Gene-editing of the monogenic diabetes associated gene *hnf1ba* to model multisystemic MODY5 disease mechanisms in zebrafish (DR)

Rosemary Hoff

This thesis is submitted in partial fulfillment of the requirements for the degree of Master of Science

Department of Biological Science

University of Bergen

December 2021.

Acknowledgment

The work presented in this thesis was conducted at the Department of Biological Science at University in Bergen (UiB). This thesis was a part of the project; Gene-editing monogenic diabetes genes to decipher disease mechanisms in multiple organ systems, led by prof. Lise Bjørkhaug Gundersen.

First and foremost, I would like to express my deepest gratitude to my main supervisors Prof. Ståle Ellingsen and Prof. Lise Bjørkhaug Gundersen for providing this project and for making this year a fantastic experience. Through your excellent supervision, patience, and guidance I have become a more independent student, and I have enjoyed every single day working on this project.

To my co-supervisor and lab-hero, Dr. Elsa Denker, I sincerely thank you for every single (tiny or big) thing you have done for me, or with me, throughout this year. For your great counselling and sharing of knowledge, it has truly been a pleasure to work with you.

Further, I would like to thank Prof. Jon Vidar Helvik and the rest of his lab for making me feel welcome, and for teaching me techniques and providing chemicals. Especially, Phd student Christine Horne for your devotion and interest in my work.

Additionally, I would like to thank the team for making this a great place to be, and to Morten Barvik for your friendship during countless hours of studying.

To my family and friends, thank you for believing in me and cheering me on. And finally, to my rock in life, Andreas, thank you for holding the fort and keeping my two boys happy and fed. I could not have done this without your endless love and support, and I am eternally grateful.

Bergen, December 2021

Rosemary Hoff

Table of Contents

Acknowledgment 1
Abbreviations 5
Nomenclature 6
Abstract 7
1 **Introduction** 8
1.1 Diabetes mellitus 8
1.1.1 Pancreas function 9
1.2 Classification of diabetes 11
1.2.1 Type 1 Diabetes 12
1.2.2 Type 2 Diabetes 12
1.2.3 Monogenic diabetes by MODY 13
1.3 MODY5 15
1.3.1 HNF1B (Homo Sapiens) 16
1.3.2 Animal models targeting *HNF1B* 19
1.4 Zebrafish (Danio Rerio) 20
1.5 CRISPR-Cas9 21
1.6 Aim of the study 22
2 **Materials and Methods** 23
2.1 Ethics Statement and zebrafish handling 23
2.2 Preparation of the samples 23
2.2.1 Fixation of 24 hpf embryos or 96 hpf larvae for whole-mount analyses 23
2.2.2 Fixation of 96 hpf embryos for analyses on cryosections 24
2.3 *In Situ* Hybridization 25
2.3.1 RNA probe preparation for ISH utilizing two approaches 25
In Situ Hybridization Protocols 33
2.3.2 ISH Protocol for Whole-mount 33
2.3.3 Cryosections ISH Protocol 35
2.4 Immunohistochemistry 37
2.4.1 Whole-Mount IHC Protocol 38
2.4.2 Cryosections IHC Protocol 39
2.5 Imaging and image processing 40
2.5.1 Zeiss Lumar.V12 Stereo microscopy 40
2.5.2 Leica fluorescence microscopy 40

2.5.3 Zeiss Axio Scan.Z1 Slide scanner 40
2.5.4 Olympus FV3000 Fluorescence Confocal Microscopy 41
2.5.5 ImageJ software 42
2.5.6 Adobe Photoshop software 42
2.6 Quantitative image analysis 42
2.6.1 ImageJ 42
2.6.2 Statistical analysis 43
2.7 CRISPR-Cas9 and Genotyping 43
2.7.1 CRISPR construct design 43
2.7.2 sgRNA generation 43
2.7.3 Microinjection 46
2.7.4 DNA extraction for genotyping 47
2.7.5 Genotyping and T7 mutation screen 48
3 **Results** 51
3.1 Multiple protein sequence alignment of HNFs 51
3.2 Developing kidneys and pancreas in 24 hpf transgenic zebrafish line Tg(wt1b:EGFP) 52
3.2.1 EGFP expression is localized in pronephros in 24 hpf zebrafish larvae 53
3.2.2 Pancreatic progenitors in 24 hpf zebrafish larvae 54
3.2.3 Localization pattern of *hnf1ba* expression in 24 hpf zebrafish larvae 55
3.3 Characterization of pancreas in 96 hpf WT zebrafish larvae 56
3.3.1 Localization of *hnf1ba* and *gcga* expression in 96 hpf zebrafish larvae 56
3.3.2 Localization of pancreas in 96 hpf transgenic zebrafish Tg(wt1b:EGFP) cross section 58
3.3.3 Pancreas in 96 hpf transgenic zebrafish Tg(wt1b:EGFP) whole-mount 59
3.4 Gene editing of the *hnf1ba* gene in zebrafish 64
3.4.1 Mutation screening of injected individuals by T7 assay 65
3.4.2 Characterization of mutation efficiency in injected embryos 66
3.5 Analysis of zebrafish CRISPR injected larvae (F0 generation) 67
3.5.1 Indels formed in microinjected individuals 67
3.5.2 Characterization of the endocrine pancreas in mosaic 96 hpf CRISPR zebrafish larvae shows abnormal pancreas development 68
3.5.3 Statistical analysis of alpha- and beta-cell numbers indicate significant pancreas hypoplasia in *hnf1ba* CRISPR injected zebrafish 71
4 **Discussion** 73
4.1 HNF protein domains and MODY associated mutations 73
4.2 Characterization of developing kidneys and pancreas in WT 24 hpf zebrafish 74

4.2.1 Transgenic EGFP in zebrafish pronephros at 24 hpf, a valuable tool for future *hnf1b* knockout characterization in zebrafish ... 74
4.2.2 Pancreatic progenitors expressing insulin and glucagon in 24 hpf zebrafish embryo ... 74
4.2.3 Localization pattern of *hnf1ba* expression in 24 hpf zebrafish embryo ... 75
4.3 Characterization of developing kidneys and pancreas in WT 96 hpf zebrafish ... 76
4.3.1 Localization of *hnf1ba* and *gcga* expression in 96 hpf zebrafish larvae ... 76
4.3.2 Cross section of 96 hpf zebrafish shows exocrine and endocrine pancreas structure ... 76
4.3.3 Characterization of endocrine pancreas in whole-mount 96 hpf zebrafish ... 77
4.4 Gene editing the *hnf1ba* gene in zebrafish ... 78
4.4.1 Mutation screening of injected individuals by T7 assay ... 78
4.4.2 Characterization of mutation efficiency and type in CRISPR injected embryos ... 78
4.5 Phenotype analysis of 96 hpf zebrafish CRISPR injected larvae ... 80
(F0-generation) ... 80
4.5.1 CRISPR-Cas9 induced mutations show endocrine pancreas hypoplasia in mosaic 96 hpf zebrafish larvae ... 80
5 **Conclusion and future aspects** ... 82
5.1 Future aspects ... 83
6 References ... 84
7 Appendix ... 89

Abbreviations

DR – Danio Rerio

GCG (HS) – Glucagon (Human) Protein

Gcga/Gcgb (DR) – Glucagon paralog A/B (zebrafish) Protein

HNF1A - Hepatocyte Nuclear Factor 1 Homeobox A (Human) Gene

HNF1A - Hepatocyte Nuclear Factor 1 Alpha (Human) Protein

HNF4A - Hepatocyte Nuclear Factor 4 Homeobox A (Human) Gene

HNF4A - Hepatocyte Nuclear Factor 4 Alpha (Human) Protein

HNF1B (HS) – Hepatocyte Nuclear Factor 1 Homeobox B (Human) Gene

HNF1B (HS) – Hepatocyte Nuclear Factor 1 Beta (Human) Protein

hnf1ba (DR) - Hepatocyte nuclear factor 1 beta paralog A (zebrafish) Gene

Hnf1ba (DR) – Hepatocyte nuclear factor 1 beta paralog A (zebrafish) Protein

Hnf1bb (DR) – Hepatocyte nuclear factor 1 beta paralog B (zebrafish) Protein

HS – Homo Sapiens

Hpf – Hours Post Fertilization

IHC - Immunohistochemistry

INS (HS) – Insulin (Human) Protein

ISH – *In situ* hybridization

MODY5 – Maturity Onset Diabetes of the Young type 5

O/N – Overnight incubation

PA – Pancreas

PH/PB/PT – Pancreas Head, Body, Tail

PCT/PST – Proximal Convoluted/Straight Tubules

Tg(wt1b:EGFP) – Transgenic Wilms Tumor 1 b Enhanced Green Fluorescent Protein

Nomenclature

This thesis will abbreviate all proteins and genes according to ordinary rules for nomenclature regarding the species mammals, murines and zebrafish (ncbi.nlm.nih.gov & zfin.org). All proteins regarding humans will be written in capital letters and their respective genes will be written in italic capital letters. All proteins regarding murines will have the first letter in capital and their respective genes written in italic in addition. Lastly, proteins regarding zebrafish will be abbreviated first letter capital, and their respective genes in lower case italic letters.

Abstract

Maturity-Onset Diabetes of the Young (MODY) is a rare form of monogenic diabetes, characterized by autosomal dominant inheritance, non-obesity, and early-onset diabetes due to reduced pancreatic beta-cell secretion. Precision medicine increasingly relies on an accurate interpretation of the consequence of genetic mutations in diseases like MODY. Clinical investigations of patients with MODY5 associated mutations in the gene encoding hepatocyte nuclear factor 1-beta (*HNF1B*), have revealed disease in not only the pancreas, but one or more additional organ systems, including the kidney, liver, and the reproductive system.

In this thesis, using CRISPR-Cas9 mediated knockout technology, we targeted the paralog of the human *HNF1B, hnf1ba* in zebrafish. To understand the expression patterns of *hnf1ba* in wild-types, we performed a spatial localization using *in situ* hybridization conducted on 24 and 96 hpf larvae and confirmed that *hnf1ba* is expressed in the pancreas and pronephros (early kidneys. Triple immuno- histochemistry was used to explore the two predominating cell types in the endocrine pancreas, the insulin-producing beta-cells and the glucagon-producing alpha-cells, as well as the exocrine pancreas targeted by EGFP in the Tg(wt1b:EGFP) zebrafish line. The results of the pancreas immunostaining showed a significant reduction in both beta- and alpha-cell numbers in CRISPR-Cas9 injected zebrafish larvae.

Further studies are needed to truly understand the vast scope of the transcription factor HNF1B, and its molecular mechanisms connected to MODY5 disease. This thesis provides several avenues for further research regarding the establishment of zebrafish as a stable MODY5 model.

1 Introduction

1.1 Diabetes mellitus

According to the World Health Organization (2021), recent global estimates indicate that 1 in 11 adults suffer from diabetes mellitus, which accounts for approximately 463 million people worldwide. In 2020, 4.3 million lives were lost to the silent pandemic that involves diabetes. Diabetes not only accounts for the top 10 causes of deaths globally but also has no disregard for socio-economic status nor national boundaries. It also increases the risk of patients developing life-threatening complications. Despite the rising numbers of diabetes cases each year, there is a positive message: With early diagnosis and appropriate treatment and care, diabetes can be managed, and its complications prevented.

Diabetes is characterized by long-term elevated glucose levels in the blood circulation, resulting in chronic hyperglycemia. The pathophysiology of diabetes often involves insufficient production and secretion of insulin from the pancreatic beta-cells, or a defect in insulin action where body tissues and cells show a decreased response to insulin. Insulin is an anabolic hormone essential for maintaining normal levels of circulating glucose and is essential for the regulation of carbohydrate, lipid, and protein metabolism (Wilcox, 2005). The physiological blood glucose concentrations in healthy individuals remain stable (within 4-8 mmol/l after a meal). A patient diagnosed with diabetes will have elevated plasma glucose concentrations ≥ 11.1 mmol/l, or a fasting glucose concentration of ≥ 7.0 mmol/l. A diagnosis can also be set if plasma glucose concentrations are ≥ 11.1 mmol/l two hours after an oral glucose tolerance test (WHO, 2006, IDF).

The resulting long-term effects of diabetes can be severe, and if left undiagnosed or untreated, can result in acute and life-threatening diseases such as cardiovascular diseases, ketoacidosis, or nonketotic hyperosmolar syndrome (WHO, 2021). Further, diabetes can lead to dysfunction or damage of vital organs like the kidneys, eyes, nerves, heart, and blood vessels. Thus, the importance of an early and precise diagnosis is essential for providing the proper treatment in order to delay or prevent these serious long-term complications, and ultimately improve the patient's quality of life.

1.1.1 Pancreas function

The pancreas is a digestive gland located posteriorly to the stomach, between the duodenum and the spleen (See Figure 1). It consists of two functional compartments/tissues called the endocrine and exocrine pancreas. The human endocrine pancreas consists of approx. 1 million cell clusters scattered throughout the pancreas known as the islets of Langerhans. Each islet consists of four major cell types that are responsible for producing the following hormones; alpha-cells, which produce glucagon (35 %), beta-cells, which produce insulin (54 %), delta-cells, which produce somatostatin, and gamma-cells, which produce pancreatic polypeptides (11 %) (Lawlor et al., 2017).

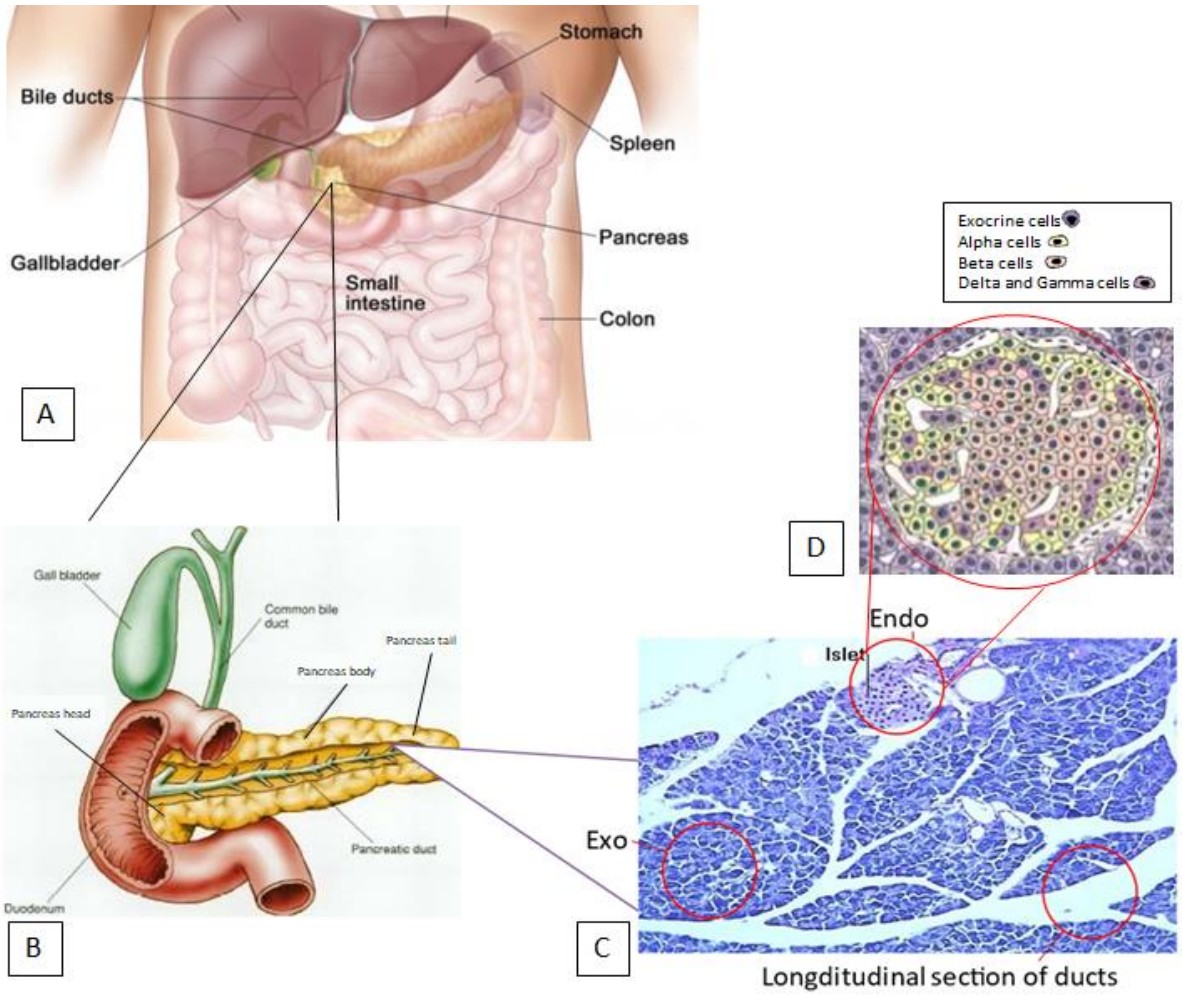


Figure 1: Anatomical illustration of pancreas. A: Location of pancreas in the abdominal cavity, situated under the liver and stomach, in near proximity to the gallbladder and small intestine. **B:** The pancreas anatomy is divided into parts, namely pancreas head, body, and tail. The pancreatic duct goes through the organ leading towards the common bile duct. **C:** Longitudinal section of pancreatic tissue stained with hematoxylin. Exocrine as well as endocrine cells are circled in red. **D:** The different cells of the endocrine islet of Langerhans are illustrated in colors. The cells are arranged in an established pattern with the beta cells (pink) in the center surrounded by alpha (yellow) and delta (light purple) cells. The islets of Langerhans are defined and limited by exocrine acinar cells (purple). Image adapted from Lise Bjørkhaug Gundersen.

Hormone secreting cells (or islet clusters) make up only 1-2% of the mass of the pancreas. The exocrine pancreas constitutes the major part of the organ and contains acinar- and ductal-cells producing and secreting digestive enzymes and bicarbonate ions, respectively. These secretions are released into small ducts that empty into the pancreatic duct which leads to the small intestine. Thus, both the pancreatic endocrine and exocrine gland functions in the digestive system.

Because glucose is a major fuel for cellular respiration and a key source of carbon skeletons for biosynthesis, maintaining blood glucose concentrations near-optimal setpoint is a critical bioenergetic and homeostatic function. An optimal metabolic balance depends on blood glucose concentrations regulated by the two antagonistic hormones insulin and glucagon. When blood glucose rises above the stable equilibrium, the release of insulin from pancreatic beta-cells triggers uptake of glucose from the blood, decreasing the overall blood glucose concentration (See Figure 2). When blood glucose drops below the optimal setpoint, the release of glucagon from pancreatic alpha-cells promotes the release of glucose from the liver into the blood, increasing the overall blood glucose concentration. This mechanism involves a negative feedback loop regulated by the two endocrine hormones (and involving the pancreas/liver). Due to their opposing effects, the combined activity of these two hormones tightly controls glucose homeostasis

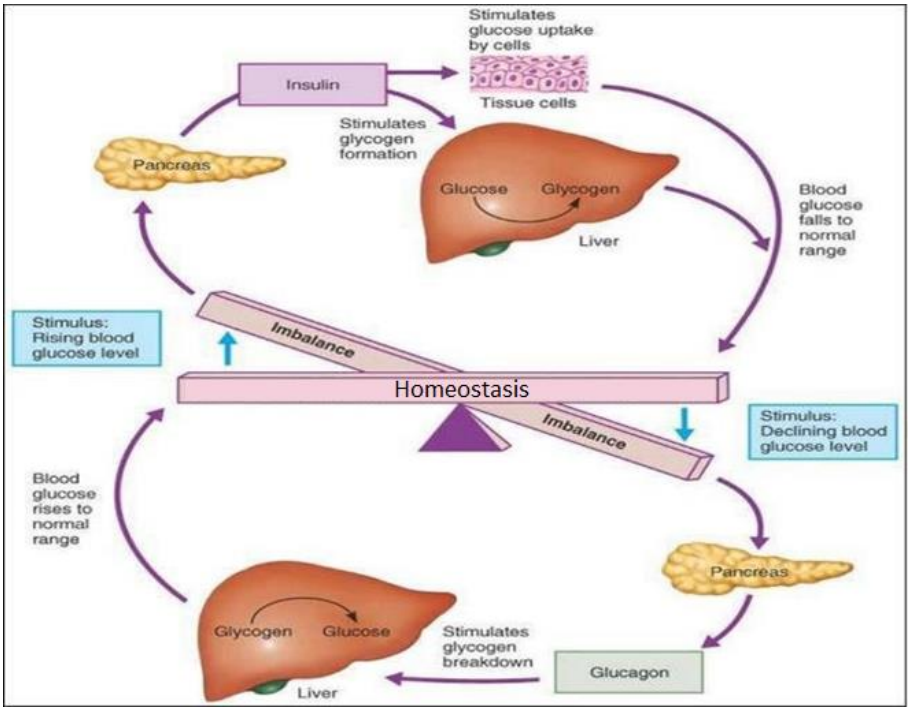


Figure 2:
Glucose homeostasis.
Illustration of normal negative feedback loop concerning blood glucose level maintenance, by the liver and pancreas. When glucose levels drop, glucagon is released from the liver, and when glucose levels rise, insulin is released from the pancreas. Illustration adapted from Pearson Education, Inc. (2015).

1.2 Classification of diabetes

Diabetes, caused by a deficiency of insulin or a decreased response to insulin, is a complex disease that presents in numerous forms. Hence, providing a precise diagnosis can be comprehensive. Different subgroups exist depending on whether diabetes is related to genetic factors, or triggered by environmental and lifestyle-related factors (Molven & Njolstad, 2011). The most prevalent forms of diabetes include type 1 diabetes (T1D) and type 2 diabetes (T2D). Worldwide, T1D accounts for approx. 5-10 % of all diabetes cases while T2D accounts for nearly 90 % of all cases. Other forms of diabetes also exist, including gestational diabetes, monogenic diabetes, and other disease-associated forms of diabetes (secondary forms), as illustrated in Figure 3. Of the monogenic forms of diabetes, Maturity-Onset Diabetes of the Young (MODY) is the most common and accounts for approx. 1-2 % of all diabetes worldwide (Nkonge et al., 2020). Other specific forms of diabetes also caused by monogenic mutations include for instance mitochondrial diabetes and neonatal diabetes (Rubio-Cabezas et al., 2014; Yamagata et al., 1996).

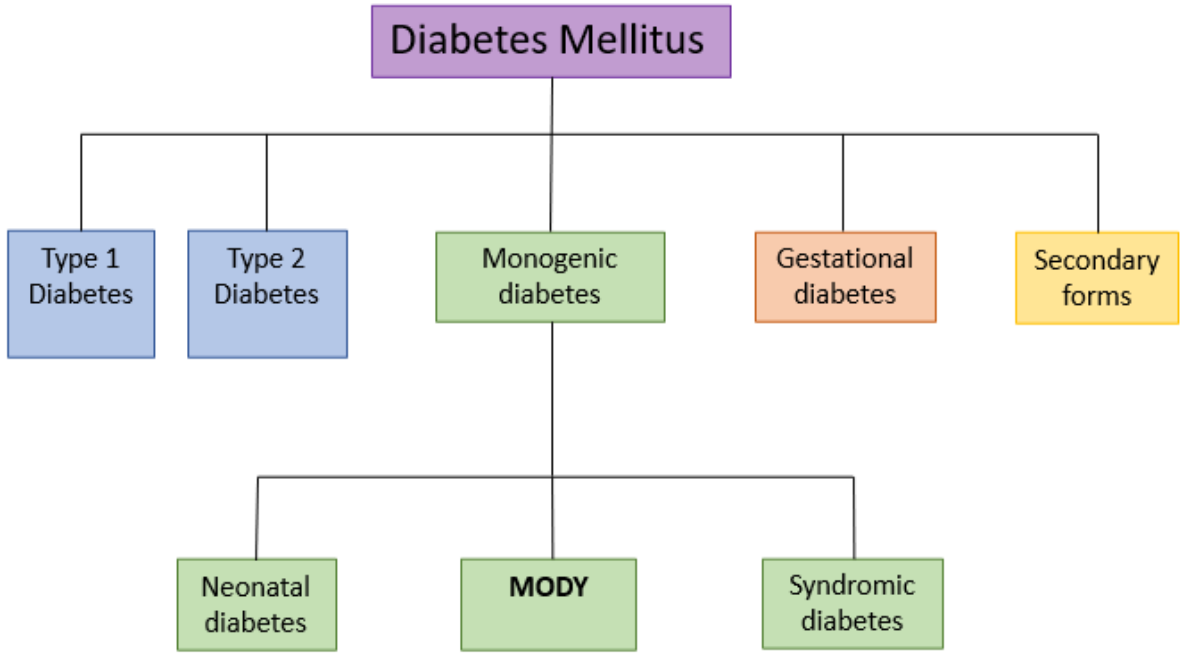


Figure 3: Illustration of the different types of diabetes. Schematic presentation of the different subgroups of diabetes mellitus.

1.2.1 Type 1 Diabetes

Type 1 diabetes, T1D is a chronic autoimmune disease characterized by loss of insulin-producing pancreatic beta-cells. Patients with T1D are therefore dependent on insulin treatment as progressive pancreatic beta-cell destruction leads to insulin deficiency. T1D linked to genetic causes is often associated with DQ8 and HLA class II genes that control the immune responses and more than 40 non-HLA susceptibility gene markers have so far been identified (Eisenbarth, 2007), (Steck & Rewers, 2011). Viral and environmental factors also play a role by either acting as a direct trigger of pancreatic islet destruction or by triggering the immune system for auto-destruction of the islets (general and systematic pathology, 5. ed 2009). Disease debut is often of early-onset, where it predominantly occurs in children and young adults (<40 years of age), and most patients are dependent on exogenous insulin injections for survival (Atkinson et al., 2014), (Harjutsalo et al., 2010). T1D is a result of a complex interaction between genetic and environmental factors and this type of diabetes is so far not preventable (WHO, 2019).

1.2.2 Type 2 Diabetes

Type 2 diabetes, T2D, the most common form of all subtypes. It has had a dramatic increase in incidences worldwide in the last 30 years and is continuing to rise (fn.no). T2D is characterized by insulin resistance or impaired insulin secretion, or both. At an early stage of insulin resistance, the pancreatic beta-cells respond by increasing insulin production and secretion. If the patient is unaware of the chronic compensation for a long period of time, the insulin maintenance fails, leading to progressive hyperglycemia and eventually diabetes (DeFronzo, 1992), (Olokoba et al., 2012). As a result, patients can go undiagnosed for many years, which increases the risk of developing late diabetic complications. The disease often has late-onset (>40 years of age) due to the gradual progression of pancreatic beta-cell failure. T2D is closely linked to obesity and inactivity and can thus often be managed by diet and physical activity alone. However, some patients also need medication such as metformin in order to decrease the amount of glucose produced by the liver (ADA, 2021). Hereditary predisposition to T2D has been linked to >100 genetic loci, identified through genome-wide association study (GWAS), and is shown to have a stronger link to family lineage than T1D (Florez, 2008), and (Dean & McEntyre and ADA, 2021). Genes that have been linked to T2D include, for instance, those encoding the hepatocyte nuclear factor 4 alpha (HNF4A), the insulin hormone and receptor (INS and INSR), as well as the glucose transporter (GLUT2) and the metabolic enzyme glucokinase (GCK). These are all active components in normal liver, kidney, and pancreas function (The genetic landscape of diabetes, Dean & McEntyre, 2004).

1.2.3 Monogenic diabetes by MODY

Maturity-onset diabetes of the young, or simply MODY, is a monogenic form of diabetes caused by mutations in a single gene. MODY is an autosomal dominantly inherited disorder characterized by early-onset diabetes in young adults (<35 years of age), non-obesity, and progressive decrease in pancreatic insulin secretion (Haldorsen et al., 2008). In some cases, additional non-genetic factors that affect insulin sensitivity, such as infection, pregnancy, and in rare cases obesity, may trigger the onset of diabetes and increase the severity of hyperglycemia in patients with MODY (Fajans et al., 2001). The estimated prevalence of MODY is 108 per million cases worldwide and is considered a rare disease (Shields et al., 2010).

According to an Australian community-based study, 0.24 % of individuals with diagnosed diabetes <35 years of age had MODY, with one in four being previously undiagnosed (Davis et al., 2017). These findings are supported by another study including children <15 years of age and diagnosed with autoantibody-negativeT1D, found to carry a mutation in a MODY gene. One-third of these children were previously undiagnosed with MODY (Johansson et al., 2017). This illustrates the challenges that prevail with the high number of MODY patients being misdiagnosed with either T1D or T2D. Hence, the prevalence of MODY is expected to be higher than previously thought (Murphy et al., 2008). Another diagnostic challenge is differentiating between MODY subtypes, which is required by genetic analyses, and can be a costly affair in low-income countries (WHO). Providing a correct diagnosis of MODY subtype is important for receiving suitable genetic counseling and appropriate treatment as early as possible, to prevent developing late diabetic complications (Hattersley et al., 2009), (Sagen et al., 2002).

Genetic causes of MODY

To date, mutations in at least 14 different genes are found to cause MODY, where each affected gene defines specific MODY subtypes (MODY1-14). MODY genes play essential roles in glucose sensing and regulation of pancreatic beta-cell insulin secretion and metabolism (See Figure 4). Some also play a role in normal pancreatic development, as they are expressed during early embryonic stages of pancreas development. For instance, the genes encoding the hepatocyte nuclear transcription factors affect the expression of downstream targets, and their normal regulation of multiple genes is essential for proper glucose uptake and metabolism. Heterozygous mutations in MODY genes have shown to affect beta-cell generation, which will

further affect the normal function of the pancreatic beta-cell and lead to diabetes development (Fajans et al., 2001).

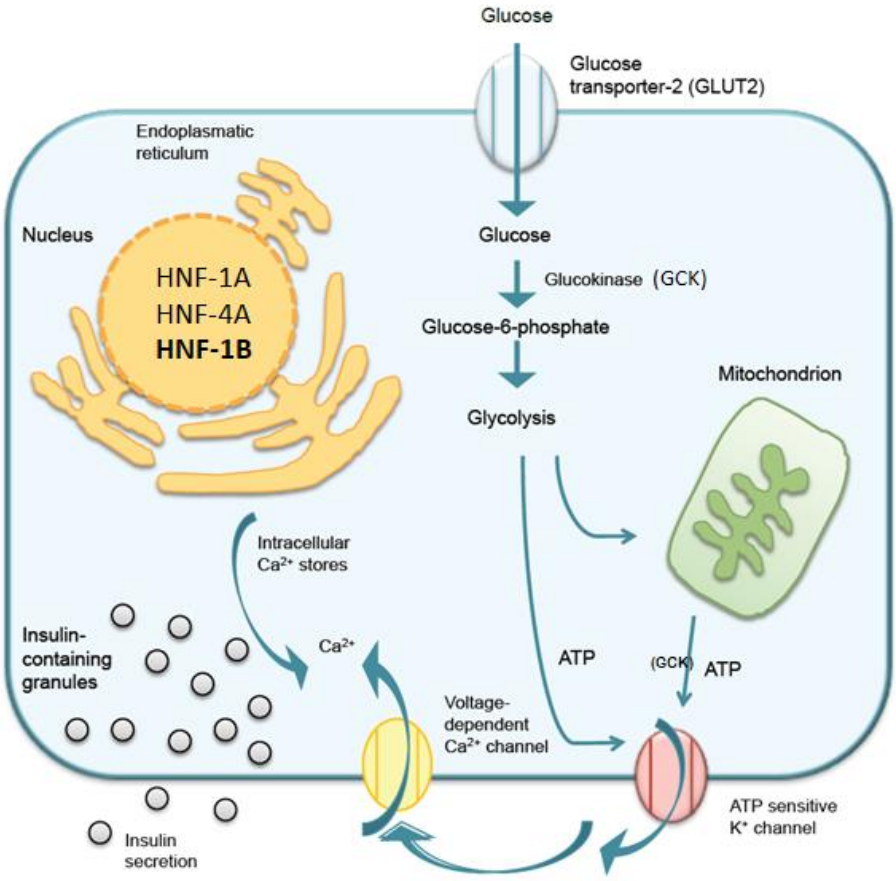


Figure 4: Schematic figure of the pancreatic beta-cell displaying some of the proteins affected in Maturity Onset Diabetes of the Young (MODY). Glucose enters the beta-cell via the glucose transporter membrane protein (GLUT2) and the "glucose sensor" of the cell, the metabolic enzyme glucokinase (GCK), phosphorylates glucose to form glucose-6-phosphate. Glucose-6-phosphate is a substrate further involved in the glycolysis and the Krebs cycle where ATP is generated. The increase in intracellular ATP closes the ATP-sensitive potassium channels, which results in a depolarization of the cell membrane. This again opens the voltage-dependent Ca2+ channels which lead to an influx of extracellular Ca2+ ions. Mobilization of intracellular Ca2+ ions result in the release of insulin-containing granules over the membrane and into the circulation where insulin again can stimulate the uptake and metabolism of glucose. This figure is adapted and modified from Fajans et al. 2001.

The most common forms of MODY are MODY3 and MODY2, accounting for 70% of all MODY cases, and are caused by mutations in the hepatocyte nuclear factor-1 alpha gene (*HNF1A* by MODY3) and the glucokinase gene (*GCK* by MODY2) (Ellard et al., 2008), (Sagen et al., 2006). Mutations in the *GCK* gene, for instance, result in impaired phosphorylation and metabolism of glucose and are associated with the mild form MODY2. Mutations in the *HNF1A* gene result in reduced insulin secretion from the pancreatic beta-cell and is the most common

form of all MODYs (Bjorkhaug et al., 2003). Mutations in the hepatocyte nuclear factor-4 alpha gene (*HNF4A* by MODY1) and hepatocyte nuclear factor-1 beta gene (*HNF1B* by MODY5) each account for 5-10% of all MODY cases. The transcription factors HNF1A, HNF4A, and HNF1B function in the nucleus of the pancreatic beta-cell and regulate the transcription of for instance the *INS* gene, either directly (HNF1A and HNF1B) or indirectly (HNF4A) (Fajans et al., 2001).

1.3 MODY5

The hepatocyte nuclear factors (HNFs) are involved in diverse expression patterns and functions in the developing embryo. Abnormalities in the liver, kidney, and the reproduction system, and function, is for instance part of the clinical phenotype in MODY5 (See Figure 5). Such multi-organ disease, is assumed due to *HNF*s involvement in auto-regulatory and cross-regulatory circuits, likely acting in a combinatorial manner to determine the early development of these affected organ tissues (Lau et al., 2018). Although diabetes is a hallmark by *HNF1B* mutations in MODY5, the most consistent clinical feature is severe renal disease. Renal cysts and diabetes syndrome (RCAD) in carriers of heterozygous mutations in the *HNF1B* gene is associated with disease such as severe renal cysts and defects, genital malformations, as well as pancreas hypoplasia, and beta-cell dysfunction (De Vas et al., 2015; Edghill et al., 2008; Haumaitre et al., 2006).

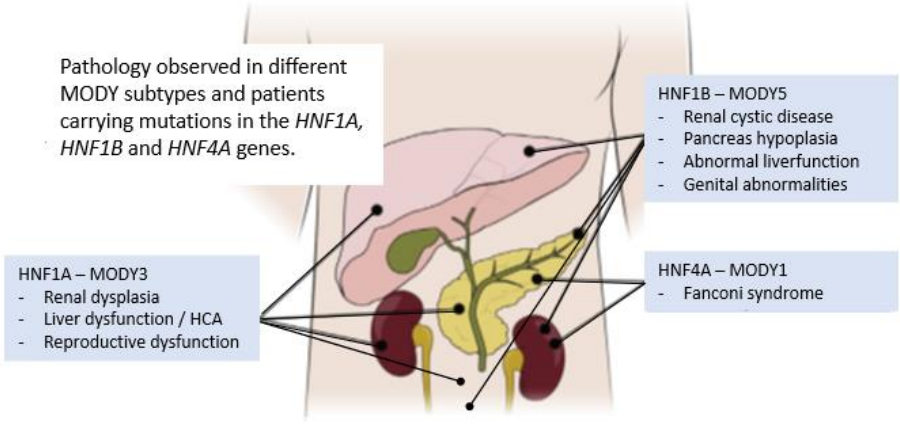


Figure 5: Anatomical drawing over MODY subtypes and their associated pathologies. Pathology observed in liver, kidneys and pancreas by MODY1, MODY3 and MODY5.

As the HNF1B transcription factor acts upstream of other transcription factors, such as GLIS-3, which is involved in duct morphogenesis and endocrine cell development, a decreased expression of HNF1B has been associated with cystic ducts in the kidneys and pancreas (De Vas et al., 2015).

HNF1B has further been shown to be involved in pancreas morphogenesis and the generation of endocrine progenitors. In addition, studies have shown a reduction in and abnormal differentiation of acinar cells in *HNF1B* haploinsufficiency in both humans and murines (De Vas et al., 2015; Haumaitre et al., 2006). Early deletion of *HNF1B* has further been shown to lead to reduced numbers of pancreatic multipotent progenitor cells (MPCs), as a result of decreased proliferation of cells, in addition to increased apoptosis (De Vas et al., 2015). Thus, lack of HNF1B which plays a crucial role in the regulatory networks that control pancreatic MPC expansion, generation of endocrine precursors, acinar cell identity as well as duct morphogenesis, suggests a crucial role of HNF1B in pancreas development and in MODY5 disease development (De Vas et al., 2015).

1.3.1 HNF1B (Homo Sapiens)

HNF1B gene and protein isoforms

The *HNF1B* gene, also known as Transcription Factor 2 (*TCF-2*), is located on chromosome 17q12 and encodes the 557 amino acid protein hepatocyte nuclear factor-1 beta (HNF1B). The gene consists of 9 exons and produces three alternatively isoforms (1, 2 and 3), all identical in their N-terminal end, but differ in the C-terminal region of HNF1B isoform 3 (the transactivation domain), illustrated in Figure 6. The HNF1B isoform 1 and 2 contains all nine exons, while isoform 3 ends after exon 6, resulting in a reading frame shift of exon 9 (Hojny et al., 2020). Transcript variant 2 uses an alternative in-frame splice site in the central coding region, resulting in an internal segment lacking in HNF1B isoform 2 compared to isoform 1. It is believed that all three isoforms are essential in embryonic pancreas and beta-cell development and function, as well as in kidney nephron development (Lau et al.,2018). In adult tissue, HNF1B acts as a classic transcription activator for the regulation and expression of multiple genes implicated in cell cycle regulation, apoptosis, and glucose metabolism (Hojny et al., 2020). HNF1B is also found to act as a regulator of the expression of genes associated with stem and progenitor cells (Yu et al., 2015).

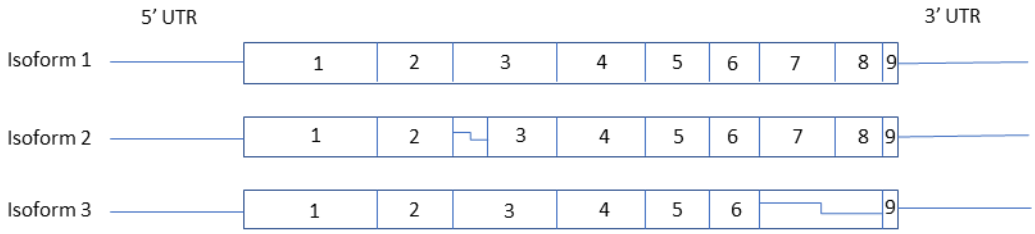


Figure 6: Schematic illustration of HNF1B isoforms. Isoform 1 is the canonical transcript, Isoform 2 lacks an internal segment at the 5' end of exon 3, Isoform 3 lacks exons 7 and 8.

Protein structure and function of HNF1B

HNF1B is a well-characterized transcription factor, highly conserved between vertebrates. It is also a homeodomain containing protein (Cubuk & Yalcin Capan, 2021) (Kim et al., 2014). Its protein structure consists of four functional domains; a dimerization domain (amino acids 1-31), a DNA-binding domain composed of an atypical POU-specific (POU_S) domain (amino acids 90-187) and a POU-homeodomain motif (POU_H) (amino acids 235 to 311), as well as a transactivation domain (amino acids 312 to 557) (illustrated in Figure 7).

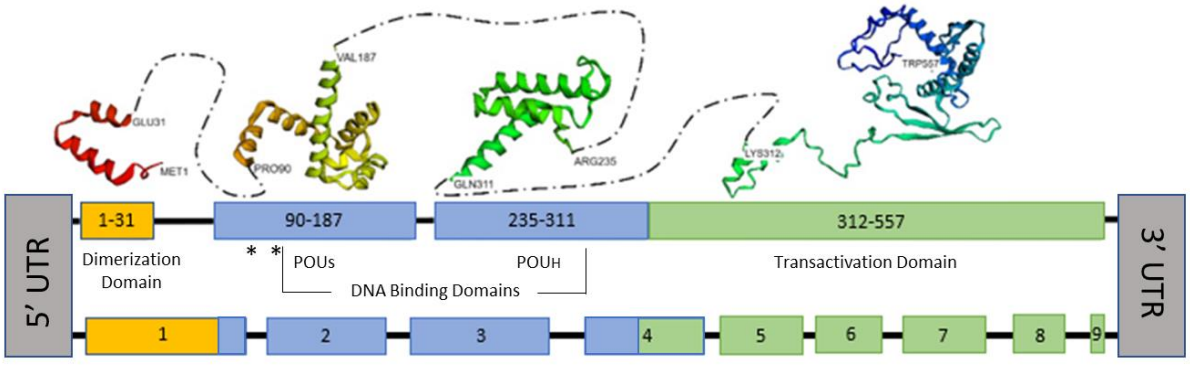


Figure 7: Schematic figure of *HNF1B* and its functional domains. HNF1B is comprised of a dimerization domain, a DNA binding motif (POU specific domain and POU homeodomain), as well as a transactivation domain in the C-terminal. Adapted from Cubuk & Capan, 2021.

The HNF1B transcription factor is predicted to regulate more than 100 target genes using known transcription factor binding site motifs from the JASPAR Predicted Transcription Factor Targets dataset.

The HNF-transcription factors (HNF-TFs), firstly described in the liver and hence the name hepatic nuclear factors, are part of a regulatory network controlling normal tissue development and metabolic functions in pancreatic beta-cells (Mitchell & Frayling, 2002; Yamagata et al., 1996).

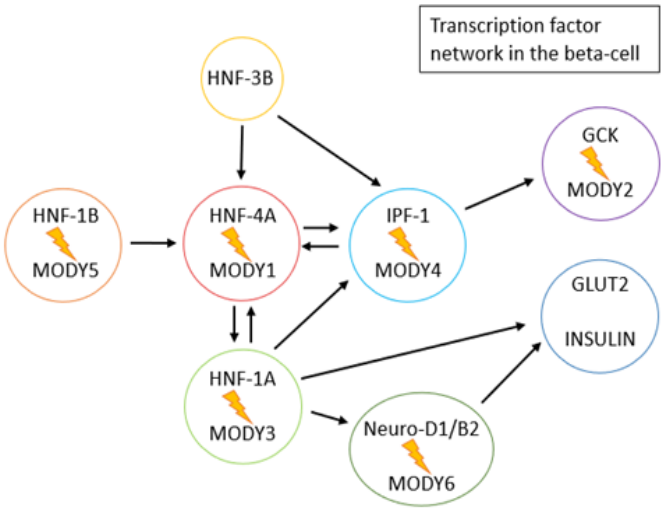


Figure 8: Schematic presentation of selected transcription factors involved in a regulatory network in the pancreatic beta-cell. Pancreatic islet development and function is dependent on this network of transcriptional regulation. For instance, HNF1B is an important regulator of HNF4A which in turn is cooperatively regulating the expression of other transcription factors to activate the GLUT2 and INS gene. Illustration adapted from Mitchell & Frayling, 2002.

HNF1B is closely related to HNF1A (See Figure 9). They are from the same homeodomain-containing family that binds to the same palindromic consensus DNA sequence to activate transcription (Cubuk & Yalcin Capan, 2021; Lau et al., 2018). For this, HNF1B operates either as a homodimer, or a heterodimer bound to HNF1A, through interaction at the dimerization domains (N-terminus). Dimeric formation by all three different isoforms of HNF1BA has been found to be tissue-specific dependent (Lau et al., 2018; Mendel et al., 1991).

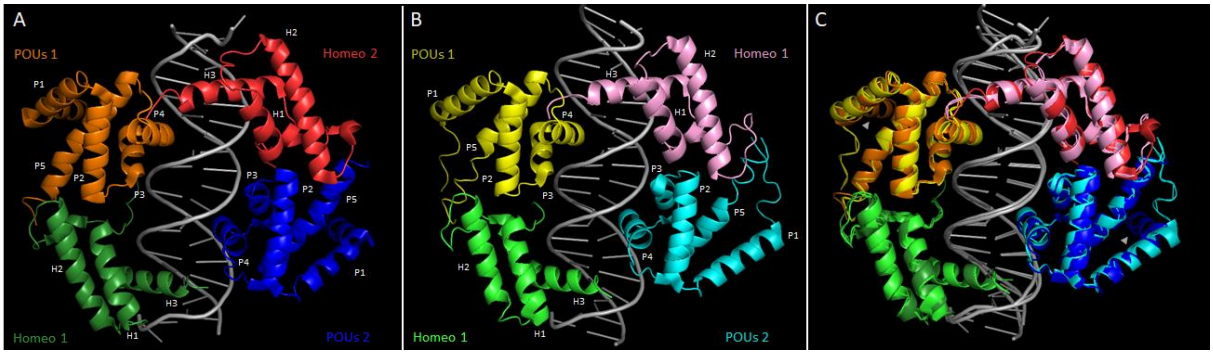


Figure 9: Homodimers of HNF1B or HNF1A bound to DNA.

A: Two HNF1B monomers (Homeo1/POUs1 bound to Homeo2/POUs2) bound to dsDNA as a dimer through its DNA binding region. The three alpha-helices of each homeodomain and five alpha-helices in each POUs domain are labeled H1-3 (green and red) and P1-5 (orange and blue), respectively **B:** Two HNF1A monomers

(Homeo1/POUs1 bound to Homeo2/POUs2) bound to dsDNA as a dimer through its DNA binding region. The three alpha-helices of each homeodomain and five alpha-helices in each POUs domain are labeled H1-3 (light green and pink) and P1-5 (yellow and cyan), respectively. **C:** Alignment of HNF1B and HNF1A homodimers. White arrows indicate longer alpha-helices at HNF1B POUs domain P1. The 3D structure of the HNF1B/HNF1A/DNA complex (PDB ID:1IC8) was retrieved from the protein data bank (Chi et al., 2002) and visualized using the molecular visualization software, PyMOL (Schrödinger).

1.3.2 Animal models targeting *HNF1B*

As an approach for improved understanding of MODYs and the underlying mechanisms for disease development, animal studies related to targeting individual *HNF* genes have mainly been performed in rodents and rodent tissue, and less so in other model systems (Lau et al., 2018).

Limitations with such rodent models exist, however, when comparing disease phenotypes in humans versus rodent MODY models. For instance, the fact that heterozygous mutations cause MODY disease in humans, and not in rodents, suggest a clear species-specific difference in genotype-phenotype correlation (Lau et al., 2018). To illustrate, while the rodent model for human MODY5 heterozygous knockout of HNF1B (+/-) lacks a disease phenotype, the complete knockout of HNF1B (-/-) leads to entire epiblast deletion and death after the blastocyst stage (De Vas et al., 2015). Further, renal-specific inactivation of HNF1B leads to polycystic kidney disease (Barbacci et al., 1999; Gresh et al., 2004). In addition, while studies of complete HNF1A (-/-) knockout in mice have shown that these mice only survive approx. 1 month after birth, and HNF4A (-/-) knockout in mouse embryo has also shown to be lethal (Molven & Njolstad, 2011; Pontoglio et al., 1996).

The zebrafish is an established model organism for several pathophysiological conditions which are related to human diseases. For instance, zebrafish exhibit conserved functions of the pancreas, as well as liver and adipose tissue(Kamel & Ninov, 2017). This can be used to study the role of conserved genes associated with the risk of metabolic disease, in relation to humans. Compared to rodent model systems, additional advantages of zebrafish include rapid embryo development, and mature digestive organs 4 days post fertilization. The optical transparency of fertilized embryos and larvae also uniquely permits *in vivo* imaging (Feierstein et al., 2015; Herrgen et al., 2010).

Though rodent models have increased our understanding of the role of HNF genes, gaps still exist in terms of complete understanding of how mutations in these genes lead to MODY disease phenotypes. Therefore, there is an urgent need for further investigations to complement the findings regarding the underlying molecular mechanisms resulting in disease phenotypes. Zebrafish might thus provide a new and improved animal model for investigating MODY gene (HNF) mutation causing disease in pancreas and other affected organs, which ultimately can increase our understanding of the molecular mechanisms underlying MODY.

1.4 Zebrafish (Danio Rerio)

Zebrafish is a freshwater teleost originating from South-Asia and amongst the vertebrate model organisms, uniquely well suited for rapid and targeted generation of mutant lines. Zebrafish is thus, widely used as a model system in the study of developmental biology, human genetics, and human pathology. One of the reasons for this is the species' transparent embryos that enables detailed investigation and imaging of vertebrate development (Dooley & Zon, 2000; Howe et al., 2013). Moreover, zebrafish are inexpensive and easy to maintain in large numbers enabling replicative analyses and high-throughput screening.

The zebrafish genome has more than 26 000 protein coding genes. The zebrafish genome went through a whole genome duplication, which was followed by gradual loss of redundant genes, and several common and important developmental mechanisms have been identified in zebrafish which are similar in mammals (Jorgens et al., 2012; Meyer & Schartl, 1999). However, for effective modelling of human genetic disease it is important to understand the extent to which zebrafish genes and gene structures are related to orthologous human genes. Approximately 70 % of zebrafish genes have a clear human orthologue, and for genes associated with human disease, the number is 82 % (Howe et al., 2013).

The zebrafish exhibits conserved functions of the pancreas, and as a model for pancreatic disease development, zebrafish has given useful insight into diseases such as T2D and obesity, metabolic pathogenesis and diabetic complications such as vascular disease and nephropathy (Jorgens et al., 2012; Zang et al., 2018; Zang et al., 2017). The paralog of the MODY5 associated gene *HNF1B, hnf1ba* is highly conserved in zebrafish (73 %) and abundantly expressed in the zebrafish pancreas (Gong et al., 2004). In this thesis, we will investigate the *HNF1B* as well as the *Glucagon* (*GCG)* paralogs *hnf1ba* and *glucagon a (gcga)* in zebrafish and observe the CRISPR-Cas9 knockout effect on *hnf1ba*.

1.5 CRISPR-Cas9

In recent years, the ability to generate targeted mutations in living organisms has been greatly enhanced with the discovery and adaptation of a method based on Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). CRISPR is named after a series of tandem repeat sequences found in many bacteria and archaea (Makarova & Koonin, 2015). These tandem repeat sequences are flanked by a conserved set of genes named CRISPR associated protein (Cas), which are similar to nuclease encoding genes. CRISPR-Cas protein is a part of the adaptive bacterial defense system against viruses, as these repeated segments often match short segments in phage genomes. The bacterial immune system will upon infestation, cleave the intruding phage DNA on exactly the corresponding repeated CRISPR array (Li et al., 2016).

Inspired by this molecular mechanism, the CRISPR-Cas9 system has been developed for genome engineering, which enables precise genome editing. The CRISPR-Cas 9 system utilizes a short synthetic RNA called guide RNA sequence (gRNA or sgRNA) comprised of a scaffold sequence necessary for Cas9-binding and a user-defined crRNA sequence, an approx. 20 nucleotide spacer that defines an exact genomic target sequence to be modified. Cas9 has a domain which binds to a short sequence of DNA called the protospacer adjacent motif (PAM). The gRNA recognizes the complementary sequence on the genomic DNA and binds to it, while Cas9 will induce a double stranded DNA break (DSB) three nucleotides adjacent to the PAM sequence (See Figure 11).

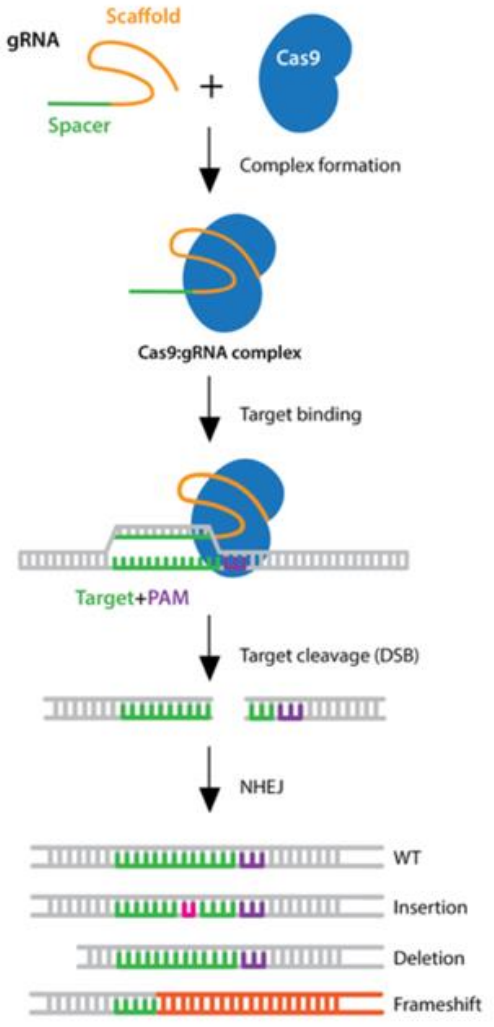


Figure 10: guide-RNA binding to Cas9-enzyme forming the known CRISPR-Cas9 complex. Cas9:gRNA complex binds to the target in the genome and makes a double stranded cut 3 nucleotides from the PAM-sequence. Repair response in the nucleus is here a Non-Homologous End-Joining (NHEJ) which can result in ideal repair leading to Wild-Type phenotype, insertion and deletion of nucleotides and/or large alterations leading to frameshift.
Source: https://www.addgene.org/guides/crispr/

A DSB will be repaired by a set of enzymes from an endogenous repair mechanism called non-homologous end-joining (NHEJ). The repair response mechanism of DSBs is imprecise, and will often introduce errors in the sequence repaired, resulting in short insertions or deletions (indels). As a result, frameshift mutations or nonsense mutations can be introduced, causing an inactivation or other functional changes in the encoded protein. An alternative repair method is homology directed repair (HDR), which uses a homologous DNA template supplied by the researcher to generate alleles containing the desired alteration. An exogenous sequence might also be added in the repair method called homology independent repair (Li et al., 2016).

1.6 Aim of the study

The aim of the project is to generate a zebrafish MODY5 model by gene editing the *hnf1ba* gene and to investigate the development of pancreas pathogenesis in *hnf1ba* zebrafish mutants. Sub-aims are to:

1. Investigate *hnf1ba* and *gcga* expression in zebrafish larvae by *in situ* hybridization.

2. Design a CRISPR-Cas9 construct targeting the *hnf1ba* gene in zebrafish.

3. Characterize the zebrafish endocrine pancreas by fluorescence immunohistochemistry of insulin- and glucagon-producing cells in WT and CRISPR-injected F0 individuals.

2 Materials and Methods

The following section shows the materials and methods regarding zebrafish handling, preparation of samples, *in situ* hybridization on whole-mount embryos and cryo section, immunohistochemistry on whole-mount and cryo section, imaging and image processing, quantitative image analysis, CRISPR-Cas9 sgRNA preparations, microinjection, TOPO-TA cloning and genotyping.

2.1 Ethics Statement and zebrafish handling

Adult and larvae zebrafish were held and maintained by the Zebrafish Facility at Institute for Bioscience (UiB) according to standard practice (Alestrom et al., 2020). Experimental use of zebrafish was approved by the Norwegian Food Safety Authority (FOTS ID 25543).
The zebrafish lines utilized in this thesis were wild-type (WT), Casper transparent line and transgenic Green Fluorescent Protein (GFP) line Tg(wt1b:EGFP) imported from Fritz Lipmann Institute, Jena, Germany. The Tg(wt1b:EGFP) line was chosen because of its fluorescent tags of the exocrine pancreas and kidneys. The Tg(wt1b:EGFP) has normal pigmentation so for some applications, the Casper line was chosen for its transparency due to lack of pigmentation.

To obtain fertilized embryos for *in situ* hybridization (ISH) and immunohistochemistry (IHC) and CRISPR microinjection experiments, the WT zebrafish males and females were held in a breeding tank and newly fertilized embryos were collected. Viable embryos were sorted and transferred to a Petri dish containing E3 medium. Embryos were raised in optimal temperature at 28.5 under standard conditions.

2.2 Preparation of the samples

2.2.1 Fixation of 24 hpf embryos or 96 hpf larvae for whole-mount analyses

Manually dechorionated 24 hpf embryos or 96 hpf larvae were fixed in 4 % Paraformaldehyde (PF) for 24 h at 4°C. Then, the samples were washed 2x 5 min in DEPC-treated 1xPBS and either used directly for immunostaining or dehydrated with increasing EtOH concentrations : 75 % EtOH (5 min), 90%EtOH (5 min) and 100 % EtOH (2 x 5 min). Samples were then either directly used for *in situ* hybridization or stored at -20°C for further *in situ* hybridization or immunostaining applications.

2.2.2 Fixation of 96 hpf embryos for analyses on cryosections

Fixation

96 hpf Zebrafish larvae were fixed in 4 % PF for 48 h at 4 °C. The fixated larvae were then washed with DEPC-treated 1xPBS (2x 5 min), cryopreserved in 25 % sucrose in 1xPBS for 24 h then embedded in Tissue-Tek or stored in 25 % sucrose + 25 % TissueTek® in 1xPBS at -20°C.

Embedding

The larvae were embedded in hand-made rectangular mold made of aluminum foil. A small amount of TissueTek® was first poured and set to slightly harden on dry ice, in order to prevent the samples to sink at the very bottom. Three individuals were then laid aligned on this first layer, with their dorsal side facing upwards and their anterior towards a pre-determined side of the mold (marked with a blue dot, see Figure 12) and covered with TissueTek®. The mold was transferred to dry ice for hardening and wrapped in parafilm to store at -80 C.

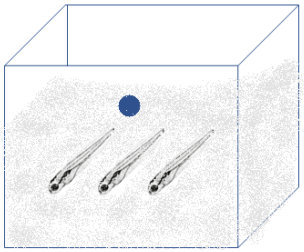


Figure 11: Embedding multiple zebrafish larvae. Three 96 hpf zebrafish larvae were lined up inside an aluminum mold. To generate cross-section, the side of the block marked with the blue dot will be the cutting side.

Cryo sectioning

10 µm-thick cross-sections were prepared using a Leica CM3050 S Research CryoStat Microtome. The whole larvae were cut, and sections were collected sequentially onto a series of 4 room-tempered glass slides (ThermoScientific, J1800AMNZ). After sectioning, the slides were heated at 65 °C for 10 min to fix the sections to the glass slide before it was stored at -80 °C.

2.3 *In Situ* Hybridization

2.3.1 RNA probe preparation for ISH utilizing two approaches

A flowchart (See Figure 12) illustrates the workflow conducted to synthesize the RNA probes to use for *in situ* hybridization (ISH) by two different approaches. The two methods utilizes either a cloned product as a template (Method 1 - in blue) or a PCR-generated template (Method 2 - in orange) to make probes targeting *hnf1ba* and *gcga*. For both methods, total RNA was first extracted from 24 or 96 hpf zebrafish embryos/larvae, followed by a reverse transcription to generate a cDNA library. The *hnf1ba* and *gcga* fragments were amplified by PCR from the cDNA libraries, either for further cloning or for direct use as a template. For Method 1, the PCR product was inserted by TA-cloning into the pCRII TOPO-vector, that contains SP6 and T7 promoters on each side of the insert. For Method 2, the primers themselves contained the promoter sequences (SP6 on Forward and T7 on the Reverse primer), eliminating the need for the cloning procedure. DIG-labeled ISH probes were then synthesized either from the linearized plasmid or the PCR product. For ISH on whole-mount of 24 hpf embryos, the first method was used, while for ISH on cryosections of 96 hpf larvae, the second method was used.

The following sections details the steps for both alternative methods.

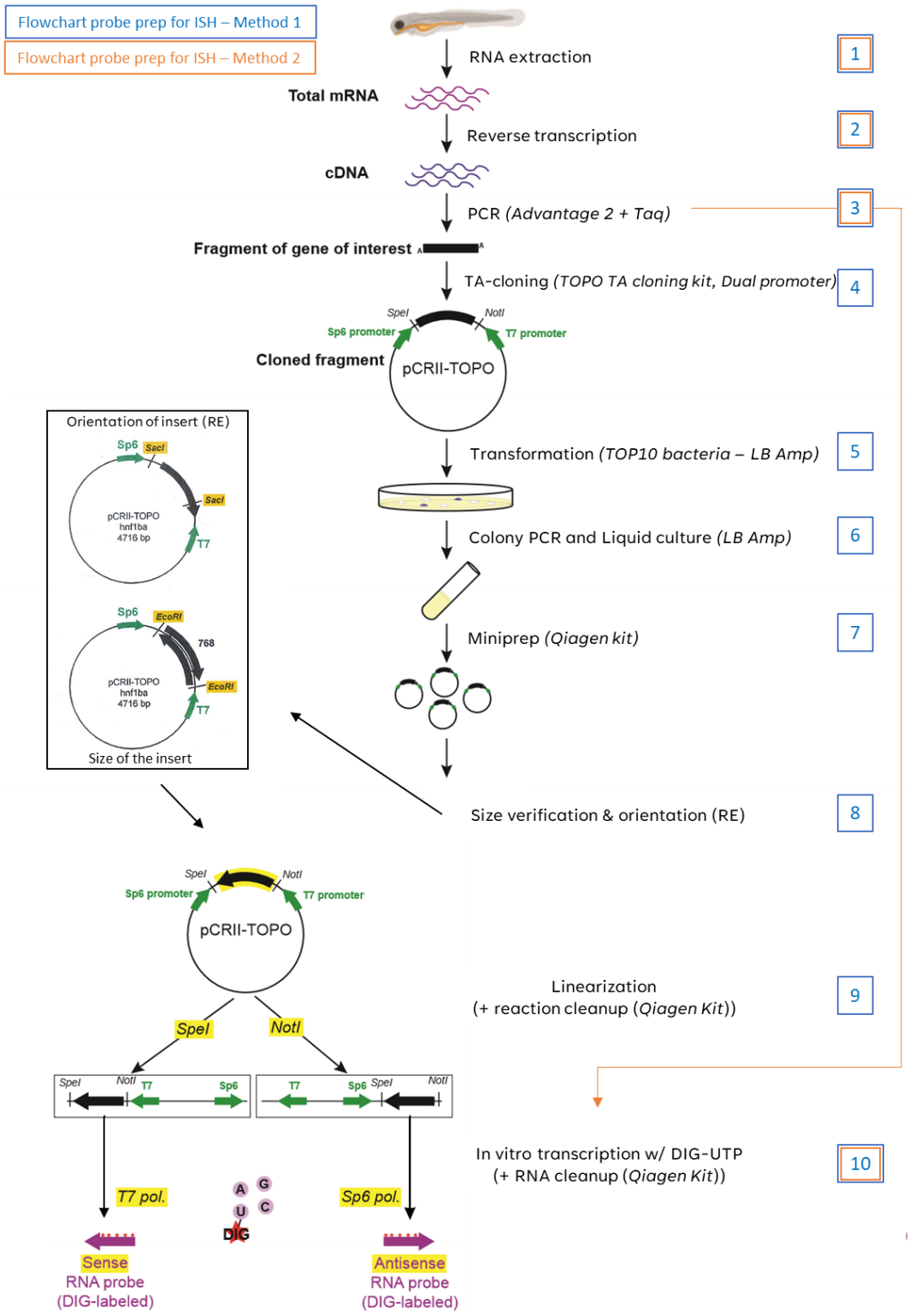


Figure 12: Flowchart of RNA probe synthesis performed in two separate ways. The blue pathway describes template preparations for probe synthesis by using TOPO-cloning and restriction enzymes (Method 1), while the orange pathway describes template preparations for probe synthesis using a DNA template generated by PCR and modified PCR (Method 2). Illustration modified from Elsa Denker.

1. RNA extraction of zebrafish embryos and larvae

RNA extraction was performed using the RNeasy Mini kit from Qiagen following the manufacturer's protocol for tissue, with minor changes described.

Dechorionated WT 24 hpf embryos or 96 hpf larvae were pooled into 1.5 ml Eppendorf tubes on ice (approx. 50 individuals in each tube) to obtain a maximum amount of tissue at 30 mg. Tissue was then lysed by vortexing in 600 µl Buffer RLT with added β-mercaptoethanol. Debris were removed by centrifugation in an Eppendorf microcentrifuge for 3 min at maximum speed and the supernatant was carefully pipetted to a tube containing 1 volume of 70 % ethanol and mixed well. The samples were then transferred to RNeasy Mini spin columns and centrifuged for 15 seconds at >8000 x g. Wash steps were performed as described in the protocol, and the total RNA was eluted in 30 µl RNase-free water.

Total RNA concentration and quality (260/280 nm and 260/230 absorbance ratios) were measured using a NanoDrop Spectrophotometer (ThermoFisher, ND-ONE-W), and RNA was immediately used in cDNA synthesis or stored at -80 ℃.

2. Reverse Transcription

Two cDNA libraries were generated from 1,5 µg of the total RNA using the Invitrogen SuperScript III Reverse Transcriptase (#18080-044, Invitrogen) cDNA synthesis kit:

Mix number one was prepared on ice as shown below, added to the RNA and incubated for 5 min at 65 °C before being transferred back to ice again for 1 min.

Mix # 1	
1 µl	Oligo(dT)20 primer (50 µM)
1 µl	dNTP (10 mM)
Water q.s.p 13µl	
13 µl	Intermediate total volume

Mix # 2	
4 µl	5X First-Strand Buffer
1 µl	DTT (0.1 M)
1 µl	RNaseOUT (40 U/µl)
1 µl	SuperScriptTM III RT (200 U/µl)
7 µl	Intermediate total volume

Mix number two was then added to the RNA reaction mix #1, for a final total volume of 20 µl. The cDNA synthesis reaction was incubated for 60 min at 50 °C, followed by 15 min at 70 °C for inactivation of the RT-enzyme and left for cooling at 4 °C.

3. Advantage PCR amplification

To amplify the desired cDNA for further RNA probe synthesis, 3 primer pairs were utilized.

Method 1	Primer sequences (5'-3')	Location (exon)	Probe size (nt)
Hnf1ba_DR_F	AAGATGCAAGGTGTCCGGTA	4-5	663
Hnf1ba_DR_R	AATGTGCTCCTGTTGCTTGG	9' UTR	663
Method 2			
Hnf1ba_DR_SP6-F	ATTTAGGTGACACTATAGAAAAGATGCAAGGTGTCCGGTA	4-5	703
Hn1ba_DR_T7-R	TAATACGACTCACTATAGGGAATGTGCTCCTGTTGCTTGG	9' UTR	703
Gcga_DR_SP6-F	ATTTAGGTGACACTATAGAAAGACTTCGTTCAGTGGCTCA	3	660
Gcga_DR_SP6-R	TAATACGACTCACTATAGGGTTTACCTGAGCGTTTCGTGC	6' UTR	660

The 20 nt T7 sequence is located at the 5' extremity of the reverse primer, and 20 nt Sp6 sequence is located at the 5' extremity of the forward primer.

Advantage 2 PCR kit (TaKaRa Clontech 639206-CLI - AH Diagnostics) was used to amplify the cDNA with a high-fidelity polymerase.

5 µl	cDNA
5 µl	10X buffer from the kit
1 µl	50X dNTP from the kit
1 µl	Primer F (10 µM)
1 µl	Primer R (10 µM)
1 µl	50X Advantage Polymerase
36 µl	MilliQ water
50 µl	Total

The cDNA was amplified using the following PCR program:

95 °C	1 min	
95 °C	30 sec	40 cycles
60 °C	1 min	
4 °C	infinity	

The amplified PCR products were purified using Qiagen MinElute PCR purification kit (#28004) as described by the manufacturer protocol and eluted in 10 µl MilliQ water.

Amplified cDNA was either TOPO-cloned or used as templates in step 10 for *In vitro* transcription (See Figure 12).

4. TOPO-TA cloning into pCRII vector

To ensure proper ligation using TA-cloning, the PCR product was added 3' A-overhangs using regular NEB Taq DNA polymerase (#B9014S, NEB) (for 10 min at 72 °C:

8 µl	purified *hnf1ba* PCR product
2 µl	10x Taq Buffer
1 µl	dNTP (10 mM)
1 µl	Taq DNA polymerase (5 U/µl)
7 µl	MilliQ water
20 µl	Total

The 3' A-tagged PCR product was used immediately for TOPO TA-cloning into pCRII-TOPO vector from Invitrogen (#K457501) in a ratio of 4:1 insert and vector. Salt solution was added to stabilize the reaction, and the ligation mix was left at RT for 1 h before transformation:

4 µl	PCR product
1 µl	Salt solution from the kit
1 µl	pCRII TOPO Vector
6 µl	Total

5. Transformation of TOPO vector into OneShot TOP10 Competent cells

One vial of OneShot TOP10 Competent Cells from Invitrogen (#K457501) was used to transform the *hnf1ba* ligated vector by heat-shock.

The cells were first thawed gently on ice before adding 2 µl TA-cloned vectors and left on ice for 30 min. The cells were then heat-shocked in an equilibrated water bath at 42 °C for 30 seconds, before cooling the transformed E.Coli bacteria on ice for 2 min. A 250 µl room tempered Super Optimal broth with Catabolite repression (SOC) medium was then added and the cells were incubated at 37 °C for 1.5 h whilst shaking at 250 rpm. 50 µl transformed competent cells were then spread onto LB plates containing 100 µg/ml Ampicillin and incubated O/N (~18h) at 37 °C. For blue/white screening of colonies, the plates had been pre-plated with 120 µl X-gal (20 mg/ml). Bacterial colonies harboring an intact plasmid without insert will have the ability to transcribe β-galactosidase to hydrolyze X-gal resulting in a blue pigment (a dimer of 5-bromo-4-chloro-indoxyl). A white colony will therefore be a bacterial strain containing a plasmid with insert.

6. Colony PCR and Liquid LB culture

A colony PCR was used to screen for colonies with an insert of the correct size.
White, clearly defined colonies of good size were circled, numbered, and transferred to a new systematically marked LB-plate using an inoculation loop or a pipette tip. Material for PCR was collected by dipping the same loop in a designated well in a 96 well PCR plate containing 100 µl MilliQ water. The PCR plate was sealed with adhesive foil, vortexed and spun before incubating at 95 °C for 10 min to lyse the bacteria, before it was used in a BioTaq PCR (Bioline BIO-21060 kit, Nordic Biosite) using primers on both side of the insert on the plasmid:

1 µl	Bacterial lysate
2 µl	NH4 buffer
1.2 µl	MgCl2
0.2 µl	dNTP 50 mM
0.4 µl	M13 F primer 10 µM (from kit)
0.4 µl	M13 R primer 10 µM (from kit)
14.4 µl	H2O
0.4 µl	BioTaq pol
20 µl	Total

The PCR reaction was run on a Thermocycler:

95 °C	1 min	
95 °C	20 sec	30 cycles
55 °C	20 sec	
72 °C	20 sec	
72 °C	10 min	
4 °C	infinity.	

The size of the PCR products was checked by electrophoresis on 1 % agarose gel using 1X TAE buffer and 1X GelRed (#41003, Biotium) and running at 90 V for 40 min.

The colonies with the correct inserts were picked from the new numbered LB agar plate and inoculated in bacteria culture tubes containing 2 ml liquid LB medium containing 100 µg/ml Ampicillin before incubation O/N at 37 °C.

7. Miniprep plasmid extraction

To extract the transformed plasmid with desired insert, LB cultures were transferred to 2 ml Eppendorf tubes and centrifuged at 13000 rpm for 3 min. Pellets were resuspended, lysed and DNA purified using QIAprep Spin Miniprep kit (QIAGEN #27104) according to the manufacturer's protocol. To ensure sufficient DNA concentration, columns were eluted twice with the same 35 µl of MilliQ water.

The concentrations were quantified by Nanodrop and the purified plasmid was stored at -20 °C until further use. Leftover liquid culture of bacteria with correct insert was stored as glycerol stocks at -20 °C freezer by adding 50 % glycerol.

8. Size verification and orientation of the *hnf1ba* plasmid insert

To confirm the size of the insert and determine its orientation, minipreps were digested with different restrictions enzymes. EcoRI has two sites incorporated in the pCR2 vector, close to the insertion sites of the insert, and was used to determine the size of the insert, while SacI has one cut site incorporated downstream of SP6 promoter and one cut site in the insert, generating asymmetric bands that differ depending on insert orientation (See Figure 12). The reactions were incubated for 1 h at 37 °C:

2 µl	Template (miniprep)
2 µl	10X SmartCut buffer
0.5 µl	RE (25 U) EcorI or SacI
15.5 µl	H_2O .
20 µl	Total

Further, the digested products were analyzed on 1 % agarose gel. The orientation of the insert was a crucial information to determine restriction enzyme to use for linearization and which promoter to use, to generate sense or antisense probes.

9. Linearization and purification of plasmid insert

Two linearization reactions (one with NotI and one with SpeI) were prepared for each plasmid, and left O/N at 37 °C to ensure complete linearization.

30 µl	template (2.5 µg)
5 µl	10X SmartCut buffer
2 µl	RE (25 U) of NotI or SpeI
13 µl	Nuclease-Free water
50 µl	Total

5 µl of digested products (+ 1µl 6X Loading buffer R0661, Thermo Scientific) were tested on 1 % agarose gel to assure correct size and purity. Qiagen MinElute Reaction Cleanup Kit (#28204-06) was used to purify the plasmid after linearization. The protocol was followed as described by the manufacturer. The purified insert was eluted in 30 µl nuclease-free water.

10. *In vitro* transcription with DIG-UTP

To amplify the desired sense and antisense RNA probes with Digoxigenin labeled UTP, an *in vitro* transcription was performed on the corresponding purified linearized plasmid: SpeI- and NotI-linearized plasmids were used with T7 and Sp6, respectively.

13 µl	Purified plasmid (1µg) in water
2 µl	10X Dig labelling mix
2 µl	10X Transcription buffer
2 µl	RNA polymerase (40 U) of T7 for SpeI cut insert, and SP6 for NotI cut insert
1 µl	RNase inhibitor
20 µl	Total

The RNA synthesis was incubated O/N at 37 °C and stopped by adding 2 µl RNase-free EDTA (0.2 M) and RNAs were further kept on ice. The probes were purified with Qiagen RNeasy cleanup Kit (#74204), following the manufacturer's protocol. The purified probes were analyzed by NanoDrop to determine their quality and yield and run on 1 % agarose gel (1X TAE, RNase-free) besides an RNA RiboRuler High Range ladder to check their size and integrity. The probes were stored at -80°C until further use.

In Situ Hybridization Protocols

In situ hybridization was used to visualize the expression patterns of the *hnf1ba* (24 hpf whole-mount and 96 hpf on sections) and *gcga* (96 hpf on sections) genes in zebrafish. The methods described are a revised protocol from (Sandbakken et al., 2012) using urea instead of formamide (Sinigaglia et al., 2018).

2.3.2 ISH Protocol for Whole-mount

ISH day 1: Pre-treatment and hybridization

Due to the fact that the RNA probe is highly sensitive for RNases, all equipment and surfaces were wiped with both Sigma RNaseZap and 70% ethanol. All solutions were made with Diethylpyrocarbonate (DEPC)-treated water that is free of RNase. Most of the buffers and solutions were made on forehand, though some solutions needed to be freshly made and were therefore made during the protocol.

Pre-treatment of embryos

Fixed 24 hpf embryos (See section 2.2.1) were transferred to small baskets in 24-well Nunc-trays and rehydrated by immersing in

70% EtOH for 5 min

50% EtOH for 5 min

30 % EtOH for 5 min

1x PBSTw (DEPC-treated) for 5 min

The embryos were then washed 4 times in 1xPBSTw for 5 min each before re-fixating with 4 % paraformaldehyde for 20 min.

Proteinase K treatment

After washing in 1xPBSTw the embryos were permeabilized by incubation with Proteinase K (30 U/mg) diluted in Proteinase K Buffer (10 µl/ml) for 7 min. The embryos were then washed in 1x PBSTw before post-fixating in 4 % paraformaldehyde for 20 min. Lastly, the embryos were washed in 1x PBSTw and 2xSSC two times each.

Hybridization of RNA probe

The embryos were transferred to an Eppendorf tube and immersed in a 250 µl hybridization buffer (Hyb-) and pre-hybridized for 2 h at 65 °C. After pre-incubation the Hyb- solution was replaced by Hyb+ solution (hybridization buffer with RNA probe tagged with DIG), and incubated O/N at 65 °C.

ISH day 2: Post-hybridization treatment

Post-hybridization wash

The embryos were transferred again to baskets and excess probe was removed by successive washes in 2x SSCTw (2 x 5 min), 4M Urea in 2x SSCTw (2 x 15 min then 30 min at 65 °C), and finally 2x SSCTw (2 x 15 min), first at 65 °C then at 37 °C to prepare tissue for RNase treatment.

RNase treatment

To degrade all RNA that is not hybridized, embryos were incubated with RNase A (0,02 mg/ml) was used for 30 min at 37 °C, followed by a wash step in only RNase buffer for 30 min at 65 °C.

Immunohistochemical detection

The larvae were washed in blocking buffer (2xSSCTw + 0.05 % Triton X-100 + 2 % Blocking solution) for 2 h at RT to block the tissue for unspecific binding of antibody, then in 1x Maleate buffer 2 x 5 min. Samples were then incubated with an anti-DIG antibody coupled with alkaline phosphatase (Fab fragments, 1:2000) o/n at RT.

ISH day 3 Visualization with NBT/BCIP:

Excess DIG-antibody was washes using 1x maleate buffer (2 x 30 min, then 2 x 10 min) before 2 x 10 min wash in visualization buffer. The embryos were then incubated in freshly made chromogen substrate NBT/BCIP in a dark box. NBT/BCIP (or p-nitroblue tetrazolium chloride 5-bromo-4-chloro-3-indolyl phosphate #ab7468), detects the alkaline phosphatase-conjugated antibodies attached to the digoxigenin labeled UTP in the RNA probes. The chromogen substrate becomes purple when the alkaline phosphatase removes a phosphate from the substrate.

The optimal incubation time for this probe was determined to be O/N at 4 °C, giving a good staining but a still low background. Embryos were washed in a stop buffer 3 x 10 min and stored in 50 % glycerol in stop buffer until mounting. The embryos were mounted on a glass slide in the same solution, using an adhesive ring as a spacer and covered with a #1,5 cover glass and sealed with nail polish.

2.3.3 Cryosections ISH Protocol

ISH day 1: Pre-treatment and hybridization

Slides with sectioned tissue were collected from -80 °C and left to air dry at RT for about 1 h and then baked for 45 min at 65 °C.

The tissue was then rehydrated in glass cuvettes with ethanol series diluted in DEPC water, first 95 % then 70 % and lastly 50 %, each for 1 min. The tissue was then washed for 1 min in 2xSSC before treating it with 10 µg/ml Proteinase K for 3,5 min for permeabilization, followed by a 5 min post-fixation in 4 % paraformaldehyde and two 2 min washes in 1xPBS.

To prepare the tissue for hybridization, a dehydration step was performed, initiating a wash containing 2xSSC for 1 min before increased concentrations of ethanol followed; first 50 % then 70 % and 95 % before the tissue was immersed 2 times in 100 % ethanol, each for 1 min. The tissue was then left to dry at RT for 1 h. A hybridization solution (Hyb-) was then prepared using and heated to 65 °C in a cabinet. Meanwhile, two probe mixes were prepared containing 1 µl (antisense or sense) probe as well as 62 µl DEPC water. Probes were denatured by 5 min

incubation at 100 °C followed by immediate transfer to ice. Each of the probe mixes was then added to 500 µl of the Hyb- solution, further referred to as Hyb+. From this, 120 µl of Hyb+ was applied to each of the glass slides and covered with Hybri-Slips to prevent the solution from evaporating. The glasses were incubated O/N at 65 °C in an airtight box held moist with on bench paper soaked with 2xSSC .

ISH day 2 Post-hybridization treatment

Excess RNA probes were flushed with 2xSSC followed by a thorough washes in a plastic cuvette, first in 2xSSC (2 x 30 min at RT), then in 4M Urea in 2xSSC (30 min at 65 °C) then 2xSSC (2 x 10 min at 37 °C).

To remove non-hybridized RNAs, the tissue was treated with 0.02 mg/ml RNase A for 20 min at 37 °C, followed by a wash in RNase buffer for 20 min at 65 °C.

To prepare the tissue for immunohistochemical detection the tissue was immersed in a Blocking buffer containing 2xSSC, 0,05 % Triton X-100 and 2 % Blocking reagent for 2,5 h, then washed with 1x Maleate buffer for 2 x 5 min. 120 µl Antibody solution with Anti-DIG-AP (1:2000) was applied to each slide and incubation performed O/N at RT in a humidified airtight box.

ISH day 3 Visualization with NBT/BCIP:

Excess antibodies were flushed away using 1x Maleate buffer followed by a more thorough wash for 2 x 10 min in 1x Maleate buffer. Visualization buffer was then added for 10 min before 200 µl freshly dissolved chromogen substrate NBT/BCIP was applied to each slide and covered immediately with cover wells.

The slides were incubated in a light-sealed incubation box. The optimal time for both probes was determined to be 4 h at RT and O/N at 4 °C. When the staining was strong and background still low, the reaction was stopped by taking off the cover wells and washing for 30 min in the Stop buffer.

The slides were then mounted with ProLong Glass medium containing NucBlue (P36981) and using #1,5 cover glass on glass slides (J1800AMNZ Thermo Scientific). The edges were sealed off using nail polish.

2.4 Immunohistochemistry

To detect the pancreas in zebrafish larvae, immunohistochemistry (IHC) targeting the endocrine insulin- and glucagon-producing cells as well as the EGFP transgene in the Tg(wt1b:EGFP) zebrafish line expressed in the exocrine pancreas, was performed on whole-mount 24/96 hpf and sectioned 96 hpf larvae. To investigate cell structures, antibodies against Synaptic Vesicle 2a were also used (See Table 1).

Table 1: Primary and secondary antibodies

Antibody	Primary ABI	Dilution ABI	Secondary ABII	Dilution ABII	Fluorescence
Glucagon	Mouse	1:100	Donkey Anti-mouse	1:400	Alexa 647 Far red
Insulin	Rabbit	1:100	Chicken Anti-rabbit	1:100	CF 568 Red
Synaptic Vesicle 2a	Rabbit	1:200	Chicken Anti-rabbit	1:100	CF 568 Red
GFP	Goat	1:100	Donkey Anti-goat	1:400	Alexa 488 Green

Table 2: Primary and secondary antibodies suppliers

Antibody	Cat. No #	Supplier
Anti-Glucagon (mouse)	MFCD00162325	Sigma
Anti-GFP (goat)	MBS448039	MyBioSource
Anti-Insulin (rabbit)	GTX128490	Gene-Tex
Anti- SV2A	HPA007863	Sigma
Anti-rabbit Alexa Flour Plus 488	A32731	ThermoFisher
Anti-mouse Alexa Flour Plus 555	A32727	ThermoFisher
Anti-goat Alexa Flour Plus 647	A32760	ThermoFisher
Donkey anti-Mouse	A16160	ThermoFisher
Chicken anti-Rabbit	A16023	ThermoFisher
Donkey anti-Goat	A14342	ThermoFisher

2.4.1 Whole-Mount IHC Protocol

IHC day 1:

The larvae were rehydrated (from 100 % EtOH, via 75 and 25 %, 5 min each) and washed in 100 % PBX (1xPBS containing 1 % Triton X-100 (T8787, Sigma)), before an incubation in Blocking buffer (100 % PBX containing 1X Casein (#B6429, Sigma)). After 2 h incubation, the blocking buffer was replaced with the Primary antibodies diluted in the blocking buffer, and incubated O/N at 4 °C with rocking agitation.

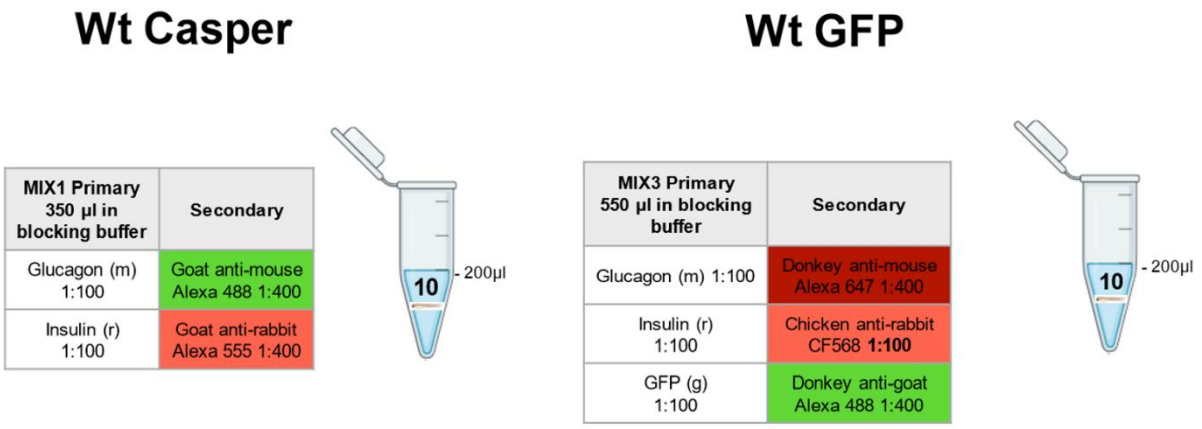


Figure 13: The illustration describes the dilution of the primary antibodies and larvae per tube. Primary antibodies were paired according to the host. 1:100 dilution in 350 µl sample equals 3.5 µl of primary antibody.

IHC day 2:

The incubated embryos were washed in PBX before a 2 h incubation in blocking buffer under rocking agitation at RT. The blocking buffer was replaced with secondary antibodies solutions diluted in blocking buffer and incubated for yet another 2 h in the dark, while in rocking agitation at RT. The secondary antibodies were chosen according to the hosts used for the primary antibodies and with different coupled fluorochromes. The different combinations are presented in Figures 13 and 14.

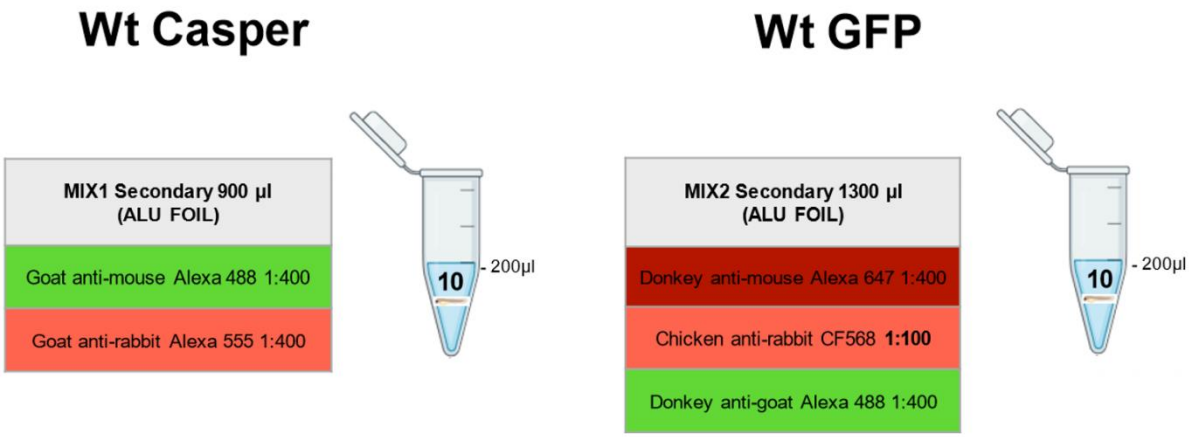


Figure 14: The illustration describes the dilution of the secondary antibodies and larvae per tube. Secondary antibodies were paired according to the host. 1:400 dilution in 900 µl mix equals 2.25 µl of secondary antibody.

The staining was finished after a thorough PBX wash and then the embryos were transferred to a 1:1 PBS/Glycerol solution for preservation.

IHC day 3:

Ring-shaped stickers were stacked to form a well on top of a glass slide where the embryo could be mounted inside. ProLong Glass was used as a mounting medium, containing NucBlue as a DNA dye. For subsequent imaging, the larvae were mounted with the dorsal side up for a dorsal view or with the right side up for a lateral view.

2.4.2 Cryosections IHC Protocol

IHC day 1:

Cryo sections of 96 hpf larvae were dried in a cabinet for 10 min at 65 °C and permeabilized in PBS-TX (PBS + 0.05 % Triton-X100) in a glass cuvette for 10 min. Each slide was then wiped on the back with a Kimtech wipe as well as around the sections with a Whatman paper to ensure hydrophobicity and minimize the risks of leaks during antibody incubation then placed in an airtight, humidified, and leveled chamber for O/N incubation at RT with 200 µl of primary antibody solution diluted in blocking solution (PBS-TX-Cas : PBS + 0.05 % Triton-X100 + 1X Casein), One slide was incubated with blocking solution only as a negative control.

IHC day 2:

The primary antibodies were carefully flushed with PBS-TX and the slides were then washed in a glass cuvette in PBS-TX 3 x 15 min. . The same drying procedure as for the primary antibodies was then used to prepare for incubation with 200 µl secondary antibody diluted in PBS-TX-Cas. The slides were incubated in a light-sealed box for 45 min at RT.

To remove nonspecific binding of antibodies and stop the reaction, the slides were flushed and washed with PBS-TX for 5 min 2 times and then in PBS for 5 min 3 times to remove Triton X-100 which could give background fluorescence.

The glass slides were mounted with ProLong Glass medium containing NucBlue and #1,5 cover glass, sealed with nail polish and left to harden O/N at 4 °C.

2.5 Imaging and image processing

2.5.1 Zeiss Lumar.V12 Stereo microscopy

The Lumar.V12 fluorescence Stereo microscope (Zeiss) was used to sort GFP-positive larvae resulting from the crossing of the Tg(wt1b:EGFP) transgenic line, at 20x magnification, using the GFP excitation/emission cube.

2.5.2 Leica fluorescence microscopy

The DM6000B fluorescence microscope (Leica) was used to study, assess and select the sample with best fluorescence quality in both whole-mount and section IHC staining. The ISH sections were also quality assessed using Bright field function at 20x magnification. Scale bars were made using the Leica application suite X (LasX) Software.

2.5.3 Zeiss Axio Scan.Z1 Slide scanner

The Axio Scan.Z1 Slide scanner (Zeiss) was used to generate tiled pictures of whole sections and have overview of entire slides, both for the IHC fluorescence sections and for the ISH sections. ZEN 2.3 Slidescan Software was used to browse through the different sections as well as for scale bars.

2.5.4 Olympus FV3000 Fluorescence Confocal Microscopy

The Olympus FV3000 confocal microscope was used to image whole-mount and section IHC staining, as well as characterization of live zebrafish embryo transgenic GFP signal. Photo Multiplier Tube (PMT) settings used in this thesis include detection of one wavelength at a time (VBF), average noise canceling 2 times per frame, and sequential scan per frame to avoid one wavelength exiting several fluorophores. Dye and detectors were distributed in 4 groups according to Table 3.

Table 3: Dye & Detector settings for confocal imaging of IHC staining

Wavelength	Emitting color	Fluorophore	Excited color	Pseudo-coloring
405	Purple UV	NucBlue	Blue	White
488	Blue	GFP and Alexa 488	Green	Green
561	Green	Alexa 555 and 568	Red	Red
640	Red	Alexa 647	Far Red	Blue

The laser scanner settings (LSM) were set to Galvano type in a one-way scanning mode, and the image was scanned in 1024x1024 pixels in a 1:1 aspect ratio. Every channel was individually adjusted for saturation and offset, and the top and bottom positions were defined for scan in Z-mode.Images were studied at 20-60x magnification and exported as Tagged Image File Formats (TIFF).

2.5.5 ImageJ software

ImageJ was utilized for processing TIFF images obtained from the confocal microscopy as well as the slide scanner. The levels were adjusted for each channel to reduce noise and background as well as the brightness and contrast function. Images are saved in RGB mode as TIFF.

2.5.6 Adobe Photoshop software

Photoshop was used to rotate, crop and correct the white balance of ISH pictures.

2.6 Quantitative image analysis

2.6.1 ImageJ

Area measurement function (1024 pixels pr. 212.132 µm) was used to measure the area of both projections as well as z-sections. The tool “free-hand selection” was utilized to draw around the stained cells, in which the program calculates the area measurements in μm^2. In addition, was Dapi staining, and the “multi-tool” option utilized to manually count cells in the Z-section image to obtain an average number of both alpha and beta cells in the structure.

2.6.2 Statistical analysis

Statistical analysis was performed using Microsoft Excel. Student t-tests were two-tailed with equal variances, and f-tests were performed to validate the t-tests. Averages are reported as standard deviation, and significant values ($p<0.05$) were marked with asterisks.

2.7 CRISPR-Cas9 and Genotyping

2.7.1 CRISPR construct design

In order to target the *hnf1ba* gene for disruption in zebrafish using the CRISPR/Cas9 system, five custom single guide RNA (sgRNA) were designed to target exon 1 (1a, 1b, and 1c), exon 2 (2a) and exon 3 (3a) (See Table 4).

Table 4: sgRNA designed to target hnf1ba gene in zebrafish. Both hnf1ba_ex1c and hnf1ba_ex3a were recommended as highly specific gRNAs by several software online.

Oligo name	Target sequence	PAM	Generated by
hnf1ba_ex1a_DR	CAGCGATACGGATTCAAAGC	CGG	Design ID idtdna
hnf1ba_ex1b_DR	GGGAAGTTATCCGGAGATGA	NGG	E-CRISP
hnf1ba_ex1c_DR	GGTAACTCCAGAGTCCAACA	AGG	Benchling, CRISPRscan, CRISPOR, IDT
hnf1ba_ex2a_DR	ATCATGCGGGCCGCGCGCCACGG	NGG	CRISPOR
hnf1ba_ex3a_DR	GGCGCCACCATGTTAGACAA	AGG	Benchling, CRISPRscan, CRISPOR, IDT

The guides were designed using different software and selected based on the gRNA scores for predicted efficiency and specificity. All oligos were ordered from Sigma Aldrich and dissolved to 100 µM concentrations.

2.7.2 sgRNA generation

sgRNAs were generated using the method described in Gagnon et al. 2014. This method consists in synthesizing a DNA template by annealing two partially overlapping oligonucleotides, that after filling-in with a T4 DNA polymerase result in a fully double-stranded template (See Figure 15). The constant oligo contains the sequence necessary for the

characteristic scaffold for binding to the Cas9 protein (See also Figure 11). The gene-specific oligo contains the desired target region (in black, also called the spacer region) as well as the promoter sequence for the RNA polymerase (in blue) and an overlap sequence to anneal to the Constant oligo. The resulting sgRNA synthesized during *in vitro* transcription contains the gene-specific target region followed by Cas9 endonuclease-recruiting sequence from the constant oligo.

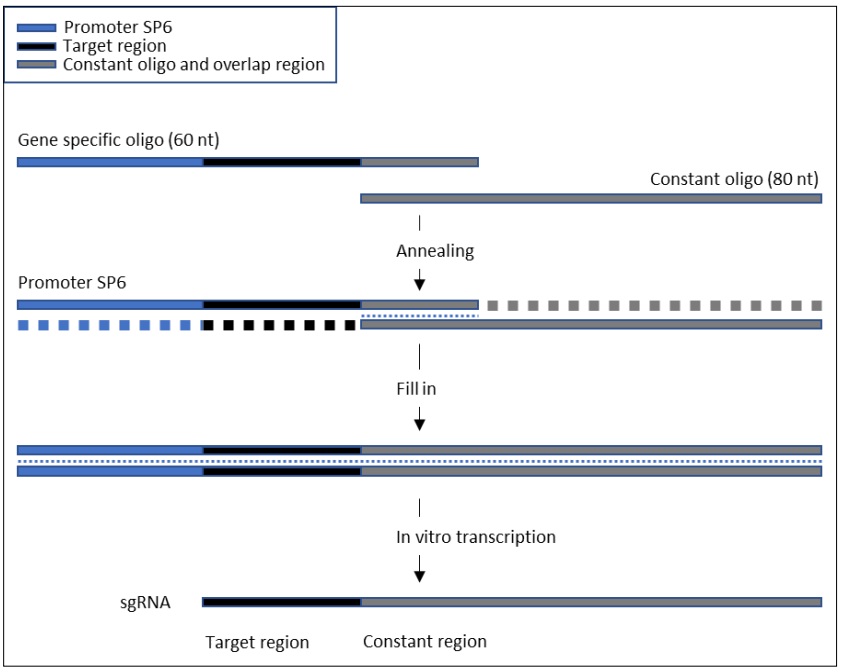


Figure 15: Diagram of cloning-independent generation of sgRNA. Gene-specific oligo (in black), the SP6 promoter region (blue), and Constant oligo sequence (grey). sgRNA are transcribed, containing the desired target and the hairpin forming constant region. Gagnon et al. 2014.

The designed sgRNAs (See Table 5) contained a promoter sequence in the 5' end followed by the gene specific target, and an overlap sequence in the 3' end to attach to the constant oligo.

Table 5: Gene specific oligos and Constant oligo.

Oligo name	Promoter (SP6) - Target - Overlap
Gene specific oligo 1a	ATTTAGGTGACACTATA-CAGCGATACGGATTCAAAGC-GTTTTAGAGCTAGAAATAGCAAG
Gene specific oligo 1b	ATTTAGGTGACACTATA-GGGAAGTTATCCGGAGATGA-GTTTTAGAGCTAGAAATAGCAAG
Gene specific oligo 1c	ATTTAGGTGACACTATA-GGTAACTCCAGAGTCCAACA-GTTTTAGAGCTAGAAATAGCAAG
Gene specific oligo 2a	ATTTAGGTGACACTATA-ATCATGCGGGCCGCGCGCCACGG-GTTTTAGAGCTAGAAATAGCAAG
Gene specific oligo 3a	ATTTAGGTGACACTATA-GGCGCCACCATGTTAGACAA-GTTTTAGAGCTAGAAATAGCAAG
Constant Oligo	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAA-CTTGCTATTTCTAGCTCTAAAAC

The annealing protocol used, followed by an oligo-anneal program on the thermocycler :

1 µl Gene-specific Oligo (100 µM)
1 µl Constant Oligo (100 µM)
8 µl MilliQ H_2O
10 µl Total

95 °C	5 min
95-85 °C	-2 °C/sec
85-25 °C	-0,1 °C/sec
4 °C	Infinity

Further, a fill-in program utilizing T4 polymerase, was used to make a double-stranded DNA template:

0.2 µl 100X NEB BSA (20 mg/ml)

0.5 µl T4 NEB DNA Pol

2 µl 10X NEB Buffer2

2.5 µl dNTP (10 µM)

4.8 µl MilliQ H2O

10 µl Total

The reaction was incubated for 20 min at 12 °C.

The generated product was purified using Qiagen PCR Cleanup Kit as described by the manufacturer, however by adding 300 µl buffer PB to the PCR reaction mix and eluting in 30 µl MilliQ H_2O.

The size of the construct (reflecting the actual annealing of the oligos) was verified by 1 % agarose gel electrophoresis at 100V for 30 min. A GeneRuler DNA Ladder 0,5 µg/ µl (from Thermo-Scientific #SM0331) was used as a molecular weight (MW) marker.

In vitro transcription

To generate the sgRNA, the Ambion MEGA script SP6 Kit was utilized (on ice):

0,5 µl ATP
0,5 µl GTP
0,5 µl CTP
0,5 µl UTP
0,5 µl 10X Buffer
0,5 µl SP6 enzyme
1 µl Nuclease-Free water
1 µl Template (0,1 - 0,2 ug)
5 µl Total

The reaction was incubated O/N at 37 °C.

The template DNA was the next day, degraded by adding TURBO DNase from the kit to the transcribed sgRNA:

14 ul Nuclease-Free water
1 ul TURBO DNase
5 + 15 µl Total

and incubated for 15 min at 37°C, before the tube was placed on ice.

To purify the newly synthesized sgRNA template, Amicon Ultra centrifuge filter was used following the manufacturer's protocol and eluting in 10 µl RNase-free water, and the RNA concentration measured using Nanodrop.

2.7.3 Microinjection

To create CRISPR-mutants, microinjection was performed on one-cell stage fertilized zebrafish embryos. Prior to the injection, the injection mix was prepared as follows:

Injection mix:

1.5 µl gRNA (200 ng)
1 µl KCl (0.3 M)
0.6 µl Cas9 protein (600 ng)
6.9 µl MilliQ-water
10 µl Total

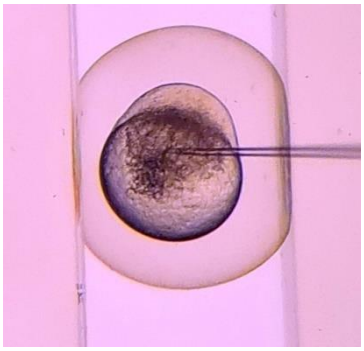


Figure 16: Microinjection of one-cell stage fertilized zebrafish embryo. Optimal injection technique injecting 1-2 nl into the yolk, close to the one-cell floor.

The mix was heated for 5 min at 37 °C and transferred to ice until use.

Newly fertilized zebrafish eggs were harvested and laid in an agarose (3%) wedged-shaped tray, formed with a plastic mold. The needles were prepared by heating capillary tubes (Borosilicate capillary with OD/ID: 1.0 mm/0.58 mm) in a Microcapillary puller at 98-100 °C and were filled with 3 µl injection mix using a 0.5-10 µl Eppendorf Microloader. After carefully attaching the filled needle into the injection apparatus (InjectMan4), the tip of the needle was carefully opened and 1-2 nl of the injection mix was injected into the yolk (close to the cell floor). The eggs were then transferred to a Petri dish, added E3-medium and bred at 28.5 C.

2.7.4 DNA extraction for genotyping

From whole embryos of WT and microinjected individuals

For DNA extraction of 48 hpf zebrafish embryos the Qiagen DNeasy Blood & Tissue Kit was used, following manufacturers protocol with minor changes:

The larvae were dechorionated and transferred to 1.5 ml microcentrifuge tubes on ice (one larvae in each tube). E3 medium was completely removed before adding 180 µl of ATL buffer and 20 µl Proteinase K (following Alt.1a for Tissue in the manufacturer's protocol), vortexed and incubated at 56 °C for 1.5 h until the larvae was completely lysed.

To extract the DNA, a mix of 200 µl AL buffer and 200 µl ethanol (96-100%) was then added, vortexed and spun down before transferring the reaction to a Dneasy Mini spin column placed in a 2 ml collection tube. Centrifugation was performed in an Eppendorf microcentrifuge at 8000 rpm for 1 min. Wash steps were performed as described in the manufacturer's protocol. The DNA was eluted twice for increased DNA yield in a 1.5 ml microcentrifuge tube reusing the same 50 µl AE-buffer.

The quality and concentration of the samples were measured and analyzed using Nanodrop, and stored at -20 °C until further use.

From tail-cut of WT and microinjected individuals

To be able to correlate genotype and phenotype, the tails of the larvae that were going to be fixed for IHC analysis were sampled before fixation. The 96 hpf larvae were euthanized with MS-222 and left in a Petri dish of ice water until the reflexes were gone. Then, the tail was cut direct posterior of the cloaca and anal fin without disturbing the rest of the larvae. The tail was transferred to a tube containing 20 µl 50 mM NaOH and the rest of the larva was transferred to another tube containing 4 % PF for fixation and IHC. The tubes were marked identically to trace which tail belongs to which individual. Tail-cuts were further heated to 95 °C in a PCR machine for 20 min and then vortexed to dissolve the tissue. If the tissue was not completely dissolved the process was repeated with 10 µl additional 50 mM NaOH and heated for 20 more min. When the tissue was completely homogenized, equal volume of 100 mM TrisHCl (pH 8.0) was added to the lysate. The genomic DNA was measured using NanoDrop and kept in the -20 °C freezer for storage.

2.7.5 Genotyping and T7 mutation screen

Advantage 2 PCR

Method 1: PCR-reaction was performed on the genomic area of all three exons to amplify the area around the target sequences, using gene-specific primers (See Table 6).
Method 2: PCR-reaction was performed on the genomic area of exon 1 surrounding the 1c target using gene specific primers with M13-tail.
PCR protocols described earlier (See section 2.3.1, *3.*).

Table 6: Primer table

Gene specific primer	**Primer sequence (5'-3')**
hnf1ba_DR_ex1a/b/c/d-F	CGCATTTCCCTCTCTAGACC
hnf1ba_DR_ex1a/b/c/d-R	GCTCTTTGAGAATTGGCGGT
hnf1ba_DR_ex2aF	GGGAACAATTGACCGCACTT
hnf1ba_DR_ex2aR	GGCACTGAAAACCGTCACAT
hnf1ba_DR_ex3aF	TTCCCGGCCACAATACTCATG
hnf1ba_DR_ex3aR	TTGAAACGGTTGCGTCTGAG
Hnf1ba_DR_ex1c-M13F	GTAAAACGACGGCCAG-CGCATTTCCCTCTCTAGACC
Hnf1ba_DR_ex1c-M13R	CAGGAAACAGCTATGACGCTCTTTGAGAATTGGCGGT

Amicon filter PCR purification and Gel electrophoresis

The PCR product was purified using Amicon Ultra centrifugal filters, 0.5 ml, 50 K (# UFC505096) from Merck Millipore according to manufacturer's protocol and eluted in 10 µl NF-water. To make sure the bands had the expected band lengths, 1µl of the PCR product in 4 µl MilliQ water was analyzed on a 1% agarose gel (1xTAE).

T7 endonuclease I Digestion assay

PCR products were subjected to the T7 assay, to screen for mutations in the amplified region. Initially, a short annealing program was run on thermocycler program OligoAnneal (described in section 3.5.2):

1 µl	Purified PCR product (200 ng)
16 µl	MilliQ water
2 µl	NEB buffer 2
19 µl	Total

After the annealing program 1 µl T7 endonuclease I (10 U, *NEB M0302L*) was added to the PCR product and incubated for 15 min at 37 °C.

The T7 endonuclease I cuts DNA fragments where mismatches are detected, and if the PCR product contains a mixture of mutated and WT sequences (mosaic F0 or heterozygous individuals), there will be mismatches after their hybridization. As a reult, one would expect three visible bands on an agarose gel: two digested bands representing the mutant alleles and one uncut band representing the WT allele.

Gel electrophoresis

To visualize the T7 digestion all 20 µl product was run on a 2.5 % agarose gel electrophoresis for 40 min at 100 V. Samples positive in T7-assay were further TOPO-cloned into pCR4 vector.

TOPO TA cloning pCR4 vector for sequencing

The TOPO TA pCR4 vector was used to clone PCR products to containing M13 sites as well as SP6 and T7 promotor sites. A pCR4 TOPO vector also contains restriction enzyme sites that can be utilized to cut out the insert and analyze the fragment size to confirm correct insert. In this method the pCR4 was used for preparation of PCR products for sequencing (M13-tail).

The same protocol was used as for TA cloning in the pCRII vector (section 2.3.1, *4.*).
Further, the transformation into competent cells, Colony PCR and liquid LB, as well as Miniprep plasmid extraction was done as previously described (See section 2.3.1, *5.*, *6.*, *7.*).

Sanger Sequencing

DNA sequencing was performed using BigDye v.3.1 (provided from the sequencing lab at UiB), primers used were either M13 or gene-specific primers containing M13 (See Table 6):

1-2 µl	Plasmid/PCR (200 ng)
1 µl	Sequencing buffer
1 µl	Sequencing primer (3.2 µM)
6-5 µl	MilliQ water
1 µl	BigDye v.3.1
10 µl	Total

Thermocycler sequence program used:

96 °C	5 min	
96 °C	10 sec	25 cycles
50 °C	5 sec	
60 °C	4 min	
4 °C	infinity.	

3 Results

The following section contains results from multiple protein alignment of HNFs, characterization of developing kidneys and pancreas in 24hpf embryo as well as 96 hpf larvae by IHC, spatial localization of *hnf1ba* and *gcga* by ISH, investigation and characterization of endocrine pancreas in CRISPR-Cas9 injected larvae and statistical analysis.

3.1 Multiple protein sequence alignment of HNFs.

The zebrafish protein sequences representing Hnf1ba and -bb were aligned to human HNF1B, mouse Hnf1b, and rat Hnf1b, to examine the degree of conservation. Compared to human HNF1B (557 aa), the Hnf1ba (557 aa) and Hnf1bb (532 aa) in zebrafish are 81 and 71 % identical, respectively. For the murines, rat (557 aa) and mouse (558 aa) Hnf1b are 97 and 96 % identical with HNF1B, respectively.

Comparing zebrafish Hnf1ba with human and murine sequences, the amino acid differences are divided into conservative, semi-conservative and non-conservative, based on their biochemical properties. The aligned domains show various degrees of conservation. The dimerization domain and the POU specific domain are highly conserved but contain a few conservative and semi-conservative amino acid replacements. A non-conservative amino acid replacement is observed, L98Q, indicating a Leucine (an aliphatic residue) replacing a Glutamine (an amide residue) in the zebrafish Hnf1ba. The Nuclear Localization Signal and POU homeodomain are also highly conserved (See Figure 17). The transactivation domain, however, shows numerous variations between zebrafish and humans, and non-conservative amino acid differences are observed throughout the transactivation domain.

When aligning zebrafish Hnf1ba versus zebrafish Hnf1bb to HNF1B, the protein sequence of Hnf1bb has larger variances compared to HNF1B, than Hnf1ba. Hnf1bb also contains a start codon 4 amino acids earlier than the other species. The two mutations related specifically to RCAD and MODY5 found in HNF1B are marked in red, R112fsdel and Q136X.

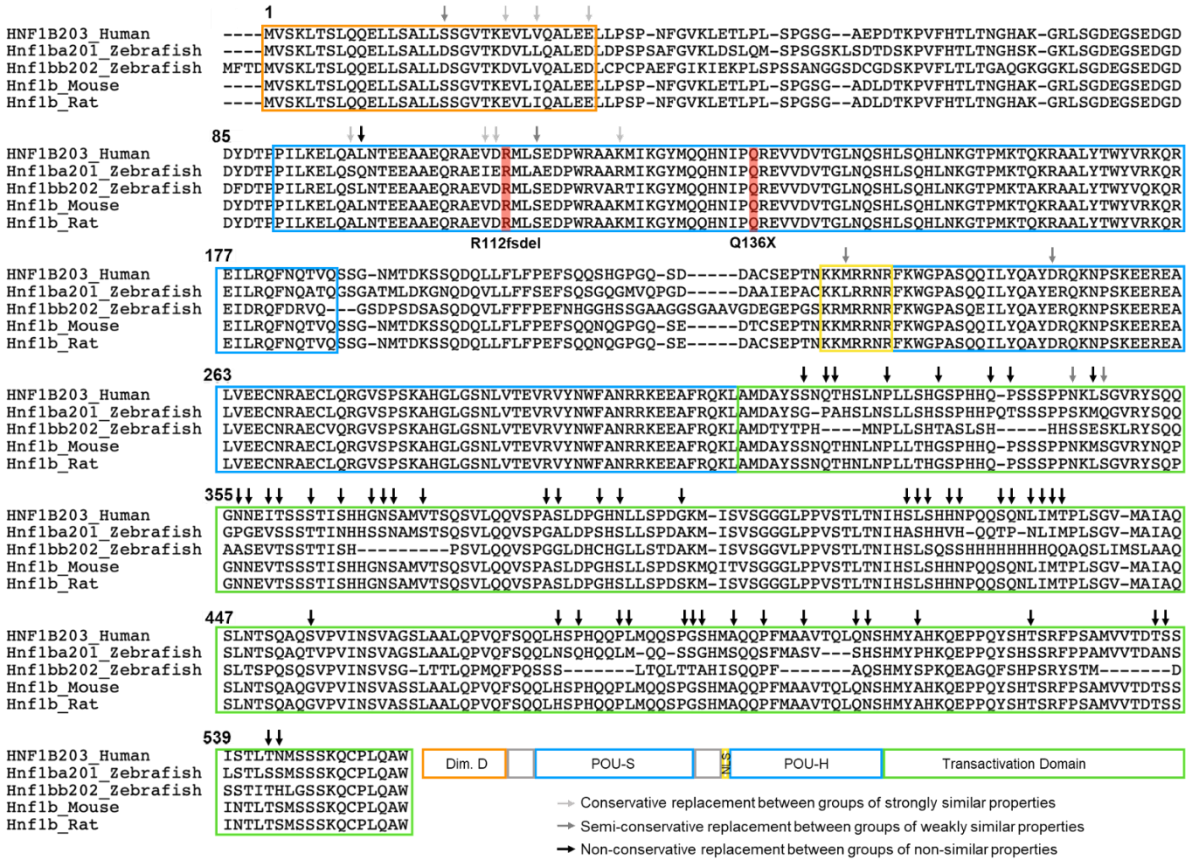


Figure 17: Multiple protein sequence alignment of human HNF1B to zebrafish Hnf1ba and Hnf1bb, mouse and rat Hnf1b. The different domains are colored, (Dimerization domain orange, POU specific and POU homeodomain are blue, Nuclear Localization Signal yellow, Transactivation domain green) and arrows indicate amino acid change of strongly (light grey arrows) and weakly (dark grey arrows) similar biochemical properties. Black arrows indicate amino acid substitutions of non-similar properties.

3.2 Developing kidneys and pancreas in 24 hpf transgenic zebrafish line Tg(wt1b:EGFP)

In order to increase our understanding of the development of MODY disease phenotype in *hnf1ba* mutants, we studied the transgenic zebrafish Tg(wt1b:EGFP) as a potentially valuable tool for studying kidney and pancreas development.

3.2.1 EGFP expression is localized in pronephros in 24 hpf zebrafish larvae

The Tg(wt1b:EGFP) zebrafish line has been generated for EGFP-expression in the developing kidneys and exocrine pancreas and has a promotor from *wilms tumor* gene (wt1b) that drives EGFP expression (Perner et al., 2007). Confocal imaging confirmed a strong EGFP signal in the pronephros of 24 hpf zebrafish embryos (See Figure 18). More specifically, EGFP expression can be observed in the glomerulus (G) leading to the neck (N) anterior, and further posterior in the proximal convoluted tubules (PCT) and the proximal straight tubules (PST).

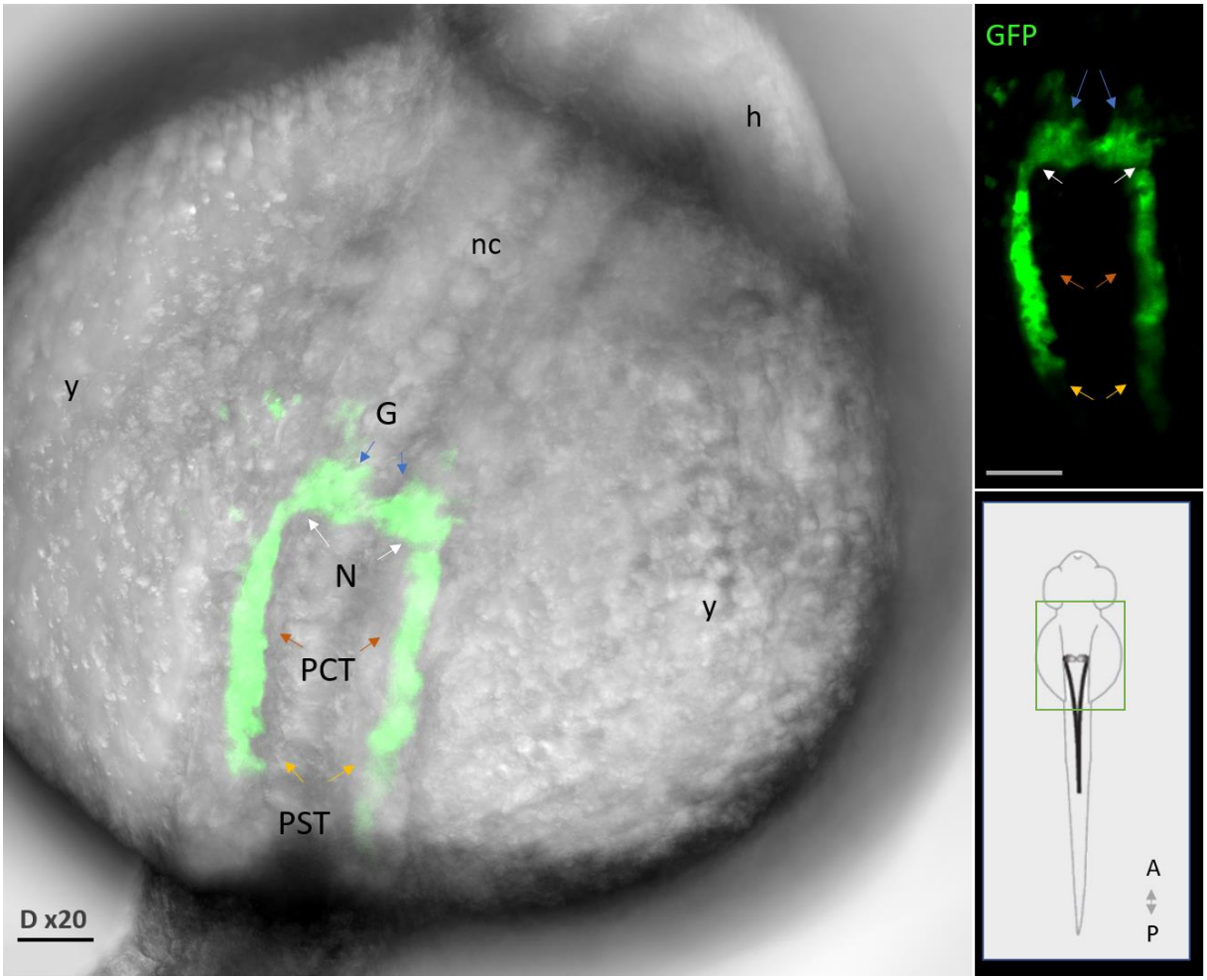


Figure 18: Location of glomerulus and pronephric duct in 24 hpf transgenic zebrafish larvae by EGFP expression. Tg(wt1b:EGFP) characterization of EGFP expression in 24 hpf larvae embedded in liquid agarose and MS222, for live imaging analysis. The pronephros are shown in green, and fluorescence was detected by confocal microscopy at 20x magnification. Blue arrows point to the glomerulus (G), white arrows to the necks (N), orange arrows point to the proximal convoluted tubules (PCT) and (posterior) the yellow arrows pointing to the two proximal straight tubules (PST). The illustration at the bottom right corner shows where the signal is located in the embryo. Dorsal view, with anterior up. Scale bar: 50 µm. h: head, nc: notochord, y: yolk.

3.2.2 Pancreatic progenitors in 24 hpf zebrafish larvae

To characterize the pancreatic organ at 24 hpf stage, Tg(wt1b:EGFP) embryos were double immunostained using glucagon and insulin specific antibodies. The results show a small structure of presumed pancreatic progenitors co-expressing both glucagon and insulin (See Figure 19).The cells have large nuclei, and the insulin staining appears stronger than that of glucagon. The pancreatic progenitor cell cluster is located on the right embryo side, anterior of the glomerulus (not depicted here, see previous section).

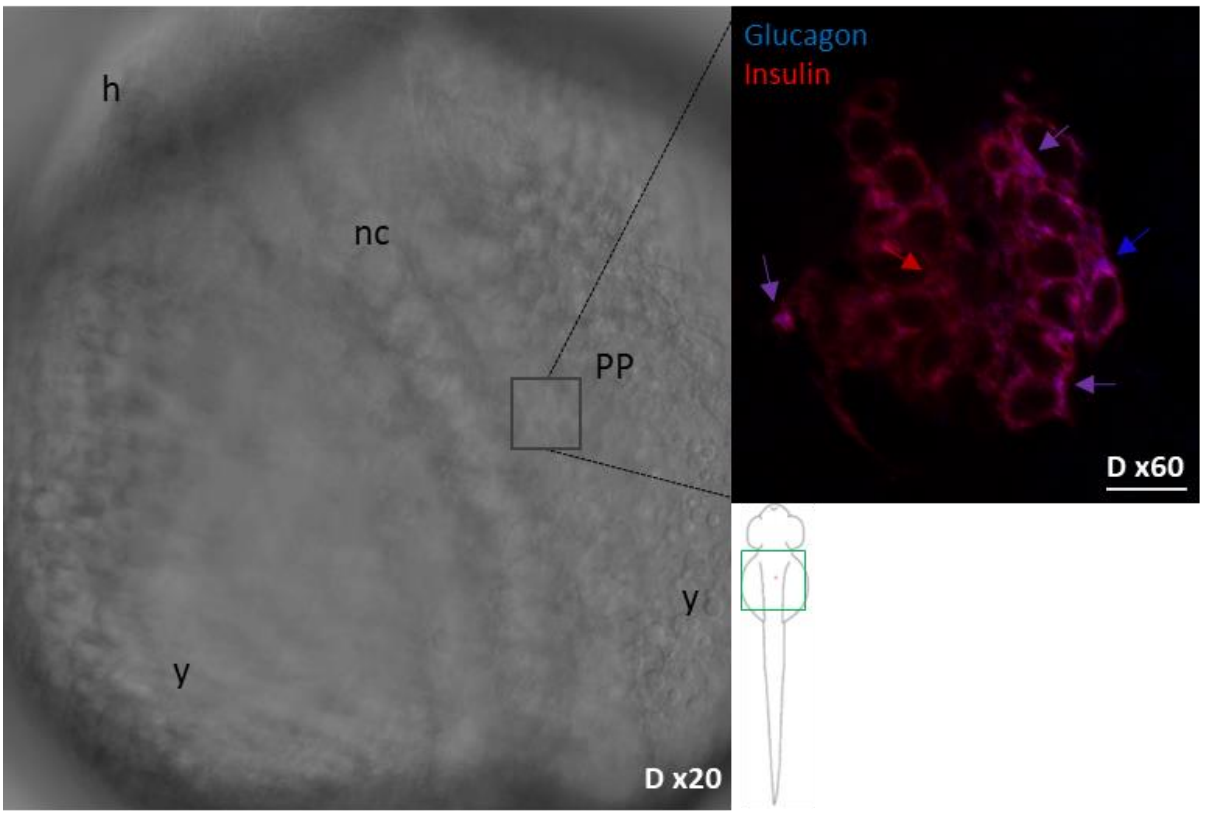


Figure 19: Location of pancreatic progenitors in 24 hpf zebrafish larvae. Embryos were immunostained using anti-insulin and anti-glucagon antibodies, and fluorescence confocal microscopy analysis revealed a cell cluster comprised of both insulin and glucagon producing cells. The co-localization is illustrated with arrows, where blue arrow points to glucagon producing cells, red arrow points to insulin producing cells and purple arrow indicates co-staining. Dorsal view with anterior up. Scale bar: 10 µm. PP: pancreatic progenitor, h: head, nc: notochord, y: yolk.

3.2.3 Localization pattern of *hnf1ba* expression in 24 hpf zebrafish larvae

For spatial localization of *hnf1ba* gene expression, *in situ* hybridization was performed on WT 24 hpf zebrafish larvae. Expression of *hnf1ba* was detected throughout the pronephros. The staining starts at the proximal convoluted tubule (PCT), all the way to the pronephric duct (PD), and further posterior to the cloaca (See Figure 20, B). When staining was analyzed in stereo microscope, the pronephric tubules were seen as two parallel stripes in anterior-posterior direction, also showing PCT, PST and PD (data not shown). At 24 hpf, the pancreas progenitors were not easily detectable, though the blue arrow indicates approx. where the structure of pancreas development starts (Ref. previous section). The sense probe showed no staining.

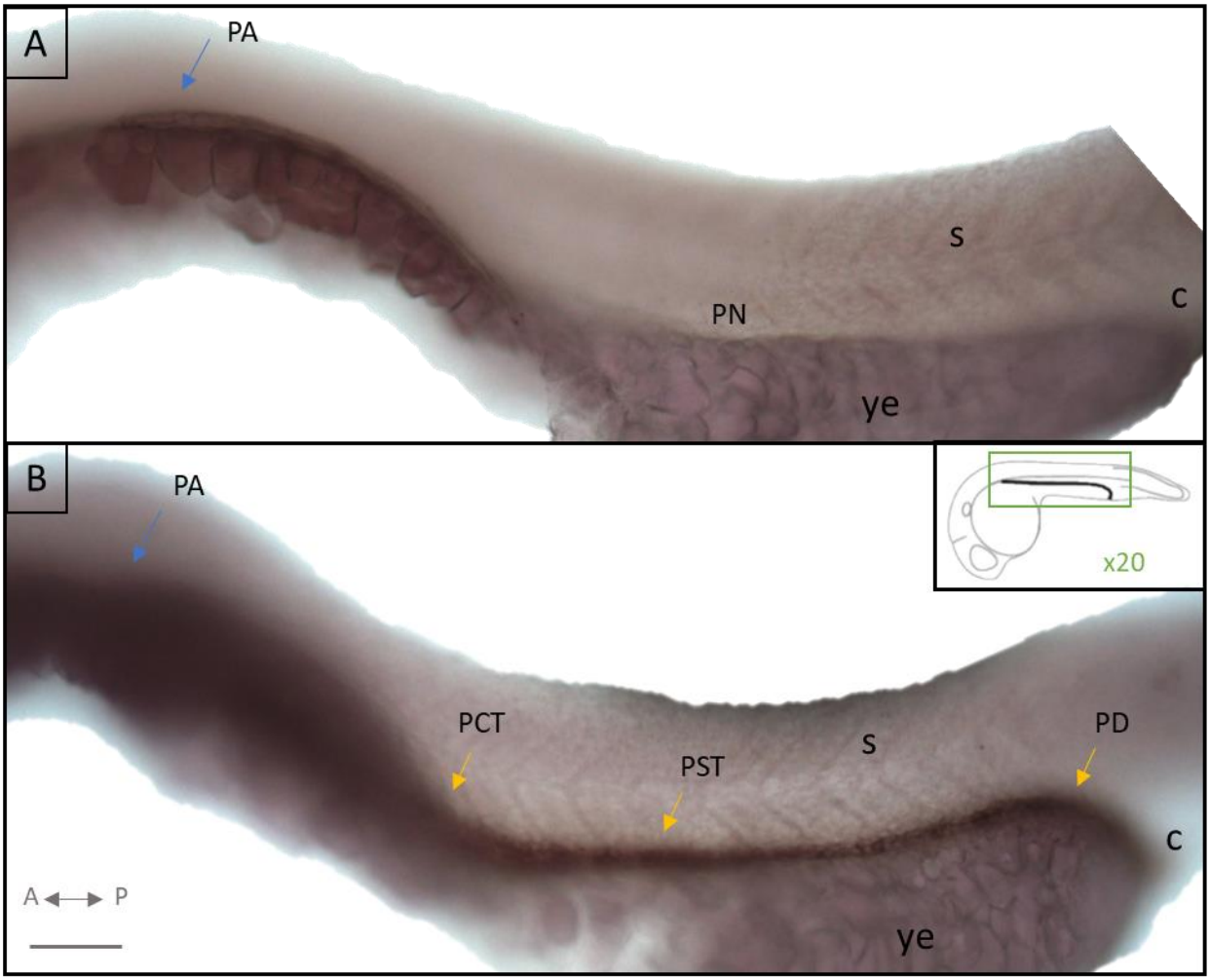


Figure 20: Expression of *hnf1ba* in 24 hpf zebrafish by *in situ* hybridization with RNA probe targeting *hnf1ba* and visualized with NBT/BCIP. Staining in the pronephros (PN). **A**: illustrates ISH in whole-mount using a sense probe, **B**: illustrates ISH in whole-mount using antisense probe. Bright-Field x20 magnification. Blue arrow indicates approx. where the pancreas progenitor (PA) is situated at 24 hpf. Yellow arrows show some of the different segments of the pronephros (PN). Lateral view with anterior to the left, yolk is partially removed. c: cloaca, PCT: proximal convoluted tubule, PST: proximal straight tubule, PD: pronephric duct, s: somites, ye: yolk extension. Scale bar: 50 µm.

3.3 Characterization of pancreas in 96 hpf WT zebrafish larvae

3.3.1 Localization of *hnf1ba* and *gcga* expression in 96 hpf zebrafish larvae

To study the localization of *hnf1ba* in more detail, *in situ* hybridization was conducted on 96 hpf WT zebrafish larvae. Frontal cryosection results show expression of *hnf1ba* (Figure 21 A,C) as well as *gcga* (Figure 21 D). The results show *hnf1ba* is expressed in the pronephros, intestine, and pancreas, and *gcga* is expressed in the pancreas. Further, in Figure 21 A, the expression of *hnf1ba* is observed on both sides of the larvae body, more specifically the neck of the pronephros. There is also weak staining in the liver. Figure 21 B shows the same WT individual, on a section more posterior (approx. 10 µm) of image A, with no detectable *gcga* expression (estimation of where the pronephros are situated are circled, though no pancreas is visible).

One section posterior to section in Figure 21 B, medial to the pectoral fin, Figure 21 C shows *hnf1ba* is still well expressed in the pronephros, most probably the proximal convoluted tubules, though there is also staining in the intestine (on the left) and pancreas (to the right, orange circle). Image D shows the cells expressing *gcga* and that they are organized in a circle on the right-hand side, ventral of the pectoral fin. In the two neighboring sections (Figure 21 C and D), a weak staining of *hnf1ba* (Figure 21 C) is observed in the same area as where *gcga* is stained in Figure 21 D. This weak staining of *hnf1ba* is believed to be the exocrine pancreas but might also include pancreatic endocrine tissue. Figure 21 E and F include images of sense probes of *hnf1ba* and *gcga*, respectively, which showed no detectable staining.

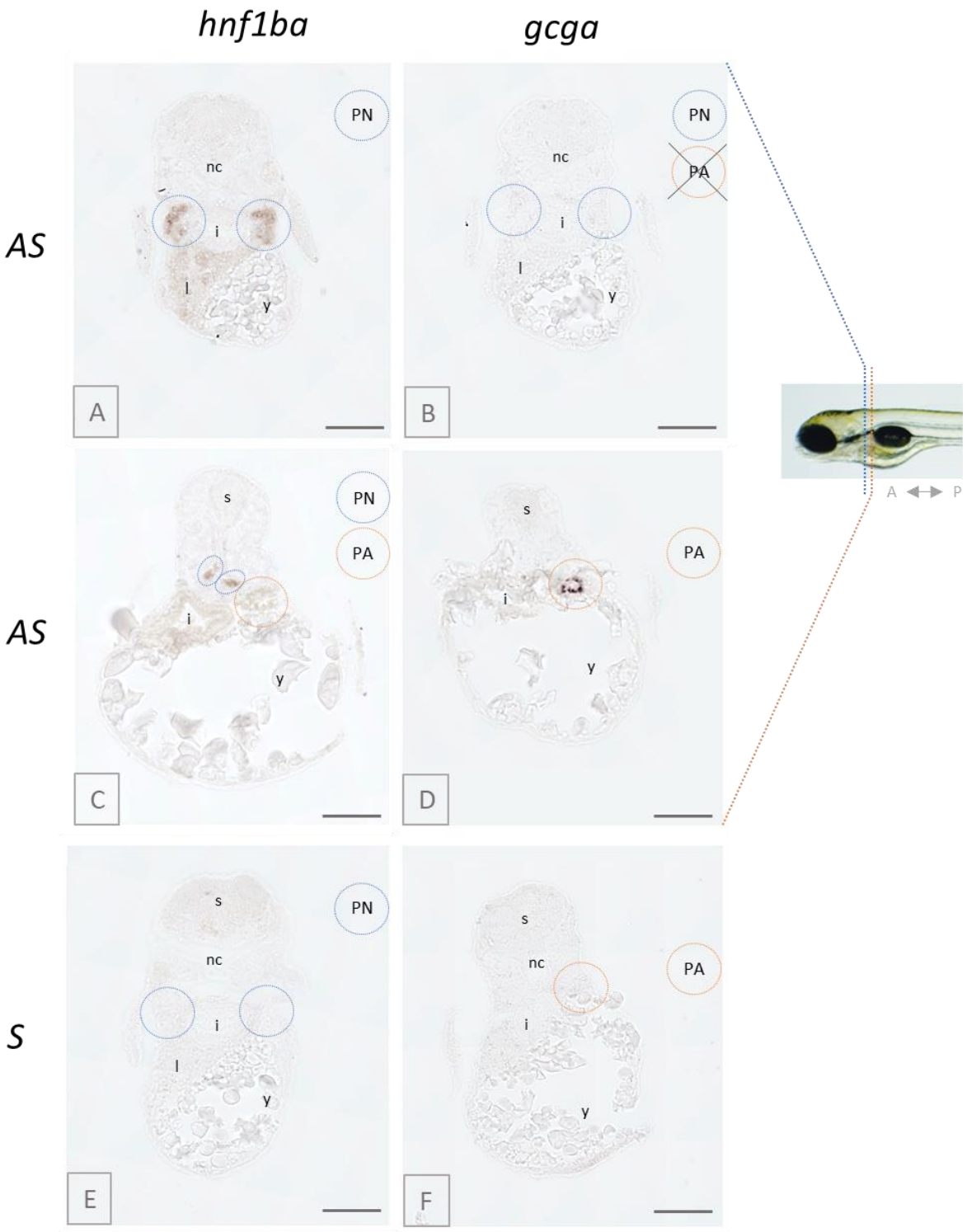
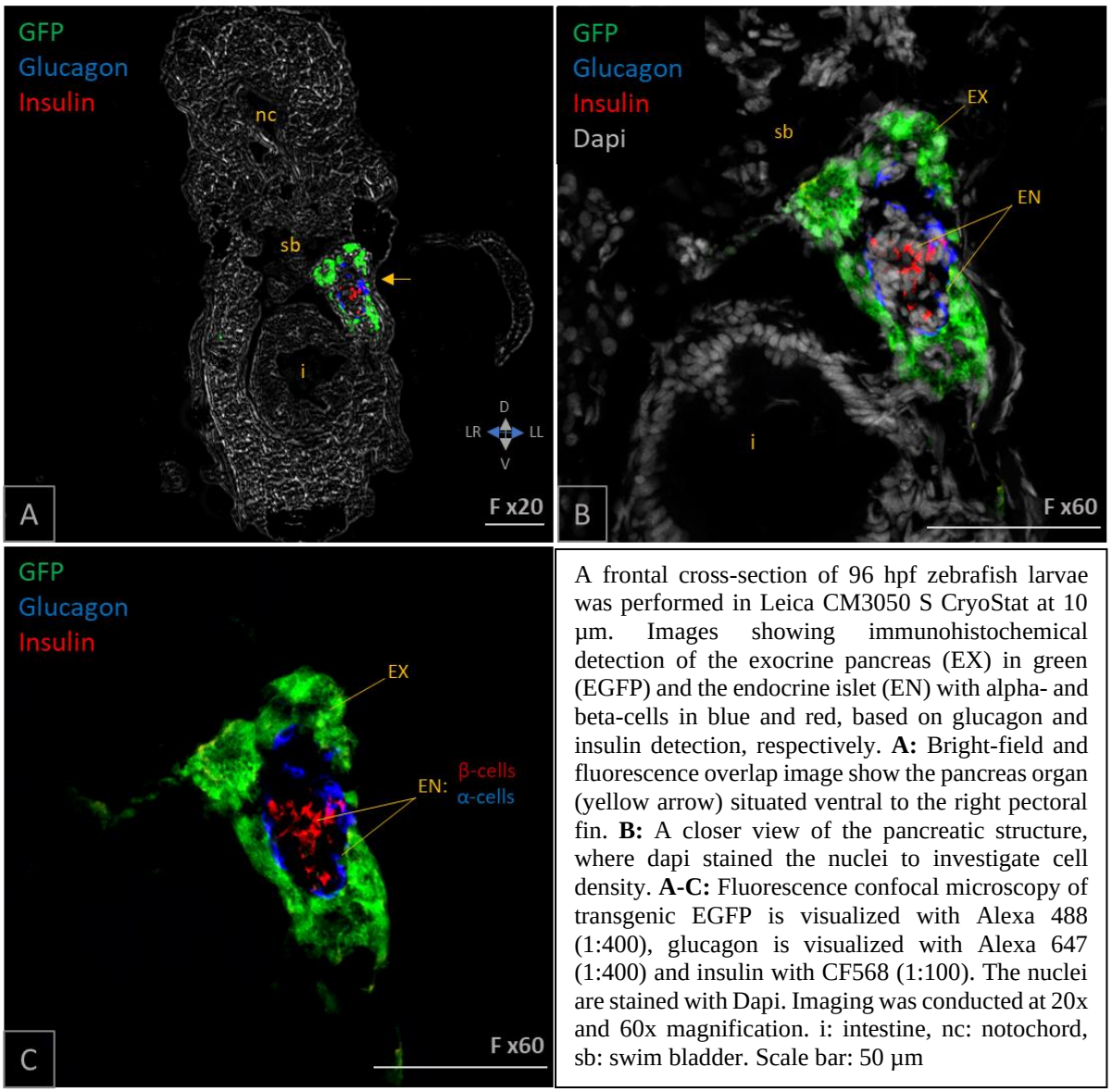


Figure 21: Localization of *Hnf1ba* and *Gcga* expression in 96 hpf zebrafish. Figure shows a lateral view of a 96 hpf zebrafish larvae. Blue stapled line indicates the section position in image A-B relative to the position in sections C-D which the orange line indicates. A-D: Expression of *Hnf1ba* and *Gcga* in 96 hpf zebrafish larvae detected by *in situ* hybridization (ISH) and stained with NBT/BCIP. A-B: good *hnf1ba* staining detected in the pronephros, more specifically the neck, and no *gcga* staining in the corresponding section. C-D: strong *hnf1ba* staining detected in the pronephros and weaker staining in the intestine and pancreas. Corresponding section show strong *gcga* staining in the pancreas. Both Antisense (AS, top pictures) and Sense (S, bottom pictures) probes are depicted. i: intestine, l: liver, nc: notochord, PA: pancreas, PN: pronephros, s: somites, y: yolk. Scale bar: 100µm.

3.3.2 Localization of pancreas in 96 hpf transgenic zebrafish Tg(wt1b:EGFP) cross section

Cryo cross sections of 96 hpf larvae were taken to investigate the endocrine pancreas as a whole structure in context to the EGFP labeled exocrine pancreas. For this, the endocrine cells were stained for glucagon and insulin, while Dapi stained the nuclei. Results show that the pancreas is situated on the right lateral side direct medial to the pectoral fin (See Figure 22, A). The exocrine pancreas surrounds the endocrine islet. The islet is oval shaped with the insulin producing cells at the center, and the glucagon producing cells surrounding the insulin-producing cells (See Figure 22, B-C).



A frontal cross-section of 96 hpf zebrafish larvae was performed in Leica CM3050 S CryoStat at 10 µm. Images showing immunohistochemical detection of the exocrine pancreas (EX) in green (EGFP) and the endocrine islet (EN) with alpha- and beta-cells in blue and red, based on glucagon and insulin detection, respectively. **A:** Bright-field and fluorescence overlap image show the pancreas organ (yellow arrow) situated ventral to the right pectoral fin. **B:** A closer view of the pancreatic structure, where dapi stained the nuclei to investigate cell density. **A-C:** Fluorescence confocal microscopy of transgenic EGFP is visualized with Alexa 488 (1:400), glucagon is visualized with Alexa 647 (1:400) and insulin with CF568 (1:100). The nuclei are stained with Dapi. Imaging was conducted at 20x and 60x magnification. i: intestine, nc: notochord, sb: swim bladder. Scale bar: 50 µm

Figure 22: Localization of exocrine and endocrine pancreatic tissue in 96 hpf zebrafish.

3.3.3 Pancreas in 96 hpf transgenic zebrafish Tg(wt1b:EGFP) whole-mount

At 96 hpf, the pancreas has grown to form a scoop-like shape at the right side of the body in zebrafish larvae (See Figure 23 A). At lateral view, the pancreas is visible at 96 hpf in larvae by formation of a head, body, and tail (from anterior towards posterior), and located directly ventral of the swim bladder.

Triple immunostaining with anti-GFP, anti-insulin and anti-glucagon antibodies, and imaging using fluorescence confocal microscopy, provides a clearer view of the pancreas structure at x60 magnification (See Figure 23 B)**.** The lateral right view of 96 hpf larvae (whole mount) reveals the endocrine part of the pancreas structure, containing only one islet of Langerhans. The endocrine cells are hidden inside the pancreas head, with alpha cells surrounding the beta cells, which are arranged in a cluster in the middle.

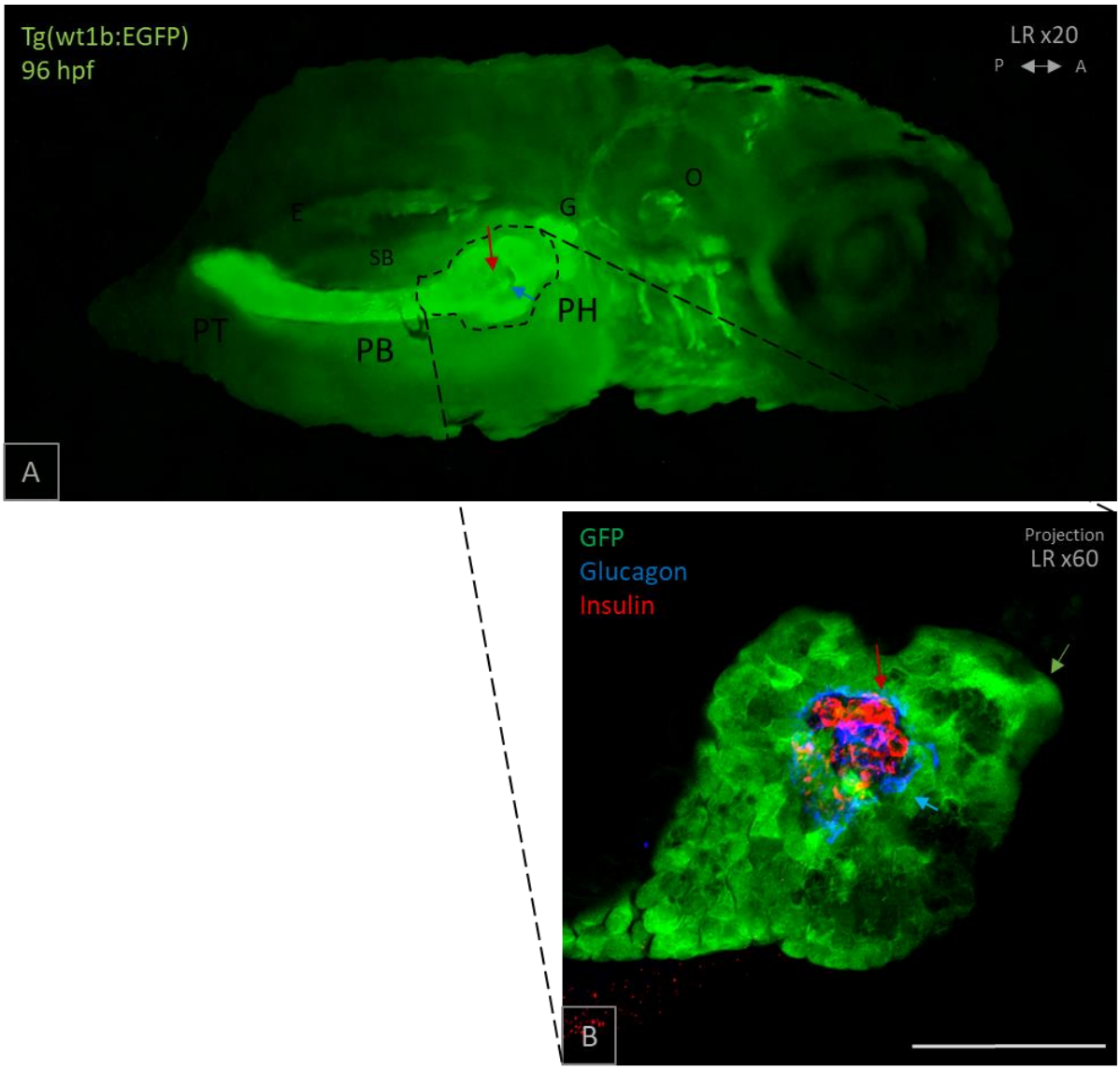
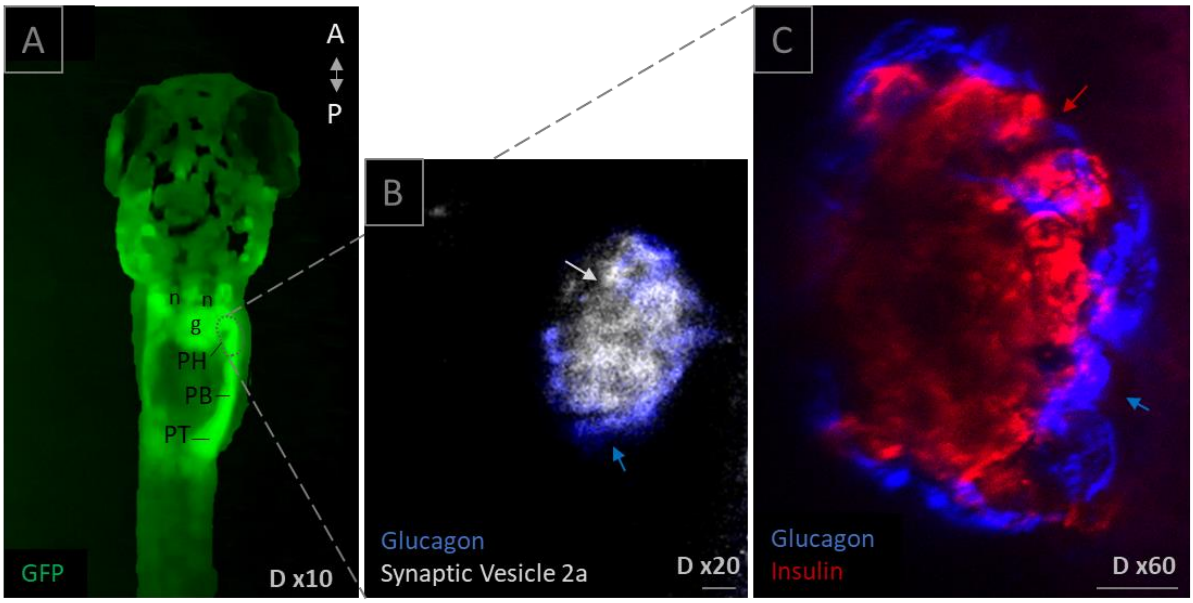


Figure 23: Pancreas location at 96 hpf in Tg(wt1b:EGFP) zebrafish larvae. A: EGFP expression in live 96 hpf larvae detected by fluorescence stereomicroscope at x20 magnification. In lateral view, the pancreas is situated on the right side of the body with the pancreas head (PH) anterior followed by the pancreas body (PB) and pancreas tail (PT), in a posterior direction. The glomerulus is anterior to the pancreas (G). The pancreas is situated ventral of the swim bladder (sb). o: otholite, e: eye. **B:** Zebrafish pancreas is shown from the right lateral side (whole mount). The exocrine pancreas is immunostained with Alexa 488 against EGFP, and the endocrine pancreas is immunostained with Alexa 647 against the glucagon-producing alpha cells and CF568 against the insulin-producing beta cells. The exocrine head is in focus, embracing the islet of Langerhans. The endocrine cells have smaller bodies with seemingly larger nuclei in the cell center, and more elongated cells surrounding these. Scale bar: 50 µm.

Dorsal view of 96 hpf transgenic zebrafish larvae shows the pancreas is situated on the right side of the larvae body, with pancreas head (PH) anterior, followed by pancreas body (PB) and tail (PT). The glomerulus (g) seems to be situated dorsally with the pronephric neck (n) anteriorly. To detect synaptic-like microvesicles and insulin granules in insulin-secreting cells of the endocrine pancreas, the larvae were double immunostained with anti-SV2a parallel to anti-glucagon and confirmed specific staining of both the anti-insulin and anti-glucagon (See Figure 24 B). The results show glucagon staining surrounding the synaptic vesicles similar to the anti-insulin. The endocrine part of the pancreas was immunostained with antibodies targeting both glucagon- and insulin-producing cells (See Figure 24 C). The results show that the glucagon-producing alpha-cells (blue) are surrounding the insulin-producing beta-cells (red) forming an oval cluster of endocrine cells in a dorsal view as well.



Figur 26: Dorsal location and characterization of the Langerhans islet in 96 hpf zebrafish.

A: The exocrine pancreas is situated on the right side of the zebrafish body, posterior to the pectoral fin at 96 hpf stage (circled), accompanied by the glomerulus (g) and pronephric neck (n) dorsoanterior. **B:** Image shows glucagon expressing alpha-cells (blue arrow) surrounding pre-synaptic vesicles (white arrow) at 20x magnification. **C:** Image shows glucagon expressing alpha-cells (blue arrow) surrounding insulin expressing beta-cells (red arrow) at 60x magnification. The double IHC is visualized using chicken anti-mouse Alexa 647 against anti-glucagon and donkey anti-rabbit CF568 against anti-insulin and anti-SV2a. Fluorescence confocal microscope images are z-sections from the projection series. **A-C:** Dorsal view with anterior up, scale bar: 10 µm.

Double immunostaining using fluorescent antibodies and confocal imaging on whole-mount WT zebrafish was done to investigate the average size of the endocrine pancreas organ by expression of insulin- and glucagon-producing cells. Figure 25 shows 3 replica WT islets of Langerhans typically with the glucagon-producing alpha-cells surrounding the insulin-producing beta-cells. Visually, the WT endocrine pancreas seems to be on average oval-shaped, containing a higher number of beta cells versus alpha cells. By comparing the three WT individuals, there are some degree of variations considering the total size of the islet, as well as when comparing alpha- and beta-cells ratio. To make an estimate of the organ structure, the area of both projections, as well as the central z-section image, was calculated using the ImageJ area measurement function after image adjustments (See section 3.4.4 and 3.6.1). In addition, dapi staining was used to count cells in the z-section image to obtain an average number of both alpha- and beta-cells in the structure (See Figure 25).

The images in Figure 25 (Z-projection), show area measurements performed on Z-projections of insulin- and glucagon-producing cell structures and also the total area of the structure. This type of imaging was used to compare the total structure composition of the WT endocrine pancreas. On average, the beta-cell structure is approx. 1100 μm^2 and the alpha-cell structure is approx. 1400 μm^2 for WT islets in zebrafish.

The Z-section results in Figure 25 (Z-section), show counted number of alpha- and beta-cells performed in the central section of the islet.

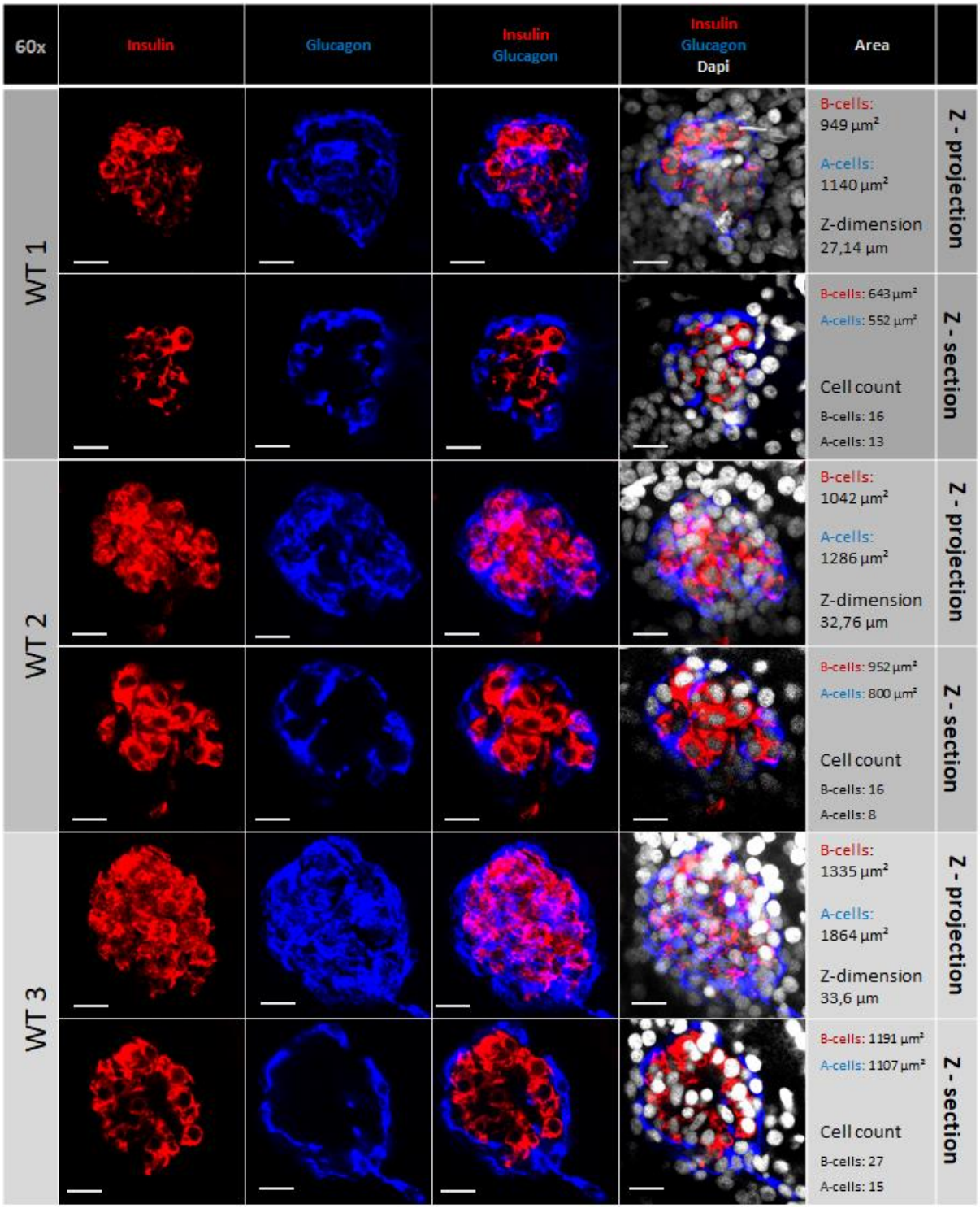


Figure 25: Characterization of WT endocrine pancreas. The beta-cells expressing insulin are immunostained with Red CF568 Chicken Anti-rabbit anti-insulin (red) and the alpha-cells expressing glucagon are immunostained with FarRed Alexa 647 Donkey Anti-mouse anti-glucagon (blue). Measurements of both beta- and alpha-cell areas are conducted in both projection as well as single image z-section. The total distance marked as Z-dimension shows the length of the structure in µm. Cell counts of both beta- and alpha-cells are conducted in the single image z-sections using dapi stained nuclei as markers. The images are performed on a fluorescence confocal microscope at 60x magnification. Scale bar: 10 µm.

Statistical analysis comparing area measurements of the Z-sections, shows a ratio of 9/8 in favor of beta-cells vs alpha-cells, which on average is around 11 % more beta-cells than alpha-cells in WT endocrine pancreas. By counting the number of cells in the most central z-section in the cluster projection, the ratio is slightly higher with an average of 38 % glucagon-producing cells versus 62 % insulin-producing cells (See Figure 25, Z-sections). The error bars indicate standard deviations, which describes a higher variation in beta-cell numbers (std.dev. 6.351) than alpha-cell numbers (std.dev. 3.606). The ratio between alpha- and beta-cells indicates that there is more beta- than alpha-cells in the islet cluster.

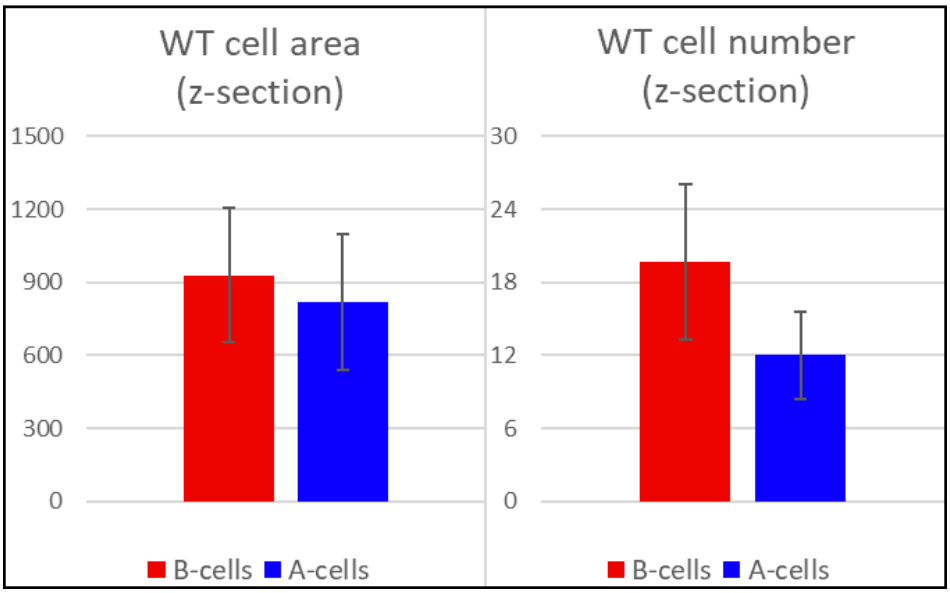


Figure 26: WT endocrine pancreas area and cell number based on alpha- and beta-cells. The columns represent beta-cells (B-cells for short) and the alpha-cells (A-cells for short) area and cell number ratios in WT individuals (n=3) measured on Z-sections. The standard deviations from average cell number calculation of alpha-cells were 3.606, and for the beta-cells 6.351. The beta-cells, which are seemingly clustered, have the largest area and cell number average (See also Figure 25).

3.4 Gene editing of the *hnf1ba* gene in zebrafish

In order to generate *hnf1ba* zebrafish mutants by CRISPR-Cas9 knockout, five different sgRNAs were tested to find the most specific and mutation efficient construct possible. The different sgRNAs were directed towards *hnf1ba* exon 1, 2 and 3 (See Figure 27).

Figure 27: Schematic presentation the *hnf1ba* gene, aligned to the sgRNA constructs. All five constructs were designed with the purpose to knock out the homeodomain by targeting exon 1, 2 or start of exon 3.

In Table 7, the calculated survival rates (numbers in percentage) of the transgenic Tg(wt1b:EGFP) line and the Casper line after microinjection are listed. The noninjected WT embryos demonstrated higher viability amongst the Casper embryos than the transgenic embryos, 85 and 54 % survival rate, respectively. The survival rate for NegCtr (injected with Cas9) was 36 % in both lines. Further, counted viable cells injected with construct 1c were 50 % of the Casper embryos versus 40 % of the transgenic embryos. For construct 3a, a survival rate at 61 % of the Casper embryos to 64 % of the transgenic embryos was shown. To summarize, the use of construct 3a resulted in an overall higher embryo survival rate than construct 1c. All surviving injected individuals at 96 hpf appeared to have had a normal development. This viability count was performed for each of the five constructs tested, and data can be viewed in the Appendix.

Table 7: Microinjection and embryo viability estimates. The CRISPR-Cas9 construct 1c and 3a was injected into transgenic as well as Casper embryos. Survival rate was calculated for WT (non-injected) embryos, negative control embryos injected only with Cas9 protein, and construct 1c and 3a injected embryos. Survival numbers are given in percent (%) of total embryos.

	Type	Total embryos	Healthy (24 h)	Dead (24 h)	Healthy (48 h)	Dead (48 h)	Survival rate (%)
Casper	Non-injected	13	7	6	7	0	54
	NegCtr (Cas9)	74	41	33	27	14	36
	sgRNA ex1c	98	73	26	49	24	50
	sgRNA ex3a	87	70	17	53	17	61
Tg(wt1b:EGFP)	Type	Total embryos	Healthy (24 h)	Dead (24 h)	Healthy (48 h)	Dead (48 h)	Survival rate (%)
	Non-injected	40	37	3	34	3	85
	NegCtr (Cas9)	50	37	13	18	19	36
	sgRNA ex1c	80	48	32	32	16	40
	sgRNA ex3a	50	38	12	32	6	64

3.4.1 Mutation screening of injected individuals by T7 assay

To identify mutation rates for sgRNAs, a T7 endonuclease assay based on amplified PCR product was performed. sgRNAs 1a, 1b, and 2a were initially found to not induce mutations with sufficient efficiency, while 1c and 3a were selected for further analysis. *Hnf1ba* exons 1 and 3, targets of sgRNAs 1c and 3a, were amplified using a high-fidelity polymerase. Further, the PCR products were analyzed by agarose gel electrophoresis and compared to WT PCR product of amplified *hnf1ba* exon 1 with expected size 338 bp, and 258 bp size for exon 3. Any additional gel bands of smaller (or larger) size, indicate T7 digestion of a heterozygous PCR fragment containing one mutated and one WT sequence, and an indication of successful CRISPR-Cas9 mutation. Figure 28 A Using construct 1c, gel electrophoresis and band analysis indicated that 10 out of 12 individuals contained mutations, which gives a rough estimate of 83 % mutation efficiency based on identification by the T7-assay (See bands indicated by a blue arrow). Figure 28 B For construct 3a, 11 out of 12 individuals indicated a mutated character, which gives a rough estimate of 91 % mutation efficiency based on the T7-assay (See bands indicated by a yellow arrow). The PCR product with the strongest mutated band intensity, for both constructs, was further TOPO-cloned for sequencing. More specifically for construct 1c, products based on individuals 1, 3, 6, 8, and 9 were chosen, and for construct 3a, products based on individuals 1, 4, 6, 10, and 11.

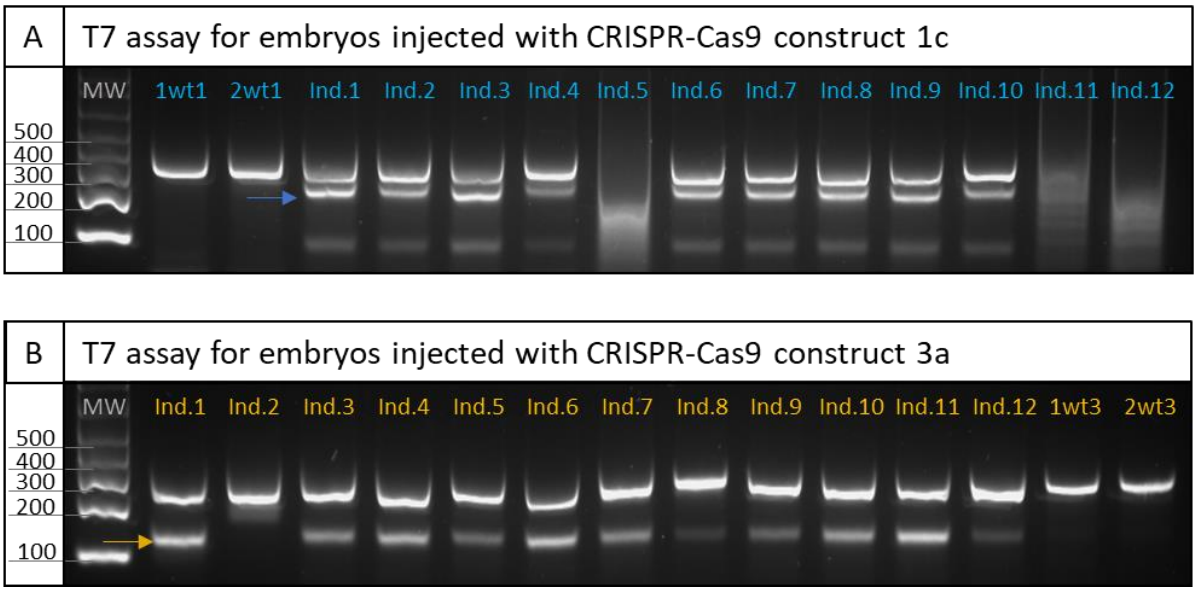


Figure 28: Agarose gel electrophoresis analysis of T7-assay treated PCR products based on different CRISPR constructs. A+B: 12 individuals injected with construct 1c (blue font) and 3a (yellow font) and 2 non-injected WT individuals, analyzed by agarose gel electrophoresis of amplified PCR products using primers targeting exon 1 (338 bp) and exon 3 (258 bp), respectively. Only one band was expected for WT samples, indicating a negative T7 assay, as well as no apparent polymorphisms. The individuals with the strongest mutated band were selected for further TOPO-cloning and transformation. **A:** T7 assay of construct 1c indicating Ind.1-10 as positive for mutations (indels) and Ind.11-12 are of unknown character. PCR products from individuals chosen for further cloning were Ind.1, Ind.3, Ind.6, Ind.8, and Ind.9. **B:** T7 assay of construct 3a showing Ind.1 and Ind.3-12 as positive for mutations (indels) and Ind.2 had seemingly only wild-type band size. The individuals chosen for further cloning were Ind.1, Ind.4, Ind.6, Ind.10, and Ind.11.

3.4.2 Characterization of mutation efficiency in injected embryos

Cloned PCR products, based on amplified *hnf1ba* exon 1 and 3 sequence, were analyzed by DNA sequencing to characterize the type of mutations induced by CRISPR-Cas9. (See Appendix for selected and mutated colonies). Sequence alignment and interpretation was performed using the Seaview5 analysis program. Construct 3a had 1 out of 10 clones that were mutated, hence an efficiency of 10 %. Table 8 shows, based on construct 1c, that 9 out of 10 clones were mutated, which gives a CRISPR-Cas9 construct efficiency of 90 %. Table 7 further highlights all the different indels formed in the respective mutated individuals, where 7 out of 10 clones harbored mutations that will potentially cause a frameshift in the genome.

Table 8: Sequence analysis of individual clones shows a high degree of mutations in CRISPR-Cas9 construct 1c injected embryos. Excerption of mutated sequences aligned to genomic *hnf1ba* exon 1 (stapled lines indicate left out sequence). The gRNA indicates where Cas9 is targeted to cut the genomic sequence. The protein effect describes what the induced mutation is expected to result in (amino acid deletion/frameshift mutation/early stop mutation). A total of 9 out of 10 clones were mutated.

Sample	Indels formed in CRISPR-Cas9 construct 1c injected embryos	Protein effect
Hnf1ba ex 1	-GCTTCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGATGTGCTTCTGCAAGCTTTGGAGGACCTGGAC- - - -	Normal (WT)
gRNA 1c	- - - - - - - - - - - - - - - - -TGTTGGACTCTGGAGTTACC- - - - - - - - -	
Individual 1		
Clone 1	-GCTTCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGA - -	Normal (WT)
Clone 2	-GCTTCTGAGCG - - - - - - - - - - CTCTGGAGTTACCAAAGA - -	Frameshift
Individual 2		
Clone 3	-GCTTCTGAGCGCCT- - - - GGACTCTGGAGTTACCAAAGA - -	Frameshift
Clone 4	-GCTTCTGAGCGCCT - - - - - - - - - - TGGAGTTACCAAAGA - -	Frameshift
Clone 5	-GCTTCTGAGCGCCTTGTnn- - - - - - -GGAGTTACCAAAGA - -	2 aa deleted, no frameshift
Individual 3		
Clone 6	-GCTTCTGAGCGCCT- - - - TGGACTCTGGAGTTACCAAAGA - -	1 aa deleted, no frameshift
Clone 7	-GCTTCTGAGCGCCTTGTTTGGACTCTGGAGTTACCAAAGA - -	Frameshift, early stop
Individual 4		
Clone 8	-GCTTCTGAGCGCCTTGTTAGGACTCTGGAGTTACCAAAGA - -	Frameshift, early stop
Clone 9	-GCTTCTGAGCGCC- -TTGAC - -	Frameshift
Clone 10	-GCTTCTG- - - - - - - - - - - - - ACTCTGGAGTTACCAAAGA - -	Frameshift

3.5 Analysis of zebrafish CRISPR injected larvae (F0 generation)

3.5.1 Indels formed in microinjected individuals

Exon 1 of tail-cut extracted DNA from CRISPR-Cas9 construct 1c injected individuals, was amplified by Adv.2 PCR and analyzed by the T7 assay. Amongst the individuals positive (mutated) for the T7 assay, the sequencing of PCR amplified product showed the presence of mutations in 22 out of 24 injected individuals (See Appendix Figure A2). In Table 9, a selection containing 6 of the 22 mutations positive individual sequences are aligned to the *hnf1ba* WT genomic sequence. Nucleotide sequence around the area of mutations shown as dotted lines or with the letter “n”, to highlight deletions and base changes, downstream of the sgRNA target sequence.

Table 9: Nucleotide sequence of tail-cut amplified exon 1 of CRISPR-Cas9 construct 1c injected F0 embryos. Excerption of mutated sequences aligned to genomic *hnf1ba* exon 1. The sgRNA sequence indicates where the Cas9 binds to genomic DNA. A total of 24 individuals were sequenced, though this excerpt includes the 6 individuals that were further analyzed for pancreas tissue profile by immunohistochemistry and confocal microscopy (See Figure 29 and Appendix). The WT and NegCtr contains no mutations in exon 1. All individuals have in exon 1, mutated sequences of one nucleotide upstream of PAM (TGG), though with different alterations. *n indicates two or more different nucleotides in the same position, due to mosaic effect in the same individual.

Sample	Indels formed in CRISPR-Cas9 construct 1c injected embryos
Hnf1ba ex 1	- - - -TCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGATGTGCTTCTGCAAGCTTTGGAGGACCTGGAC- - - -
gRNA 1c	- - - - - - - - - - - - - - - -TGTTGGACTCTGGAGTTACC- - - - - - - - -
WT	- - - -TCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGATGTGCTTCTGCAAGCTTTGGAGGACCTGGACC -
NegCtr	- - - -TCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGATGTGCTTCTGCAAGCTTTGGAGGACCTGGACC -
Individual 1	- - - -TCTGAGCGCCTTGnnnGnnnCnGGAnTTnnnnnnAGG - - nnnTTnnGnAnGnTTTGGAnnAnCnnnnC -
Individual 2	- - - -TCTGAGCGCCTTGnnnGnnnCnGGnnTnnCnAnGnnn - - TTnTTnnnCnnGnnnnGGAnnnnnnnCnnC -
Individual 3	- - - -TCTGAGCGCCTTGnnnnnnnnnnGnnnnTnnCAAAGA - - GnnnnnTnnnnnnnnnnnGGAnnnnnnnnn -
Individual 4	- - - -TCTGAGCGCCTTGnnnGnnTCnGGAnTTACCAAAnAT – GTGCTTCnGCAAGCTTTGGAnGACCnGGnCC -
Individual 5	- - - -TCTGAGCnCCTTGnCnGnnnnTnnnAnnnAnnnGC - - - - - - - TTCTGCTTnnnnnGGAnGnGnnGGACC -
Individual 6	- - - -TCTGAGCGCCTTGnnGGAnnCnGnnnnnnnnAAnAn - - nnnnnTnnnnnnnnnTTGGAnnnnnnnnnnC -

3.5.2 Characterization of the endocrine pancreas in mosaic 96 hpf CRISPR zebrafish larvae shows abnormal pancreas development

The results of mutated individuals 1-6, previously sequenced (See Table 9), immunostained with anti-insulin and anti-glucagon show that alpha- and beta-cells are both distinctly different between WT and mutants (See Figure 29). Different in terms of both the formation as well as the ratio between the two predominant endocrine cell types; alpha- and beta-cells. Figure 29 shows, from a lateral view, a comparison of the endocrine pancreas in WT and NegCtr to CRISPR-cas9 injected mutants, as well as individuals 1-6, and also includes the constituent nucleotide sequence and area measurements.

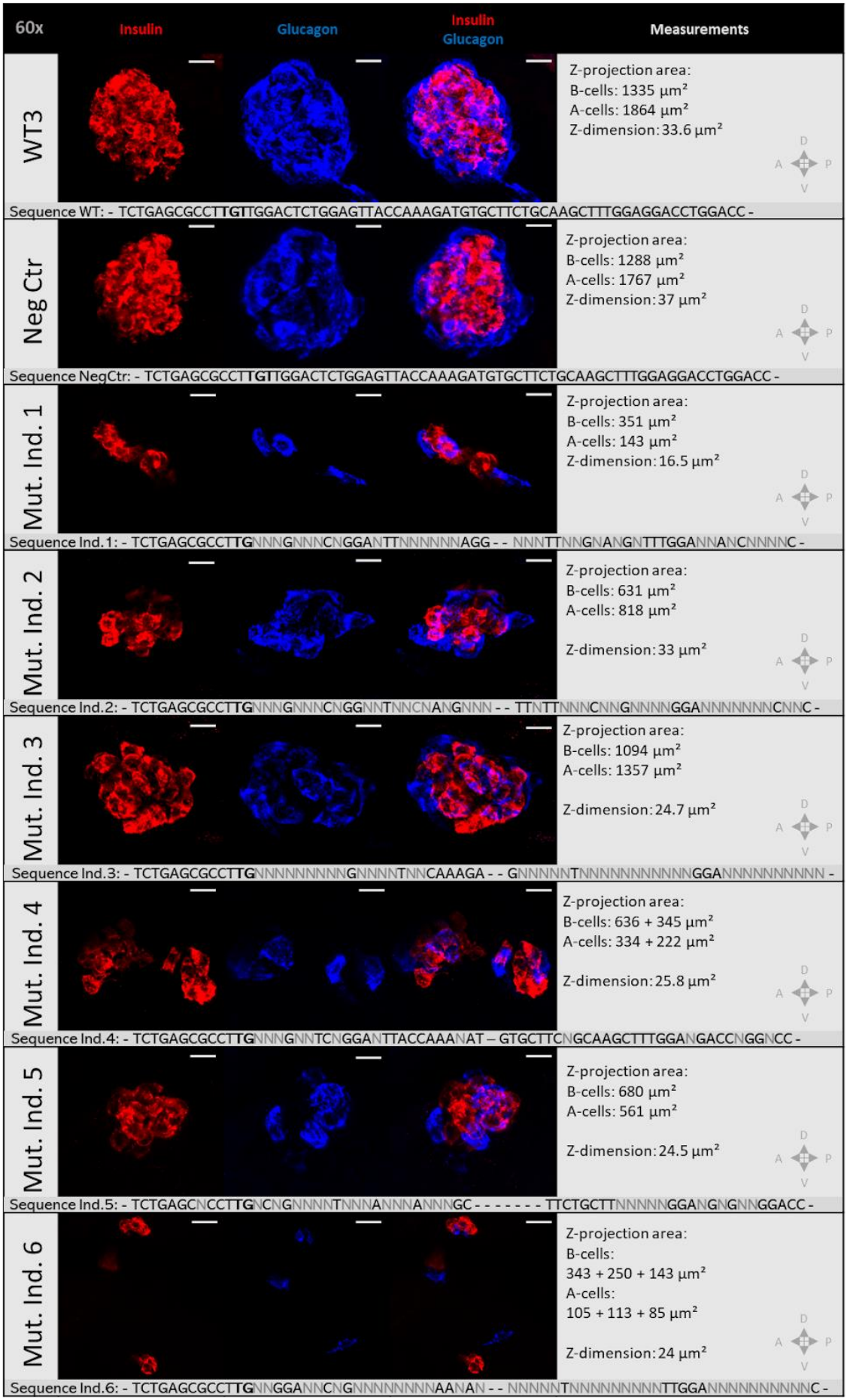

60x
Insulin
Glucagon
Insulin
Glucagon
Measurements
WT3
Z-projection area:
B-cells: 1335 µm²
A-cells: 1864 µm²
Z-dimension: 33.6 µm²
Sequence WT: - TCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGATGTGCTTCTGCAAGCTTTGGAGGACCTGGACC -
Neg Ctr
Z-projection area:
B-cells: 1288 µm²
A-cells: 1767 µm²
Z-dimension: 37 µm²
Sequence NegCtr: - TCTGAGCGCCTTGTTGGACTCTGGAGTTACCAAAGATGTGCTTCTGCAAGCTTTGGAGGACCTGGACC -
Mut. Ind. 1
Z-projection area:
B-cells: 351 µm²
A-cells: 143 µm²
Z-dimension: 16.5 µm²
Sequence Ind.1: - TCTGAGCGCCTTGNNNGNNNCNGGANTTNNNNNNAGG - - NNNTTNNGNANGNTTTGGANNANCNNNNC -
Mut. Ind. 2
Z-projection area:
B-cells: 631 µm²
A-cells: 818 µm²
Z-dimension: 33 µm²
Sequence Ind.2: - TCTGAGCGCCTTGNNNGNNNCNGGNNTNNCNANGNNN - - TTNTTNNNCNNGNNNNGGANNNNNNNCNNC -
Mut. Ind. 3
Z-projection area:
B-cells: 1094 µm²
A-cells: 1357 µm²
Z-dimension: 24.7 µm²
Sequence Ind.3: - TCTGAGCGCCTTGNNNNNNNNNGNNNNTNNCAAAGA - - GNNNNNTNNNNNNNNNNNGGANNNNNNNNNN -
Mut. Ind. 4
Z-projection area:
B-cells: 636 + 345 µm²
A-cells: 334 + 222 µm²
Z-dimension: 25.8 µm²
Sequence Ind.4: - TCTGAGCGCCTTGNNNGNNTCNGGANTTACCAAANAT – GTGCTTCNGCAAGCTTTGGANGACCNGGNCC -
Mut. Ind. 5
Z-projection area:
B-cells: 680 µm²
A-cells: 561 µm²
Z-dimension: 24.5 µm²
Sequence Ind.5: - TCTGAGCNCCTTGNCNGNNNNTNNNANNNANNNGC - - - - - - - TTCTGCTTNNNNNGGANGNGNNGGACC -
Mut. Ind. 6
Z-projection area:
B-cells:
343 + 250 + 143 µm²
A-cells:
105 + 113 + 85 µm²
Z-dimension: 24 µm²
Sequence Ind.6: - TCTGAGCGCCTTGNNGGANNCNGNNNNNNNNAANAN - - NNNNNTNNNNNNNNNTTGGANNNNNNNNNNC -

Figure 29: Endocrine pancreas show hypoplasia in mosaic 96 hpf CRISPR injected individuals. The beta-cells expressing insulin are immunostained with Red CF568 Chicken Anti-rabbit anti-insulin (red) and the alpha-cells expressing glucagon are immunostained with FarRed Alexa 647 Donkey Anti-mouse anti-glucagon (blue). The total distance marked as Z-dimension shows the length of the structure in µm. Measurements and cell counts of both beta- and alpha-cells are conducted in the single image z-sections using Dapi stained nuclei as markers for cell number. Below each individual image is the respective mutated sequence included. N indicate two or more different nucleotides in the same position. The images are performed on a fluorescence confocal microscope at 60x magnification. Scale bar: 10 µm.

To analyze the entire endocrine structure in WT and NegCtr individuals, the area measurements of projections show roughly 1300 μm^2 for the beta-cells and 1800 μm^2 for the glucagon-producing alpha-cells (See Figure 29). Compared to Mut. Ind. 1, the projection area of beta- and alpha-cells are 73 % and 92 % smaller, respectively, than average compared to WT and NegCtr set to 100 %. The difference is most severe regarding the beta-cells area, with just a few glucagon-producing cells scattered. Mut. Ind. 2 has a larger endocrine structure than Mut. Ind. 1, though compared to WT and NegCtr the beta-cell area is approx. 52 % smaller and the area of glucagon-producing cells are 55 % smaller. Mut. Ind. 3, on the other hand, has a much more normal sized endocrine structure than the other mutants, with a rather large beta-cell cluster, though not completely rounded. Comparing the beta-cell structure to the other WT individuals in Figure 25, no abnormality is observed or measured in this Mut. Ind. 3. Comparing the alpha-cells, a difference is however observed. The WT individuals have glucagon-producing cells surrounding the entire insulin-producing cell cluster, while Mut. Ind. 3 has only a small belt of glucagon-producing cells. Mut. Ind. 4 has formed two smaller clusters or islets, seemingly arising from the same root. The beta-cell structures are approx. 50 and 74 % smaller compared to WT and NegCtr average, considering the two islets separately. In total, the two clusters combined are only about 25 % smaller than a WT beta-cell structure and could be considered of rather normal areal size (See WT1, Figure 25). The alpha-cell area in Mut. Ind 4 are 82 and 88 % smaller (two clusters) than WT average, dispersed around the two beta-clusters. Mut. Ind. 5 has beta-cells in a round cluster, though 48 % smaller area than WT. When investigating the alpha-cells, which are 69 % smaller in size and differ from the WT and NegCtr, these are scattered and non-covering. Lastly, Mut. Ind. 6 has the entire endocrine pancreas scattered in three small clusters isolated from each other. Comparing the area measurements to WT and NegCtr, the three beta-cell structures of Mut. Ind. 6 are 74, 81 and 89 % smaller, and the alpha-cells are 94, 93 and 95 % smaller in size. Combined, the three islets make up around 57 % of insulin-producing cells, and 17 % of glucagon-producing cells, compared to WT (100%).

Another interesting observation is that the alpha-cells are not necessarily surrounding the beta-cells in the mutated individuals. The Z-dimensions are reduced in all of the mutated individuals compared to the WT3 and NegCtr, indicating smaller depth of clusters in total. Compared to the WT1 in Figure 25, some of the mutated individuals are of average depth. The nucleotide sequences of every individual indicate a large proportion of mutated sequence. To explore more closely, statistical analyses were performed on z-sections of the mutated individuals and compared to the WT and NegCtr.

3.5.3 Statistical analysis of alpha- and beta-cell numbers indicate significant pancreas hypoplasia in *hnf1ba* CRISPR injected zebrafish

To determine if the area and number of endocrine specific pancreatic cells are significantly different in the mutants compared to the WT, a student t-test was performed on average cluster/cell area and cell number. The groups have equal variance determined by f-test, and any p-value < 0.05 after paired t-test is marked with an asterisk indicating significance. In diagram A, alpha-cell area in WT versus mutant individuals were analyzed and indicate a significant difference in cell area (p:0.046). The standard deviation bars also indicate some variation within the different individuals, considering both WT and mutants alpha-cell area. In diagram B, alpha-cell numbers in WT and mutant individuals were also significantly different, showing a drastically lower number of glucagon-producing cells in the mutants compared to WT (p: 0.008). The standard deviation indicates less variations in the mutants than the WT. In diagram C, beta-cells area in WT and mutant individuals also show a drastic reduction in insulin-producing cells in the mutants, though this is not supported statistically (p: 0.148). The standard deviation is wide indicating a large spread of cell area amongst the mutated individuals. In diagram D, the beta-cells are counted in both WT and mutant central z-section and found to be statistically significant (p: 0.036).

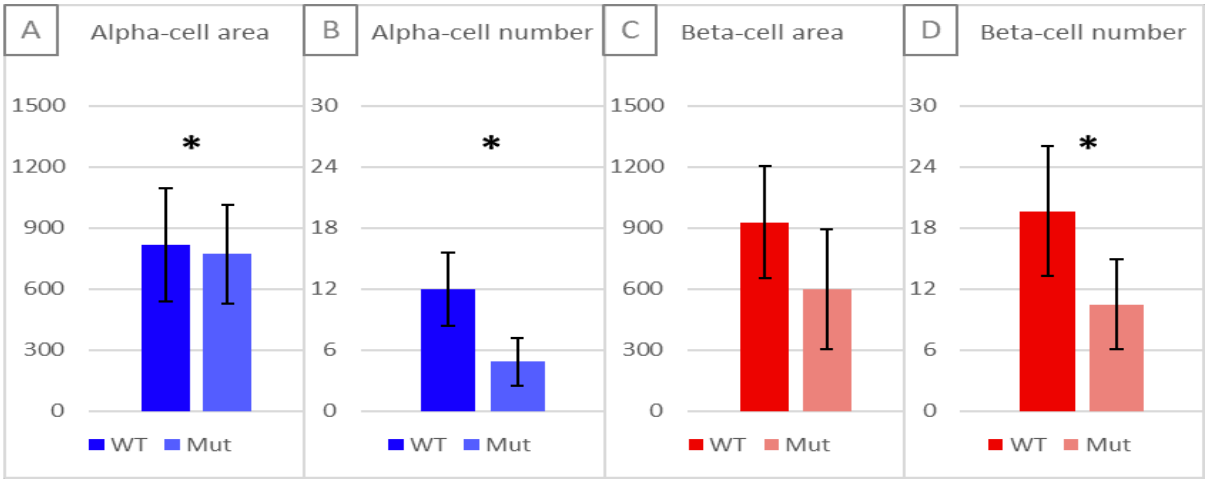


Figure 30: Statistical analysis of average alpha- and beta-cells area and numbers in WT and Mut. Ind. Student t-test analysis of average cell area and cell number indicates significant changes in WT versus mutated individuals. **A:** Show that average cell area in alpha-cells is significantly reduced in mutant individuals compared to WT (p: 0.046). **B:** Show that average alpha cell numbers are significantly reduced in mutant individuals compared to WT (p: 0.008). **C:** Show that average beta-cell areas are not significantly reduced between WT and mutant individuals (p: 0.148). **D:** Show, however, a significant reduction in average beta-cell number in mutants versus WT (p: 0.036).

4 Discussion

4.1 HNF protein domains and MODY associated mutations

In our study, the *hnf1ba* gene is knocked out after exon 1, which most likely will affect all Hnf1ba isoforms. To compare, a study investigating the human genotype-phenotype correlation of the closely related and HNF1B "functional partner" HNF1A, found that the prevalence of pathogenic missense *HNF1A* mutations associated with MODY3 disease, predominate in the dimerization domain and DNA binding domain, while pathogenic truncating mutations predominate in the transactivation domain, causing MODY3 by haploinsufficiency (Bellanne-Chantelot et al., 2008). Further, the majority of the pathogenic missense mutations were located in exon 1-6, affecting all three HNF1A isoforms, causing earlier onset and more severe phenotypes than those located in later exons (Harries, 2006).

To assess the conservation of orthologous protein sequence of Hnf1ba it was aligned to Hnf1bb (zebrafish), HNF1B (human), Hnf1b (mouse) and Hnf1b (rat). A high degree of conservation was seen for the dimerization domain as well as the POU homeodomain for all species. Hnf1bb has larger variances in the transactivation domain than the other orthologues, but the dimerization domain, as well as the two DNA binding domains are highly conserved. The transactivation domain is highly conserved between human, mouse and rat, however not for the two zebrafish paralogs. The associated mutations R112fsdel and Q136X, are located in a sequence area which were focused on in this study due to that they severely affect the HNF1B protein and is shown to cause severe phenotypes in patients (Haumaitre et al., 2006; Ulinski et al., 2006). The mutations are located in residues R112 and Q136, which are conserved in all species. R112fsdel result in the synthesis of a truncated HNF1B protein with an intact dimerization domain but lacking part of the POU_S domain and the entire POU homeodomain (POU_H), as well as the C-terminal transactivation domain (Haumaitre et al., 2006). The nonsense mutation Q136X, associated with early-onset diabetes as well as chronic renal failure, results in a premature stop in the HNF1B protein, and a protein lacking most of the DNA binding domain and the whole transactivation domain (Ulinski et al., 2006).

4.2 Characterization of developing kidneys and pancreas in WT 24 hpf zebrafish

4.2.1 Transgenic EGFP in zebrafish pronephros at 24 hpf, a valuable tool for future *hnf1b* knockout characterization in zebrafish

The transgenic Tg(wt1b:EGFP) zebrafish line has EGFP in the exocrine pancreas and kidneys which is expressed from the promoter of the *wilms tumor 1b* gene (*wt1b*). In 24 hpf larvae, the EGFP is expressed only in the pronephros (See Figure 18), while in 96 hpf the EGFP is also expressed in the pancreas (See Figure 21). The zebrafish pronephric structure shares structure conservation with the human nephron, which is comprised of the glomerulus, PCT, PST, DE (makes up the thick ascending limb in nephrons), DL (makes up the distal convoluted tubule in nephrons) and PD (makes up the collecting duct in nephrons) (Poureetezadi & Wingert, 2016). As developmental renal disease has been associated with *HNF1B* mutations in MODY5 patients, and inactivation of Hnf1b in murines has been shown to lead to tubular defects, the Tg(wt1b:EGFP) line will be a valuable tool in further investigations of the *hnf1ba* knockout effect to see correlating phenotypes between zebrafish and that reported in humans (Haumaitre et al., 2006).

4.2.2 Pancreatic progenitors expressing insulin and glucagon in 24 hpf zebrafish embryo

Immunostaining with anti-insulin and anti-glucagon in the Tg(wt1b:EGFP) at 24 hpf, show a small structure of presumed pancreatic progenitors co-expressing both glucagon and insulin (See Figure 19). Investigating the double immunostained islet cluster, there is evidence of both insulin- and glucagon-producing cells, as well as coexpression, which could indicate pancreas progenitor cells developing or differentiating. According to Kimmel and Meyer (Kimmel & Meyer, 2010), the endocrine precursors emerge within the pre-pancreatic endoderm during early somite stages, and insulin positive beta-cells increase in number to form a cluster in the anterior trunk region at around 24 hpf. The reason for the EGFP (in the exocrine pancreas) not being present, is due to the pre-pancreatic ventral bud not forming until 34 hpf. The ventral bud will then expand and surround the endocrine islet to form the head of the pancreas and further posterior to form the body and tail of the pancreas. Based on previous studies using ISH, Kinkel & Prince (Kinkel & Prince, 2009) found that the cells forming a cluster initially only contained beta-cells, however at 24 hpf, they also found alpha-cells and delta-cells, proposing that the

differentiated cells migrate to the dorsal side of the gut, forming the pancreatic islet. A different theory launched, is that endocrine progenitor cells originate from the epithelial ducts and express multiple hormones before becoming terminally differentiated. This has been proposed for both porcine (Lukinius et al., 1991), murines (Herrera et al., 1991) and humans (De Krijger et al., 1992; Lukinius et al., 1991), and as a common precursor stem cell for pancreatic hormone producing cells. Based on our study, cells expressing both insulin and glucagon in the outer layers of the cluster, could indicate a common precursor stem cell, first expressing insulin before differentiating into glucagon-producing cells, and hence support the latter theory.

4.2.3 Localization pattern of *hnf1ba* expression in 24 hpf zebrafish embryo

The essential transcription factor Hnf1ba is strongly linked to embryonic development of several organs, including the kidneys and pancreas. The *hnf1ba* gene expression in distinct organs of WT developing zebrafish embryos was thus visualized by ISH (See Figure 20). The paralogues of *hnf1ba,* which are *hnf1bb* and *hnf1a*, have an overall nucleotide identity of 66 and 47 %, respectively. If the overall percentage of identity is above 85 % the RNA probe used in ISH may cross hybridize to the paralogues, thus we expect no cross-reactions for our *hnf1ba* probe (Nielsen & Knudsen, 2002). The pronephros are described as segments containing the different parts illustrated in Figure 20. Rostral and central segment accounts for PCT, PST as well as Distal Early tubule (DE), and the caudal segment includes the Distal Late tubule (DL) and PD of the pronephric structure. We found that *hnf1ba* is expressed in the rostral, central and caudal segment of the pronephros, which correlates with Wingert & Davidson (2011) findings. They also characterized the expression of *hnf1ba* versus *hnf1bb* in zebrafish larvae. They found that the *hnf1ba* expression was located in the rostral, central and caudal segment of the pronephros, while hnf1bb was located in the rostral and central segment. This indicates that our *hnf1ba* probe is specific. As the pancreas is difficult to spot at 24 hpf, due to unspecific staining in the yolk as well as the small endocrine progenic structure at this larval stage, *hnf1ba* expression could not be assessed to target the pancreas at 24 hpf (Kimmel & Prince, 2009).

4.3 Characterization of developing kidneys and pancreas in WT 96 hpf zebrafish

4.3.1 Localization of *hnf1ba* and *gcga* expression in 96 hpf zebrafish larvae

The expression of *hnf1ba* were found located in the pronephros, intestine, liver and pancreas, while *gcga* were found located in the pancreas (See Figure 21). This indicates that the RNA probe targeting *gcga* is a highly specific probe. Considering the *hnf1ba* probe, our findings do correlate with previous studies of zebrafish pancreatic- and entero-endocrine cells that showed expression of *hnf1ba*, *hnf1bb* and *hnf1a* in the intestine and *hnf1ba* and *hnf1a* in the pancreas as well as in the pronephros and liver (Lancman et al., 2013; Lavergne et al., 2020; Naylor & Davidson, 2014). By assessing the time of which the staining occurred (24 hours), could indicate that *hnf1ba* is low expressed. It could also imply that the probe needs a further tRNA precipitation to be more effective. By comparing the two single-stained sections of the same individual (See Figure 21 C-D), the *hnf1ba* staining is overlapping near the pancreas area where the *gcga* is stained, which indicates that *hnf1ba* most likely also is expressed in the pancreas.

4.3.2 Cross section of 96 hpf zebrafish shows exocrine and endocrine pancreas structure

A cross section immunostained with anti-insulin, anti-glucagon and anti-GFP confirms the rounded shape of the endocrine islet, with insulin-producing beta-cells in the center forming a cluster, and the glucagon-producing alpha-cells surrounding it (See Figure 22). The exocrine pancreas is enclosing the islet in an elongated shape as shown by the EGFP expression in Tg(wt1b:EGFP). The Dapi staining provides visualization of higher cell density within the islet. It also shows cells in the section that are not stained with either glucagon or insulin that could indicate other cell types in the outer layer of the insulin cluster, such as delta-cells.

4.3.3 Characterization of endocrine pancreas in whole-mount 96 hpf zebrafish

The exocrine part of the zebrafish pancreas constitutes most of the organ at 96 hpf (See Figure 23). The endocrine part of the zebrafish, however, is locally restricted into one sole islet cluster. In our study, there are no evident signs of additional islets forming at 96 hpf larval stage, though into zebrafish adulthood, several more islets do appear (Tiso et al., 2009). The anatomy of the Langerhans islet includes forming a rounded cluster with insulin producing beta cells in the middle, and glucagon producing alpha cells as a layer around the beta cells. The islet is thus organized similarly to the mammalian islet. The exocrine pancreas embraces the islet forming a scoop like shape. Compared to the 24 hpf larvae, which basically only demonstrated insulin expressing beta-cells (and a few glucagon expressing alpha-cells), the exocrine pancreas and the rest of the glucagon expressing alpha-cells seems to have developed between 24 and 96 hpf. Acinar-cells have small nuclei and large cell bodies, which is what was observed at 96 hpf in zebrafish exocrine tissue in our study. The anti-SV2a targeted presynaptic vesicles in the endocrine cells, as done previously by Lancman et al. (Lancman et al., 2013), which confirms that the staining is specific to the endocrine beta-cells. We do not find that SV2a is expressed in the alpha-cells.

The results from characterizing WT endocrine pancreas shows consistency in higher numbers of beta cells vs alpha cells ratio (See Figure 23). The standard deviations indicate, however, a wide spread in WT cell type specific number, where a span from 16 to 27 beta-cells and 8 to 15 alpha-cells in the central Z-section, was observed. This could indicate the need to obtain more data on WT endocrine pancreas including several replicas. The cell specific and structural morphology of islets in the WT individuals are comparable. Comparing the Z-projections, all individuals have the same characteristic structure with a large, dense cluster of beta-cells surrounded by a thick layer of alpha-cells. The alpha-cells are mainly situated at the boundary of the Islet. The glucagon-producing alpha-cells are elongated in shape compared to the insulin-producing beta-cells, which are larger and more rounded, supported by comparing the Z-projection area ratio against the Z-section ration.

4.4 Gene editing the *hnf1ba* gene in zebrafish

4.4.1 Mutation screening of injected individuals by T7 assay

We detected a slight reduction in viability of both Casper and Tg(wt1b:EGFP) lines when injected with the different CRISPR constructs (NegCtr, 1c or 3a). The standard conditions for optimal breeding were upheld and the embryos were kept under the microscope for a minimum period of time. The injected embryos undergo a great deal of stress when the chorion and yolk are punctured. Further, the CRISPR-Cas9 complex could also contribute to the death rate observed (See Table 8). Since the survival rate between the injected and non-injected Tg(wt1b:EGFP) are comparable one can assume that neither construct have a particularly toxic effect. The same could apply for the Casper line as well, though embryos injected with construct 1c has a slightly less survival rate than those injected with construct 3a. In both zebrafish lines, the injected embryos with Cas9 protein do have the lowest survival rate of all, which could indicate some toxic effect of the Cas9 protein alone.

The T7 assay indicates the efficiency rate of mutations induced by the CRISPR-sgRNAs by digesting heterozygous PCR fragments containing both WT and mutated sequence. Construct 1c was estimated to induce mutations in 83 % and construct 3a was estimated to induce mutations in 91 % of embryos injected (See Figure 28). For construct 1c, the individuals analyzed had a stronger intensity mutant band, which could indicate more copies of the mutated sequence in the sample.

4.4.2 Characterization of mutation efficiency and type in CRISPR injected embryos

Results of the T7-assay indicates that only one of the tested sgRNA had the predicted and sufficient efficiency according to sequence analysis of cloned PCR products. Construct 1c were shown to cause mutations in 90 % of injected individuals, with a 70 % predicted ability to cause a frameshift mutation in 9 out of 10 clones in 4 different individuals (See Table 9). Construct 3a caused mutations in 1 out of 10 clones indicating a 10 % efficiency and hence not a sufficient number of mutants.

One major challenge concerning CRISPR-Cas9 gene editing, is to design a sgRNA that has high specificity and efficiency but without any off-target activity. We experienced difficulties concerning the *hnf1ba* gene in finding a suitable target site, hence some sgRNAs designed were

not successful. The editing efficiency of different sgRNA is predicted by the different sgRNA design software, but do not always reflect how they work in live embryos. One of the factors not considered by the prediction algorithms used in this experiment is chromatin structure. Our sgRNA was designed on a 1D model of the DNA laid out as a straight sequence. In reality, this strand of DNA is twisted and assembled in a 3D environment into the chromosomes found in the cell nucleus. Due to this natural conformation of genomic DNA, not all segments of the DNA strand are equally accessible to binding of Cas9/CRISPR complexes. Our sgRNAs could refer to an area in the zebrafish chromosome with a structure that prevents, or limits, the ability for the larger Cas9 complex to attach. The research on chromatin effects on sgRNA is just emerging, but there is growing evidence that chromatin state can affect editing efficiencies, with a preference for an open chromatin structure (Chuai et al., 2018; Verkuijl & Rots, 2019). Both specificity as well as efficiency was probably the issue concerning the sgRNA's 1a, 1b, and 2a that did not induce mutations in the zebrafish, or sgRNA 3a that was not efficient enough. Technical errors are also something to bear in mind when injecting embryos with CRISPR-Cas9 sgRNA, such as the amount of injection mix (too much could have toxic effects), or mechanical induced ruptures of the embryo. To increase the probabilities for a successful CRISPR induced mutation, Kroll et al.(Kroll et al., 2021), suggest using several sgRNAs simultaneously to target multiple sites in the target gene to assure nearly complete biallelic knockout in the injected embryo (90 % or more).

Characterizing CRISPR-Cas9 induced mutations by cloning and colony PCR is a useful way to investigate all the different mutations induced in one mosaic individual. This method of sequence preparation can also give an indication of the predicted outcome on the protein upon mutation type, to see if the mutations induced cause frameshift or nonsense mutations. The method is time-consuming, and not suited to screen many injected embryos. Another faster approach to determine successfully induced CRISPR-Cas9 mutation, and which could be used as a pre-scanning of relevant individuals before further TOPO-cloning, is to sequence the PCR product directly without cloning individual PCR fragments. By this, one will not be able to predict the specific mutation, though it is an indication of efficiency of the CRISPR construct.

4.5 Phenotype analysis of 96 hpf zebrafish CRISPR injected larvae (F0-generation)

4.5.1 CRISPR-Cas9 induced mutations show endocrine pancreas hypoplasia in mosaic 96 hpf zebrafish larvae

The CRISPR-Cas9 designed sgRNA 1c was utilized to target the *hnf1ba* gene exon 1 in transgenic Tg(wt1b:EGFP) zebrafish embryos. Based on the preliminary microinjection results, which showed a high mutation efficiency of the 1c sgRNA (90 %), individual injected embryos showed a high mutation frequency (22 out of 24 individuals, See Appendix). The results indicate that a mosaic knockout effect on the *hnf1ba* gene presumably can cause haploinsufficiency (See Figure 29 and Table 9). The difficulties in analyzing mosaic F0 individuals are the variable phenotypic penetration. However, mutants with an overt developmental defect can be identified in heterogeneous populations even if they are not complete knockouts. This indicates that the disruption of a single locus may be adequate as incomplete removal of WT allele does not impair detection of the phenotype (Burger et al., 2016; Kroll et al., 2021). MODY5 patients are displaying haploinsufficiency, hence a MODY5 phenotype might likely be caused in zebrafish even without a high degree of biallelic mutations.

The six different mutant individuals that were studied harbored different variations in the nucleotide sequence, and also demonstrated different degrees of pancreas hypoplasia (See Figure 29). Injected, developing embryos up to 96 hpf did not show any signs of morphological defects and appeared normal. Mut. Ind.1 was one of the individuals with the most severe hypoplasia of all individuals analyzed, demonstrating both alpha- and beta-cells, as well as the total islet structure dramatically different from WT and NegCtr. The endocrine pancreas experienced a loss of 73 % insulin-producing cells and 92 % of glucagon-producing cells. The diameter size was half of the WT, which confirms that the entire structure of the islet is smaller. Mut. Ind.2 and 3 also presented with pancreas hypoplasia, though not as severe as the first individual. Mut. Ind.4 harbored a type of branching of the endocrine pancreas, creating two coherent clusters. The glucagon- production in Mut.Ind.4, however, is most likely affected by the large reduction in alpha-cells in the structure (82 and 88 % decrease in the two branches) compared to WT. Mut. Ind.5 is rather similar to Mut. Ind.2 in structural size of Islet clusters, however with the beta-/alpha-cell area ratio reversed. This could indicate that the induced mutation has had a greater impact on the number of alpha-cells than the beta-cell formation. Mut. Ind. 6 has the entire endocrine pancreas fragmented in three small clusters isolated from

each other, with beta-cells in clusters and alpha-cells separated from the clusters. A near complete lack of alpha-cells was also observed (in total, 17 % glucagon-producing cells). Relevant for our mutants, Lancman et al. (Lancman et al., 2013) identified, in a hypomorphic *hnf1ba*^{S430} zebrafish mutant, pancreas hypoplasia and partial loss of Hnf1ba protein function. Both the exocrine and the endocrine pancreas were severely reduced, and they further found that Hnf1ba plays an essential role in regulating beta-cell number and pancreas specification (Lancman et al., 2013). We identified a significant reduction in beta-cell number, correlating with Lancman et al. (Lancman et al., 2013) findings, who found a significant correlation of reduced beta-cell number in *hnf1ba* mutant at 80 and 125 hpf. However, we did not find a significant reduction in beta-cell area in the mutants, though a trend showed reductions in size. On average, the results indicate a beta-cell reduction of approx. 50 % in the mutants compared to the WT. The standard deviations in beta-cell area are quite high due to the large spread in shapes and sizes in both WT and mutant which could explain why the difference is not significant. We found, in addition, a significant reduction in alpha-cell number as well as a significant reduction in alpha-cell area. This could indicate both a reduction in alpha-cell number but also in alpha-cell size and arrangement in the mutated individuals. The standard deviation in the alpha-cell area did indicate less variations in the mutants than the WT. A new aspect is the loss of alpha-cells, which are not mentioned as related phenotype caused by Hnf1ba loss of function and need to be further assessed.

As pancreas hypoplasia and pancreas dysfunction by reduced insulin secretion are phenotypes associated with MODY5 by *HNF1B* mutations, our *hnf1ba* gene targeted zebrafish models displaying pancreas hypoplasia may be a valuable tool for further generation of a stable F1 mutant of heterozygous *hnf1ba* alleles for studies related to MODY5 disease. If, however, the F0 generation displays *hnf1ba* mutant gonads, creating such a stable F1 mutant might be challenging.

5 Conclusion and future aspects

In order to improve our understanding of the causes for MODY5 disease phenotypes, a zebrafish MODY5 model could provide an important tool for identifying the underlying molecular disease mechanisms. In this thesis we investigated whether the zebrafish *hnf1ba* gene, exposed to CRISPR-Cas9 mediated knockout editing, could represent such a MODY5 animal model by exhibiting MODY5-like phenotypes. Initially, we detected *hnf1ba* expression in the pronephros, pancreas, liver and intestine. A further characterization of the endocrine pancreas identified a similar construction of the Langerhans islet to that seen in humans, with insulin-producing beta-cells forming a cluster and glucagon-producing alpha-cells surrounding the cluster. Lastly, we found that CRISPR-Cas9 microinjected zebrafish larvae did exhibit phenotypes such as pancreas hypoplasia, which could indicate haploinsufficiency. The F0-generation harboring *hnf1ba* mutations may, however, potentially be of mosaic phenotype since the degree and type of mutations varies between cells/tissue. Analyses of beta- and alpha-cell numbers were found to be statistically significantly reduced in the *hnf1ba*-edited individuals compared to WT individuals.

There are no reports on a clear phenotype-genotype correlation in human patients harboring *HNF1B* mutations. The reason for this is most likely due to few individuals studied/available, as MODY5 is a rarer form of MODY disease. Some of the mutant individuals presented in this thesis had severe phenotypes while others were mildly affected, in terms of pancreas hypoplasia. It is thus of importance to establish a stable F1-generation to further investigate mutation characteristics, *hnf1ba* expression and localization, and coherent cell/tissue specific phenotypes in both pancreas, kidney, and other *hnf1ba* expressing organs.

5.1 Future aspects

In order to establish a new F1 generation, our injected mosaic F0 individuals must be assessed for mutations in the gonads by genotyping F1 embryos. The F0-individuals will be 4 months old and mature, as of November 2021, and will be genotyped in order to establish a stable mutant line. Further, a double ISH (or FISH) could tell more whether *hnf1ba* and *gcga* are coexpressed, which could give some answers as to why the alpha-cells were seemingly more affected by the *hnf1ba* knockout. An interesting and informative method to further assess any colocalization in the pancreas, would be to perform a quadruple IHC/ISH, by combining FISH of *hnf1ba* expression and IHC of insulin, glucagon and EGFP (in the exocrine pancreas). There was also an indication that *hnf1ba* was expressed in the exocrine pancreas, as well as in the pronephros, which also needs further addressing due to the related human phenotypes of MODY5 patients. The value of having a transgenic zebrafish expressing EGFP in the kidneys allows for a deeper understanding of the organ development and if it is affected by *hnf1ba* haploinsufficiency similar to that seen in humans. Since *hnf1bb* is a paralogous gene connected to the HNF1B ortholog, it is of interest to assess if hnf1bb is co-expressed with hnf1ba, or if *hnf1bb* is expressed elsewhere. Of interest is also whether *hnf1bb* expression is affected by *hnf1ba* knockout. This applies also for the paralog *hnf1a, which* is an ortholog of *HNF1A*. HNF1A is known to dimerize with HNF1B in a huge network of transcription factors, this would be interesting to see if this applies also for the hnf1ba and hnf1a in zebrafish, and whether it is affected by hnf1ba knockout. Further, one could expand in the investigation of disease phenotype, possibly creating an hnf1bb knockout, as well as an hnf1a knockout with multiple combinations of the three mutants in stable F1 or F2-generations. Due to that MODY5 causes diabetes in humans, the MODY5 zebrafish models should also be investigated for impaired insulin secretion by glucose measurements to look for correlating diabetic effects in the stable F1/F2 generations.

Lastly, what would be interesting to know more about regarding the larval stage, would be the organogenesis of the pancreas. From MODY5 patients we assume HNF1B plays a major part in the developing embryo, leading to abnormal organogenesis resulting in pancreas hypoplasia. By mapping pancreatic structure development in zebrafish from late somite stage to 120 hpf, one could potentially see the development of hormone-producing progenitors and how they are affected by *hnf1ba* knockout.

6 References

Alestrom, P., D'Angelo, L., Midtlyng, P. J., Schorderet, D. F., Schulte-Merker, S., Sohm, F., & Warner, S. (2020). Zebrafish: Housing and husbandry recommendations. *Lab Anim*, *54*(3), 213-224. https://doi.org/10.1177/0023677219869037

Atkinson, M. A., Eisenbarth, G. S., & Michels, A. W. (2014). Type 1 diabetes. *Lancet*, *383*(9911), 69-82. https://doi.org/10.1016/S0140-6736(13)60591-7

Barbacci, E., Reber, M., Ott, M. O., Breillat, C., Huetz, F., & Cereghini, S. (1999). Variant hepatocyte nuclear factor 1 is required for visceral endoderm specification. *Development*, *126*(21), 4795-4805. https://www.ncbi.nlm.nih.gov/pubmed/10518496

Bellanne-Chantelot, C., Carette, C., Riveline, J. P., Valero, R., Gautier, J. F., Larger, E., Reznik, Y., Ducluzeau, P. H., Sola, A., Hartemann-Heurtier, A., Lecomte, P., Chaillous, L., Laloi-Michelin, M., Wilhem, J. M., Cuny, P., Duron, F., Guerci, B., Jeandidier, N., Mosnier-Pudar, H., . . . Timsit, J. (2008). The type and the position of HNF1A mutation modulate age at diagnosis of diabetes in patients with maturity-onset diabetes of the young (MODY)-3. *Diabetes*, *57*(2), 503-508. https://doi.org/10.2337/db07-0859

Bjorkhaug, L., Sagen, J. V., Thorsby, P., Sovik, O., Molven, A., & Njolstad, P. R. (2003). Hepatocyte nuclear factor-1 alpha gene mutations and diabetes in Norway. *J Clin Endocrinol Metab*, *88*(2), 920-931. https://doi.org/10.1210/jc.2002-020945

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L. M., Catena, R., Jinek, M., Robinson, M. D., & Mosimann, C. (2016). Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development*, *143*(11), 2025-2037. https://doi.org/10.1242/dev.134809

Chuai, G., Ma, H., Yan, J., Chen, M., Hong, N., Xue, D., Zhou, C., Zhu, C., Chen, K., Duan, B., Gu, F., Qu, S., Huang, D., Wei, J., & Liu, Q. (2018). DeepCRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biol*, *19*(1), 80. https://doi.org/10.1186/s13059-018-1459-4

Cubuk, H., & Yalcin Capan, O. (2021). A Review of Functional Characterization of Single Amino Acid Change Mutations in HNF Transcription Factors in MODY Pathogenesis. *Protein J*, *40*(3), 348-360. https://doi.org/10.1007/s10930-021-09991-8

Davis, T. M., Makepeace, A. E., Ellard, S., Colclough, K., Peters, K., Hattersley, A., & Davis, W. A. (2017). The prevalence of monogenic diabetes in Australia: the Fremantle Diabetes Study Phase II. *Med J Aust*, *207*(8), 344-347. https://doi.org/10.5694/mja16.01201

De Krijger, R. R., Aanstoot, H. J., Kranenburg, G., Reinhard, M., Visser, W. J., & Bruining, G. J. (1992). The midgestational human fetal pancreas contains cells coexpressing islet hormones. *Dev Biol*, *153*(2), 368-375. https://doi.org/10.1016/0012-1606(92)90121-v

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

DeFronzo, R. A. (1992). Pathogenesis of type 2 (non-insulin dependent) diabetes mellitus: a balanced overview. *Diabetologia*, *35*(4), 389-397. https://doi.org/10.1007/BF00401208

Dooley, K., & Zon, L. I. (2000). Zebrafish: a model system for the study of human disease. *Curr Opin Genet Dev*, *10*(3), 252-256. https://doi.org/10.1016/s0959-437x(00)00074-5

Edghill, E. L., Oram, R. A., Owens, M., Stals, K. L., Harries, L. W., Hattersley, A. T., Ellard, S., & Bingham, C. (2008). Hepatocyte nuclear factor-1beta gene deletions--a common cause of renal disease. *Nephrol Dial Transplant*, *23*(2), 627-635. https://doi.org/10.1093/ndt/gfm603

Eisenbarth, G. S. (2007). Update in type 1 diabetes. *J Clin Endocrinol Metab*, *92*(7), 2403-2407. https://doi.org/10.1210/jc.2007-0339

Ellard, S., Bellanne-Chantelot, C., Hattersley, A. T., & European Molecular Genetics Quality Network, M. g. (2008). Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young. *Diabetologia*, *51*(4), 546-553. https://doi.org/10.1007/s00125-008-0942-y

Fajans, S. S., Bell, G. I., & Polonsky, K. S. (2001). Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N Engl J Med*, *345*(13), 971-980. https://doi.org/10.1056/NEJMra002168

Feierstein, C. E., Portugues, R., & Orger, M. B. (2015). Seeing the whole picture: A comprehensive imaging approach to functional mapping of circuits in behaving zebrafish. *Neuroscience*, *296*, 26-38. https://doi.org/10.1016/j.neuroscience.2014.11.046

Florez, J. C. (2008). Clinical review: the genetics of type 2 diabetes: a realistic appraisal in 2008. *J Clin Endocrinol Metab*, *93*(12), 4633-4642. https://doi.org/10.1210/jc.2008-1345

Gong, H. Y., Lin, C. J., Chen, M. H., Hu, M. C., Lin, G. H., Zhou, Y., Zon, L. I., & Wu, J. L. (2004). Two distinct teleost hepatocyte nuclear factor 1 genes, hnf1alpha/tcf1 and hnf1beta/tcf2, abundantly expressed in liver, pancreas, gut and kidney of zebrafish. *Gene*, *338*(1), 35-46. https://doi.org/10.1016/j.gene.2004.05.003

Gresh, L., Fischer, E., Reimann, A., Tanguy, M., Garbay, S., Shao, X., Hiesberger, T., Fiette, L., Igarashi, P., Yaniv, M., & Pontoglio, M. (2004). A transcriptional network in polycystic kidney disease. *EMBO J*, *23*(7), 1657-1668. https://doi.org/10.1038/sj.emboj.7600160

Haldorsen, I. S., Vesterhus, M., Raeder, H., Jensen, D. K., Sovik, O., Molven, A., & Njolstad, P. R. (2008). Lack of pancreatic body and tail in HNF1B mutation carriers. *Diabet Med*, *25*(7), 782-787. https://doi.org/10.1111/j.1464-5491.2008.02460.x

Harjutsalo, V., Lammi, N., Karvonen, M., & Groop, P. H. (2010). Age at onset of type 1 diabetes in parents and recurrence risk in offspring. *Diabetes*, *59*(1), 210-214. https://doi.org/10.2337/db09-0344

Harries, L. W. (2006). Alternate mRNA processing of the hepatocyte nuclear factor genes and its role in monogenic diabetes. *Expert Rev Endocrinol Metab*, *1*(6), 715-726. https://doi.org/10.1586/17446651.1.6.715

Hattersley, A., Bruining, J., Shield, J., Njolstad, P., & Donaghue, K. C. (2009). The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatr Diabetes*, *10 Suppl 12*, 33-42. https://doi.org/10.1111/j.1399-5448.2009.00571.x

Haumaitre, C., Fabre, M., Cormier, S., Baumann, C., Delezoide, A. L., & Cereghini, S. (2006). Severe pancreas hypoplasia and multicystic renal dysplasia in two human fetuses carrying novel HNF1beta/MODY5 mutations. *Hum Mol Genet*, *15*(15), 2363-2375. https://doi.org/10.1093/hmg/ddl161

Herrera, P. L., Huarte, J., Sanvito, F., Meda, P., Orci, L., & Vassalli, J. D. (1991). Embryogenesis of the murine endocrine pancreas; early expression of pancreatic polypeptide gene. *Development*, *113*(4), 1257-1265. https://www.ncbi.nlm.nih.gov/pubmed/1811941

Herrgen, L., Ares, S., Morelli, L. G., Schroter, C., Julicher, F., & Oates, A. C. (2010). Intercellular coupling regulates the period of the segmentation clock. *Curr Biol*, *20*(14), 1244-1253. https://doi.org/10.1016/j.cub.2010.06.034

Hojny, J., Bartu, M., Krkavcova, E., Nemejcova, K., Sevcik, J., Cibula, D., Fryba, V., Plincelnerova, L., Dundr, P., & Struzinska, I. (2020). Identification of novel HNF1B mRNA splicing variants and their qualitative and semi-quantitative profile in selected

healthy and tumour tissues. *Sci Rep, 10*(1), 6958. https://doi.org/10.1038/s41598-020-63733-x

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. C., Koch, R., Rauch, G. J., White, S., . . . Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498-503. https://doi.org/10.1038/nature12111

Johansson, B. B., Irgens, H. U., Molnes, J., Sztromwasser, P., Aukrust, I., Juliusson, P. B., Sovik, O., Levy, S., Skrivarhaug, T., Joner, G., Molven, A., Johansson, S., & Njolstad, P. R. (2017). Targeted next-generation sequencing reveals MODY in up to 6.5% of antibody-negative diabetes cases listed in the Norwegian Childhood Diabetes Registry. *Diabetologia*, *60*(4), 625-635. https://doi.org/10.1007/s00125-016-4167-1

Jorgens, K., Hillebrands, J. L., Hammes, H. P., & Kroll, J. (2012). Zebrafish: a model for understanding diabetic complications. *Exp Clin Endocrinol Diabetes*, *120*(4), 186-187. https://doi.org/10.1055/s-0032-1304565

Kamel, M., & Ninov, N. (2017). Catching new targets in metabolic disease with a zebrafish. *Curr Opin Pharmacol*, *37*, 41-50. https://doi.org/10.1016/j.coph.2017.08.007

Kim, E. K., Lee, J. S., Cheong, H. I., Chung, S. S., Kwak, S. H., & Park, K. S. (2014). Identification and Functional Characterization of P159L Mutation in HNF1B in a Family with Maturity-Onset Diabetes of the Young 5 (MODY5). *Genomics Inform*, *12*(4), 240-246. https://doi.org/10.5808/GI.2014.12.4.240

Kimmel, R. A., & Meyer, D. (2010). Molecular regulation of pancreas development in zebrafish. *Methods Cell Biol*, *100*, 261-280. https://doi.org/10.1016/B978-0-12-384892-5.00010-4

Kinkel, M. D., & Prince, V. E. (2009). On the diabetic menu: zebrafish as a model for pancreas development and function. *Bioessays*, *31*(2), 139-152. https://doi.org/10.1002/bies.200800123

Kroll, F., Powell, G. T., Ghosh, M., Gestri, G., Antinucci, P., Hearn, T. J., Tunbak, H., Lim, S., Dennis, H. W., Fernandez, J. M., Whitmore, D., Dreosti, E., Wilson, S. W., Hoffman, E. J., & Rihel, J. (2021). A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *Elife*, *10*. https://doi.org/10.7554/eLife.59683

Lancman, J. J., Zvenigorodsky, N., Gates, K. P., Zhang, D., Solomon, K., Humphrey, R. K., Kuo, T., Setiawan, L., Verkade, H., Chi, Y. I., Jhala, U. S., Wright, C. V., Stainier, D. Y., & Dong, P. D. (2013). Specification of hepatopancreas progenitors in zebrafish by hnf1ba and wnt2bb. *Development*, *140*(13), 2669-2679. https://doi.org/10.1242/dev.090993

Lau, H. H., Ng, N. H. J., Loo, L. S. W., Jasmen, J. B., & Teo, A. K. K. (2018). The molecular functions of hepatocyte nuclear factors - In and beyond the liver. *J Hepatol*, *68*(5), 1033-1048. https://doi.org/10.1016/j.jhep.2017.11.026

Lavergne, A., Tarifeno-Saldivia, E., Pirson, J., Reuter, A. S., Flasse, L., Manfroid, I., Voz, M. L., & Peers, B. (2020). Pancreatic and intestinal endocrine cells in zebrafish share common transcriptomic signatures and regulatory programmes. *BMC Biol*, *18*(1), 109. https://doi.org/10.1186/s12915-020-00840-1

Lawlor, N., George, J., Bolisetty, M., Kursawe, R., Sun, L., Sivakamasundari, V., Kycia, I., Robson, P., & Stitzel, M. L. (2017). Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. *Genome Res*, *27*(2), 208-222. https://doi.org/10.1101/gr.212720.116

Li, M., Zhao, L., Page-McCaw, P. S., & Chen, W. (2016). Zebrafish Genome Engineering Using the CRISPR-Cas9 System. *Trends Genet*, *32*(12), 815-827. https://doi.org/10.1016/j.tig.2016.10.005

Lukinius, A., Wilander, E., Stridsberg, M., Eriksson, B., & Oberg, K. (1991). Electron-microscopical immunocytochemical study of a pancreatic islet amyloid polypeptidoma. *Endocr Pathol*, *2*(3), 169-175. https://doi.org/10.1007/BF02915458

Makarova, K. S., & Koonin, E. V. (2015). Annotation and Classification of CRISPR-Cas Systems. *Methods Mol Biol*, *1311*, 47-75. https://doi.org/10.1007/978-1-4939-2687-9_4

Mendel, D. B., Hansen, L. P., Graves, M. K., Conley, P. B., & Crabtree, G. R. (1991). HNF-1 alpha and HNF-1 beta (vHNF-1) share dimerization and homeo domains, but not activation domains, and form heterodimers in vitro. *Genes Dev*, *5*(6), 1042-1056. https://doi.org/10.1101/gad.5.6.1042

Meyer, A., & Schartl, M. (1999). Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol*, *11*(6), 699-704. https://doi.org/10.1016/s0955-0674(99)00039-3

Mitchell, S. M., & Frayling, T. M. (2002). The role of transcription factors in maturity-onset diabetes of the young. *Mol Genet Metab*, *77*(1-2), 35-43. https://doi.org/10.1016/s1096-7192(02)00150-6

Molven, A., & Njolstad, P. R. (2011). Role of molecular genetics in transforming diagnosis of diabetes mellitus. *Expert Rev Mol Diagn*, *11*(3), 313-320. https://doi.org/10.1586/erm.10.123

Murphy, R., Ellard, S., & Hattersley, A. T. (2008). Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab*, *4*(4), 200-213. https://doi.org/10.1038/ncpendmet0778

Naylor, R. W., & Davidson, A. J. (2014). Hnf1beta and nephron segmentation. *Pediatr Nephrol*, *29*(4), 659-664. https://doi.org/10.1007/s00467-013-2662-x

Nielsen, H. B., & Knudsen, S. (2002). Avoiding cross hybridization by choosing nonredundant targets on cDNA arrays. *Bioinformatics*, *18*(2), 321-322. https://doi.org/10.1093/bioinformatics/18.2.321

Nkonge, K. M., Nkonge, D. K., & Nkonge, T. N. (2020). The epidemiology, molecular pathogenesis, diagnosis, and treatment of maturity-onset diabetes of the young (MODY). *Clin Diabetes Endocrinol*, *6*(1), 20. https://doi.org/10.1186/s40842-020-00112-5

Olokoba, A. B., Obateru, O. A., & Olokoba, L. B. (2012). Type 2 diabetes mellitus: a review of current trends. *Oman Med J*, *27*(4), 269-273. https://doi.org/10.5001/omj.2012.68

Perner, B., Englert, C., & Bollig, F. (2007). The Wilms tumor genes wt1a and wt1b control different steps during formation of the zebrafish pronephros. *Dev Biol*, *309*(1), 87-96. https://doi.org/10.1016/j.ydbio.2007.06.022

Pontoglio, M., Barra, J., Hadchouel, M., Doyen, A., Kress, C., Bach, J. P., Babinet, C., & Yaniv, M. (1996). Hepatocyte nuclear factor 1 inactivation results in hepatic dysfunction, phenylketonuria, and renal Fanconi syndrome. *Cell*, *84*(4), 575-585. https://doi.org/10.1016/s0092-8674(00)81033-8

Poureetezadi, S. J., & Wingert, R. A. (2016). Little fish, big catch: zebrafish as a model for kidney disease. *Kidney Int*, *89*(6), 1204-1210. https://doi.org/10.1016/j.kint.2016.01.031

Rubio-Cabezas, O., Hattersley, A. T., Njolstad, P. R., Mlynarski, W., Ellard, S., White, N., Chi, D. V., Craig, M. E., International Society for, P., & Adolescent, D. (2014). ISPAD Clinical Practice Consensus Guidelines 2014. The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatr Diabetes*, *15 Suppl 20*, 47-64. https://doi.org/10.1111/pedi.12192

Sagen, J. V., Njolstad, P. R., & Sovik, O. (2002). Reduced prevalence of late-diabetic complications in MODY3 with early diagnosis. *Diabet Med*, *19*(8), 697-698. https://doi.org/10.1046/j.1464-5491.2002.00688_1.x

Sagen, J. V., Odili, S., Bjorkhaug, L., Zelent, D., Buettger, C., Kwagh, J., Stanley, C., Dahl-Jorgensen, K., de Beaufort, C., Bell, G. I., Han, Y., Grimsby, J., Taub, R., Molven, A., Sovik, O., Njolstad, P. R., & Matschinsky, F. M. (2006). From clinicogenetic studies of maturity-onset diabetes of the young to unraveling complex mechanisms of glucokinase regulation. *Diabetes*, *55*(6), 1713-1722. https://doi.org/10.2337/db05-1513

Sandbakken, M., Ebbesson, L., Stefansson, S., & Helvik, J. V. (2012). Isolation and characterization of melanopsin photoreceptors of Atlantic salmon (Salmo salar). *J Comp Neurol*, *520*(16), 3727-3744. https://doi.org/10.1002/cne.23125

Shields, B. M., Hicks, S., Shepherd, M. H., Colclough, K., Hattersley, A. T., & Ellard, S. (2010). Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia*, *53*(12), 2504-2508. https://doi.org/10.1007/s00125-010-1799-4

Sinigaglia, C., Thiel, D., Hejnol, A., Houliston, E., & Leclere, L. (2018). A safer, urea-based in situ hybridization method improves detection of gene expression in diverse animal species. *Dev Biol*, *434*(1), 15-23. https://doi.org/10.1016/j.ydbio.2017.11.015

Steck, A. K., & Rewers, M. J. (2011). Genetics of type 1 diabetes. *Clin Chem*, *57*(2), 176-185. https://doi.org/10.1373/clinchem.2010.148221

Tiso, N., Moro, E., & Argenton, F. (2009). Zebrafish pancreas development. *Mol Cell Endocrinol*, *312*(1-2), 24-30. https://doi.org/10.1016/j.mce.2009.04.018

Ulinski, T., Lescure, S., Beaufils, S., Guigonis, V., Decramer, S., Morin, D., Clauin, S., Deschenes, G., Bouissou, F., Bensman, A., & Bellanne-Chantelot, C. (2006). Renal phenotypes related to hepatocyte nuclear factor-1beta (TCF2) mutations in a pediatric cohort. *J Am Soc Nephrol*, *17*(2), 497-503. https://doi.org/10.1681/ASN.2005101040

Verkuijl, S. A., & Rots, M. G. (2019). The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. *Curr Opin Biotechnol*, *55*, 68-73. https://doi.org/10.1016/j.copbio.2018.07.005

Wilcox, G. (2005). Insulin and insulin resistance. *Clin Biochem Rev*, *26*(2), 19-39. https://www.ncbi.nlm.nih.gov/pubmed/16278749

Yamagata, K., Oda, N., Kaisaki, P. J., Menzel, S., Furuta, H., Vaxillaire, M., Southam, L., Cox, R. D., Lathrop, G. M., Boriraj, V. V., Chen, X., Cox, N. J., Oda, Y., Yano, H., Le Beau, M. M., Yamada, S., Nishigori, H., Takeda, J., Fajans, S. S., . . . et al. (1996). Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature*, *384*(6608), 455-458. https://doi.org/10.1038/384455a0

Yu, D. D., Guo, S. W., Jing, Y. Y., Dong, Y. L., & Wei, L. X. (2015). A review on hepatocyte nuclear factor-1beta and tumor. *Cell Biosci*, *5*, 58. https://doi.org/10.1186/s13578-015-0049-3

Zang, L., Maddison, L. A., & Chen, W. (2018). Zebrafish as a Model for Obesity and Diabetes. *Front Cell Dev Biol*, *6*, 91. https://doi.org/10.3389/fcell.2018.00091

Zang, L., Shimada, Y., & Nishimura, N. (2017). Development of a Novel Zebrafish Model for Type 2 Diabetes Mellitus. *Sci Rep*, *7*(1), 1461. https://doi.org/10.1038/s41598-017-01432-w

WHO, IDF, 2021: https://www.who.int/health-topics/diabetes#tab=tab_1

7 Appendix

ISH solutions and chemicals

Antibody solution

1.2 ml of antibody solution covered 10 slides.

x ml antibody (volume depends on concentration)

200 ml 5x maleate buffer

100 ml 10x blocking solution

3 ml Triton X-100

MilliQ® water to a total volume of 1 ml

DEPC treated water

0.5ml DEPC

1 l MilliQ® water

The solution was vigorously mixed. before incubation in RT (room temperature) overnight. The solution was autoclaved the next day.

Ethanol series

Ethanol series were made with absolute ethanol and either DEPC treated MilliQ®water or MilliQ® water to the desired concentration.

Hybridization solution without probe (Hyb-)

3.6 ml TEN buffer

10 ml 50% Dextran sulphate

100 µl Tween 20

25 ml 8M Urea (final concentration 4M)

OBS: Urea must be added to the solution directly before use. The rest of the components can be mixed on forehand and kept at -20 °C in aliquots. Aliquots of 2740 µl Hyb- without urea must be added 5000 µl urea right before use. Aliquots of 1370 µl Hyb- without urea must be added 2500 µl of urea right before use.

Hybridization solution with probe (Hyb+)

500 µl provided solution for approximately four glasses (120 µl per glass):

387 µl Hyb-

50 µl 10% Blocking solution

These two components were preheated to 65 °C and mixed using a vortex.

1 µl of the desired probe was mixed with 62 µl DEPC treated water before boiling it at 100 °C for 5 minutes and then directly put back on ice again. This was important to keep the probe in an uncoiled state. All the components were then mixed vigorously by using a vortex.

Proteinase K

10 mg Proteinase K was dissolved in 5 ml Proteinase K buffer. and then aliquots of 250 ml each were stored at -20 °C until use. Each aliquot of 250 ml was diluted in 50 ml Proteinase K buffer before use to provide a final concentration of 10 mg/ml.

Proteinase K buffer

100 ml 1 M Tris-HCl pH 8.0

100 ml 0.5 M EDTA

800 ml DEPC treated MilliQ® water

RNase buffer

10 ml 1 M Tris-HCl pH 7.5

2 ml 0.5 EDTA

100 ml 5 M NaCl

MilliQ® water was added to a total volume of 1 l.

Staining solution (NBT/BCIP)

Dissolve 1 (NBT/BCIP) tablet in 10 ml of MilliQ® water by vortexing.

Stop buffer

10 ml 1 M Tris-HCl pH 7.5

2 ml 0.5 M EDTA

30 ml 5 M NaCl

MilliQ® water was added to a total volume of 1 l.

Ten buffer

0.5 ml 1M Tris-HCl pH 7.5

3 ml 5 M NaCl

0.1 ml 0.5 M EDTA

Visualization buffer

20 ml 5 M NaCl

100 ml 1 M Tris-HCl pH 9.5

100 ml 0.5 M $MgCl_2x6H_2O$

MilliQ® water was added to a total volume of 1 l.

0.1 M Triethanolamine (TEA) pH 8.0

TEA buffer was made by adding 665 µl of TEA in 50 ml DEPC treated MilliQ® water. pH was adjusted to 8.0 by using saturated HCl.

0.1 M Tris Tween pH 8.2

50 ml 1 M Tris pH 8.2 and 450 ml MilliQ water. Add 50 µl of Tween20 (0.1%).

0.5 M MgCl2

101.6 g $MgCl_2x6H_2O$ was dissolved in MilliQ® water to a total volume of 1 liter and the solution was later sterilized by autoclaving.

1 mg/ml RNase A

50mg of RNase A was dissolved in 45 ml 0.01 M NaOAc pH 5.2 by heating the solution to 100 °C for 15 minutes. It was then slowly cooled down to room temperature and pH was adjusted by adding 5 ml of 1 M Tris-HCl pH 7.4. This gave a concentration of 1 mg/ml and aliquots of 1 ml were stored in -20 °C. Each aliquot was diluted in 50 ml of RNase buffer before use to provide a final concentration of 0.02 mg/ml Rnase A.

1 M Tris-HCl pH 7.5. 8.0 and 9.5

133.1 g Tris was dissolved in 800 ml DEPC treated MilliQ® water and pH was adjusted to the desired pH by using concentrated HCl. The volume was later adjusted to 1 liter with DEPC treated dH_2O. The solution was sterilized by autoclaving.

1x Maleate

200 ml 5xMaleate

800 ml MilliQ® water

1xPBS

200 ml 5xPBS

800 ml DEPC treated MilliQ® water

2xSSC

100 ml 20xSSC

900 ml DEPC treated MilliQ® water or MilliQ® water.

2xSSC, 0.05% Triton X-100, 2% Blocking reagent

40 ml 2xSSC

25 µl Triton X-100

10 ml 10 % Blocking reagent

3 M NaOAc

408.3 g sodium acetate trihydrate (NaOAc) was diluted in MilliQ® water, and pH was adjusted to 5.2 before the volume was adjusted to 1 liter with MilliQ® water. The solution was made sterile by autoclaving.

4% Paraformaldehyde (PF)

40 g PF was dissolved in 500 ml MilliQ® water and heated up to 650 °C while stirring. The solution was kept at 650 °C and stirring for 15 minutes, with frequent monitoring. A few drops of NaOH were added until the solution turned blank. and the solution was then cooled down on ice. 200 ml of 5xPBS in DEPC treated MilliQ® water was added, and the pH was adjusted to 7.4 with NaOH. The volume was adjusted to 1 l with MilliQ® water. The final solution was portioned in aliquots of 250 ml containers and stored at -20 °C until use.

5 M NaCl

292 g of NaCl was dissolved in MilliQ® water to a final volume of 1 l. The solution was made free of RNase by adding 0.5 ml of DEPC that incubated overnight and was autoclaved the next day.

5x Maleate buffer

58 g maleic acid was measured and 850 ml MilliQ® water was added. pH was adjusted to 7.5 with sodium hydroxide pellets. It was important to carefully introduce the pellets one at the time as it causes an exothermic reaction. The perfect number of pellets was approximately 40-42 g (45 g pellets exceeded the buffer capacity). 43.8 g NaCl was added. and the volume was adjusted with MilliQ® water to 1 l.

5xPBS in DEPC treated MilliQ® water

40 g NaCl

1 g KCl

7.2 g Na_2HPO_4

1.2 g KH_2PO_4

Approximately 800 ml MilliQ® water was added and the pH was adjusted to 7.4. Volume was adjusted to 1 l with MilliQ® water after correct pH. and 0.5 ml of DEPC was added to the solution and incubated over night before autoclaving.

8M Urea in solution

12 g urea dissolved in MilliQ® water or DEPC treated water to a total volume of 25 ml

or

2.4 g urea dissolved in MilliQ® water or DEPC treated water to a total volume of 5 ml

or

1.2 g urea dissolved in MilliQ® water or DEPC treated water to a total volume of 2.5 ml

10% Blocking solution

15 g Blocking Reagent

30 ml 5x Maleate buffer

DEPC treated MilliQ® water to a total volume of 150 ml.

The solution was dissolved by careful mixing and autoclaved before use. 12 ml and 0.5 ml aliquots were made, and the solution were stored at -20 °C.

20xSSC in DEPC treated MilliQ® water

175.3 g NaCl

88.2 g Tri-sodium citrate-dihydrate

The salts were dissolved in 800 ml MilliQ® water. and the pH was adjusted to 7. The volume was adjusted to 1 l with MilliQ® water. The solution is made RNase free by adding 0.5 ml DEPC for incubation overnight and autoclaved the next day.

50 % urea in 2xSSC

25 ml urea (12g urea dissolved in 25 ml of MilliQ® water) was mixed with 25 ml 2xSSC

50 % Dextran sulphate

5 g of Dextran sulphate were dissolved in 5 ml DEPC treated MilliQ® water by using an incubator at 65 °C. Volume was adjusted to 10 ml with DEPC treated MilliQ® water.

70% glycerol

35 ml glycerol was mixed with 15 ml 1xPBS for 2-3 hours using a slow mixer to prevent bubbles.

Table A1: Chemicals, product number and manufacturer used in ISH.

Chemical	Product #	Manufacturer
Absolute ethanol	32221	Sigma-Aldrich Merck
Anti-Digoxigenin-AP. Fab-fragments	11093274910	Sigma-Aldrich Merck
Blocking Reagent (Roche)	11096176001	Sigma-Aldrich Merck
Dextran sulphate	D8906	Sigma-Aldrich Merck
Diethyl pyro carbonate (DEPC)	D5787	Sigma-Aldrich Merck
EDTA (0.5 M)	324506	Sigma-Aldrich Merck
Glycerol	G5516	Sigma-Aldrich Merck
HCl	H1758	Sigma-Aldrich Merck
KCl	60131	Sigma-Aldrich Merck
KH_2PO_4	P5655	Sigma-Aldrich Merck
Maleic acid	M0375	Sigma-Aldrich Merck
$MgCl_2x6H_2O$	M2670	Sigma-Aldrich Merck
NaCl	31434	Sigma-Aldrich Merck
NaOAc	71188	Sigma-Aldrich Merck
NaOH solution	415413	Sigma-Aldrich Merck
NaOH pellets	S8045	Sigma-Aldrich Merck
Na_2HPO_4	71640	Sigma-Aldrich Merck
NBT/BCIP Ready-to-Use tablets	11697471001	Sigma-Aldrich Merck
Paraformaldehyde (PF)	P6148	Sigma-Aldrich Merck
Proteinase K	P2308	Sigma-Aldrich Merck
RNase A	R4875	Sigma-Aldrich Merck
RNaseZAP ™	R2020	Sigma-Aldrich Merck
Sodium hydroxide pellets	30620	Sigma-Aldrich Merck
Triethanolamine (TEA)	90279	Sigma-Aldrich Merck
Triton X-100	T8787	Sigma-Aldrich Merck
Tris	252859	Sigma-Aldrich Merck
Tri-Sodium citrate-dihydrate	71402	Sigma-Aldrich Merck
Tissue-Tek® O.C.T. ™ Compound	4583	Sakura®
Tween® 20	P9416	Sigma-Aldrich Merck
Urea	U5378	Sigma-Aldrich Merck
Urea	U5378	Sigma-Aldrich Merck

Colony PCR of cloned injected individuals

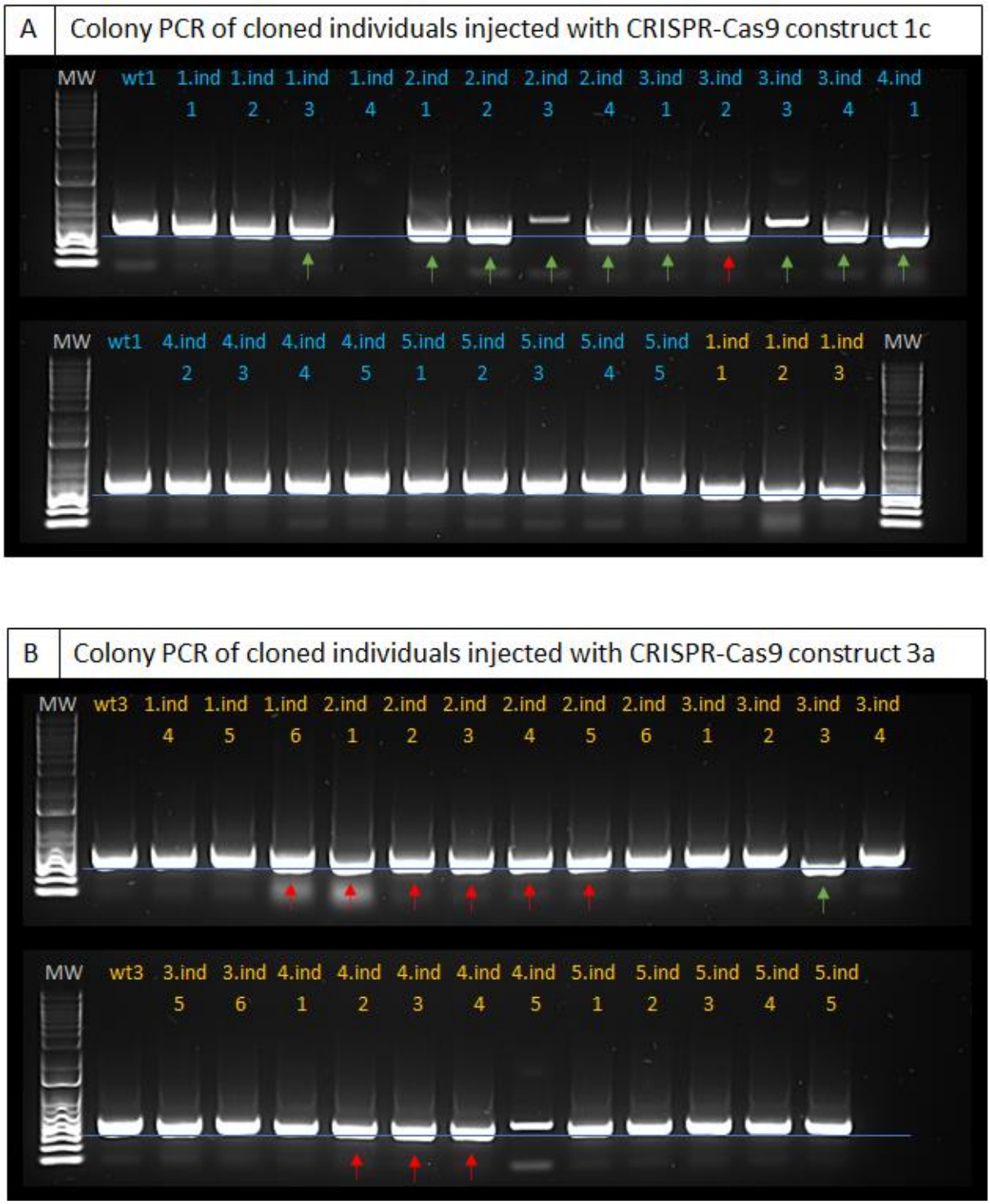


Figure A1: Agarose gel electrophoresis of colony PCR performed on pCR4 TOPO TA cloned and transfected CRISPR-Cas9 individuals positive for T7 assay. The plasmid insert for constructs 1c (blue font) and 3a (yellow font) will have a wt band at 473 bp and 393 bp, respectively, which include the M13 sequence. Any band smaller (deletion) or much larger (insertion) is selected for sequencing. The blue line was used to assess minor size differences. **A:** Green arrows indicate bands with minor changes worth sequencing. The individuals chosen for further sequencing of construct 1c were Ind.1-4 some, including several different clones.

B: Green arrows indicate bands with minor changes worth sequencing. The individuals chosen for further sequencing of construct 3a were Ind.1-4 some, including several different clones.

Viability of microinjections

Table A3: Viability rate (%) of sgRNA 1a, 1b and 2a injected embryos.

Type	Total embryos	Healthy (24 h)	Dead (24 h)	Healthy (48 h)	Dead (48 h)	Survival rate (%)
Wt	143	137	6	137	0	96
hnf1ba ex1a	52	45	7	45	0	87
hnf1ba ex1b	43	37	6	35	2	81
hnf1ba ex2a	121	110	11	110	0	91

CRISPR-Cas9 injected zebrafish larvae, tail-cut sequences

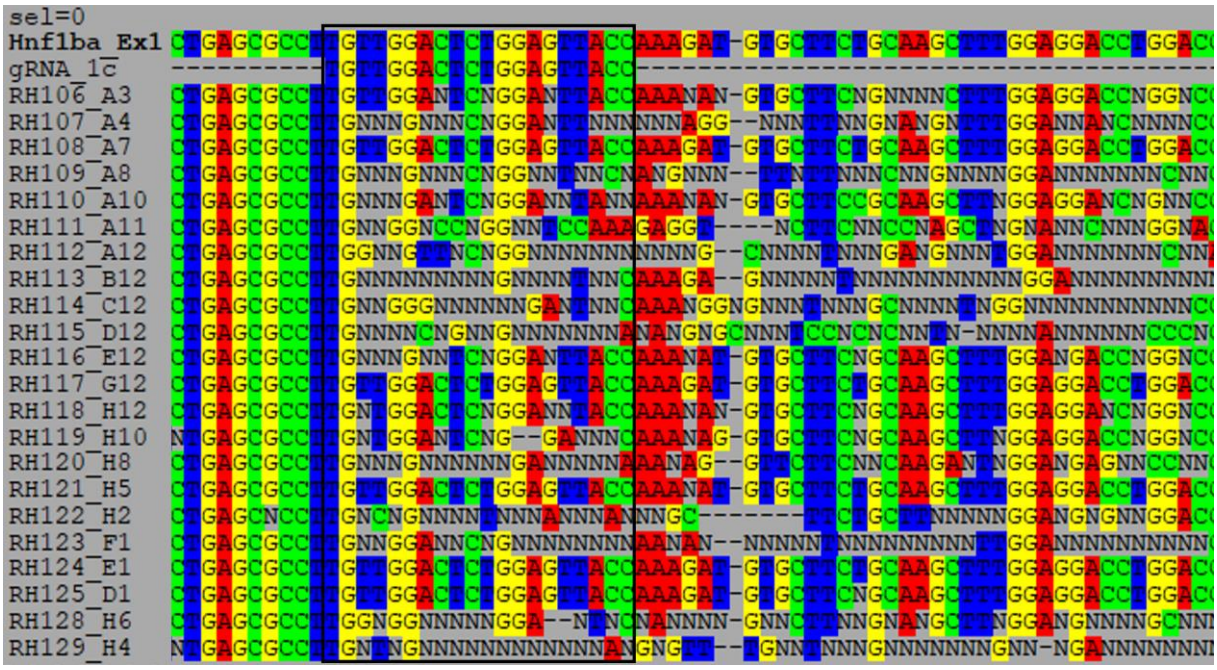


Figure A2: Sanger sequences of PCR amplified tail-cut DNA of CRISPR-Cas9 sgRNA 1c injected zebrafish embryos. A4 (Mut. ind.1), A8 (Mut.ind.2), B12 (Mut.ind.3), E12 (Mut.ind.4), H2 (Mut.ind.5) and F1 (Mut.ind.6) were chosen for confocal imaging.