the “pancreatic juices” drain into the duodenum to during digestion (Bastidas-Ponce et al., 2017). The endocrine function of the pancreas is carried out by structures called the Islets of Langerhans. These islets are made up of five types of cells: α-cells which secrete glucagon, β-cells which secrete insulin, δ-cells which secrete somatostatin, PP-cells which secrete pancreatic polypeptide, and ε-cells which secrete ghrelin (Figure 1.2). The Islets of Langerhans form tight clusters and are found throughout the pancreas. The islets are closely associated with blood vessels in the
The mature pancreas has both exocrine and endocrine function. Acinar cells (blue) secrete digestive enzymes into the pancreatic ducts. The pancreatic ducts are branched tubes lined with epithelial ductal cells (purple) that connect to the common bile duct and drain into the duodenum. Structures called the Islets of Langerhans make up the endocrine component of the pancreas. Islets are made up of five cell types which each secrete a different hormone: α-cells (glucagon), β-cells (insulin), δ-cells (somatostatin), PP-cells (pancreatic polypeptide), and ε-cells (ghrelin). The cells of the islet function in regulating blood glucose levels. Created with BioRender.com.
pancreas, which helps control the release of hormones that regulate blood glucose levels (Pan and Wright, 2011). The two main hormones involved in glucose homeostasis are insulin and glucagon. When blood glucose levels are high, β-cells release insulin to stimulate glucose uptake by skeletal muscles and trigger glycogen synthesis in the liver. When blood glucose levels are low, α-cells release glucagon to stimulate glucose production and release in the liver (Aronoff et al., 2004). Impaired glucose regulation leads to diabetes mellitus, a disease that impacts millions of people world-wide.

1.3 Endoderm, liver, and pancreas development

The main focus of this thesis is studying the role of TBX3 in endoderm, liver, and pancreas development. The development of all three tissues has been extensively studied in several model organisms. To better understand how the work presented in this thesis fits into these areas, we will review what is known about endoderm, liver, and pancreas development.

1.3.1 Endoderm development

Gastrulation is a process that establishes the three germ layers: the ectoderm, mesoderm, and endoderm. This thesis focuses on both liver and pancreas development, organs which are derived from the inner most germ layer called the endoderm. Additionally, this thesis examines the possible role of the transcription factor TBX3 in establishing the hepatic and pancreatic domains of the foregut endoderm. As such, we will begin with a brief overview of endoderm development, prior to liver and pancreas development.
1.3.1.1 Endoderm formation

During gastrulation, epiblast cells migrate to the interior of the embryo through a structure at the posterior end called the primitive streak (Lawson et al., 1986). These cells go through a transient bipotent intermediate stage called the mesendoderm (Lickert et al., 2002), before settling in the middle layer to become the mesoderm or the innermost layer to become the definitive endoderm (DE). As the DE develops, it is divided into three regions: the foregut, midgut, and hindgut. The foregut gives rise to the esophagus, trachea, lungs, thyroid, stomach, liver, pancreas, and gallbladder, while the midgut gives rise to the small intestine, and the hindgut gives rise to the colon (Zorn and Wells, 2009). Cells that migrate through the primitive streak first become anterior endoderm (foregut), while cells migrating in later are part of the posterior endoderm (hindgut) (Lawson et al., 1987). The DE cells undergo a series of migration and extension movements to form an anterior-posterior axis. The anterior and posterior ends then fold and move towards each other to form the primitive gut tube (Zorn and Wells, 2009).

1.3.1.2 Nodal signaling in endoderm development

*Nodal*, a member of the transforming growth factor-β family, is crucial for many events in endoderm development. Nodal signaling is responsible for primitive streak formation and maintenance (Conlon et al., 1994). Evidence shows that the Wnt pathway maintains Nodal signaling. In the mouse, *Wnt3a* stimulates *Nodal* expression during primitive streak formation (Ben-Haim et al., 2006). Once the cells migrate through the primitive streak, *Nodal*, regulates mesendoderm differentiation. Studies across many species show that low levels of *Nodal* induce mesoderm, while high *Nodal* is required for DE
specification (Green and Smith, 1990; Shen, 2007). The anterior region of the primitive streak produces high levels of Nodal to drive DE development.

Nodal signaling mediates DE differentiation by regulating expression of various transcription factors. Nodal induces expression of Mixer family transcription factors which are important for DE development by suppressing a mesoderm fate (Shen, 2007; Zorn and Wells, 2009). In both Xenopus and mouse, loss of Mixer genes results in impaired DE formation and expanded mesoderm formation (Hart et al., 2002; Kofron et al., 2004). Nodal signaling also induces expression of Foxa2, Sox17, and Eomesodermin, which are important for DE development, as mutations in all three genes result in impaired endoderm development. Not only do these transcription factors control endodermal gene expression, they also influence cellular movements of during anterior-posterior axis formation (Arnold et al., 2008; Dufort et al., 1998; Kanai-Azuma et al., 2002), tying Nodal signaling to DE morphogenesis as well.

Nodal signaling distinguishes between a mesoderm and endoderm fate through close interactions with other signaling pathways, including fibroblast growth factor (FGF) and bone morphogenetic protein (BMP). As mentioned previously, high levels of Nodal signaling induce DE, while lower levels induce mesoderm. Nodal acts in a dose-dependent manner in conjunction with FGF signaling to distinguish these fates. In the presence of FGF and low nodal signaling, the mesodermal gene Xenopus Brachyury (Xbra), homologous to Brachyury in mice, is expressed. However, high levels of nodal signaling repress Xbra expression (Latinkic et al., 1997). Additionally, in zebrafish, BMP and FGF signaling inhibit endoderm differentiation by reducing the number of endodermal precursors. FGF inhibits downstream effectors of Nodal that are required for endoderm differentiation, (Poulain et al., 2006), highlighting the interplay between
multiple signaling pathways during endoderm development. Nodal signaling plays many crucial roles during endoderm development; it is involved in primitive streak formation, morphogenetic movements in body axis formation, and distinguishing between endoderm and mesoderm fates.

1.3.1.3 Endoderm patterning

Once the endoderm is specified and the primitive gut tube forms, gradients of the signaling pathways Wnt, FGF, BMP, and retinoic acid pattern the gut tube into the foregut, midgut, and hindgut (Figure 1.3). The three domains are distinguished by broad expression pattern of transcription factors. The foregut at the anterior end gut tube is defined by Sox2 expression, Pdx1 is expressed in the midgut, and Caudal homeobox (Cdx) genes are expressed in the hindgut at the posterior end of the gut tube (Figure 1.3). The signals that pattern the endoderm come from adjacent mesodermal tissues (Kumar et al., 2003; Wells and Melton, 2000), demonstrating the close relationship between the two germ layers. As development continues, each region receives different levels of the signals. This results in expression of lineage-specific transcription factors in distinct domains, permitting each section of the gut tube to develop into different organs.

Wnt signaling is high in the posterior gut tube and promotes an intestinal fate while repressing a foregut fate (Cervantes et al., 2009; Gregorieff et al., 2004). Repressing β-catenin, a downstream effector of canonical Wnt signaling, results in ectopic liver and pancreas (foregut-derived organs) development in the posterior gut tube. Additionally, β-catenin directly represses Hhex, a foregut specific gene, to posteriorize the gut tube (McLin et al., 2007). Similarly, FGF signaling suppresses a foregut fate while promoting a posterior fate. Mesoderm-derived FGF4 signals act on neighboring endoderm cells to induce expression of posterior gut tube markers Pdx1 and CdxB, and suppresses
Figure 1.3: Anterior-posterior axis patterning in the endoderm during gut tube formation

Signals from surrounding mesodermal tissues pattern the gut tube in an anterior-posterior fashion. WNT, FGF, and BMP signals promote an intestinal fate in the mid- and hindgut at the posterior end of the gut tube while inhibiting an anterior fate. Conversely, retinoic acid signaling establishes the foregut domain in the anterior gut tube and suppresses a posterior fate. These signals induce transcription factors which broadly pattern the gut tube. The foregut is characterized by SOX2 expression, while the midgut and hindgut express PDX1 and Caudal homebox (CDX) genes, respectively. Different levels of signaling activity induce lineage-specific transcription factors in distinct domains in the gut tube, specifying these regions to develop to different endodermal organs. Adapted from Zorn and Wells, 2009. Created with Biorender.com.
foregut markers *Hex* and *Nkx2.1*. FGF4 is also important for establishing the midgut-hindgut boundary, illustrating its pivotal role in gut tube patterning (Dessimoz et al., 2006). BMP signaling also inhibits anterior gut tube development while promoting a more posterior fate. Enhanced BMP signaling diminishes the anterior domain and expands the posterior gut domain. Conversely, low BMP signaling expands the anterior domain into the posterior regions of the gut tube (Tiso et al., 2002). On the other hand, retinoic acid (RA) signaling promotes an anterior fate. Inhibiting RA signaling results in a loss of pancreatic and liver domains, and ectopic expression of posterior endoderm markers. Overexpression of RA has the opposite effect, with anterior markers expanding to posterior regions (Stafford and Prince, 2002). Taken together, the Wnt, FGF, and BMP signaling pathways pattern the gut tube in an anterior-posterior fashion by promoting mid- and hindgut fates. The anterior gut tube receives RA signals instead and patterns anterior endoderm into the foregut.

Understanding gut tube patterning has been instrumental in developing protocols to differentiate pluripotent stem cells (PSCs) to DE and other endoderm-derived cell types. Studies from model organisms have allowed us to mimic the appropriate levels and timing of signals to efficiently direct PSCs to primitive streak (Gadue et al., 2006), mesendoderm (Tada et al., 2005), and DE (D’Amour et al., 2005). PSC-derived DE can be further differentiated to endodermal lineages, including hepatocytes and pancreatic β-cells, which is discussed in Section 1.6.4 of this thesis.

### 1.3.2 Liver development

The liver arises from the foregut domain of the endoderm. Prior to hepatic specification, there are two populations of progenitors that give rise to the liver bud: two lateral liver
progenitor domains and a ventral midline liver domain. Fate-mapping studies show that these populations exist in spatially distinct locations in the embryo, which come together to form the region of foregut that becomes the liver. Interestingly, the lateral domains can only give rise to cells of the liver and the pancreas, while the ventral midline domain can give rise to the liver bud as well as other endodermal midline tissues (Tremblay and Zaret, 2005). Once these domains come together, transcription factors that act as pioneer factors prepare this region for hepatic specification.

Pioneer transcription factors bind to regions of the chromatin to facilitate gene transcription. One way they do this is by binding target sites, even in tightly packed heterochromatin, and facilitate opening them up to allow other factors to bind and activate transcription (Zaret and Carroll, 2011). FoxA and Gata4 transcription factors are bound to the liver-specific Albumin enhancer prior to hepatic specification (Bossard and Zaret, 1998; Cirillo et al., 2002; Gualdi et al., 1996), making this region open for transcription after the endoderm receives hepatic specification signals (Figure 1.4A). Mice lacking FoxA family members Foxa1 and Foxa2 fail to undergo hepatic specification, as there is no detectable liver bud or expression the hepatic marker Afp. These phenotypes are not seen in Foxa1 and Foxa2 single mutants (Lee et al., 2005), showing that both transcription factors are required to establish hepatic competence prior to specification.

1.3.2.1 Hepatic specification

Hepatic specification begins around E8.5 in the mouse, or 23 days of gestation in humans (Gordillo et al., 2015). The ventral foregut lies next to the cardiac mesoderm and the septum transversum mesenchyme, which secrete fibroblast growth factor (FGF) and bone morphogenetic protein (BMP) signals, respectively, to induce hepatic
specification (Figure 1.4B). Ventral foregut cells from mouse embryos cultured with cardiac mesoderm express hepatic genes *Albumin* and *Afp*, while cells cultured without cardiac mesoderm do not, suggesting the cardiac mesoderm secretes signals necessary for hepatic specification. (Gualdi et al., 1996). Specifically, Fgf1 and Fgf2 are expressed by the cardiac mesoderm and can induce expression of *Albumin*, *Afp*, and *Ttr* in the adjacent ventral foregut endoderm (Jung et al., 1999). Furthermore, Fgf8 promotes cellular outgrowth of the liver bud (Jung et al., 1999). These findings suggest that FGF signals from the cardiac mesoderm are needed for proper hepatic specification and morphogenetic movements during liver development. Additionally, FGF signaling acts via the mitogen activated protein kinase (MAPK) pathway to influence liver development. Phosphorylated-ERK, the active form, is expressed in lateral hepatic progenitors and in the ventral foregut endoderm, correlating with *Fgf* expression in the adjacent tissues. The FGF/MAPK pathway induces *Afp* and *Albumin* expression, and is needed for hepatic differentiation and outgrowth during liver bud formation (Calmont et al., 2006). These studies establish the important role of the FGF pathway in hepatic specification.

BMP signaling is also critical to for hepatic induction. The septum transversum mesenchyme (STM) secretes Bmp4 to the adjacent ventral foregut endoderm to specify the hepatic endoderm. *Bmp4* null embryos exhibit delayed formation of the liver bud, and BMP inhibition prevents *Albumin* expression in the ventral foregut endoderm (Rossi et al., 2001). Additionally, Bmp4 maintains expression of Gata4 in the ventral foregut endoderm, allowing for hepatic gene expression upon receiving inductive FGF signals (Rossi et al., 2001). Though Bmp4 is primarily responsible for hepatic induction, Bmp2 can act redundantly (Chung et al., 2008; Rossi et al., 2001). BMP signaling works in conjunction with the FGF pathway to induce liver gene expression in the hepatic
endoderm, and is required for proper cell movement and morphogenesis during liver bud formation. In zebrafish embryos, Bmp and Fgf signaling were blocked in the region that gives rise to the liver prior to hepatic specification, resulting in a depleted expression of the hepatic specification markers hhex and prox1 (Shin et al., 2007). Both Fgf and Bmp signaling are required for expression of Hex and Albumin in the hepatic endoderm in chick embryos (Zhang et al., 2004), demonstrating that the role of these pathways in hepatic specification and development is conserved across multiple species.

Interestingly, Wnt signaling has been implicated in hepatic specification in zebrafish. Wnt2bb mutant embryos do not have any detectable hepatic tissue, and have little to no expression of hex or prox1. Wnt2bb is expressed from the lateral plate mesoderm, which is adjacent to the hepatic endoderm (Ober et al., 2006). Although there is little evidence for Wnt signaling in hepatic specification in other species, the study highlights the importance of mesoderm-derived signals during hepatic induction and specification.

Additionally, the transcription factor Hnf1β is required for hepatic specification. Hnf1β−/− embryos do not form a liver bud or express early hepatic markers including Afp, Albumin, Prox1, and Hex. Both Foxa1 and Foxa2 are expressed in the ventral foregut of Hnf1β−/− embryos, albeit at reduced levels, suggesting that Hnf1β may be required to sustain their expression. Finally, FGF signals do not induce Albumin or Ttr expression in ventral foregut endoderm explants from Hnf1β−/− embryos (Lokmane et al., 2008), suggesting that Hnf1β plays a key role in both establishing competence and hepatic specification. Once the hepatic endoderm receives inductive signals and is specified, a set of transcription factors facilitates a series of morphogenetic movements to form the liver bud.
Figure 1.4: Overview of liver development

Schematic summary of different stages of liver development.

(A) Prior to specification, pioneer factors Foxa1 and Fox2 bind to the enhancer of the liver-specific gene Albumin, opening it up for transcription.

(B) The cardiac mesoderm and septum transversum mesenchyme (STM) secrete fibroblast growth factor (FGF) and bone morphogenetic protein (BMP) signals, respectively, to the neighboring ventral foregut endoderm to specify the hepatic endoderm (red).

(C) The liver bud forms when the specified hepatic endoderm thickens and elongates away from the gut lumen. The cells, called hepatoblasts, become columnar in shape, then delaminate and migrate into the STM as the basal lamina breaks down.

(D) The liver bud undergoes rapid proliferation and expansion when the hepatoblasts receive Wnt, hepatocyte growth factor (HGF), and FGF signals.

(E) Hepatoblasts receive different signals and differentiate to either hepatocytes or cholangiocytes, the epithelial lining of the bile duct. Most hepatoblasts receive Oncostatin M (OSM), HGF, and Dexamethasone (Dex) signals to differentiate and mature into functional hepatocytes. A small subset of hepatoblasts receive Notch and Wnt signals to differentiate to cholangiocytes. Adapted from Gordillo et al., 2015. Created with BioRender.com.
1.3.2.2 Liver bud formation

At E9.5 in the mouse, or 26 days of gestation in humans, the specified hepatic endoderm begins to thicken and change shape, becoming a pseudostratified layer of cells called hepatoblasts. Hepatoblasts are bipotential progenitors that differentiate into either hepatocytes or cholangiocytes, epithelial cells that line intrahepatic bile ducts. Hepatoblasts express fetal markers (Af$\text{p}$), as well as both hepatocyte ($Hnf4\alpha$, $Albu$min, $Ck18$) and cholangiocyte markers ($Ck19$) (Gordillo et al., 2015). Liver bud formation can be separated into three stages. First, the hepatic endoderm thickens as it elongates away from the gut lumen and the cells take on a columnar shape. This occurs at the same time when hepatic gene expression is induced, and the ventral foregut closes. Next, the hepatoblasts form a pseudostratified layer, with nuclei being in different apical and basal positions within the cell. As these hepatoblasts divide, they undergo “interkinetic nuclear migration” (INM), with the nuclei moving to the apical side of the cell. Finally, the basement membrane degrades and the hepatoblasts delaminate and migrate into the STM to form the liver bud (Bort et al., 2006). The cellular changes that occur during liver bud formation are carefully controlled by several transcription factors (Figure 1.4C).

The transcription factor $Hex$ is responsible for proper patterning of the ventral endoderm and growth of the liver bud (Bort et al., 2004; Martinez Barbera et al., 2000). Hepatic genes are expressed in $Hex^{-/-}$ embryos but the liver bud fails to grow due to a reduction in proliferating cells (Bort et al., 2004). Additionally, hepatoblasts in $Hex^{-/-}$ embryos fail to undergo INM to form a pseudostratified layer, do not form a diverticulum, and do not delaminate into the STM (Bort et al., 2006). These findings suggest that $Hex$ is not required for hepatic specification, but rather for liver bud formation and growth.
Members of the GATA transcription factor family have also been implicated in liver bud formation and growth. *Gata4/5/6* are involved in liver bud formation in zebrafish (Holtzinger and Evans, 2005), and *Gata4/6* plays a similar role in the mouse (Zhao et al., 2005), demonstrating a conserved function across species. *Gata6*−/− embryos have a small liver outgrowth, but the hepatoblasts fail to delaminate. Hepatic gene expression appears normal in *Gata6*−/− embryos, likely due to *Gata4* compensating for the lack of *Gata6*. *Gata6*, but not *Gata4*, is expressed in delaminating hepatoblasts, explaining why there is a defect in liver bud migration, but not specification in *Gata6*−/− embryos (Zhao et al., 2005). Although *Gata4* and *Gata6* act redundantly in hepatic specification, they have distinct roles in hepatoblast migration.

The transcription factor Prox1 is also required for hepatoblasts to delaminate and migrate into the STM as the liver bud grows. *Prox1*−/− embryos have small liver buds with fewer proliferating hepatoblasts compared to wildtype embryos (Sosa-Pineda et al., 2000). Instead of migrating into the surrounding mesenchyme, hepatoblasts remain clustered around the hepatic diverticulum because the basal lamina surrounding the liver bud is not degraded. Alternatively, the migration defect may be explained by elevated levels of E-cadherin, which is normally downregulated in migrating hepatoblasts. Interestingly, hepatic gene expression is not perturbed in *Prox1*−/− embryos (Sosa-Pineda et al., 2000). Taken together these findings suggest that Prox1 mediates hepatoblast delamination and migration, not hepatic specification. The transcription factor Tbx3 also plays a role in hepatoblast migration as *Tbx3*−/− embryos have a similar phenotype as the *Prox1*−/− embryos. The role of Tbx3 in liver development is discussed in further detail in Section 1.5.3 of this thesis.
Similarly, the transcription factors Onecut1 (also known as Hnf6) and its close homolog Onecut2 (Oc2), play a role in hepatoblast migration. *Hnf6* and *Oc2* are expressed in early liver development and act redundantly. Mice deficient for both genes have a similar phenotype to *Prox1*−/− embryos in that they had small liver buds, a delay in basement membrane degradation, and the hepatoblasts fail to migrate into the STM. Additionally, *Hnf6* and *Oc2* regulate genes involved in cell migration and adhesion, including E-cadherin. (Margagliotti et al., 2007). After the hepatoblasts delaminate into the surrounding mesenchyme, they undergo rapid proliferation to form the liver.

### 1.3.2.3 Liver bud expansion

Around E10.0 in the mouse (31 days of gestation in humans), hepatoblasts migrate into the STM and the undergo rapid proliferation. (Gordillo et al., 2015). This process is controlled by several signaling pathways (Figure 1.4D). Although Wnt signaling initially plays an antagonistic role to liver development (see Section 1.3.1.3), it is crucial for hepatoblast proliferation during liver bud expansion (Tan et al., 2006). Inhibiting β-catenin in chick embryos leads to small livers (Suksaweang et al., 2004). Similarly, in zebrafish, β-catenin drives proliferation and growth in the developing liver (Goessling et al., 2008; McLin et al., 2007). Nuclear β-catenin levels are highest between E10.0 - E12.0 in the mouse, the same time as peak liver bud expansion (Micsenyi et al., 2004). Mice with a liver-specific deletion of β-catenin have reduced liver size due to fewer proliferating hepatoblasts (Tan et al., 2008). Additionally, Wnt signaling is activated during liver regeneration in the mouse and promotes hepatocyte proliferation (Goessling et al., 2008). These studies demonstrate the role conserved role of Wnt signaling in regulating proliferation and growth during liver development.
The hepatocyte growth factor (HGF) and FGF pathways also control hepatoblast proliferation by regulating β-catenin activity. Nuclear β-catenin promotes proliferation during liver regeneration (Monga et al., 2001) and in hepatocellular carcinoma (Nhieu et al., 1999). In rat hepatocytes, the HGF receptor Met physically interacts with β-catenin. When HGF binds Met, both Met and β-catenin become phosphorylated, and the Met-β-catenin complex dissociated, allowing for β-catenin translocation into the nucleus (Monga et al., 2002). Similarly, FGF signaling mediates β-catenin activation in hepatoblasts to regulate proliferation. Fgf10 is produced by hepatic stellate cells and binds to the Fgf2rb receptor, which is expressed on the surface of hepatoblasts, inducing a signaling cascade that activates β-catenin. Fgf10−/− and Fgfr2b−/− mouse embryos have small livers, and hepatoblast proliferation in Fgf10−/− embryos is impaired (Berg et al., 2007). Taken together, these studies demonstrate that the Wnt, HGF, and FGF pathways regulate hepatoblast proliferation via β-catenin during liver bud expansion.

1.3.2.4 Hepatocyte differentiation and maturation

Once liver bud expansion occurs, hepatoblasts begin to differentiate to either hepatocytes or cholangiocytes at E13.5 (around 56 days of gestation in humans) (Gordillo et al., 2015). Signaling via HGF, Oncostatin M (OSM), and glucocorticoids drive the majority of hepatoblasts towards a hepatocyte fate (Figure 1.4E). HGF is produced by sinusoidal mesenchymal cells of the developing liver, and is crucial for regulating proper morphology and size during liver development via the c-Met receptor, which is expressed on hepatocytes (Schmidt et al., 1995). HGF is also critical for maintaining signaling pathways and expression of genes involved in hepatocyte differentiation (Michalopoulos et al., 2003). During liver regeneration, HGF is upregulated upon injury and acts as a potent mitogen to induce hepatocyte proliferation (Michalopoulos and
DeFrances, 1997). Additionally, OSM produced by fetal hematopoietic cells in the liver, binds to its receptor gp130, which is expressed on the hepatocyte surface, to facilitate maturation. Hepatocytes from gp130−/− embryos are less functional compared to hepatocytes from wildtype embryos (Kamiya et al., 1999). OSM signaling regulates liver gene expression and maturation via the STAT3 pathway. Blocking this pathway using a dominant negative form of STAT3 attenuates hepatic gene expression and differentiation (Ito et al., 2000). Additionally, OSM signaling mediates the Ras signaling pathway in maturing hepatocytes. Specifically, OSM regulates K-Ras to ensure proper localization of the adherens junctions proteins E-cadherin and ZO-1. (Matsui et al., 2002). The glucocorticoid Dexamethasone also regulates E-cadherin and ZO-1 expression in fetal hepatocytes, demonstrating that this pathway works in conjunction with OSM to facilitate hepatocyte maturation (Matsui et al., 2002). Dexamethasone also works in combination with the HGF pathway to stimulate hepatocyte maturation (Kamiya et al., 2001). Both OSM and HGF signaling are required for hepatocyte maturation, but they act through distinct pathways and at different stages of development. OSM signaling is active in the developing mouse embryo at E14.0 and E18.0, but is not active postnatally, while HGF signaling continues after birth. (Kamiya et al., 2001). In addition to mediating hepatocyte function and cellular structure, these pathways induce expression of transcription factors necessary for hepatocyte maturation.

Two important transcription factors required for hepatocyte maturation are C/ebpα and Hnf4α. C/ebpα expression is first detected when hepatoblasts begin differentiating, suggesting it is important for hepatocyte development (Nagy et al., 1994; Shiojiri et al., 2004). C/ebpα−/− mice are born, but die shortly after birth, and have significantly smaller livers that do not function properly (Flodby et al., 1996). Interestingly, hepatocytes from
C/ebpα mice have increased expression of biliary markers (Tomizawa et al., 1998). C/ebpα binds to the promoter of Hnf6, which promotes a cholangiocyte fate, and negatively regulates its expression, providing evidence that C/ebpα promotes a hepatocyte fate by suppressing cholangiocyte differentiation (Rastegar et al., 2000). The STAT3 pathway is inactive in livers of C/ebpα−/− embryos (Burgess-Beusse and Darlington, 1998). However, this phenotype is rescued upon addition of OSM (Mackey and Darlington, 2004), suggesting the possibility that OSM signaling controls C/ebpα expression during hepatocyte maturation.

The transcription factor Hnf4α is also critical for proper hepatocyte differentiation, morphology, and function. Although both hepatoblast and liver formation occurs normally in Hnf4α−/− embryos, hepatocyte differentiation is severely impaired. Several hepatic transcription factors are still expressed in Hnf4α−/− embryos, except for Hnf1α and pregnane-x-receptor, suggesting that Hnf4α is upstream of these factors in a signaling cascade that regulates hepatocyte differentiation (Li et al., 2000). Hnf4α−/− embryos, have a severe defect in the liver morphology and function. Hepatocytes from Hnf4α−/− embryos are small, round, and loosely connected in contrast to hepatocytes from wildtype embryos, which are large, flat, and have tight associations. Additionally, livers from Hnf4α−/− embryos fail to express genes associated with glucose metabolism, and do not express cell adhesion and cell junction molecules including E-cadherin and ZO-1 (Parviz et al., 2003). HNF4α occupies a majority of actively transcribed genes in adult hepatocytes, further demonstrating its significant role in regulating liver function (Odom et al., 2004). To study further its role in hepatocyte maturation, Hnf4α was deleted in mouse embryos after the liver specification stage. These embryos have reduced hepatic
gene transcription compared to wildtype embryos (Kyrmizi et al., 2006). Hnf4α is crucial for regulating hepatic gene expression, liver function, and epithelial morphogenesis during hepatocyte differentiation. The transcription factor Tbx3, also plays a role in maintaining the transcriptional regulatory network required for proper hepatoblast fate decisions. This is discussed in further detail in Section 1.5.3 of this thesis.

The work done in model organisms to study liver development has been monumental in establishing protocols to differentiate pluripotent stem cells (PSCs) to hepatocytes. PSC-derived hepatocytes are a powerful tool for studying human liver development and uncovering any human-specific differences that have been previously unknown. The evolution of hepatocyte differentiation protocols is discussed in Section 1.6.4 of this thesis.

1.3.3 Pancreas development

The pancreas is derived from two distinct regions of the foregut endoderm that bud off and come together. Pancreas development occurs in two stages, correlating with the appearance of endocrine cells. In the primary transition, a few glucagon and insulin positive cells appear early in development, while only endocrine cells appearing in the secondary transition contribute to the mature islet (Herrera, 2000). During primary transition, the pancreatic endoderm is induced, pancreatic multipotent progenitor cells (MPCs) undergo rapid proliferation and expansion between days E9.0-E12.5. Secondary transition begins around E13.5, where pancreatic epithelium forms a complex branched network that is segregated into different domains, and the MPCs differentiate to mature pancreatic cell types (Bastidas-Ponce et al., 2017; Pan and Wright, 2011).
Figure 1.5: Overview of pancreas development

Schematic summary of different stages of pancreas development.

(A) Specification of dorsal pancreatic bud. The notochord secretes factors blocking sonic hedgehog signaling (Shh) in the adjacent dorsal pancreatic endoderm (purple), allowing for expression of pancreatic genes Pdx1, Ptf1a, Isl1, and Mnx1.

(B) Epithelial pancreatic branches are segregated into the tip (blue) and trunk (purple) domains. Notch signaling promotes Nkx6.1 expression in trunk cells, while Notch inhibition allows expression of Ptf1a in tip cells.

(C) Bipotential trunk cells differentiate to ductal cells (light purple) upon receiving Notch signals. Notch inhibition drives trunk cells to endocrine progenitors (green) which are characterized by Ngn3 expression. Endocrine progenitors then differentiate into the five endocrine cell types found in the pancreas: α-cells, β-cells, δ-cells, PP-cells, and ε-cells.

(D) Endocrine cells (green) undergo epithelial-mesenchymal transition (EMT) and delaminate from the trunk into the surrounding mesenchyme. Differentiated endocrine cells receive various signals that facilitate aggregation to form Islets of Langerhans, the functional endocrine component of the pancreas. Adapted from Bastidas-Ponce et al., 2017. Created with BioRender.com.
1.3.3.1 Dorsal and ventral pancreas specification

The pancreas develops from two regions of the foregut endoderm to form the dorsal and ventral pancreatic buds. Pdx1 is considered a master regulator of pancreas development, as it is one of the earliest genes expressed in pancreatic endoderm. Pdx1 expression is detected in the pancreatic epithelium around E8.5 (Ahlgren et al., 1996). Pdx1⁻/⁻ embryos form a rudimentary dorsal pancreatic bud and no ventral bud. However, the bud fails to expand and these mice do not form a pancreas (Jonsson et al., 1994; Offield et al., 1996). A small number of insulin- and glucagon-producing cells are detected in Pdx1⁻/⁻ embryos. However these embryos have far fewer insulin and glucagon positive cells than wildtype mice (Ahlgren et al., 1996). Lineage tracing studies show that multipotent progenitor cells (MPCs) expressing Pdx1 give rise to all three pancreatic lineages (Gu et al., 2002). These findings suggest Pdx1 is needed for ventral but not dorsal pancreatic induction, and is required for subsequent growth, differentiation, and maturation of MPCs.

Ptf1a is also important in early pancreas development. Ptf1a expression is detected in the pancreatic epithelium at E9.5, and Ptf1a positive cells give rise to all duct, acinar, and pancreatic endocrine cells (Kawaguchi et al., 2002). Similar to Pdx1 null embryos, Ptf1a⁻/⁻ embryos form a rudimentary dorsal pancreatic bud and no ventral bud. Ptf1a also acts in fate determination, as cells that should become pancreatic cells adopt a duodenum fate in Ptf1a⁻/⁻ embryos (Kawaguchi et al., 2002). Additionally, Ptf1a⁻/⁻ mice fail to develop pancreatic exocrine cells, and have fewer pancreatic endocrine cells which are mislocalized to the spleen (Krapp et al., 1998). These studies demonstrate that Ptf1a is important at both specification of the ventral bud, and later in pancreas development.
Two transcription factors that are important for dorsal bud formation are Mnx1 and Isl1. Embryos lacking Mnx1 do not form a dorsal pancreatic bud, while the ventral pancreatic bud appears largely normal. \( \text{Mnx1}^{-/-} \) embryos have small pancreatic islets, with fewer \( \beta \)-cells that are less functional compared to wildtype embryos (Harrison et al., 1999; Li et al., 1999). Similarly, \( \text{Isl1}^{-/-} \) embryos lack a dorsal pancreatic bud, with little effect on ventral pancreas development. These embryos completely fail to develop islets (Ahlgren et al., 1997). These findings reveal involvement of different transcription factors in dorsal and ventral bud specification, and yet again demonstrate that transcription factors have multiple roles in different stages of pancreatic development.

Signaling pathways also play a critical role in early pancreatic specification. The dorsal and ventral pancreatic epithelium have different surrounding tissues, thus receive different specification signals to induce bud formation. At E8.5, the dorsal pancreatic endoderm lies in close proximity to the notochord. The notochord expresses Activin-\( \beta \)B and FGF2, which suppress Sonic hedgehog (Shh) signaling in the dorsal pancreatic endoderm and allow for \( Pdx1 \) expression. (Hebrok et al., 1998; Kim et al., 1997) (Figure 1.5A). Retinoic acid (RA) is also critical for dorsal pancreatic development. In Xenopus, RA blocks Shh signaling in the pancreatic endoderm, and is needed for further pancreatic development (Chen et al., 2004). Additionally, mice lacking retinaldehyde dehydrogenase 2, which is involved in RA synthesis, do not express \( Pdx1 \) in the pancreatic endoderm, and do not form a dorsal pancreatic bud (Martín et al., 2005; Molotkov et al., 2005). At E9.0 in the mouse, the dorsal aorta fuse and displace the notochord, putting the dorsal pancreatic endoderm in direct contact with endothelial cells. Endothelial cells secrete signals that induce \( Pdx1 \) expression in the dorsal
pancreatic endoderm, and *Insulin*, *NeuroD*, and *Pax6* later in endocrine pancreas differentiation (Lammert et al., 2001).

The ventral pancreatic bud develops from the lateral domains of the ventral foregut endoderm, which also gives rise to the liver, extrahepatic bile ducts (Tremblay and Zaret, 2005). As described in Section 1.3.2.1, the ventral foregut endoderm lies is close proximity to the cardiac mesoderm and septum transversum mesenchyme, which secrete FGF and BMP signals, respectively, to specify the liver. The region of ventral foregut endoderm that do not receive these signals becomes ventral pancreatic bud (Deutsch et al., 2001; Rossi et al., 2001). The molecular mechanisms distinguishing liver versus ventral pancreas development are discussed in greater detail in Section 1.4 of this thesis. Additionally, multipotent progenitors expressing both *Pdx1* and *Sox17* are present in the ventral foregut endoderm. The Notch effector Hes1 downregulates Sox17 in a subset of these cells, permitting *Pdx1* expression, and therefore development of the ventral pancreas (Spence et al., 2009). Interestingly, *Hes1*−/− embryos form ectopic pancreatic tissue (Sumazaki et al., 2004), suggesting that Hes1 regulates ventral pancreatic bud development by preventing a biliary fate through repression of Sox17.

**1.3.3.2 Pancreatic multipotent progenitor cell expansion**

As the dorsal and ventral pancreatic endoderm are specified, they thicken and evaginate into the surrounding mesenchyme, forming the pancreatic buds. Both pancreatic buds are comprised of multipotent pancreatic progenitors (MPCs). The ventral bud rotates, coming into contact with the dorsal bud, and the two buds fuse to form the pancreas (Slack, 1995). The buds undergo a complex series of morphological changes to form the highly branched pancreatic duct epithelium (Villasenor et al., 2010). This transient stage allows MPCs to proliferate and generate a sufficient progenitor pool, which is critical for
appropriate size of the final pancreas organ (Stanger et al., 2007). In addition to their role in pancreas specification and fate determination, Pdx1 and Ptf1a are also important for maintaining MPC identity (Burlison et al., 2008). The transcription factor Sox9 also marks MPCs. Sox9 maintains the MPC pool by regulating proliferation and keeping MPCs in an undifferentiated state (Seymour et al., 2007). Additionally, Sox9 directly regulates the expression of the transcription factors Hnf1β, Hnf6, and Foxa2, which maintain MPC identity (Lynn et al., 2007). These transcription factors are crucial for generating and maintaining the MPC pool.

Additionally, Fgf10 signaling controls proliferation and the size of the MPC pool during branching morphogenesis. Fgf10−/− embryos form pancreatic buds, but they do not grow or undergo differentiation and subsequent branching (Bhushan et al., 2001). Fgf10 regulates MPC expansion by inducing Notch signaling to keep MPCs in an undifferentiated state (Hart et al., 2003; Norgaard et al., 2003). Mice lacking Hes1, a downstream effector of Notch, have fewer MPCs, leading to accelerated endocrine differentiation and pancreatic hypoplasia (Jensen et al., 2000a). Additionally, overexpression of Notch inhibits both exocrine and endocrine development, suggesting that this pathway is important for preventing MPC differentiation (Hald et al., 2003; Murtaugh et al., 2003). Notch signaling is key for maintaining MPC identity and generating the MPC pool.

1.3.3.3 Pancreatic trunk versus tip differentiation

At E11.5, pancreatic epithelium is divided into two domains: the tip domain which gives rise to acinar cells, and the trunk domain which gives rise to endocrine and ductal cells (Bastidas-Ponce et al., 2017). The Nkx6 factors (Nkx6.1 and Nkx6.2) and Ptf1a play an antagonistic role in segregating these domains. Nkx6.1 expression is to restricted to the
trunk while *Ptf1a* is in the tip (Figure 1.5B). Loss of Nkx6 factors results in *Ptf1a* expression in the trunk domain, and in an endocrine-to-acinar cell fate switch (Schaffer et al., 2010). Conversely, loss of *Ptf1a* results in an acinar-to-endocrine fate switch (Dong et al., 2008). The expression of *Nkx6.1* in the trunk and *Ptf1a* in the tip is regulated by Notch signaling (Figure 1.5B). Notch drives *Nkx6.1* expression in trunk MPCs to prevent an acinar fate (Afelik et al., 2012). Additionally, Notch represses *Ptf1a* expression, and loss of Notch signaling enhances pancreatic exocrine cell differentiation (Esni et al., 2004). Finally, several studies (Apelqvist et al., 1999; Jensen et al., 2000a) implicate Notch signaling in determining an endocrine versus exocrine fate, demonstrating another crucial role for this pathway in pancreas development.

1.3.3.4 Endocrine specification

Trunk cells contain bipotential progenitors that give rise to pancreatic endocrine or ductal cells (Figure 1.5C). *Neurogenin3* (*Ngn3*) is a marker of endocrine progenitors (Apelqvist et al., 1999; Jensen et al., 2000b). Lineage tracing experiments reveal that Ngn3+ cells give rise to all endocrine cells but not ductal cells (Gu et al., 2002). *Ngn3* expression occurs in two distinct phases, the timing of which correlates with the onset of the primary and secondary phases of endocrine development (Villasenor et al., 2008). Additionally, Ngn3 controls expression of endocrine-specific genes *Isl1*, *Pax4*, *Pax6*, and *NeuroD* (Gradwohl et al., 2000). *Ngn3*− mice do not have pancreatic endocrine cells (Gradwohl et al., 2000), and expression of *Ngn3* under the *Pdx1* promoter leads to ectopic formation of endocrine cells (Schwitzgebel et al., 2000). Up to this point in pancreas development, the Notch pathway has favored an endocrine fate. However, during endocrine specification, Notch signaling plays an inhibitory role and promotes a ductal cell fate (Qu et al., 2013) in part by regulating expression levels of the ductal cell marker.
Sox9 (Shih et al., 2012). Notch represses Ngn3 via Hes1 to prevent an endocrine fate (Lee et al., 2001). Additionally, disrupting the Notch pathway results in increased Ngn3 expression and enhanced endocrine differentiation (Apelqvist et al., 1999). The relationship between Notch signaling and Ngn3 expression is critical for proper endocrine specification.

1.3.3.5 Endocrine cell differentiation

After endocrine specification, transcription factors specific to different endocrine lineages are expressed, driving differentiation to different islet cell types (Figure 1.5C). As development progresses, Pdx1 and Nkx6.1 become increasingly restricted to β-cells. Pdx1 maintains β-cell function and identity. Loss of Pdx1 leads to a decrease in insulin-positive cells and attenuated β-cell specific gene expression (Ahlgren et al., 1998). Additionally, Pdx1 regulates β-cell proliferation, which is required to generate the appropriate number of β-cells in islets (Gannon et al., 2008). Nkx6.1 also plays a critical role in β-cell development. Nkx6.1−/− mice have fewer β-cell precursors and no β-cell differentiation (Sander et al., 2000). Nkx6.1 specifies β-cells prior to Ngn3 expression (Nelson et al., 2007). Finally, Nkx6.1 drives β-cell proliferation and regulates the glucose responsiveness function of β-cells (Schisler et al., 2008). These studies show that both Pdx1 and Nkx6.1 have important roles in multiple stages of pancreas development, particularly during β-cell maturation.

Additionally, Pax4 is needed for β- and δ- cell differentiation, while Arx is required for α-cell differentiation. The balance between levels of these transcription factors is key in regulating endocrine fates. Pax4−/− mice have more α-cells and fewer β- and δ-cells (Sosa-Pineda et al., 1997). Conversely, Arx−/− mice have more β- and δ-cells and fewer
α-cells. (Collombat et al., 2003) Finally, Pax4 and Arx double knockout mice have significant increase in the number of δ-cells (Collombat et al., 2005). Pax4 and Arx negatively regulate each other (Collombat et al., 2003) to maintain the proper levels for normal differentiation of these endocrine cell types.

Another transcription factor that regulates endocrine lineage specification is Nkx2.2. Nkx2.2 is first expressed in the pancreatic epithelium in early development, and then later becomes restricted to the endocrine lineage (Sussel et al., 1998). Nkx2.2−/− embryos do not have β-cells, and have fewer α- and PP-cells. Although islets exist in these embryos, none of the endocrine hormones are produced and Nkx2.2−/− mice are hyperglycemic (Sussel et al., 1998). Additionally, the loss of Nkx2.2 results in the expansion of ε-cell population and a reduced number of β-cells (Prado et al., 2004). Nkx2.2 is also upstream of Nk6.1 (Sander et al., 2000), and directly represses Arx (Papizan et al., 2011), thus promotes a β-cell over α-cell fate. These studies demonstrate that Nkx2.2 is required for both proper differentiation and function of β-cells.

Once lineage specification occurs, transcription factors MafA and MafB are responsible for α- and β-cell maturation. MafB is expressed in immature α- and β-cells (Artner et al., 2006) and needed for β-cell maturation (Artner et al., 2007). As the cells mature, MafB becomes restricted to α-cells and is needed for proper α-cell function (Artner et al., 2006). Mature β-cells make the switch from MafB to MafA expression (Nishimura et al., 2006). Islet development is not affected in MafA−/− mice, but they do develop diabetes and have impaired glucose tolerance after birth, suggesting that MafA is more important for β-cell function rather than development (Zhang et al., 2005). The balance between these two factors is key in distinguishing between and α- and β-cell fate.
1.3.3.6 Islet formation and maturation

Once endocrine cells have differentiated, they delaminate from the pancreatic trunk epithelium and migrate into the surrounding mesenchyme. Though the exact mechanism of how the delamination occurs remains unknown, there is strong evidence that it occurs through an epithelial-to-mesenchymal transition (EMT) (Figure 1.5D). EMT is characterized by loss of cell polarity and adhesion, and changes in cytoskeletal structure. Snail2, a potent inducer of EMT, is detected in Ngn3-positive cells, suggesting EMT is initiated shortly after endocrine specification. Additionally, E-cadherin, and epithelial marker, expression is lost while N-cadherin, a mesenchymal marker, is upregulated in cells just outside the pancreatic epithelium. The change from E- to N-cadherin is a hallmark of EMT, and correlates with Snail2 expression, suggesting that it may be involved in this transition (Rukstalis and Habener, 2007). Cells emerging from the trunk epithelium take on the characteristic shape of cells undergoing EMT, and express markers, such as Vimentin and Snail family transcription factors, associated with mesenchymal cells (Cole et al., 2009). Ngn3 may be responsible for triggering EMT and delamination of endocrine cells from the trunk epithelium. Upon Ngn3 expression, cells being losing polarity, down regulate E-cadherin, break down the basal lamina, and delaminate into the mesenchyme. Additionally these cells express mesenchymal markers Vimentin and N-cadherin, and Ngn3 controls Snail2 expression in the pancreas (Gouzi et al., 2011). Taken together, these studies provide compelling evidence that endocrine cells undergo EMT while delaminating from the trunk epithelium.

After endocrine cells delaminate from the trunk epithelium, they migrate into the surrounding mesenchyme and cluster together to form pancreatic islets. The mechanism of how these cells find each other and aggregate remains unclear. Integrins mediate
pancreatic endocrine cell migration. Inhibiting specific integrins prevents endocrine cell emergence from the trunk (Cirulli et al., 2000), and migration after delamination (Yebra et al., 2003). The GTPase Rac1 mediates endocrine islet migration by regulating E-cadherin expression and cell spreading. Blocking Rac1 results in increased E-cadherin levels, and impaired actin remodeling in endocrine cells (Greiner et al., 2009). Cell adhesion molecules have also been implicated in aggregation during islet assembly. Disrupting E-cadherin expression in β-cells prevents them from properly integrating into islet structures (Dahl et al., 1996). Additionally, the cell adhesion molecule N-CAM is required for proper endocrine cell positioning within the islet, as loss of N-CAM results in mislocalization of α-cells in developing islets (Esni et al., 1999). Interactions between the islets and surrounding endothelial cells and neurons also influences islet structure and function (Bastidas-Ponce et al., 2017). β-cells continue to mature and become fully functional postnatally.

**1.3.3.7 Differences in human pancreas development**

Most of our knowledge of pancreas development is based on studies done in model organisms. Studies done on pancreas development in human embryos (Jeon et al., 2009; Lyttle et al., 2008; Piper et al., 2004) reveal similarities and differences between the two species. While the timing of some transcription factors is slightly delayed in humans compared to mouse development, the expression occurs in the same general chronological and spatial patterns (Jennings et al., 2015). For example, in the mouse, *Pdx1* expression is first detected when the dorsal pancreatic endoderm is adjacent to the notochord (Kim et al., 1997). However, in humans, *PDX1* expression is not detected until the gut tube has closed and the dorsal pancreatic is no longer near the notochord and dorsal aorta (Jennings et al., 2013). One major difference between human and mouse
development is that endocrine cells are produced in a single stage in humans rather than in two distinct waves as seen in the mouse (Jennings et al., 2013). Another significant difference is that NKKX2.2 is not detected until after endocrine specification (Jennings et al., 2013), while in the mouse it is first detected during bud formation (Sussel et al., 1998). Finally, the mature islet architecture differs between mice and humans. In the mouse, β-cells make up the majority (75%) of the islet and are concentrated in the center, with the other endocrine cells surrounding them. In humans, β-cells represent about 50% of the islet, and the different endocrine cell types are distributed homogenously throughout the islet (Brissova et al., 2005). These observations note that while the most steps of pancreas development are similar between species, there are some human-specific differences. Studying these differences can provide insight into human β-cell biology, which can be applied improving generation of pluripotent stem cell derived β-cells for use in a therapeutic context.

1.4 Liver versus pancreas development

The liver and pancreas are closely related and develop from the same region of the endoderm. This thesis explores the potential role of the transcription factor TBX3 in endoderm patterning and distinguishing between a liver and pancreas fate. Next, we will discuss what is known about liver versus pancreas development.

The liver and ventral pancreas both develop from the lateral endoderm of the ventral foregut. Within the lateral endoderm, the ventral pancreatic endoderm domain lies caudal to the hepatic domain (Tremblay and Zaret, 2005). The transcription factor Hex plays a role in establishing the liver and ventral pancreas domains in the foregut endoderm by controlling the proliferation and positioning of the ventral foregut endoderm (Figure 1.6A). Ventral pancreas specification fails in Hex\(^{−/−}\) embryos because the
Figure 1.6: Liver versus ventral pancreas specification

The liver and the ventral pancreas develop from the same part of the ventral foregut endoderm.

(A) The transcription factor Hex controls proliferation and positioning of the ventral foregut endoderm past the cardiac mesoderm.

(B) The cardiac mesoderm and septum transversum mesenchyme (STM) secrete fibroblast growth factor (FGF) and bone morphogenetic protein (BMP) signals, respectively, to the neighboring ventral foregut endoderm to specify the hepatic endoderm (red), that expresses liver-specific genes Albumin, Afp, and Ttr. BMP signals prevent expression of pancreatic genes Pdx1 and Ptf1a in the hepatic endoderm. These genes are expressed in the cells do not receive FGF signals from the cardiac mesoderm to specify the ventral pancreatic epithelium (purple). Adapted from Bort et al., 2004. Created with BioRender.com.
endoderm is not positioned beyond the cardiac mesoderm, thus all cells receive FGF signals and are specified to liver (Bort et al., 2004). In wildtype embryos, Pdx1 is expressed in the ventral endoderm that extends past the cardiac mesoderm, and this region becomes the ventral pancreatic bud. Since the endoderm is not positioned correctly in Hex<sup>−/−</sup> embryos, ventral pancreatic bud does not form. Explant studies using Hex<sup>−/−</sup> ventral foregut endoderm show that it is still competent to induce pancreatic gene expression (Bort et al., 2004). This provides evidence that ventral foregut endoderm cells can adopt either a liver or pancreas fate depending on their position in the developing embryo.

The liver and ventral pancreas are both specified around the same time at E8.5. The region of the ventral foregut endoderm adjacent to the cardiac mesoderm expresses the hepatic gene Albumin while the region of the ventral endoderm extending past the cardiac mesoderm expresses Pdx1, demonstrating that there are distinct, non-overlapping liver and pancreas domains (Deutsch et al., 2001) (Figure 1.6B). Ventral foregut explants from 2-6 somite embryos (prior to the onset of FGF signaling and hepatic specification) express pancreatic genes including Pdx1, Ptf1a, and Isl1. When these explants are cultured with FGF or cardiac mesoderm, pancreatic genes are no longer present while Albumin is strongly expressed, suggesting that pancreas is the default fate of the ventral foregut endoderm (Deutsch et al., 2001). The cardiac mesoderm induces a hepatic fate in the ventral foregut directly next to it, while the section of the ventral foregut that extends past the cardiac mesoderm escapes repressive FGF signals and develops into the ventral pancreas.

BMP signaling also plays a role in patterning the liver versus pancreatic domains in the ventral foregut endoderm. BMP signals are required for liver induction. When BMP
signaling is blocked, *Pdx1* expression is activated at the expense of *Albumin* expression in the ventral foregut (Rossi et al., 2001). BMP signals make cells competent to receive FGF signals that induce a hepatic fate and exclude a pancreatic fate. Overexpressing Bmp2b in zebrafish results in cells that normally contribute to the ventral pancreas to adopt a liver fate instead (Chung et al., 2008), providing further evidence that BMP signaling promotes hepatic specification by suppressing a pancreatic fate.

Patterning and specification of the liver and ventral pancreas is a tightly regulated process. Hex controls proliferation and proper positioning of the ventral foregut endoderm. This ensures that the presumptive hepatic endoderm receives FGF and BMP signals to become the liver, while the ventral pancreas endoderm escapes these signals and develops into the ventral pancreatic bud. Further studies into transcription factors and signals that segregate the liver and ventral pancreas domains will not only enhance our knowledge of endoderm patterning, but can also help refine differentiation protocols to generate functionally mature hepatocytes and β-cells from pluripotent stem cells.

1.5 Characteristics of TBX3

TBX3 is a transcription factor that has many roles in development and disease. In particular, Tbx3 has been well-studied during mouse liver development. However, there have been instances in which development in the mouse differs from human development. One goal of this thesis is to determine whether TBX3 functions similarly in mouse and human liver development. Additionally, this thesis examines the role of TBX3 in distinguishing hepatic versus pancreatic domains during foregut endoderm patterning, and the role of TBX3 in pancreatic differentiation from pluripotent stem cells. Here we discuss the various functions of TBX3.
1.5.1 Structure of TBX3

TBX3 is a member of the T-box family of transcription factors, which is involved in many developmental processes. T-box transcription factors are characterized by the highly conserved T-box DNA binding domain, which is 180-200 amino acids in length and contains the consensus half site 5’-AGGTGTGAAA-3’, known as the T-box binding element (Papaioannou, 2014). The T-box family is divided into five subfamilies: T, Tbx1, Tbx2, Tbx6, and Tbr1. Tbx3 is a member of the Tbx2 subfamily, along with Tbx2, Tbx4, and Tbx5 (Naiche et al., 2005). In particular, Tbx3 and Tbx2 have highly similar sequences and have similar functions in development (Gibson-Brown et al., 1998; Lüdtke et al., 2016; Singh et al., 2012; Zirzow et al., 2009). There is some degree of redundancy between Tbx2 and Tbx3, but this is dependent on biological context.

TBX3 is located on chromosome 12, has seven exons, and is 723 amino acids in length. The T-box domain is located in exons 1-4 (Khan et al., 2020). An isoform with an additional 20 amino acids in the T-box domain, termed TBX3+2a, has been identified (Figure 1.7A). Whether or not these isoforms have similar or distinct functions remains contested (Fan et al., 2004; Hoogaars et al., 2008). TBX3 has two repression domains: one is within the T-box DNA binding domain (amino acids 123-300) and the second is at the C-terminal end (amino acids 567-623), the latter being the dominant repression domain (Carlson et al., 2001; Khan et al., 2020). TBX3 also has an activation domain (amino acids 423-500), however it appears that TBX3 acts primarily as a repressor (Carlson et al., 2001) (Figure 1.7B).

The crystal structure of TBX3 protein has been identified. The structure of the T-box domain is similar to that of Xbra, the Xenopus homolog of Brachyury, which had been previously crystallized. However, TBX3 binds differently to DNA than Xbra does. T-box
Figure 1.7: Structure and function of TBX3

(A) TBX3 is located on chromosome 12 and has 7 exons. The T-box DNA binding domain is located in exons 1-4. An isoform with an extra 60 bp encoding exon 2a exits, but its function remains controversial.

(B) The TBX3 protein is 723 amino acids long. The T-box DNA binding domain is represented in the magenta boxes. TBX3 has one activation domain (orange box), and two repression domains (green boxes), with the C-terminal repression domain being dominant.

(C) In mouse liver development, Tbx3 is expressed in the hepatoblasts and maintains expression of Hnf4α and C/ebpα to promote a hepatocyte fate, and represses Hnf6 expression to suppress a cholangiocyte fate. Adapted from Khan et al., 2020 and Zong and Stanger, 2012. Created with BioRender.com.
factors are thought to bind to the consensus half site as monomers, then dimerize to become fully functional. The crystal structure of the TBX3/DNA complex shows that TBX3 binds to consensus sequences but does not dimerize (Coll et al., 2002). This highlights differences in how members of the T-box family bind to DNA.

1.5.2 Functions of TBX3

Tbx3 is dynamically expressed during early development in mouse embryos, and drives mesendoderm differentiation as it correlates with markers T, EOMES, and SOX17 in Xenopus embryos (Weidgang et al., 2013). Additionally, it regulates EOMES expression and drives endoderm differentiation (Kartikasari et al., 2013). Tbx3 has also been detected in several tissues during development, including the heart, limbs, lungs, mammary gland (Douglas and Papaioannou, 2013; Gibson-Brown et al., 1998; Lüdtke et al., 2016; Singh et al., 2012). Tbx3 is also expressed in the developing liver, which is discussed in further detail in Section 1.5.3. Although Tbx3 expression has been observed in both the developing and adult pancreas in mice, its function has yet to be determined (Begum and Papaioannou, 2011; Zhou et al., 2007). In mouse embryos, Tbx3 expression is mainly restricted to the pancreatic mesenchyme surrounding the epithelial branches. The Tbx2 expression pattern generally overlaps with that of Tbx3 in embryos. In the adult mouse, Tbx2 is expressed in pancreatic islet while Tbx3 is restricted to exocrine tissue, suggesting a possible functional difference between the two transcription factors (Begum and Papaioannou, 2011). In human embryos, TBX3 is expressed in SOX9+/PTF1A+ tip progenitors that give rise to exocrine component of the pancreas (Villani et al., 2019). This differs from the mouse, as Tbx3 is not detected in any pancreatic progenitor cells, highlighting a potential species-specific difference between human and mouse. Additionally, neither TBX2 or TBX3 is expressed in adult
human β-cells. However, TBX2 and TBX3 expression are induced in both type 1 and type 2 diabetes, raising the possibility that they play a role in disease progression (Kaestner et al., 2019; Russell et al., 2019, https://hpap.pmacs.upenn.edu/). This is not seen in the mouse, further suggesting possible species-specific differences in TBX3 expression and function during pancreas development and disease.

In humans, TBX3 is implicated in Ulnar-Mammary Syndrome (UMS), a rare genetic disorder that causes defects in limb development, mammary and apocrine gland hypoplasia, and genital abnormalities. It is an autosomal dominant disorder caused by a wide variety of loss-of-function mutations in one allele of TBX3. These include missense, frameshift, and early truncation mutations. The specific mutation determines the severity of the phenotype in the patient (Bamshad et al., 1997, 1999). A mouse model was developed to further study UMS. However, unlike in humans, a UMS-like phenotype is only seen when there are mutations in both alleles of Tbx3, suggesting a possible difference in the requirement of TBX3 between species. Interestingly, yolk sac and liver defects are seen in Tbx3−/− mice (Davenport et al., 2003). This has not been observed in human with UMS, since there are no patients with homozygous mutations in TBX3.

In addition to its roles in development, TBX3 acts as a cell cycle regulator in the p53 pathway. p14ARF represses MDM2, resulting in increased levels of p53, which leads to cell cycle arrest. Disruption of this pathway leads to uncontrolled cell proliferation. TBX3 directly represses CDKN2A, which encodes p14ARF (Carlson et al., 2002; Lingbeek et al., 2002), Additionally, TBX3 directly suppresses p21, a downstream target of p53 that mediates cell cycle arrest (Willmer et al., 2016). TBX3 overexpression results in uncontrolled proliferation in several types of cancers including liver and breast cancers (Renard et al., 2007; Yarosh et al., 2008). Additionally, TBX3 overexpression promotes
tumor invasiveness and migration in breast cancer and melanoma by repressing E-cadherin and promoting epithelial-to-mesenchymal transition (Krstic et al., 2016, 2019; Rodriguez et al., 2008). TBX3 contributes to tumor progression through many different mechanisms, highlighting its potential as a therapeutic target for treating cancer patients. Tbx3 is also important for mouse embryonic stem cells (mESC) maintenance. Tbx3 is expressed in mESCs, and is required to maintain self-renewal and pluripotency (Ivanova et al., 2006; Niwa et al., 2009; Russell et al., 2015). Additionally, Tbx3 enhances the quality of induced pluripotent stem cells from mouse fibroblasts (Han et al., 2010). Conversely, TBX3 is not expressed in human embryonic stem cells (hESCs), is not required for self-renewal and maintenance, and hESCs lacking TBX3 are still able to differentiate (Esmailpour and Huang, 2012). This difference in the role of TBX3 in mESCs and hESCs can be explained by pluripotent state. mESCs are thought to be in a naïve state, while human ESCs are in a primed state. This difference in pluripotency state calls for different culture conditions and gene expression patterns to maintain pluripotency in mESCs versus hESCs (Nichols and Smith, 2009). This includes Tbx3 which is expressed in naïve mESCs but not primed hESCs (Davidson et al., 2015). This difference allowed us to utilize a human pluripotent stem cell system to generate viable TBX3 knockout cells and study the function of TBX3 during human liver development.

1.5.3 TBX3 in liver development

Tbx3 plays several roles during mouse liver development. Tbx3 is expressed in hepatoblasts, controls liver bud outgrowth, and regulates fate decisions later in development. Tbx3−/− embryos form small liver buds which do not develop into a functionally a mature liver. Additionally, the basement membrane does not degrade and hepatoblasts fail to migrate into the septum transversum mesenchyme. Hepatoblasts
from $Tbx3^{−/−}$ embryos have elevated levels of E-cadherin and laminin, and decreased expression of $Prox1$ in the liver bud (Lüdtke et al., 2009). $Tbx3^{−/−}$ embryos display a similar phenotype to $Prox1^{−/−}$ embryos. As mentioned in Section 1.3.2.2, $Prox1$ is required for hepatoblast delamination and migration into the surrounding mesenchyme as the liver bud forms. These findings suggest that while $Tbx3$ is not required for hepatic specification, it is necessary for proper growth and progression of liver development, in part through maintenance of $Prox1$ expression (Lüdtke et al., 2009). $Tbx3$ is also critical for proper fate decisions during hepatoblast differentiation as it promotes a hepatocyte fate over a cholangiocyte fate. $Tbx3$ maintains expression of hepatic transcription factors $Hnf4α$ and $C/ebpα$, while repressing $Hnf6$ which is important for cholangiocyte development (Figure 1.7C). In $Tbx3^{−/−}$ embryos there is a switch in the expression pattern of these transcription factors, with increased levels $Hnf6$ and its downstream target $Hnf1β$, and decreased $Hnf4α$ and $C/ebpα$ levels. Other hepatic markers including $Albumin$, $AFP$, $αAT$, and $c-Met$ also have reduced expression in $Tbx3^{−/−}$ embryos, resulting in more cholangiocytes and fewer hepatocytes compared to their wildtype counterparts. $Tbx3$ negatively regulates the tumor suppressor p19$^{ARF}$ (Brummelkamp et al., 2002). $Tbx3^{−/−}$ embryos exhibit a significant reduction in proliferating hepatoblasts due to increased levels of p19$^{ARF}$. Overexpressing p19$^{ARF}$ wildtype hepatoblasts results in decreased proliferation and expression of cholangiocyte markers $Ck19$ and $Ck7$, the same phenotype seen in hepatoblasts from $Tbx3^{−/−}$ embryos. Taken together, these results suggests $Tbx3$ controls hepatobiliary fate decisions of hepatoblasts by regulating proliferation via p19$^{ARF}$ (Suzuki et al., 2008). The role of $Tbx3$ in regulating hepatoblast delamination and fate decisions has been well-established in mice, but it is unclear whether TBX3 acts similarly in human liver development. To address this question, we
used pluripotent stem cells to study the function of TBX3 during this developmental process.

1.6 Pluripotent stem cells

Pluripotent stem cells (PSCs) are characterized by the ability to self-renew while maintaining an undifferentiated state, and have the capability to differentiate to all cell types of the body. PSCs include embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs), and serve as a valuable model system for studying human development, physiology, and disease. PSCs can also be used in drug discovery, and in cell-replacement therapies. This thesis utilizes PSCs to study the role of TBX3 in liver and pancreas development. We generated TBX3 knockout PSC lines and differentiated them to hepatocytes and pancreatic progenitors to study how the loss of TBX3 impacts these processes.

1.6.1 Discovery of embryonic stem cells

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of blastocyst stage embryos, which occurs prior to implantation, during development. The ICM consists of pluripotent stem cells that develop into the three germ layers (the ectoderm, mesoderm, and endoderm) and their derivatives. In 1981, two groups successfully isolated the ICM from mouse blastocysts and cultured them in vitro (Evans and Kaufman, 1981; Martin, 1981). These mouse ESCs (mESCs) can proliferate and be maintained on fibroblast feeder cells, and form embryoid bodies, teratomas, and germline competent chimeras (Ginis et al., 2004), confirming their pluripotent characteristics. The discovery of mESC provided a new platform to study development, as they could be easily differentiated to different cell types. Another advantage of mESCs is that they could be genetically manipulated to introduce specific mutations,
advancing the way we studied mammalian genetics. (Evans, 2011). The work done to identify and establish mESCs was a groundbreaking achievement in the field of developmental biology, and paved the way for the discovery and characterization of human ESCs.

Human embryonic stem cells (hESCs) were first described in 1998 (Thomson et al., 1998). Human embryos produced for in vitro fertilization were obtained at the cleavage stage and cultured to the blastocyst stage. The ICM population was removed from the blastocysts and cultured on fibroblast feeder cells for multiple passages without differentiating, while maintaining a normal karyotype. Like mESCs, hESCs form teratomas and can differentiate to all three germ layers. Prior to hESCs isolation and culture in vitro, studying human development was challenging. hESCs serve as a model system for uncovering attributes of development specific to humans that could not be understood using traditional model organisms. In addition to studying human development, hESC-derived cell types can be used in disease modeling studies, and have the potential for use in cell-replacement therapies, making them a valuable tool for advancing disease treatments.

1.6.2 Induced pluripotent stem cells

The stem cell field experienced a major advancement with the advent of induced pluripotent stem cells (iPSCs). iPSCs helped bypass some disadvantages of using hESCs. First, they avoid any possible ethical controversies of using ESCs since they are derived from patients instead of embryos. Another advantage is that they provide a novel opportunity to use patient-derived cells to study specific mutations in development and disease progression. Additionally, since the patient’s own cells are used to generate iPSCs, they can avoid potential rejection if used in stem cell-based treatments.
There were several key findings that lead to the development of iPSCs. First, was somatic cell nuclear transfer, in which the nucleus from a differentiated somatic cell is transferred to an enucleated egg, which develops into an embryo or fully development organism. This was first demonstrated in *Xenopus* (Gurdon, 1962) and later in sheep (Wilmut et al., 1997), suggesting that terminally differentiated cells contain all the necessary information needed for development. Next, came the idea of a master regulator, in which certain transcription factors have the ability to convert cells to different fates. Expressing *MyoD*, which controls myogenesis, resulted in expression of muscle-specific genes in various non-muscle cell types (Weintraub et al., 1989).

These findings raised the question whether there were transcription factors that allowed for differentiated cells to be reprogrammed to a pluripotent state similar to ESCs. Mouse embryonic and adult fibroblasts were converted to iPSCs by expression of *Oct4*, *Sox2*, *c-myc*, *Klf4* (Takahashi and Yamanaka, 2006). Soon after, human adult fibroblasts were reprogrammed to iPSCs with these same transcription factors (Takahashi et al., 2007). Interestingly, another group was able to reprogram human fibroblasts to iPSCs using a different set of factors (*OCT4*, *SOX2*, *NANOG*, and *LIN28*), demonstrating there are different combinations of factors than can reprogram cells (Yu et al., 2007). Both groups showed that iPSCs express pluripotent markers and can form teratomas, indicating they have properties similar to hESCs. The innovation of iPSC was a major breakthrough in the field.

1.6.3. Pluripotent stem cells as a model system

Both ESCs and iPSCs serve as powerful tools for studying development, physiology and disease. The knowledge gleaned about development from other model organisms has been used to develop protocols to differentiate PSCs to many cell types, including
neuronal cells, cardiomyocytes, hematopoietic cells, pancreatic β-cells, and hepatocytes. Using PSCs to study development can uncover species-specific differences and identify aspects of development unique to humans (Rowe and Daley, 2019). However, it has been difficult to generate functionally mature cell types from PSCs. Ongoing work by many groups has drastically improved these protocols over time, bringing us closer to PSC-derived cells that resemble cells found in vivo.

PSCs are also an excellent tool for human disease modelling. They can be easily genetically manipulated using various technologies such as zinc finger nucleases (Zou et al., 2009), transcription activator-like effector nucleases (TALENs) (Hockemeyer et al., 2011), or clustered regularly interspaced short palindromic repeats (CRISPR) (Maguire et al., 2019). Introducing specific mutations into genes of interest allows for study of both development and disease mechanism and progression. iPSCs from patients with a disease confer a particular advantage, as they already contain the mutation(s) of interest. The mutation can be corrected using gene editing, providing an isogenic wildtype control for disease modeling (Rowe and Daley, 2019). PSCs also serve as a valuable tool for drug discovery. They can be differentiated to different cell types, on which large-scale drug screens can be performed. This is particularly advantageous for identifying compounds that may act as therapeutic agents when used in disease lines. Furthermore, studying the effects of these compounds on PSC-derived cell types can help elucidate the mechanism of various diseases (Shi et al., 2017).

The ultimate goal is generating different cell types from PSCs that can be used in cell-replacement therapies in patients. While this technology holds great promise, there are still some safety challenges in this area. First, is the concern of tumorigenicity of PSCs. PSCs can form teratomas and can accumulate mutations that confer a growth
advantage while in culture over many passages. It is imperative to make sure that there are no undifferentiated cells present in the cells being put into patients (Yamanaka, 2020). Additionally, cells derived from ESCs may induce and immune response and result in rejection. iPSCs have an advantage in this regard since they are derived from the patient’s own cells and are less likely to be rejected by the body (Shi et al., 2017). Despite these issues, PSCs remain an increasingly promising avenue for use in clinical applications and treatments.

1.6.4 Differentiation of PSCs to hepatocytes and pancreatic β-cells

PSCs can be differentiated to all cell types of the body, including endodermal lineages such hepatocytes and pancreatic β-cells (Figure 1.8). The differentiation protocols that have been developed are based on developmental studies done in mice and other model organisms, helping us mirror the appropriate timing of signals during development.

The general PSC hepatocyte differentiation protocol involves adding Activin to generate definitive endoderm (DE), FGF2 and BMP4 to induce hepatoblast specification, and HGF, OSM, and Dexamethasone for hepatocyte maturation (Ogawa et al., 2013; Si-Tayeb et al., 2010; Takayama et al., 2012). At each stage of the protocol, the differentiating cells expressed appropriate markers (Sox17 at DE, HNF4α and AFP in hepatoblasts, and ALB in hepatocytes). Interestingly, the hepatocytes generated from these 2-dimensional (2D) based protocols resemble fetal hepatocytes rather than mature hepatocytes. Differentiating PSCs in a 3-dimensional (3D) culture system generates cells that are more similar to mature hepatocytes than those from in a 2D system (Luo et al., 2018). The aggregation of hepatocytes results in increased expression of
Differentiating PSCs to endodermal lineages

Figure 1.8: Generalized schematic of differentiation protocols to direct PSCs to endodermal lineages, such as hepatocytes and pancreatic β-cells. Activin mimics Nodal signaling to drive PSCs to mesendoderm. High activin generates definitive endoderm (DE). Adding FGF and BMP directs DE to hepatoblasts, then hepatocyte growth factor (HGF), Dexamethasone (Dex), and Oncostatin M (OSM) drive differentiation to hepatocytes. Adding cyclopamine, FGF, and retinoic acid (RA) to DE generates pancreatic endoderm. Adding more specialized molecules drives the cells to hormone-producing cells found in the pancreatic islets. Adapted from Zorn and Wells, 2009. Created with BioRender.com
cytochrome p450 enzymes (Ogawa et al., 2013) and improved albumin secretion (Gieseck et al., 2014) in the 3D differentiations. Although it is still difficult to generate fully functional hepatocytes from PSCs, differentiation protocols are continually improving and getting closer to the end goal of generating mature hepatocytes for use in therapies for liver disease and damage.

There has been great interest in generating pancreatic β-cells from PSCs for potential use in treatment or drug discovery for diabetes. Protocols for β-cell differentiation were established based on knowledge of pancreatic development in other model organisms. Although the β-cells generated from PSCs was an immense advancement, initial protocols yielded β-cells that were poly-hormonal (expressed a combination of insulin and either glucagon or somatostatin) and did not secrete insulin in response to glucose stimulation (D’Amour et al., 2006; Nostro et al., 2011). hESC-derived pancreatic progenitors were able to produce functional β-cells when transplanted into mice and allowed to mature for 3-4 months (Kroon et al., 2008; Rezania et al., 2012), but up until this point functionally mature β-cells could not be generated in vitro. It was later found that hESC-derived pancreatic progenitors enriched for the transcription factor NKX6.1 generated mature β-cells upon transplantation into mice (Rezania et al., 2013). Differentiation protocols were modified to allow for NKX6.1 enrichment in pancreatic progenitors (Nostro et al., 2015). In 2014, two groups published in vitro protocols in which PSCs were successfully differentiated to mono-hormonal insulin expressing β-like cells that secreted insulin in response to glucose (Pagliuca et al., 2014; Rezania et al., 2014). Additionally, upon transplantation of PSC derived β-like cells into mice, insulin was detected just two weeks later, and they were able to reverse diabetes in the mice. While these protocols greatly advanced the field, there is still room for improvement,
particularly in increasing yield of β-cells. The ability to generate β-cells from PSCs is not only monumental for studying β-cell biology, but is also immensely promising for identifying and developing diabetic drugs, and can potentially be used in therapeutic context for diabetes treatment.

1.7 Overview of research goals
The role of Tbx3 in liver development has been well established in mouse models. Tbx3 regulates hepatoblast delamination and migration during liver bud formation. Additionally, Tbx3 influences hepatoblast fate decisions by promoting a hepatocyte fate over a cholangiocyte fate. To date, it is unclear whether TBX3 functions similarly during human liver development. We utilized human pluripotent stem cells and established hepatocyte differentiation protocols to address this question. We generated TBX3 knockout PSC lines using CRISPR/Cas9 genome editing technology. We hypothesized that the TBX3 mutant PSCs would differentiate to hepatocytes less efficiently than wildtype PSCs, indicating that TBX3 acts similarly in mouse and human liver development. During the hepatocyte differentiations, we made a surprising observation that TBX3 mutant lines express pancreas-specific genes. This finding raised new questions of whether TBX3 distinguishes between liver and pancreas domains during endoderm patterning, and whether the loss of TBX3 impacts pancreatic differentiation. Chapter 2 of this thesis outlines generation of TBX3 mutant lines and relevant methodology to study TBX3 in liver and pancreas differentiation. Chapter 3 discusses how TBX3 impacts hepatocyte differentiation and pancreatic progenitor generation from PSCs, and highlights a potential mechanism for how TBX3 acts in these developmental findings. Finally, Chapter 4 discusses conclusions, potential impacts and implications from these studies.
CHAPTER 2: MATERIALS AND METHODS

2.1 Pluripotent stem cell culture and maintenance

PSCs were cultured in an environment of 5% CO₂, 5% O₂, and 90% N₂ on 0.1% gelatin (Sigma) and irradiated mouse embryonic fibroblasts. They were grown in human embryonic stem cell (hESC) medium consisting of DMEM/F12 (Corning) with 15% Knockout Serum Replacement (Gibco), 2mM L-glutamine (Corning), 1x Non-Essential Amino Acids (Gibco), 1x Penicillin/Streptomycin (Corning), 0.1 mM β-mercaptoethanol (Gibco), and 10 ng/mL bFGF (R&D Systems). Medium was changed daily. PSCs were grown to 80% confluency and split at a 1:12 ratio using TrypLE (Gibco) dissociation regent. Cells were replated in hESC medium with 5µM ROCK inhibitor Y-27632 dihydrochloride (Tocris).

2.2 Generation of TBX3 mutant lines using CRISPR/Cas9 genome editing

TBX3 mutant PSC lines were generated, as described previously (Maguire et al., 2019). Briefly, two guide RNAs were designed to create a deletion in the endogenous TBX3 locus. TBX3 gRNA1, 5'-GGAGTGAGCACCTCC-3', targeted the transcription site and TBX3 gRNA2, 5'-CAAAGTAAACCATGCAGCTCC-3', targeted the boundary of exon 5 and intron 4 of the TBX3 locus. The gRNAs were cloned into the gRNA_Cloning Vector (Addgene no. 41824). 0.5 µg of each plasmid with the gRNA, and 1 µg of pCas9_GFP vector (Addgene no. 44719) were transfected into one well of a 6-well plate containing 3.5 x 10⁵ cells. Cells were sorted using a BD FACS Aria II (Becton Dickinson) cell sorted and plated at a single cell density on 10 cm tissue culture dishes coated with 1:3 Matrigel (Corning) and supplemented with mouse embryonic fibroblasts. Single colonies were picked, with half being cultured for expansion, and the other half screened for the
deletion by PCR and TOPO cloning (Thermo Fisher Scientific). Clones containing the deletion were then sequenced for confirmation. All gRNA, screening, and sequencing primer sequences can be found in Table 2.1.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBX3 ATG gRNA1</td>
<td>5'-GGAGTGAGTGAGCCTCTCC-3'</td>
</tr>
<tr>
<td>TBX3 intron-exon gRNA2</td>
<td>5'-CAAAGGTAAACCATGCAC-3'</td>
</tr>
<tr>
<td>TBX3 Deletion Screening</td>
<td>F: 5'-CCTTTAAGCCGGGTCTAGCAC-3'</td>
</tr>
<tr>
<td></td>
<td>R: 5'-ATCTTTGCACTCCTACCCAC-3'</td>
</tr>
<tr>
<td>TBX3 Wildtype allele Screening</td>
<td>F: 5'-AACACACGCAGCTCGGACTTG-3'</td>
</tr>
<tr>
<td></td>
<td>R: 5'-ATCCATGATCGGCGTTGCC-3'</td>
</tr>
<tr>
<td>TBX3 Deletion Sequencing</td>
<td>5'-GTGCATTAGTTGTGATTCTGCC-3'</td>
</tr>
</tbody>
</table>

### 2.3 Hepatocyte differentiation

The hepatocyte differentiation protocol previously described (Ogawa et al., 2013) was modified as follows. PSCs were split onto 1:3 Matrigel (Corning) coated 6-well plates and cultured until they reached 80-90% confluency before starting Day 0 of the differentiation. Cells were differentiated as a monolayer in 5% O₂, 5% CO₂, and 90% N₂. Cells were cultured for 24 hours (Day 0) in RPMI medium (Corning) with 2mM L-glutamine, 50 µg/ml stabilized ascorbic acid (SAA) (Wako), 4.5×10⁻⁴ M MTG (Sigma), 100ng/mL Activin A (R&D Systems), and 2µM CHIR (Tocris). Cells were then cultured for 24 hours (Day 1) in RPMI medium with 2mM L-glutamine, 50 µg/ml SAA, 4.5×10⁻⁴ M MTG, 100ng/mL Activin A, and 5 ng/mL bFGF (R&D Systems). From Days 2-6, cells were cultured serum free differentiation (SFD) medium, with 2mM L-glutamine, 50 µg/ml
SAA, 4.5×10^{-4} M MTG, 100ng/mL Activin A, and 5 ng/mL bFGF. SFD medium consists of Iscove’s DMEM (Corning) supplemented with 25% Ham’s/F12 (Corning), 0.5% N2 supplement (Gibco), 1% B27 without retinoic acid (RA) supplement (Gibco), and 1% bovine serum albumin (BSA) (Sigma). At Day 6, purity of the definitive endoderm (DE) was assessed by flow cytometry with DE being 95-98% CXCR4+/C-kit+. Cells were then dissociated using TrypLE for 2 minutes and replated with ROCK inhibitor (Tocris) onto 1:3 Matrigel coated 6-well plates at a density of 3.5 x 10^5 cells per well. On Day 7, the medium was replaced with hepatoblast specification medium consisting of H16 base medium supplemented 2mM L-glutamine, 50 µg/ml SAA, 4.5×10^{-4} M MTG, 40 ng/mL bFGF, and 50 ng/mL BMP4 (R&D Systems). H16 base medium consists of DMEM Low glucose (Gibco) supplemented with 25% Ham’s/F12, 1% B27 with RA (Gibco), and 1% BSA. Medium was changed every other day. On Day 13 medium was replaced with hepatocyte maturation medium A consisting of H16 base medium supplemented with 2mM L-glutamine, 50 µg/ml SAA, 4.5×10^{-4} M MTG, 20 ng/mL Hepatocyte Growth Factor (HGF) (R&D Systems), 20 ng/mL Oncostatin-M (OSM) (R&D Systems), and 40 ng/ml dexamethasone (Dex) (Sigma). Medium was changed every other day. On Day 21, cells were transferred to an environment of 20% O_2 and 5% CO_2, and 90% N_2. Medium was replaced with hepatoblast maturation medium B consisting of H21 medium supplemented with 2mM L-glutamine, 50 µg/ml SAA, 4.5×10^{-4} M MTG, 20 ng/mL HGF, 20 ng/mL OSM, and 40 ng/mL Dex. H21 medium is the same as H16 except using DMEM high glucose (Gibco) as the base medium. Medium was changed every other day until Day 25. Cells were harvested on Day 0, Day 6, Day 9, Day 12, Day 15, Day 18, Day 21 and Day 25 for analysis.
2.4 Pancreatic β-cell Differentiation

The pancreatic β-cell differentiation protocol previously described (Rezania et al., 2014) was modified as follows. PSCs were split onto 1:30 Matrigel coated 6-well plates and cultured until they reached 80-90% confluency before starting Day 0 of the differentiation. Cells were grown in 20% O₂ and 5% CO₂, and 90% N₂, and differentiated as a monolayer, Cells were cultured for 24 hours in RPMI medium with 100ng/mL Activin A, and 3µM CHIR (Day 0). Cells were cultured for another 24 hours (Day 1) in RPMI medium with 100 ng/mL Activin A, 0.3 µM CHIR, and 5 ng/mL bFGF. On Day 2, cells were differentiated to definitive endoderm (DE) with serum free differentiation (SFD) medium containing 100 ng/mL Activin A. At Day 3, purity of the DE was assessed as described above. From Day 3-5, cells were cultured in DMEM/F12 medium with 1% fetal bovine serum (FBS) (Gibco), 50µg/ml stabilized ascorbic acid (SAA), 1.25 mM IWP2 (Tocris), and 50 ng/ml FGF7 (R&D Systems). Medium was changed every day. On Day 6 and 7, cells were cultured in DMEM high glucose medium with 1% B27 without RA, 1X Glutamax (Gibco), 50µg/ml SAA, 0.5% ITS-X (Gibco), 50 ng/mL FGF7, 0.5µM SANT1 (Sigma), 1µM retinoic acid (RA) (Sigma), 100nM LDN (Tocris), and 500nM Phorbol 12-myristate 13-acetate (Tocris). Medium was changed daily. From Day 8-10, cells were cultured in DMEM high glucose medium with 1% B27 without RA, 1X Glutamax, 50µg/ml SAA, 0.5% ITS-X, 2 ng/mL FGF7, 0.5µM SANT1, 0.1µM RA, 100nM LDN, and 250nM Phorbol 12-myristate 13-acetate. Medium was changed daily. On Day 11 and 12 cultured in MCDB131 base media with 20 mM D-glucose (Sigma), 2% FBS, 1X Glutamax, 0.5% ITS-X, 10µg/ml Heparin (Sigma), 10µM ZnSO₄ (Sigma), 0.5µM SANT1, 0.05µM RA, 200nM LDN, 1µM T3 (Sigma), 2µM ALK5i II (Enzo). Medium was changed daily. From Days 13-25, cells were cultured in MCDB131 base media with 20 mM D-
glucose, 2% FBS, 1X Glutamax, 0.5% ITS-X, 10µg/ml Heparin, 10µM ZnSO₄, 200nM LDN, 1µM T3, 2µM ALK5i II, 100nM g-secretase inhibitor XX (Calbiochem). Medium was changed every other day. Cells were harvested on Day 0, Day 3, Day 6, Day 8, Day 11, Day 13, and Day 25 for analysis.

2.5 Genomic Analysis

Cells were harvested on Day 12 of the hepatocyte differentiation by dissociation with 0.25% Trypsin/EDTA (Gibco) for 4 minutes. RNA was extracted as described below and sent to Genewiz for library preparation and sequencing. Cells were harvested on Day 6 and Day 8 of the pancreas differentiation as described above. RNA was extracted and sent to the Center for Applied Genomics Biorepository Core at the Children’s Hospital of Philadelphia for library preparation and sequencing. Data was analyzed using Rosalind (https://rosalind.onramp.bio/), with a HyperScale architecture developed by OnRamp BioInformatics, Inc. (San Diego, CA). Individual sample counts were normalized via Relative Log Expression (RLE) using DESeq2 R library (Love et al., 2014). Read Distribution percentages, violin plots, identity heatmaps, and sample MDS plots were generated as part of the QC step. Deseq2 was also used to calculate fold changes and p-values and perform optional covariate correction. Clustering of genes for the final heatmap of differentially expressed genes was done using the PAM (Partitioning Around Medoids) method using the fpc R library. Hypergeometric distribution was used to analyze the enrichment of pathways, gene ontology, domain structure, and other ontologies. The topGO R library was used to determine local similarities and dependencies between GO terms in order to perform Elim pruning corrections. Several database sources were referenced for enrichment analysis, including Interpro (Mitchell et al., 2019), NCBI (Geer et al., 2009), KEGG (Kanehisa and Goto, 2000; Kanehisa et
al., 2017, 2019), MSigDB (Liberzon et al., 2011; Subramanian et al., 2005), REACTOME (Fabregat et al., 2018), WikiPathways (Slenter et al., 2018). Enrichment was calculated relative to a set of background genes relevant for the experiment. Gene set enrichment analysis was performed (Mootha et al., 2003; Subramanian et al., 2005) using a gene list comparing gene expression in dorsal pancreatic bud and hepatic cord tissues dissected from human embryos using laser capture technology (Jennings et al., 2017). The gene list was sorted by a p-value of <0.05, then by fold change. The top 200 upregulated genes were the most enriched in the dorsal pancreatic bud, and the top 200 downregulated genes were most enriched in the hepatic cord. GSEA was performed on these subsets using the “gene_set” permutation and an FDR cutoff of 5%.

2.6 RNA isolation and quantitative real time polymerase chain reaction

Cells were harvested by dissociation with 0.25% Trypsin/EDTA for 4 minutes. Cellular RNA was isolated using the PureLink RNA Micro Scale Kit (Invitrogen) following the manufacturer’s protocol. Random hexamers (Invitrogen) were used with the SuperScript III Reverse Transcriptase System (Invitrogen) to synthesize cDNA from 500 ng of extracted RNA. Quantitative real time polymerase chain reaction (qRT-PCR) reactions were done in triplicate on a Roche LightCycler 480 II using SYBR Select Master Mix (Applied Biosystems). Serial dilutions of H9 embryonic stem cell genomic DNA were used to generate a standard curve, and TBP (Veazey and Golding, 2011) was used as a house keeping gene to determine relative gene expression levels. The primers that were used in this study can be found in Table 2.2
### Table 2.2. Table of forward and reverse primers used for qRT-PCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBX3</td>
<td>5'-TGAGATGTCTGGGCTGG-3'</td>
<td>5'-CTTACCAGCCACCATCCA-3'</td>
</tr>
<tr>
<td>HNF4α</td>
<td>5'-TCAACCAACCTCATCCTCTT-3'</td>
<td>5'TCCTCTCACCAGTCTGTT-3'</td>
</tr>
<tr>
<td>AFP</td>
<td>5'-GTGTGTCTCAAGAGCCACTTAC-3'</td>
<td>5'-CACCTGAAGAAGTGGTCATC-3'</td>
</tr>
<tr>
<td>SERPINA1</td>
<td>5'-AGGGAAGAGGATGTTGAGG-3'</td>
<td>5'-TCTGTTTCTGGGCTCTGGT-3'</td>
</tr>
<tr>
<td>ALBUNIN</td>
<td>5'-GTGAACACACAGCCAGGCAACA-3'</td>
<td>5'-TCCTGGGAAAGCATGCTC-3'</td>
</tr>
<tr>
<td>TBX1</td>
<td>5'-CAGTGGCATAGTCTGTTAGAAGG-3'</td>
<td>5'-CGCTACTGAGGTTGAGCAA-3'</td>
</tr>
<tr>
<td>EOMES</td>
<td>5'-TATTTCTGGCTTGGCCACAGGTC-3'</td>
<td>5'-ACTCAATCCACTGCCCACCTACA-3'</td>
</tr>
<tr>
<td>FOXA2</td>
<td>5'-GCATTTCCAATCTTTACGCGGTG-3'</td>
<td>5'-GCCCTTGCAGGCAATACACATT-3'</td>
</tr>
<tr>
<td>SOX2</td>
<td>5'-CCATCACCACAACAGCAAT-3'</td>
<td>5'-AGTCCAGGATCTCTCTCATAA-3'</td>
</tr>
<tr>
<td>PDX1</td>
<td>5'-GGAGCTGGTCTGATGTG-3'</td>
<td>5'-CACTTTAGGCGGCGGTT-3'</td>
</tr>
<tr>
<td>CDX2</td>
<td>5'-AAGGAAGGATGACGTACCCTGAGC-3'</td>
<td>5'-CACGTGGTAACCGCGTAGTC-3'</td>
</tr>
<tr>
<td>ISL1</td>
<td>5'-CAGAAGGAGGACGCGGCTCTAAT-3'</td>
<td>5'-GACTGGCTACGTCTGTTAGGTAT-3'</td>
</tr>
<tr>
<td>MNX1</td>
<td>5'-AGAAGGCGGAAACCCACACCTGTA-3'</td>
<td>5'-CCGAGACGTAAGCAAAACCCT-3'</td>
</tr>
<tr>
<td>NKX6.1</td>
<td>5'-AAGAAGCAACGGCTGGAGATTG-3'</td>
<td>5'-CCGAGTTGGGATCCAGAGCATT-3'</td>
</tr>
<tr>
<td>APOA2</td>
<td>5'-ATGATGGCTGGAGGCTGATGCCTTA-3'</td>
<td>5'-AAGCTGCTGCGGCTGATACCT-3'</td>
</tr>
<tr>
<td>TTR</td>
<td>5'-ATGGGCTACAACTGAGGAGGAT-3'</td>
<td>5'-AGATGCCAAGTGGCTCAGGTAAGA-3'</td>
</tr>
<tr>
<td>TBP</td>
<td>5'-TTGCTGAGAAGATGTGTGGGAGATG-3'</td>
<td>5'-GTGAAGGTTGGGGAGGCTGTTG-3'</td>
</tr>
</tbody>
</table>

### 2.7 Western Blot

Cells were harvested with 1.5x Laemmli buffer (75 mM Tris-HCl, 15% glycerol, 3% SDS, 3.75 mM EDTA, and 200 mM NaF). Cell lysates were boiled at 95°C for 10 minutes and protein was quantified using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Samples were aliquoted and diluted with 4x Laemmli buffer. 20 µg of protein from each sample were run on a 4%-12% Bis-Tris SDS-polyacrylamide gel (Invitrogen).
and transferred to a 0.45 µm pore size nitrocellulose membrane (Thermo Fisher Scientific). The membrane was stained with Ponceau S (Sigma-Aldrich) to ensure successful transfer. The membrane was washed with 1X TBS (Bio-Rad) with 0.1% Tween-20 (Sigma) (TBST), then blocked in 5% nonfat dry milk in TBST for 1 hour at room temperature. The membrane was probed with primary antibody diluted in 5% nonfat dry milk in TBST overnight at 4°C. The membrane was washed three times with TBST and placed in a horseradish peroxidase conjugated secondary antibody diluted in 5% nonfat dry milk in 1X TBST for 1 hour at room temperature. The membrane was washed three times with TBST. Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific) was added to the membrane and exposed onto HyBlot CL autoradiography film (Denville Scientific) to detect protein of interest. A list of antibodies used can be found in Table 2.3.

2.8 Flow Cytometry

Cells were harvested by dissociation with 0.25% Trypsin/EDTA for 3 to 5 minutes. For intracellular staining, cells were fixed with 1.6% paraformaldehyde (Electron Microscopy Science) for 30 minutes at 37°C. Cells were washed in 1X PBS (Corning), then permeabilized and stained in 1X saponin buffer (Biolegend). Primary and secondary antibodies were diluted to the appropriate concentrations in saponin buffer and cells were stained for thirty minutes each at room temperature. Following staining, cells were washed in saponin, and resuspended in FACS buffer (1X PBS (Corning) with 0.1% BSA (Sigma) and 0.1% sodium azide (Sigma)). For extracellular staining, conjugated primary antibodies were diluted to the appropriate concentration in FACS buffer and cells were stained for fifteen minutes at room temperature. Following the staining, cells were washed and resuspended in FACS buffer. All samples were analyzed on a CytoFLEX V2-Br-R2
flow cytometer (Beckman Coulter Life Sciences) and FlowJo Version 10.6.2 (Beckton Dickenson) software program. A list of antibodies used can be found in Table 2.3.

Table 2.3: Primary and secondary antibodies

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Dilution and Application</th>
<th>Source and Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBX3</td>
<td>Goat Polyclonal IgG</td>
<td>1:200 for Western Blot</td>
<td>Santa Cruz #sc-17871</td>
</tr>
<tr>
<td>b-actin</td>
<td>Mouse</td>
<td>1:2000 for Western Blot</td>
<td>Santa Cruz #sc-81178</td>
</tr>
<tr>
<td>CXCR4-APC</td>
<td>Mouse</td>
<td>1:50 for flow cytometry</td>
<td>BD #555976</td>
</tr>
<tr>
<td>CD117-PE</td>
<td>Mouse</td>
<td>1:100 for flow cytometry</td>
<td>BD #340529</td>
</tr>
<tr>
<td>HNF4a</td>
<td>Rabbit Monoclonal IgG</td>
<td>1:100 for flow cytometry</td>
<td>Cell signaling #3113</td>
</tr>
<tr>
<td>AFP</td>
<td>Mouse Monoclonal IgG1</td>
<td>1:400 for flow cytometry</td>
<td>R&amp;D Systems #MAB1368</td>
</tr>
<tr>
<td>AAT</td>
<td>Mouse Monoclonal IgG1</td>
<td>1:100 for flow cytometry</td>
<td>Santa Cruz #sc-59438</td>
</tr>
<tr>
<td>Albumin</td>
<td>Rabbit Polyclonal IgG</td>
<td>1:10000 for flow cytometry</td>
<td>DakoCytomation #A0001</td>
</tr>
<tr>
<td>PDX1</td>
<td>Rabbit Monoclonal IgG</td>
<td>1:1000 for Western Blot</td>
<td>Cell Signaling #5679</td>
</tr>
<tr>
<td>NKX6.1</td>
<td>Mouse Monoclonal IgG1</td>
<td>1:250 for flow cytometry</td>
<td>DSHB #F55A10</td>
</tr>
<tr>
<td>PDX1-biotinylated</td>
<td>Goat Polyclonal IgG</td>
<td>1:50 for flow cytometry</td>
<td>R&amp;D Systems #BAF2419</td>
</tr>
<tr>
<td>c-peptide</td>
<td>Rabbit Polyclonal IgG</td>
<td>1:100 for flow cytometry</td>
<td>Cell Signaling #4593S</td>
</tr>
<tr>
<td>GCG</td>
<td>Mouse Monoclonal IgG1</td>
<td>1:1000 for flow cytometry</td>
<td>Santa Cruz #sc-514592</td>
</tr>
<tr>
<td>SST</td>
<td>Rat Monoclonal IgG2a</td>
<td>1:400 for flow cytometry</td>
<td>Santa Cruz #sc-47706</td>
</tr>
<tr>
<td>Mouse IgG1 - Alexa488</td>
<td>Goat</td>
<td>1:400 for flow cytometry</td>
<td>Jackson Immunoresearch # 115-545-205</td>
</tr>
<tr>
<td>Mouse IgG1 - 647</td>
<td>Donkey</td>
<td>1:400 for flow cytometry</td>
<td>Jackson Immunoresearch # 715-606-151</td>
</tr>
<tr>
<td>Mouse IgG1 - PE</td>
<td>Goat</td>
<td>1:400 for flow cytometry</td>
<td>Jackson Immunoresearch #115-115-205</td>
</tr>
<tr>
<td>Rabbit IgG - Alexa 647</td>
<td>Goat</td>
<td>1:400 for flow cytometry</td>
<td>Jackson Immunoresearch #111-605-144</td>
</tr>
<tr>
<td>Rabbit IgG - Alexa 488</td>
<td>Goat</td>
<td>1:400 for flow cytometry</td>
<td>Jackson Immunoresearch #111-545-144</td>
</tr>
<tr>
<td>Rat IgG - APC</td>
<td>Goat</td>
<td>1:400 for flow cytometry</td>
<td>ThermoFisher Scientific #A21247</td>
</tr>
<tr>
<td>Streptavidin Pacific Blue</td>
<td></td>
<td>1:400 for flow cytometry</td>
<td>Invitrogen #S-11222</td>
</tr>
</tbody>
</table>
2.9 Cytochrome P450 activity assay

The cytochrome p450 activity assay was performed using P450-GloTM CYP3A4 Luciferin-IPA kit (Promega, V9002). On Day 22 of the hepatocyte differentiation, cells were cultured with 25 μM Rifampicin (Sigma) or with DMSO (Tocris) as a vehicle control. To confirm that the induced activity was specific for CYP3A4 enzyme, inhibition controls included the selective inhibitor ketoconazole at 1 μM (Sigma) in the presence of 25 μM Rifampicin. Net signal was calculated by subtracting background luminescence values (no-cell control) from Rifampicin or DMSO values. Luminescence values were normalized for 5x10^5 cells to account for differences in cell numbers between wells.

2.10 Statistical Analysis

Results from multiple experiments are expressed as the mean ± standard error of the mean (SEM). An unpaired two-tailed Student’s t test for groups with equal variance was performed to determine p values. All statistical analyses were performed on Prism version 8.4.3 for Mac (GraphPad Software). In the figures, *p<0.05, **p<0.01, ***p<0.001, and **** p≤ 0.0001, and n denotes individual experiments.
CHAPTER 3: IMPACTS OF LOSS OF TBX3 ON HEPATOCYTES AND PANCREAS DIFFERENTIATION FROM HUMAN PLURIPOTENT STEM CELLS

3.1 Introduction

The liver and pancreas play vital roles in metabolism and digestion. Both organs arise from the posterior foregut region of the developing gut tube (Deutsch et al., 2001). The gut tube is characterized by an expression pattern of different transcription factors that specify distinct domains for each of the endodermal organs. While the hepatic and pancreatic domains of the gut tube lie in close proximity to each other, they are specified by different signals. FGF from the cardiac mesoderm (Gualdi et al., 1996; Jung et al., 1999) and BMP4 from the septum transversum specify the hepatic endoderm (Rossi et al., 2001). Activin and FGF2 signals from the notochord, and retinoic acid from the lateral plate mesoderm specify the dorsal pancreatic endoderm (Hebrok et al., 1998).

The transcription factor TBX3 has been shown to be involved in development in a number of model systems. TBX3 drives mesendoderm in Xenopus and mouse embryos (Weidgang et al., 2013), and has been implicated in limb and heart development (Gibson-Brown et al., 1998; Singh et al., 2012). TBX3 is a member of the T-box gene family, with both repressor and activator domains, but is thought to act predominantly as a transcriptional repressor (Carlson et al., 2001; Naiche et al., 2005). In humans, heterozygous mutations in TBX3 result in Ulnar-Mammary Syndrome (UMS), a disorder causing defects in limb, mammary, and apocrine gland development (Bamshad et al., 1997). Tbx3 has also been implicated in liver development, as it is expressed in hepatoblasts, bipotential progenitors that give rise to hepatocytes and cholangiocytes, which comprise the liver bud. Tbx3−/− mice have small liver buds that fail to expand and mature (Lüdtke et al., 2009). Tbx3 promotes a hepatocyte fate by maintaining
expression of hepatic transcription factors in hepatoblasts, while repressing cholangiocyte-specific genes (Suzuki et al., 2008). Tbx3 has been detected in the mouse pancreas, both during development and in the adult, but its function is unknown (Begum and Papaioannou, 2011). The role of TBX3 in human liver and pancreas development remains unclear.

Pluripotent stem cells (PSCs), such as human embryonic stem cells (ESC) and induced pluripotent stem cells (iPSC) can give rise to all cell types, and are a model system for studying human development, physiology and disease. PSCs are easily manipulated using CRISPR/Cas9 genome editing technology, making them valuable for studying the role of specific genes in development and disease (Cong et al., 2013; Maguire et al., 2019). Protocols to differentiate PSCs to both hepatocytes (Ogawa et al., 2013; Si-Tayeb et al., 2010) and pancreatic β-cells (D’Amour et al., 2006; Nostro et al., 2011; Pagliuca et al., 2014; Rezania et al., 2014) have been established and improved drastically over time. However, it is still difficult to generate functionally mature terminal cell types in vitro.

Here we use PSCs to study the role of TBX3 in human liver and pancreas development. TBX3 mutant PSC lines were generated using CRISPR/Cas9 genome editing, and differentiated into hepatocytes and pancreatic progenitors. The loss of TBX3 caused a defect in hepatocyte differentiation, similar to what was seen in Tbx3−/− mouse embryos. Interestingly, the loss of TBX3 resulted in the expression of PDX1 mRNA, a master regulator for pancreas development (Ahlgren et al., 1996; Stoffers et al., 1997) during the hepatocyte differentiation. Furthermore, PSCs lacking TBX3 differentiated more efficiently to pancreatic progenitors compared to wildtype PSCs. These data suggest that TBX3 may regulate liver development through suppression of pancreatic genes. A
better understanding of how TBX3 regulates pancreatic precursor efficiency may provide a potential avenue to improve generation of in vitro-derived pancreatic β-cells for use in therapeutic contexts.

3.2 Results

3.2.1 Loss of TBX3 impairs hepatocyte differentiation in human PSCs

To study the role of TBX3 in human liver development, a TBX3 mutant line was generated in the CHOPWT4 iPSC background using CRISPR/Cas9 genome editing technology. Two guide RNAs (gRNAs) were designed to generate a 6.9 Kb deletion that includes the transcriptional start site and the entire DNA binding domain, rendering a non-functional protein (Figure 3.1A). The deletion in both alleles was verified by sequencing (Figure 3.6A) and confirmed to have a normal karyotype (Figures 3.6B). The genome-edited line is termed iPSC−, and the unedited line, designated as iPSC+/+, was used as an isogenic control (Table 3.1). To confirm the phenotype in a second genetic background, we generated TBX3 mutations in the Mel1 ESC line (Micallef et al., 2012) using the same strategy (Figure 3.6C-D), referred to as ESC− (Table 3.1).

![Table 3.1 TBX3 mutant pluripotent stem cell lines](image)

<table>
<thead>
<tr>
<th>Name</th>
<th>Genetic Background</th>
<th>Allele 1</th>
<th>Allele 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPSC+/+</td>
<td>CHOP WT4</td>
<td>Wild type</td>
<td>Wild type</td>
</tr>
<tr>
<td>iPSC−</td>
<td>CHOP WT4</td>
<td>6.9 Kb deletion</td>
<td>6.9 Kb deletion</td>
</tr>
<tr>
<td>ES+/+</td>
<td>Mel1-INS-GFP</td>
<td>Wild type</td>
<td>Wild type</td>
</tr>
<tr>
<td>ES−</td>
<td>Mel1-INS-GFP</td>
<td>6.9 Kb deletion</td>
<td>6.9 Kb deletion</td>
</tr>
</tbody>
</table>

TBX3 is not expressed in human PSCs (Esmailpour and Huang, 2012), therefore control and mutant lines were differentiated into hepatocytes using a modified protocol (Ogawa
et al., 2013) to verify loss of TBX3 at the mRNA and protein levels (Figure 3.1B). In iPSC+/+ cells, TBX3 expression was induced at day 6 (definitive endoderm stage) of differentiation, and peaked at day 12 (hepatoblast stage). TBX3 was not expressed in the iPSC−/− line during the differentiation at the mRNA or protein level (Figure 3.1C-D).

Liver marker gene expression was analyzed in the differentiated iPSC cell cultures to determine the impact of TBX3 loss on human liver development. There was no difference in HNF4α levels, but expression of hepatoblast marker α-fetoprotein (AFP) was delayed in the iPSC−/− line. Tbx3−/− mice develop a normal liver bud, but generate a small liver due to impaired migration out of the gut tube epithelium (Lüdtke et al., 2009), suggesting that Tbx3 is not required for hepatoblast specification, but rather for maturation. This explains the subtle effect observed on hepatoblast marker expression in the differentiation of the iPSC−/− line. A more severe defect was observed later in the differentiation, as expression of hepatocyte markers SERPINA1, encoding α1-antitrypsin (AAT) protein, and ALBUMIN were significantly reduced in iPSC−/− cells compared to iPSC+/+ cells (Figure 3.1E). At the protein level, there were fewer HNF4α+/AFP+ cells (Figure 3.1F-G) and AAT+/ALBUMIN+ cells (Figure 3.1H-I) from the iPSC−/− line compared to the iPSC+/+ line. End-stage iPSC−/− hepatocytes treated with rifampicin had significantly reduced CYP3A4 activity compared to iPSC+/+ hepatocytes, indicating that the TBX3 is needed for hepatocyte functionality (Figure 3.1J). Similar defects were seen in the ESC−/− line (Figure 3.7A - 3.7I). It is possible that the ESC line is more sensitive to the loss of TBX3, explain why expression levels of HNF4α and AFP are lower in this line compared to the iPSC−/− line. However, the overall phenotype of TBX3 impairing hepatocyte differentiation is the same in both PSC lines. These results indicate that the loss of TBX3...
impairs the ability of PSCs to differentiate to hepatocytes, suggesting that TBX3 is important in both mouse and human liver development.

Figure 3.1

A

TBX3

1 2 3 4 5 6 7

B

PSC D0 → DE D7 → Hepatoblast D13 → Hepatocyte D25

SOX17 FOXA2 HNF4α AFP AAT Albumin

C

Expression relative to GAPDH

780 kDa

40 kDa

D

TBX3

+/+

β-actin

E

HNF4α

FP

SERPINA1

ALB

F

D15

D25

HNF4α

IPSC

80.0%

91.7%

IPSC

42.0%

62.5%

IPSC

80.0%

91.7%

IPSC

28.3%

43.5%

IPSC

11.2%

18.1%

IPSC

Normalized CYP4A Activity

0 10000 20000 30000 40000

D25
Figure 3.1: Loss of TBX3 impairs hepatocyte differentiation in the iPSC⁺ line

(A) Schematic of endogenous TBX3 locus with sites of gRNAs indicated above exons. Schematic underneath represents TBX3 locus after 6.9 kb deletion, resulting in a mutant TBX3 PSC line.

(B) Schematic of directed differentiation protocol of PSCs to hepatocytes.

(C) Time-course analysis of TBX3 expression during hepatocyte differentiation by qRT-PCR (n = 5 per time point, per cell line).

(D) Western blot analysis of TBX3 protein in Day 12 iPSC⁺/+ and iPSC⁻/⁻ hepatoblasts.

(E) Time-course analysis of hepatoblast (HNF4α and AFP) and hepatocyte (SERPINA1 and Albumin) markers during hepatocyte differentiation by qRT-PCR (n = 5 per time point, per cell line).

(F) Analysis of hepatoblast markers HNF4α and AFP at Day 15 and Day 25 of differentiation by intracellular flow cytometry (one representative replicate shown).

(G) Time-course analysis percentage HNF4α⁺/AFP⁺ cells by intracellular flow cytometry (n = 5 per time point, per cell line).

(H) Analysis of hepatocyte markers AAT and ALBUMIN at Day 15 and Day 25 of differentiation by intracellular flow cytometry (one representative replicate shown).

(I) Time-course analysis of percentage of AAT⁺/ALBUMIN⁺ cells in iPSC⁺/+ and iPSC⁻/⁻ lines by intracellular flow cytometry (n = 5 per time point, per cell line).

(J) Rifampicin-induced CYP3A4 activity in iPSC⁺/+ and iPSC⁻/⁻ hepatocytes at Day 25 of the differentiation. Activity is normalized for 5x10⁵ cells (n = 4 per cell line).
3.2.2 Pancreas-specific genes are upregulated in *TBX3* knockout cells

*Tbx3* drives mesendoderm differentiation in *Xenopus* and mouse embryos (Weidgang et al., 2013). We examined expression of early lineage markers to ensure that the hepatocyte defect we observed is not due to impaired mesendoderm or definitive endoderm differentiation. There was no difference in expression of mesoderm marker *TBXT*, mesendoderm marker *EOMES*, or definitive endoderm marker *FOXA2* between the iPSC\(^{+/+}\) and iPSC\(^{-/-}\) lines (Figure 3.2A). This demonstrates that the loss of *TBX3* does not impact mesendoderm differentiation, and highlights a potential species-specific difference in *TBX3* function during early development. Next, we examined whether loss of TBX3 affected gut tube patterning. The developing gut tube displays a well-established expression pattern of transcription factors that specify different organ domains (Zorn and Wells, 2009) (Figure 1.1). To determine if the loss of TBX3 affected gut tube patterning, the expression of *SOX2*, *CDX2*, and *PDX1*, which specify the anterior gut tube, posterior gut tube, and pancreatic endoderm respectively, were examined. There was no impact on the expression of *SOX2*, and a minor upregulation of *CDX2* in the iPSC\(^{-/-}\) line. There was a large increase in expression of *PDX1* in the iPSC\(^{-/-}\) line compared to the iPSC\(^{+/+}\) line (Figure 3.2B). This was surprising as *PDX1* is not normally expressed when using a hepatocyte differentiation protocol. Western blot analysis did not detect PDX1 protein in the mutant line (Figure 3.2C). PDX1 protein levels are tightly regulated by ubiquitination and proteasomal degradation (Claiborn et al., 2010), providing a possible explanation for the increase in *PDX1* mRNA but not PDX1 protein in the hepatocyte differentiation. Additionally, expression of *ISL1* (Ahlgren et al., 1997) and *MNX1* (Harrison et al., 1999; Li et al., 1999), genes involved in early pancreas development, were also increased in the iPSC\(^{-/-}\) line (Figure 3.2D), suggesting
Figure 3.2

(A) Time-course analysis of mesoderm (TBXT), mesendoderm (EOMES), and endoderm (FOXA2) markers during definitive endoderm differentiation by qRT-PCR (n = 3 per time point, per cell line).

(B) Time-course analysis of anterior gut tube (SOX2), pancreatic endoderm (PDX1), and posterior gut tube (CDX2) markers during hepatocyte differentiation by qRT-PCR. (n = 5 per time point, per cell line). PDX1 expression in iPSC+/− differentiated to pancreatic endoderm for (+) control (n = 6).

(C) Western blot analysis of PDX1 protein in Day 15 iPSC+/+ and iPSC−− immature hepatocytes and control EndoC-βH1 cell line.

(D) Time-course analysis of early pancreatic markers ISL1 and MNX1 during hepatocyte differentiation by qRT-PCR (n = 5 per time point, per cell line).

Figure 3.2: Pancreatic genes are expressed in the iPSC−− line during hepatocyte differentiation

(A) Time-course analysis of mesoderm (TBXT), mesendoderm (EOMES), and endoderm (FOXA2) markers during definitive endoderm differentiation by qRT-PCR (n = 3 per time point, per cell line).

(B) Time-course analysis of anterior gut tube (SOX2), pancreatic endoderm (PDX1), and posterior gut tube (CDX2) markers during hepatocyte differentiation by qRT-PCR. (n = 5 per time point, per cell line). PDX1 expression in iPSC+/− differentiated to pancreatic endoderm for (+) control (n = 6).

(C) Western blot analysis of PDX1 protein in Day 15 iPSC+/+ and iPSC−− immature hepatocytes and control EndoC-βH1 cell line.

(D) Time-course analysis of early pancreatic markers ISL1 and MNX1 during hepatocyte differentiation by qRT-PCR (n = 5 per time point, per cell line).
a loss of repression of pancreatic genes. Similar results were also seen the ESC\textsuperscript{−−} line (Figure 3.8A-3.8D). PDX1 drives PSCs to a pancreatic fate during differentiation in part by repression of hepatic-specific genes, including \textit{TBX3} (Teo et al., 2015). The upregulation of \textit{PDX1}, \textit{ISL1}, and \textit{MNX1} in the iPSC\textsuperscript{−−} line suggests that TBX3 may be suppressing their expression in order to drive PSCs to a hepatocyte fate. Evidence of TBX3 regulating cell fates exists, as it promotes a hepatocyte over cholangiocyte fate in hepatoblasts (Suzuki et al., 2008). Further studies to the relationship between TBX3 and PDX1 are needed. Assay for transposase-accessible chromatin sequencing (ATAC-seq) experiments can identify how chromatin accessibility changes for different genes, particularly at pancreatic gene loci, with the loss of \textit{TBX3}. Additionally, chromatin immunoprecipitation sequencing (ChIP-seq) studies can identify direct targets of TBX3 at the hepatoblast stage. These studies may help to determine whether PDX1 and TBX3 cross-regulate each other to maintain lineage fidelity of pancreatic and hepatic domains respectively during foregut patterning.

3.2.3 Loss of \textit{TBX3} enhances pancreatic progenitor generation from PSCs

Considering that \textit{TBX3} loss led to de-repression of pancreatic genes during \textit{in vitro} liver differentiation, we tested whether a lack of \textit{TBX3} enhanced \textit{in vitro} pancreatic differentiation. The iPSC\textsuperscript{+/+} and iPSC\textsuperscript{−−} lines were differentiated to pancreatic progenitors using a modified version of an established protocol (Rezania et al., 2014) (Figure 3.3A). The iPSC\textsuperscript{−−} line generated a higher percentage of the PDX1\textsuperscript{+/}/NKX6.1\textsuperscript{+} pancreatic progenitor 2 (PP2) population compared to iPSC\textsuperscript{+/+} cells (Figure 3.3B-C). Though \textit{PDX1} levels increased in the iPSC\textsuperscript{−−} hepatocyte differentiation, there was no difference in \textit{PDX1} expression between iPSC\textsuperscript{+/+} and iPSC\textsuperscript{−−} cells in the pancreas differentiation. This
Figure 3.3

(A) Schematic of directed differentiation protocol of PSCs to PP2 cells.

(B) Flow cytometry analysis of the PP2 makers PDX1 and NKX6.1 at day 11 of differentiation (one representative replicate shown).

(C) Quantification of the percentage of PDX1+/NKX6.1+ and PDX1+ cells in samples examined in (B) (n = 6 per cell line).

(D) Time-course analysis of early pancreatic markers PDX1, NKX6.1, ISL1, and MNX1 during pancreatic differentiation by qRT-PCR (n = 6 per time point, per cell line).

(E) Time-course analysis hepatoblast (AFP and TTR) and hepatocyte (SERPINA1) markers during pancreatic differentiation by qRT-PCR (n = 6 per time, point per cell line).

Figure 3.3: Loss of TBX3 enhances pancreatic progenitor generation in the iPSC−/− line

(A) Schematic of directed differentiation protocol of PSCs to PP2 cells.

(B) Flow cytometry analysis of the PP2 makers PDX1 and NKX6.1 at day 11 of differentiation (one representative replicate shown).

(C) Quantification of the percentage of PDX1+/NKX6.1+ and PDX1+ cells in samples examined in (B) (n = 6 per cell line).

(D) Time-course analysis of early pancreatic markers PDX1, NKX6.1, ISL1, and MNX1 during pancreatic differentiation by qRT-PCR (n = 6 per time point, per cell line).

(E) Time-course analysis hepatoblast (AFP and TTR) and hepatocyte (SERPINA1) markers during pancreatic differentiation by qRT-PCR (n = 6 per time, point per cell line).
is likely because the pancreatic differentiation protocol has factors that drive cells towards a pancreatic identity, and virtually all cells already express PDX1 (Figure 3.3C). However, *NKX6.1* expression was higher in iPSC−/− PP2 cells compared to iPSC+/+ PP2 cells. *NKX6.1* is crucial for pancreatic progenitor identity and required for further differentiation to β-cells (Memon et al., 2018; Nostro et al., 2015; Rezania et al., 2013). Additionally, expression of *ISL1* was markedly increased in iPSC−/− cells compared to iPSC+/+ cells, while there was a subtle increase in *MNX1* expression in iPSC−/− cells (Figure 3.3D). We also examined hepatic gene expression during the pancreatic differentiation. Levels of hepatic genes were lower in the iPSC−/− line compared to the iPSC+/+ line (Figure 3.3E), suggesting the loss of *TBX3* generates a purer PP2 population. Similar results were seen in the ESC−/− line (Figure 3.9A-3.9D). These data demonstrate that a lack of TBX3 improves differentiation efficiency and purity, as the *TBX3* mutant PSC lines generated more PP2 cells with increased expression of pancreas genes and decreased expression of hepatic genes, compared to the wildtype PSC lines. This suggests a novel role for TBX3 in inhibiting pancreas differentiation.

### 3.2.5 TBX2 does not compensate for the loss of TBX3 in hepatocyte or pancreas differentiation

*TBX3* is part of the highly conserved T-box gene family. Specifically, it is part of the *TBX2* subfamily which also includes *TBX2*, *TBX4*, and *TBX5*. *TBX3* and *TBX2* are highly homologous in both their sequence and structure. Their DNA-binding domain sequences are about 95% similar, while the repression domains share 67% homology (Khan et al., 2020). Additionally, *TBX2* and *TBX3* have several overlapping functions, both in development and cancer (Gibson-Brown et al., 1998; Lingbeek et al., 2002; Singh et al., 2012). Because of the share such a high degree of similarity, *TBX2* expression levels
Figure 3.4

A

TBX2

Expression Relative to TBP

Day of Differentiation

B

HNF4α

AFP

SERPINA1

ALBUN

Expression Relative to TBP

Day of Differentiation

C

PDX1

Expression Relative to TBP

Day of Differentiation

D

WT

TBX3−/−

TBX2−/−;TBX3−/−

Isotype

E

%PDX1+ /NKX6.1+

Percent
Figure 3.4 *TBX2* does not compensate for the loss of *TBX3* in hepatocyte or pancreas differentiation

(A) Time-course analysis of *TBX2* expression during hepatocyte differentiation by qRT-PCR (n = 4 per time point, per cell line). Wildtype in blue, *TBX3*−/− in red, and *TBX2*−/− ;*TBX3*−/− in green. The same color scheme follows for the remaining panels of this figure.

(B) Time-course analysis of hepatoblast (*HNF4α* and *AFP*) and hepatocyte (*SERPINA1* and *Albumin*) markers during hepatocyte differentiation by qRT-PCR (n = 5 per time point, per cell line).

(C) Time-course analysis pancreatic marker *PDX1* during hepatocyte differentiation by qRT-PCR (n = 4 per time point, per cell line).

(D) Flow cytometry analysis of the PP2 makers PDX1 and NKX6.1 at day 11 of differentiation (one representative replicate shown).

(E) Quantification of the percentage of PDX1+/NKX6.1+ and PDX1+ cells in samples examined in (D) (n = 4 per cell line).
were examined during the hepatocyte differentiation. TBX2 is not normally expressed during the hepatocyte differentiation, but surprisingly, TBX2 levels increased in the absence of TBX3 (Figure 3.4A). To determine whether TBX2 compensates for the loss of TBX3, a TBX2 and TBX3 double knockout line termed TBX2\(^{-/-}\);TBX3\(^{-/-}\) was generated. This line was differentiated to hepatocytes to confirm the knockout as there is no TBX2 expression (Figure 3.4A), and to determine how the loss of TBX2 impacts hepatocyte differentiation. If TBX2 and TBX3 play a compensatory role in liver differentiation, the double knockout line was expected to have a more severe effect than the TBX3 single knockout alone. While hepatocyte differentiation was still impaired in the TBX2\(^{-/-}\);TBX3\(^{-/-}\) line, there was no significant difference between the double and TBX3 single knockout line (Figure 3.4B) This suggests that they do not compensate for each other, and that TBX3 plays a more significant role in hepatocyte differentiation.

Similar to the TBX3 single knockout, there was an increase in PDX1 levels in TBX2\(^{-/-}\);TBX3\(^{-/-}\) cells during the hepatocyte differentiation (Figure 3.4C). Because PDX1 expression was also increased in this line, we tested whether the loss of TBX2 impacted differentiation to PP2 cells. If TBX2 compensates for TBX3 act in the pancreas differentiation, we expect that the double knockout would generate pancreatic progenitors even more efficiently than the TBX3 single knockout line. Again, while the TBX2\(^{-/-}\);TBX3\(^{-/-}\) line generated more PP2 cells compared to wildtype cells, there was no difference in the ability of the TBX3 single knockout and TBX2\(^{-/-}\);TBX3\(^{-/-}\) lines to generate pancreatic progenitors (Figure 3.4D-E). Taken together, these results imply that TBX2 does not compensate for the loss of TBX3 in either hepatocyte or pancreas differentiation. In both these developmental processes, the loss of TBX3 appears to have a large impact on the differentiation while the loss of TBX2 does not. Therefore, we
conducted gene expression studies to gain insight into how *TBX3* may be acting in pancreas and liver development.

### 3.2.5 iPSC<sup>-</sup> cells are enriched for a pancreatic gene signature

To investigate the effect of the loss of *TBX3* on pancreatic differentiation in greater detail, gene expression was examined using RNA-seq in iPSC<sup>-</sup> versus iPSC<sup>+/+</sup> lines at days 6 and 8 of pancreas differentiation (Figure 3.5A and 3.10A). We chose day 6 as *TBX3* expression peaked at this point in the differentiation, and was comparable to *TBX3* levels in day 12 during the hepatocyte differentiation (Figure 3.5B). This stage of differentiation is representative of the gut tube (GT) endoderm when endodermal cells begin patterning to the foregut. We also chose pancreatic progenitor 1 (PP1) cells at day 8, which represents when cells first express the pancreas master regulator PDX1. At both the GT and the PP1 stages, pancreas specific genes were upregulated while hepatic genes were downregulated in iPSC<sup>-</sup> versus iPSC<sup>+/+</sup> (Figure 3.10B and Figure 3.5C). Pathway analysis showed that genes involved in metabolic pathways were downregulated in iPSC<sup>-</sup> GT cells compared to iPSC<sup>+/+</sup> GT cells (Table 3.2), and pathways relating to pancreatic development and function were upregulated in iPSC<sup>-</sup> PP1 (Table 3.3). Interestingly, genes involved in axonal guidance were also enriched in iPSC<sup>-</sup> PP1 cells. This provides further evidence that these cells have enhanced pancreatic characteristics, as β-cells share many common features with neuronal cells (Arntfield and van der Kooy, 2011). To determine if the gene expression signature of iPSC<sup>-</sup> cells in our *in vitro* differentiation was indicative of a pancreas commitment at the expense of liver fate, we used gene set enrichment analysis with a previously published gene set from primary human embryos comparing the dorsal pancreatic buds to hepatic cords (Jennings et al., 2017). Hepatic cord specific genes were significantly enriched in
Figure 3.5

A. Pancreas and Hepatocyte differentiation:

- Pancreas: D6, D8, PP1
- Hepatocyte: D12

B. Expression of TBX3 relative to TBP:

Day of Differentiation:
- Expression levels shown for Pancreas (blue) and Hepatoblast D12 (green).

C. Log2 Fold Change at Day 8 for PP1:

Genes: TTR, CYP51A1, PDX1, PTF1α, ALB, NFE2, FOXA, MOV34.

D. Dorsal Pancreatic Bud and Hepatic Cord Enrichment Scores:

- Dorsal Pancreatic Bud: NES: 1.41, FDR: 0.013
- Hepatic Cord: NES: -1.10, FDR: 0.186

E. Gene expression heatmap for IPSC+ and IPSC- in different cell types.
Figure 3.5: iPSC<sup><sup>−/+</sup></sup> PP1 cells are enriched for a pancreatic gene signature

(A) Schematic of stages from PSC differentiations collected for RNA-seq analysis.
(B) Time-course analysis of TBX3 expression by qRT-PCR during pancreatic differentiation (n = 6) and Day 12 hepatoblasts (n = 5) for comparison.
(C) Volcano plot of downregulated and upregulated genes in iPSC<sup>−/+</sup> versus iPSC<sup>+/+</sup> PP1 (Day 8) cells. p-Adj = 0.05, fold change: ≥ 1.5 and ≤ -1.5.
(D) GSEA analysis comparing normalized gene expression of samples examined in (C) to genes enriched in human fetal dorsal pancreatic bud and hepatic cord.
(E) Heat map of genes identified by covariate meta-analysis that are commonly upregulated and downregulated in iPSC<sup>−/+</sup> versus iPSC<sup>+/+</sup> cells at different developmental stages and tissues types. PP1 = pancreatic progenitor 1, PGT = Day 6 primitive gut tube, Hep = Day 12 hepatoblasts.
iPSC+/+ versus iPSC− GT cells, and pancreatic bud specific genes were significantly enriched in the iPSC− PP1 population (Figure 3.10C and Figure 3.5D). These data further confirm that the loss of TBX3 helped drive cells towards a pancreatic fate and away from a hepatic fate.

Table 3.2: Top pathways downregulated in iPSC− primitive gut tube cells

<table>
<thead>
<tr>
<th>Pathway</th>
<th>p-value</th>
<th>Gene Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycosaminoglycan Metabolism</td>
<td>1.2x10^{-4}</td>
<td>FMOD, PRELP, HS3ST3B1, HS3ST1, DSE,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS3ST5, HS6ST2, ACAN, VCAN, PAPSS2,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS3ST3A1, CSGALNACT1, CHST2</td>
</tr>
<tr>
<td>Cholesterol Biosynthesis</td>
<td>9.3x10^{-4}</td>
<td>SQULE, FDFT1, HMGCS1, HMGCR, MSM01</td>
</tr>
<tr>
<td>Metabolism of Lipids</td>
<td>1.6x10^{-3}</td>
<td>BCHE, HPGD, PLD1, FABP7, RGL1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPARGC1A, ALB, SQLE, PISD, INSIG1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEC24D, FDFT1, OSBPL10, ACOT12, ACSM3,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>STARD4, HMGCS1, ELOVL5, ABHD4, GPX2,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PRKD1, SPTSSB, HMGCR, MSM01, ARSG,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HSD17B11, GC, AHR, CYP4A11, ACOT2,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PTGS1, PLA2G16, GPD1L, PIPS1B, ABCB4,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAV1, AKR1D1, PLA2G3</td>
</tr>
<tr>
<td>Metabolism of Carbohydrates</td>
<td>4.9x10^{-3}</td>
<td>FMOD, PRELP, SLC2A1, ALDH1A1, HS3ST3B1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS3ST1, MANBA, DSE, SLC2A3, HS3ST5,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS6ST2, ACAN, PFKP, VCAN, PAPSS2, SLC2A14,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS3ST3A1, CSGALNACT1, PPP1R3C, CHST2</td>
</tr>
</tbody>
</table>

Table 3.3: Top pathways upregulated in iPSC+/− PP1 cells

<table>
<thead>
<tr>
<th>Pathway</th>
<th>p-value</th>
<th>Gene Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation of gene expression in early pancreatic precursor cells</td>
<td>5.310^{-6}</td>
<td>PDX1, PTF1A, ONECUT1, NKX6.1</td>
</tr>
<tr>
<td>Axon Guidance</td>
<td>1.7x10^{-4}</td>
<td>TN1, DPYSL4, SEMA6D, EPHB3, PLXNA1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DSCAM, RGMA, PLXNA2, PSMB8, PRKCC,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEMA5A, DPYSL5, KCNQ2, SCN3B, EPHA7,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNC5C, EPHB1, CACNB2, CACNA1D, S RGAP3,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABLIM3, SH3GL2</td>
</tr>
<tr>
<td>Interferon Signaling</td>
<td>6.2x10^{-4}</td>
<td>HLA-B, TRIM46, TRIM3, MT2A, PSMB8,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EGR1, HLA-C, STAT1, HLA-DQB1</td>
</tr>
<tr>
<td>Regulation of beta-cell development</td>
<td>5.4x10^{-3}</td>
<td>PDX1, PTF1A, ONECUT1, NKX6.1</td>
</tr>
</tbody>
</table>
Additional characterization was performed using RNA-seq analysis of iPSC+/− versus iPSC+/− day 12 hepatoblasts (Figure 3.5A). PROX1, which is required for hepatoblast migration and is downregulated in Tbx3−/− mice (Lüdtke et al., 2009), and CDH2, a mesenchymal cell marker, were downregulated in iPSC−/− hepatoblasts, suggesting a migratory defect. Hepatoblast markers HNF4α and AFP were slightly upregulated in iPSC−/− hepatoblasts, because the hepatocyte defect was not apparent until later in the differentiation. Consistent with our prior gene expression findings, PDX1, ISL1, and MNX1 were upregulated in iPSC−/− hepatoblasts (Figures 3.10D). To better understand the global changes in gene expression due to the loss of TBX3, a covariate meta-analysis was performed to identify commonly dysregulated genes in iPSC−/− cells at three developmental stages: PGT and PP1 cells from the pancreatic differentiation, and hepatoblasts from the hepatocyte differentiation. The covariate analysis showed 1398 genes similarly upregulated or downregulated in all three developmental stages (Figure 3.5E). Several pathways, including epithelial-mesenchymal transition (EMT) and TNFα signaling via NF-κB were dysregulated with the loss of TBX3 regardless of cell type (Table 3.4), suggesting a common role of TBX3 in these processes during pancreas and liver development. Taken together these data highlight potentially novel functions of TBX3, and present possible commonalities in liver and pancreas development.
Table 3.4 Genes commonly up- and down-regulated in iPSC+/ primitive gut tube, PP1, and hepatoblast cells

<table>
<thead>
<tr>
<th>Pathway</th>
<th>p-value</th>
<th>Gene Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epithelial-Mesenchymal Transition</td>
<td>5.3x10^-8</td>
<td>VEGFA, WNT5A, DKK1, DAB2, CRLF1, GREM1, AB13BP, ECM2, CADM1, VCAN, GJA1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCOLCE2, LAMC1, PVR, CDH2, LAMA3, LAMC2, MATN2, SNTB1, VEGFC, FBN1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>COL12A1, COL11A1, FBN2, MATN3, CYR61, NT5E, GADD45B, OXTR, ITGAV, TGFB1, IL6,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LOX, COL6A3, COL5A1, MMP1, FMOD, TNFRSF12A, GLIPR,</td>
</tr>
<tr>
<td>TNFα Signaling Via NF-κB</td>
<td>8.1x10^-4</td>
<td>VEGFA, LAMB3, DDX58, CD83, ACKR3, JAG1, IL7R, EFNA1, G0S2, SPHK1, EGR1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FOSL2, MARCKS, DUSP4, PTGS2, F2RL1, LDLR, PTPRE, CYR61, HBEGF, SLC16A6, GADD45B,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CSF1, EDN1, KLF4, IL6, PDE4B, IL1A, IL18</td>
</tr>
<tr>
<td>Pancreas Beta Cells</td>
<td>3.2x10^-4</td>
<td>CSK2, ISL1, PDX1, SST, PAX6, DCX, NKX6.1, AKT3, DPP4</td>
</tr>
<tr>
<td>Inflammatory Response</td>
<td>3.3x10^-3</td>
<td>RGS1, FZD5, IL7R, MSLR1, TACR1, CD82, CSF3R, SPHK1, PIK3R5, P2RY2, IL2RB,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PVR, AHR, TAPBP, LDLR, PTPRE, ITGB8, HBEGF, CSF1, EDN1, IL6, PDE4B, PTAFR, IL1A,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SLC7A1, IL18, SCN1B</td>
</tr>
</tbody>
</table>

3.3 Discussion

This study investigates the role of TBX3 in both human liver and pancreas development. Our findings demonstrate that TBX3 is needed for hepatocyte differentiation, suggesting that it functions similarly in the mouse and human. Unexpectedly, the loss of TBX3 results in pancreatic gene expression during the hepatocyte differentiation, raising the possibility that TBX3 drives liver development through repression of pancreas-specific genes. This led us to examine how TBX3 impacts pancreas differentiation. We found that TBX3 mutant PSCs generate a higher yield of pancreatic progenitors with an enhanced pancreatic gene signature. Further genomic analysis demonstrated that epithelial-mesenchymal transition (EMT) is commonly dysregulated in both pancreas
and liver development, presenting a potentially common function for TBX3 in these processes.

TBX3 is known to play a role in EMT in several different biological contexts. TBX3 promotes invasiveness of breast cancer and melanoma cells through direct regulation of EMT genes \textit{SLUG} (Krstic et al., 2019) and E-cadherin (Rodriguez et al., 2008), respectively. During liver bud expansion, hepatoblasts undergo EMT and delaminate into the surrounding mesenchyme. \textit{Tbx3\textsuperscript{-/-}} mice maintain E-cadherin expression, indicating a failure in EMT and subsequent liver bud expansion (Lüdtke et al., 2009). In pancreas development, endocrine progenitors undergo EMT and delaminate from the branching epithelium into the surrounding mesenchyme to form pancreatic islets (Gouzi et al., 2011). Although \textit{Tbx3} expression has been observed in the developing and adult mouse pancreas, its role is unclear. \textit{Tbx3} was detected in the pancreatic mesenchyme during budding and early morphogenesis (Zhou et al., 2007) near branching epithelial tubes (Begum and Papaioannou, 2011). Based on the role of TBX3 in regulating EMT in other biological contexts, it may function similarly in delamination during islet formation.

Cytoskeletal structure influences delamination (Kesavan et al., 2014) and differentiation (Mamidi et al., 2018) in the mouse, and enhances PP2 generation from PSCs (Hogrebe et al., 2020). We show that loss of \textit{TBX3} enhances PP2 generation, highlighting possible links between TBX3, cytoskeletal state, and differentiation.

This study provides insight into the biology of patterning of the hepatic and pancreatic domains during foregut development. We established that TBX3 is critical for proper hepatocyte development and maturation in humans, similar to the role of Tbx3 in mouse liver development. We also demonstrate that lowering \textit{TBX3} levels optimizes the quality and number of pancreatic progenitors derived from PSCs. These results enhance the
knowledge of early endoderm development, and provide a methodology to improve the generation of in vitro derived pancreatic progenitors, which can be further differentiated to pancreatic cell types for use in therapeutic applications.

Figure 3.6

Figure 3.6: Characterization of TBX3 mutant PSC lines
(A) Sequencing of the region to verify deletion in both alleles of the TBX3 locus in the iPSC<sup>−/−</sup> line. Bars represent locations of TBX3 gRNA1 and TBX3 gRNA2.
(B) Karyotype of iPSC<sup>−/−</sup> line.
(C) Sequencing of the region of the gRNA sequences to verify deletion in both alleles of endogenous TBX3 locus in the ESC<sup>−/−</sup> line. Bars represent locations of TBX3 gRNA1 and TBX3 gRNA2.
(D) Karyotype of ESC<sup>−/−</sup> line.
Figure 3.7

A

TBX3

Expression Relative to TBP

Day of Differentiation

ESC++

ESC−

B

80 kDa

TBX3

40 kDa

β-actin

C

MNF45

Expression Relative to GFP

Day of Differentiation

ESC++

ESC−

APF

SERPINA1

ALBUN

D

D15

D25

ESC++

ESC−

HNF4α

AFP

55.5%

78.3%

E

% HNF4α/AFP

Day of Differentiation

ESC++

ESC−

F

D15

D25

ESC++

ESC−

ALBUN

AAT

5.2%

17.8%

G

% AAT/ALBUN

Day of Differentiation

ESC++

ESC−

H

Normalized CYP3A4 Activity

D25

ESC++

ESC−
Figure 3.7: Loss of TBX3 impairs hepatocyte differentiation in the ESC⁺ line
(A) Time-course analysis of TBX3 expression during hepatocyte differentiation by qRT-PCR. (n = 3 per time point, per cell line).
(B) Western blot analysis of TBX3 protein in Day 12 ESC⁺/+ and ESC⁻/- hepatoblasts.
(C) Time-course analysis of hepatoblast (HNF4α and AFP) and hepatocyte (SERPINA1 and ALBUMIN) markers during hepatocyte differentiation by qRT-PCR (n = 3 per time, point per cell line).
(D) Analysis of hepatoblast markers HNF4α and AFP at Day 15 and Day 25 of differentiation by intracellular flow cytometry (one representative replicate shown).
(E) Time-course analysis of percentage of HNF4α⁺/AFP⁺ cells by intracellular flow cytometry (n = 3 per time point, per cell line).
(F) Analysis of hepatocyte markers AAT ad ALBUMIN at day 15 and Day 25 of differentiation by intracellular flow cytometry (one representative replicate shown).
(G) Time-course analysis of percentage of AAT⁺/ALBUMIN⁺ cells by intracellular flow cytometry (n = 3 per time point, per cell line).
(H) Rifampicin-induced CYP3A4 activity in ESC⁺/+ and ESC⁻/- hepatocytes at Day 25 of the differentiation. Activity is normalized for 5x10⁵ cells (n = 3 per cell line).
Figure 3.8: Pancreatic genes are expressed in the ESC−/− line during hepatocyte differentiation

(A) Time-course analysis of mesoderm (TBXT), mesendoderm (EOMES), and endoderm (FOXA2) markers during definitive endoderm differentiation by qRT-PCR (n = 3 per time point, per cell line).

(B) Time-course analysis of anterior gut tube (SOX2), pancreatic endoderm (PDX1), and posterior gut tube (CDX2) markers during hepatocyte differentiation by qRT-PCR (n = 3 per time point, per cell line). PDX1 expression in ESC+/+ differentiated to pancreatic endoderm for (+) control (n = 5).

(C) Western blot analysis of PDX1 protein in Day 15 ESC+/+ and ESC−/− immature hepatocytes and control EndoC-βH1 cell line.

(D) Time-course analysis of early pancreatic markers ISL1 and MNX1 during hepatocyte differentiation by qRT-PCR (n = 3 per time point, per cell line).
Figure 3.9: Loss of TBX3 enhances pancreatic progenitor generation in the ESC−/− line

(A) Flow cytometry analysis of the PP2 makers PDX1 and NXX6.1 at day 11 of differentiation (one representative replicate shown).

(B) Quantification of the percentage of PDX1+/NXX6.1+ and PDX1+ cells in samples examined in (D) (n = 5 per cell line).

(C) Time-course analysis of early pancreatic markers PDX1, NXX6.1, ISL1, and MNX1 during pancreatic differentiation by qRT-PCR (n = 5 per time point, per cell line).

(D) Time-course analysis hepatoblast (AFP, TTR, ApoA2, and ApoB) markers during pancreatic differentiation by qRT-PCR (n = 5 per time point, per cell line).
Figure 3.10: iPSC\textsuperscript{+-} PGT and hepatoblast cells are enriched for hepatic gene signature

(A) Sample correlation of heatmap correlating samples of same replicate group, developmental stage, and genotype. PGT = Day 6 primitive gut tube, PP1 = Day 8 pancreatic progenitor, Hep = Day 12 hepatoblasts.

(B) Volcano plot of downregulated and upregulated genes in iPSC\textsuperscript{--} versus iPSC\textsuperscript{++} primitive gut tube (Day 6) cells. p-Adj = 0.05, fold change: > 1.5 and < -1.5.

(C) GSEA analysis comparing normalized gene expression of samples examined in (B) to genes enriched in human fetal hepatic cords and dorsal pancreatic bud.

(D) Volcano plot of downregulated and upregulated genes in iPSC\textsuperscript{--} versus iPSC\textsuperscript{++} hepatoblasts cells. p-Adj = 0.05, fold change: > 1.5 and < -1.5.
CHAPTER 4: SUMMARY AND SPECULATIONS

4.1 Summary of Findings
In these studies, we used human PSCs lines to investigate the role of TBX3 in human liver and pancreas development. Our findings demonstrate that TBX3 is important for human liver development, in part by negatively regulating pancreatic gene expression. TBX3 mutant PSCs differentiate more efficiently to pancreatic progenitors, as these cells have enhanced pancreatic gene expression at the expense of hepatic genes. RNA-sequencing analysis determined that genes involved in epithelial-mesenchymal transition are commonly dysregulated in both liver and pancreas cells, suggesting a common role for TBX3 in these developmental processes. These findings provide insight into how TBX3 functions during liver and pancreas development, but several questions still remain. This work presents potential avenues for improving pancreatic progenitor generation from PSCs that need further investigation. These future studies can help improve PSC-derived β-cells for use in therapeutic treatments for diabetes.

4.2 Enhancing cholangiocyte differentiation from PSCs
One role for Tbx3 in mouse liver development is to drive hepatocyte differentiation and suppress a cholangiocyte fate. Tbx3−/− mouse embryos have increased expression of cholangiocyte-specific genes and decreased expression of hepatocyte genes. Part of the work in this thesis investigates the role of TBX3 in human liver development, and how it impacts hepatocyte differentiation. We find that the loss of TBX3 significantly impairs hepatic gene expression and functionality. However, whether the loss of TBX3 enhances the ability of PSCs to differentiate to cholangiocytes remains unclear.
During liver development, a small subset of hepatoblasts adjacent to the portal mesenchyme differentiate into cholangiocytes which mature to form intrahepatic bile ducts (IHBDs) (Zong and Stanger, 2012). Portal mesenchyme cells express Jagged-1, a ligand activating the Notch signaling pathway, and its receptor Notch2 is expressed on neighboring hepatoblasts. Notch activation reduces the expression of the hepatic transcription factors Hnf1α, Hnf4α, and C/ebpα while increasing the expression of Hnf1β in these hepatoblasts, promoting cholangiocyte differentiation (Tanimizu and Miyajima, 2004). Notch activation also induces expression of Sox9 (Zong et al., 2009), which promotes a cholangiocyte fate by repressing Cebp/α (Antoniou et al., 2009). This results in increased levels of the transcription factor Hnf6 (Shiojiri et al., 2004; Yamasaki et al., 2006) which is upstream of Hnf1β, and are both important for cholangiocyte differentiation. Both Hnf6−/− and Hnf1β−/− mice have fewer IHBDs compared to their wildtype counterparts (Clotman et al., 2002; Coffinier et al., 2002), demonstrating they are required for proper cholangiocyte differentiation and IHBD formation. Additionally, Hnf6, and its homolog OneCut2 (Oc2) regulate cholangiocyte development by creating a TGFβ signaling gradient, with TGFβ signaling being high near the portal vein. Upon deletion of both Hnf6 and Oc2, the TGFβ gradient in the developing liver is disrupted, resulting in hybrid cells that have characteristics of both hepatocytes and cholangiocytes (Clotman et al., 2005). These studies highlight the bipotential nature of hepatoblasts and how the delicate balance of transcription factor expression dictates a cholangiocyte or hepatocyte fate. Understanding the role of TBX3 in influencing these fate decisions in humans is important for improving protocols to drive PSCs towards a cholangiocyte fate.

Preliminary data show that TBX3 mutant lines have a slight increase in levels of cholangiocyte markers SOX9 and CK19 during a second maturation phase described in
Figure 4.1: Cholangiocyte marker expression in \(TBX3\) mutant PSCs

(A) Schematic representation of the hepatocyte differentiation protocol with second hepatocyte maturation phase.

(B) Time-course analysis of cholangiocyte markers \(SOX9\) and \(CK19\) by qRT-PCR at the end of the hepatocyte differentiation in wildtype (blue) and \(TBX3\) knockout (red) PSCs. Created with BioRender.com
the original hepatocyte differentiation protocol (Figure 4.1A-B). This suggests that not only does the loss of TBX3 impair hepatocyte differentiation, but may also enhance cholangiocyte differentiation. An explanation for the subtle increase in cholangiocyte markers is that the differentiation conditions support hepatocyte maturation rather than cholangiocyte maturation. To fully address whether cells lacking TBX3 adopt a cholangiocyte fate over a hepatocyte fate, wildtype and TBX3 mutant PSCs need to be differentiated to cholangiocytes using established protocols (Ogawa et al., 2015; Sampaziotis et al., 2015). Cells would be collected at different time points during the differentiation and assayed for cholangiocyte marker expression, including SOX9, CK19, and NOTCH2. End-stage cholangiocytes would be assayed for maturation markers, morphology, and functionality to determine if a lack of TBX3 can enhance cholangiocyte differentiation. This would have potential biological and clinical implications. First, these experiments would provide further insight into how TBX3 influences hepatoblast fate decisions during human liver development. Secondly, it may allow for enhanced generation of PSC-derived cholangiocytes that can be used for disease modelling, drug discovery and development, or for use in therapeutic contexts such as cystic fibrosis treatment.

Tbx3 maintains expression of hepatic transcription factors Hnf4α and Cebpα, allowing for hepatoblast proliferation and delamination via an epithelial-mesenchymal (EMT) like process. Conversely, Tbx3-deficient cells express cholangiocyte markers Hnf6 and Hnf1β, have reduced proliferation and do not delaminate, and differentiate to cholangiocytes (Lüdtke et al., 2009). It remains unclear whether Tbx3 promotes a hepatocyte fate over a cholangiocyte fate by regulating EMT. Further studies to
characterize the relationship between TBX3 and EMT can provide insight into the mechanism of how it regulates fate decisions during development.

4.3 TBX3 and epithelial-mesenchymal transition in pancreas differentiation

Our RNA-sequencing analysis revealed that in TBX3 mutant cells, genes involved in epithelial-mesenchymal transition (EMT) were commonly dysregulated in both the liver and pancreas differentiations. This suggests that TBX3 may have a common role in regulating EMT in during these developmental processes.

Briefly, EMT is a process by which epithelial cells lose their classic characteristics and adopt properties of mesenchymal cells. Upon EMT induction, epithelial cells lose their apical-basal polarity and cellular junctions that hold the cells together. Downregulation of epithelial cadherin (E-cadherin) is a hallmark of EMT. E-cadherin repression is mediated by several transcription factors including SNAIL, SLUG, and TWIST. As EMT progresses, the cells take on the spindle-shaped morphology of mesenchymal cells, and begin to express mesenchymal markers such as neural cadherin (N-cadherin) and vimentin. The underlying basement membrane degrades and the extracellular matrix (ECM) is remodeled, allowing cells to migrate into the surrounding stroma. EMT is involved in many developmental processes, as well as in wound healing and cancer progression (Dongre and Weinberg, 2019).

In liver development, EMT occurs during liver bud formation and hepatoblast migration. As the liver bud thickens, the basement membrane degrades and hepatoblasts migrate into the surrounding mesenchyme. This process is controlled by several transcription factors including Tbx3. Tbx3\(^{-/}\) embryos fail to downregulate E-cadherin and the basement membrane does not degrade, preventing hepatoblast migration, and thus liver
bud expansion (Lüdtke et al., 2009). The fact that EMT was dysregulated in our *TBX3* mutant hepatoblasts correlates with the phenotype observed in *Tbx3<sup>−/−</sup>* mouse embryos, providing further evidence that TBX3 acts similarly during liver development in both species. Additionally, TBX3 has been shown to advance tumor invasiveness in several types of cancer by directly regulating genes involved in EMT (Krstic et al., 2016; Rodriguez et al., 2008). TBX3 is clearly involved in EMT regulation in many different biological contexts, raising the possibility that *TBX3* may also facilitate EMT during pancreas development.

An EMT-like process is thought to mediate endocrine cell delamination during pancreas development. As endocrine progenitors differentiate and mature, they delaminate from the epithelial branches and migrate into the surrounding mesenchyme to form pancreatic islets (Gouzi et al., 2011; Rukstalis and Habener, 2007). Additionally, the actin cytoskeleton plays an important role in pancreas development. Blocking F-actin disassembly prevents β-cell delamination from the trunk epithelium, and impairs β-cell function and development (Kesavan et al., 2014). Cytoskeleton state also mediates differentiation earlier in pancreas development during differentiation to either a ductal or endocrine cell fate (Mamidi et al., 2018). Recently, manipulating actin polymerization was shown to improve β-cell differentiation from PSCs, further demonstrating the relationship between cytoskeleton and differentiation (Hogrebe et al., 2020). Hogrebe and colleagues find that blocking actin polymerization prevents premature expression of the endocrine marker *NGN3*, and allows for expression of the pancreatic progenitor marker *NKX6.1*, which is important for enhancing β-cell generation from PSCs (Nostro et al., 2015; Rezania et al., 2013). Thus, the cytoskeleton influences pancreatic differentiation by regulating the timing of transcription factor expression.
Figure 4.2: Pancreatic hormone production in TBX3 mutant PSCs

(A) Analysis of β-cell markers NKX6.1 and INSULIN at Day 25 of differentiation by intracellular flow cytometry (one representative replicate shown).

(B) Quantification of mean fluorescence intensity (MFI) from flow cytometry of Insulin, Glucagon, and Somatostatin in ESC^{+/−} and ESC^{−/−} cells at Day 25 of pancreas differentiation. MFI represented as normalized to wildtype (n = 4 per line).

(C) Quantification of mean fluorescence intensity (MFI) from flow cytometry of C-peptide, Glucagon, and Somatostatin in iPSC^{+/−} and iPSC^{−/−} cells at Day 25 of pancreas differentiation. MFI represented as normalized to wildtype (n = 4 per line).
Our findings suggest that the loss of TBX3 enhances pancreatic progenitor generation from PSCs. We postulate that this likely due to de-repression of pancreatic genes rather than changes in cytoskeletal state. However, preliminary data suggest that upon further differentiation, TBX3 mutant PSCs generate very few insulin-producing cells (Figure 4.2A). There were lower levels of glucagon (α-cells) and somatostatin (δ-cells) in TBX3 mutant PSCs as well (Figure 4.2B-C), indicating a failure in differentiation past the pancreatic progenitor stage. This raises the possibility that TBX3 plays a role later in pancreas development and differentiation. Because our RNA-sequencing data show that genes involved in EMT are dysregulated in the TBX3 mutant line during the pancreas differentiation, and since TBX3 is known to regulate EMT, we suggest a novel role for TBX3 in regulating EMT and/or cytoskeleton state during endocrine cell delamination from the pancreatic epithelium.

We propose a model in which TBX3 has a biphasic role depending on the stage of pancreas development. This phenomenon is quite common for transcription factors, especially during pancreas development. Early in development, TBX3 plays an antagonistic role, skewing the cells toward a hepatic lineage and away from a pancreatic fate (Figure 4.3A). Later, TBX3 may be required for endocrine cell delamination and differentiation by mediating EMT and the cytoskeleton state (Figure 4.3B). Although Tbx3 expression has been detected in the developing mouse pancreas, it is mainly restricted to the mesenchyme surrounding the pancreatic epithelium, and its function is unknown (Begum and Papaioannou, 2011; Zhou et al., 2007). Interestingly, Tbx2, but not Tbx3, is expressed in pancreatic endocrine tissue in the mouse (Begum and Papaioannou, 2011). In human embryos, TBX3 is expressed in SOX9+/PTF1A+ tip progenitors, which give rise to exocrine component of the pancreas (Villani et al., 2019).
Figure 4.3 Proposed biphasic role for TBX3 during pancreas development

(A) Early in development, TBX3 plays an inhibitory role by repressing pancreatic gene expression, thus skewing cells towards a hepatic fate and away from a pancreatic fate. (B) Later, TBX3 may function in endocrine cell delamination by regulating epithelial-mesenchymal transition, presenting a novel role for TBX3 in pancreas development. Created with BioRender.com

This is not seen in the mouse, suggesting a possible species-specific difference in TBX3 expression and function during pancreas development. Alternatively, this may be an example of distinct T-box family members being used for the same function in different species, with Tbx2 acting in mouse and TBX3 acting in human pancreas development.

Because TBX3 is present in pancreatic tip progenitors, it may also be present in trunk cells of the pancreatic epithelium, which gives rise to endocrine progenitors. If this is the case, TBX3 may be mediating EMT in endocrine cell differentiation and delamination from the trunk during islet formation. Further characterization of TBX3 expression in the developing human pancreas is needed. Since TBX3 mutant PSCs differentiate poorly to β-cells, further studies to determine a link between TBX3, EMT, and/or the cytoskeletal state are also required. These studies would investigate how this relationship may
impact pancreas differentiation, and provide insight into methods of improving both pancreatic progenitor and β-cell generation from PSCs.

4.4 Applications for in vitro differentiation of PSCs to pancreatic β-cells

Generating pancreatic β-cells from PSCs is an area of great interest. PSC-derived β-cells are immensely advantageous not only for studying β-cell biology, but also for clinical applications such as disease modelling and treating diabetic patients. Differentiation protocols have been continuously improving over time, bringing the field closer to generating functionally mature β-cells from PSCs. Our studies contribute to this goal by examining how TBX3, which has not been previously studied in pancreas differentiation, impacts pancreatic progenitor and β-cell differentiation from PSCs.

We propose that TBX3 expression may be used as an indicator for how well a particular PSC line differentiates to pancreatic progenitors. It is common that different PSC lines have variable differentiation efficiencies even when using the same protocols. This can be attributed to differences in genetic background such as endogenous gene expression and signaling or DNA methylation state (Ortmann and Vallier, 2017). Differentiation efficiencies are particularly variable when generating β-cells from patient-derived iPSCs, presenting a significant challenge in using these lines for disease modelling or therapeutic purposes. Levels of endogenous TBX3 can vary amongst PSC lines, which may impact their ability to differentiate to pancreatic progenitors. Because TBX3 promotes hepatic gene expression, PSC lines with high levels of TBX3 may differentiate less efficiently to pancreatic lineages. Thus, lowering endogenous TBX3 expression by manipulating signaling pathways may help improve pancreas differentiation efficiency, especially in patient-specific iPSC lines.
Tbx3 is a known downstream target of the Wnt/β-catenin pathway (Eblaghie et al., 2004; Renard et al., 2007), which plays an inhibitory role during early pancreas development (McLIn et al., 2007). The pancreas protocol utilizes a WNT inhibitor early in the differentiation to mimic gut tube anteriorization, but it is removed in later stages (Rezania et al., 2014). Keeping the WNT inhibitor in the differentiation medium longer until the pancreatic progenitor 1 (PP1) stage may reduce TBX3 levels, and improve pancreatic progenitor differentiation. However, in our differentiations, we see that TBX3 expression is high, even in the presence of the WNT inhibitor, and TBX3 levels decline after the WNT inhibitor is removed. This suggests that TBX3 expression is not under the control of WNT signaling in this context, and additional manipulations may reveal a mechanism to reduce endogenous TBX3 levels to improve pancreatic differentiation efficiency.

The BMP pathway is another a promising candidate for controlling TBX3 levels during the pancreas differentiation. Tbx3 has been implicated as a downstream target of the Bmp pathway in several developmental processes (Tumpel et al., 2002; Yang et al., 2006). Furthermore, the Bmp pathway inhibits pancreatic specification while promoting hepatic specification (Rossi et al., 2001). The protocol uses a BMP inhibitor to drive the cells to pancreatic progenitor stages (Rezania et al., 2014). Interestingly, TBX3 expression levels peak at the gut tube stage, and are drastically reduced upon addition of the BMP inhibitor to the differentiation medium. In lines with poor pancreas differentiation due to high endogenous TBX3 levels, adding the BMP inhibitor earlier in the differentiation protocol may reduce TBX3 levels prior to the gut tube stage and help skew the cells towards a pancreatic lineage, thus improving differentiation to the pancreatic progenitor stage.
Additionally, *TBX3* expression levels can be used as a readout while testing various small molecules to improve pancreas differentiation. *TBX3* expression is not only controlled by the WNT/β-catenin and BMP pathways, but is also under the influence of sonic hedgehog (Lüdtke et al., 2016) and retinoic acid signaling (Ballim et al., 2012). All four of these signaling pathways are involved in pancreas development, thus manipulating their activity could help to further improve differentiation protocols. Changes that result in lower levels of *TBX3* expression can help determine which conditions are ideal for pancreatic differentiation. Controlling the timing and levels of signaling pathway activity to lower *TBX3* levels is a promising and relatively straightforward method for enhancing pancreas differentiation from PSCs. Being able to generate a large number of pure pancreatic progenitors is important because it improves the chances of producing mature and functional β-cells from PSCs. Moreover, manipulating signaling pathways can reduce *TBX3* levels without disrupting the gene, bypassing any potential issues if TBX3 is needed later in pancreatic development.

4.5 Summary and conclusions

In summary, this thesis explores the role of TBX3 in human liver development. Using *TBX3* mutant PSC lines, we found that TBX3 is required for proper hepatocyte differentiation by maintaining hepatic gene expression, and influencing EMT. This is similar to Tbx3 function in mouse liver development, highlighting the conserved role between species. Surprisingly, we found that the loss of *TBX3* results in reduced hepatic gene expression and increased expression of pancreatic genes in the context of both hepatocyte and pancreas differentiation protocols (Figure 4.4A). This suggests that TBX3 may control liver versus pancreas fate by repressing pancreatic gene expression, uncovering a novel role for TBX3 during endoderm patterning and differentiation. We
also found that \textit{TBX3} mutant PSCs generate more pancreatic progenitors compared to wildtype PSCs, further implying that TBX3 hinders pancreas differentiation. However, \textit{TBX3} mutant PSCs differentiate less efficiently to $\beta$-cells, suggesting that it may be required later in differentiation (Figure 4.4B). Further studies into the mechanism of how TBX3 impacts pancreas differentiation and development are needed. One possibility is that TBX3 mediates EMT or influences the cytoskeletal state during development. Reducing \textit{TBX3} levels may be one way to improve both the quality and quantity of PSC-derived pancreatic progenitors, which can then be differentiated to $\beta$-cells. Generating functionally mature $\beta$-cells from PSCs has been an important goal in the pancreas biology field. Existing protocols have been continuously evolving, yet there are still improvements that can be made. Not only does the work in this thesis uncover a novel role for TBX3 in endoderm development, it also provides a new method for improving PSC-derived pancreatic progenitors, which can then be further differentiated to $\beta$-cells for use in therapeutic contexts.
Figure 4.4: Summary of loss of TBX3 in β-cell differentiation from PSCs

(A) High TBX3 levels promotes increased hepatic gene expression during the pancreatic differentiation. (B) Loss of TBX3 enhances pancreatic progenitor generation (PP2, purple cell). However, TBX3 may be required for β-cell differentiation, as TBX3 knockout cells generate β-cells less efficiently. Wildtype conditions are shown in black arrows, and effects of TBX3 knockout are shown in red arrows. Thickness of arrows indicates differentiation efficiency, with thicker arrows meaning increased efficiency. Created with BioRender.com


Collombat, P., Hecksher-Sørensen, J., Broccoli, V., Krull, J., Ponte, I., Mundiger, T.,


coverage, classification and access to protein sequence annotations. Nucleic Acids Res. 47, D351–D360.


progenitors from multiple human pluripotent stem cell lines. Stem Cell Reports 4, 591–604.


Nkx6.1 lies downstream of Nkx2.2 in the major pathway of β-cell formation in the pancreas. Development 127, 5533–5540.


Disease-Related Gene in Human Induced Pluripotent Stem and Embryonic Stem Cells. Cell Stem Cell 5, 97–110.