Exploring *hnf1ba* knockout zebrafish mutant as a model for studying MODY5 disease mechanisms in early pancreas development

Oda Fløtre

Master of Science

Integrated Teacher Programme in Science and Mathematics

Department of Biological Sciences, University of Bergen

June 2023

Acknowledgements

The presented study was conducted at the Department of Biological Science at the University of Bergen. It is part of the project: "Analysis of zebrafish *hnf-1beta* mutant to understand MODY5 disease mechanisms in pancreas development", a collaborative project between the Western Norway University of Applied Sciences and the University of Bergen.

First, I would like to express my deepest gratitude to my main supervisor, Professor Ståle Ellingsen. Your excellent guidance and valuable advice have encouraged me in all stages of the work. I am truly grateful for the opportunity to be a part of such an exciting project. My sincere thanks also go to my co-supervisor Dr Elsa Denker for sharing your knowledge and creating an engaging environment. Your assistance and technical expertise is deeply appreciated.

A special thanks goes to my co-supervisor, PhD student Eunice Kabanyana Mchaina, for devoting many hours to teaching me lab techniques and for engaging in my thesis. I am truly grateful for all the good and occasionally frustrating times we have shared.

I would also like to thank Professor Lise Bjørkhaug Gundersen for your help throughout the project. Your positive attitude and insightful comments have elevated the quality and the experience of the project. Further, I would like to thank everyone at MDB for their warm welcome and continuous support. I feel lucky to have been part of such a supportive and collaborative environment.

To my fellow students and friends, thank you for making these five years an incredible experience. Thanks to my lovely family for always believing in me and cheering me up. Lastly, I would like to thank Rasmus Kvile for your encouragement and endless support.

Bergen, June 2023
Oda Fløtre

Contents

Abstract

Maturity-onset diabetes of the young (MODY) is a monogenic form of diabetes, characterised by pancreatic beta-cell dysfunction, early onset (<25 years of age), and autosomal dominant inheritance. Heterozygous mutations in the hepatocyte nuclear factor-1 beta (*HNF1B)* gene are found to cause MODY5, a severe subtype of MODY. This subtype is associated with a wide range of clinical manifestations, including pancreas hypoplasia, renal abnormalities and genital malformations. Currently, there is no fully representative animal model for studying MODY5 disease development in the pancreas. To address this issue, we investigated a homozygous *hnf1ba* knockout zebrafish model, representing the L16fs mutation found in humans.

Our study demonstrates the essential role of *hnf1ba* in pancreas development. The knockout of *hnf1ba* leads to significant alterations in the expression levels of several genes associated with pancreas development and function. We observed increased expression levels of *foxa3, gcga, hnf1bb, insb* and *pax6b*, and decreased levels of *gcgb, insa* and *pdx1*. By immunostaining glucagon- and insulin-producing cells, we identified structural variations in the pancreas among *hnf1ba*(-/-) zebrafish larvae. The mutants exhibited a lower average number of alpha- and beta-cells compared to wild-type larvae; however, this reduction was not statistically significant. These observed variations suggest variable penetrance of the *hnf1ba* mutation, consistent with previous findings of varying phenotypes in MODY5 patients. Moreover, we observed various morphological abnormalities in the mutants, including a curved tail, enlarged yolk sac, pericardial edema and reduced heart rate, indicating a multisystemic nature similar to MODY5. We anticipate our study as a starting point towards establishing a new in vivo model of MODY5 disease. The model has the potential to provide deeper insights into the role of *hnf1ba* and hence advance our knowledge of its relevance to human health and MODY5 pathogenesis.

Selected abbreviations

aa	Amino acid
ATP	Adenosine triphosphate
BMI	Body mass index
bp	Base pair
cDNA	complementary DNA
CRISPR	Clustered regularly interspaced short palindromic repeats
Ct	Cycle threshold
DNA	Deoxyribonucleic acid
dpf	Days post fertilisation
foxa3	Fork head box A3 (zebrafish gene)
gcga/gcgba	Glucagon paralog a/b (zebrafish genes)
GCK	Glucokinase (human protein)
GLUT2	Glucose-transporter 2 (human protein)
GOI	Gene of interest
HKG	Housekeeping gene
HNF1A	Hepatocyte nuclear factor-1 alpha (human protein)
HNF1B	Hepatocyte nuclear factor-1 beta (human gene)
HNF1B	Hepatocyte nuclear factor-1 beta (human protein)
hnf1ba	Hepatocyte nuclear factor-1 beta a (zebrafish gene)
Hnf1ba	Hepatocyte nuclear factor-1 beta a (zebrafish protein)
hnf1bb	Hepatocyte nuclear factor-1 beta b (zebrafish gene)
Hnf1bb	Hepatocyte nuclear factor-1 beta b (zebrafish protein)
hpf	Hours post fertilisation
IHC	Immunohistochemistry
insa/insb	Insulin paralog a/b (zebrafish genes)
kb	Kilobase
MPC	Multipotent progenitor cell
MODY	Maturity-onset diabetes of the young
mRNA	messenger RNA
NRT	No reverse transcriptase control
NTC	No template control
pax6b	Paired box 6b (zebrafish gene)

pdx1	Pancreatic and duodenal homeobox 1 (zebrafish gene)
RGE	Relative gene expression
RNA	Ribonucleic acid
RT	Room temperature
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
T1D	Type 1 diabetes
T2D	Type 2 diabetes
TGD	Teleost-specific gene duplication event
wt	Wild-type

1 Introduction

1.1 Diabetes mellitus

1.1.1 What is diabetes mellitus?

Diabetes mellitus, or diabetes, is a chronic condition that affects over 537 million humans worldwide.[1] Projections estimate that this number will increase to 783 million by 2045, making diabetes one of the fastest-growing global health emergencies of the 21st century.[2] Causing more than 6.7 million deaths in 2021, diabetes is among the top 10 causes of mortality worldwide.[1] The disease affects individuals across socioeconomic status and nationalities, and almost one-in-two affected adults are unaware they have the condition.[3] While more knowledge is needed regarding the cause and ideal treatment of diabetes, a positive outlook towards research and treatment options offers hope for those living with the disease.

Diabetes occurs when the body cannot produce enough insulin or effectively use the insulin it produces.[4] Insulin is an anabolic hormone that regulates blood sugar levels by facilitating glucose uptake from the bloodstream into the body's cells.[5] The hormone is produced in the pancreas and is crucial for regulating carbohydrate, lipid and protein metabolism.[5] Low insulin levels result in elevated blood glucose levels, or hyperglycemia, a common effect of uncontrolled diabetes. Diabetes is diagnosed when the blood glucose concentration is ≥7.0 mmol/l after fasting for 8 hours or when the blood glucose concentration is ≥11.1 mmol/l two hours after an oral glucose tolerance test.[1,4] In healthy individuals, the blood glucose concentration remains stable within 4-6 mmol/l when fasting for 8 hours.[4]

When diabetes is uncontrolled, it can cause disabling and life-threatening complications. Long-term hyperglycemia is associated with severe damage to multiple organs like the eyes, kidneys, heart and blood vessels, and may lead to kidney failure, nerve damage and cardiovascular diseases.[6,7] Hence, early and accurate diagnosis is crucial for preventing serious complications and improving health outcomes.[1]

1.1.2 Pancreas anatomy and function

The pancreas is a gland located behind the stomach measuring 15-25 cm in length (Figure 1.1). The organ has both exocrine and endocrine functions.[8] Approximately 95% of the pancreas comprises exocrine tissue, consisting of acinar and ductal cells that secrete digestive enzymes and bicarbonate ions into the gastrointestinal tract.[9] The secretion empties through the pancreatic duct, which merges with the common bile duct into the duodenum.[10] The pancreas

also comprises endocrine cells that secrete pancreatic hormones directly into the bloodstream. Endocrine cells are scattered throughout the pancreas in cell clusters known as islets of Langerhans.[9] These islets consist of at least four different cell types, including alpha-cells that produce glucagon (34%), beta-cells that produce insulin (54%), delta-cells that produce somatostatin (10%) and gamma-cells that produce pancreatic polypeptides (~1-2%).[11] The islets are arranged with beta-cells in the centre and alpha-cells in the surrounding tissue (Figure 1.1).

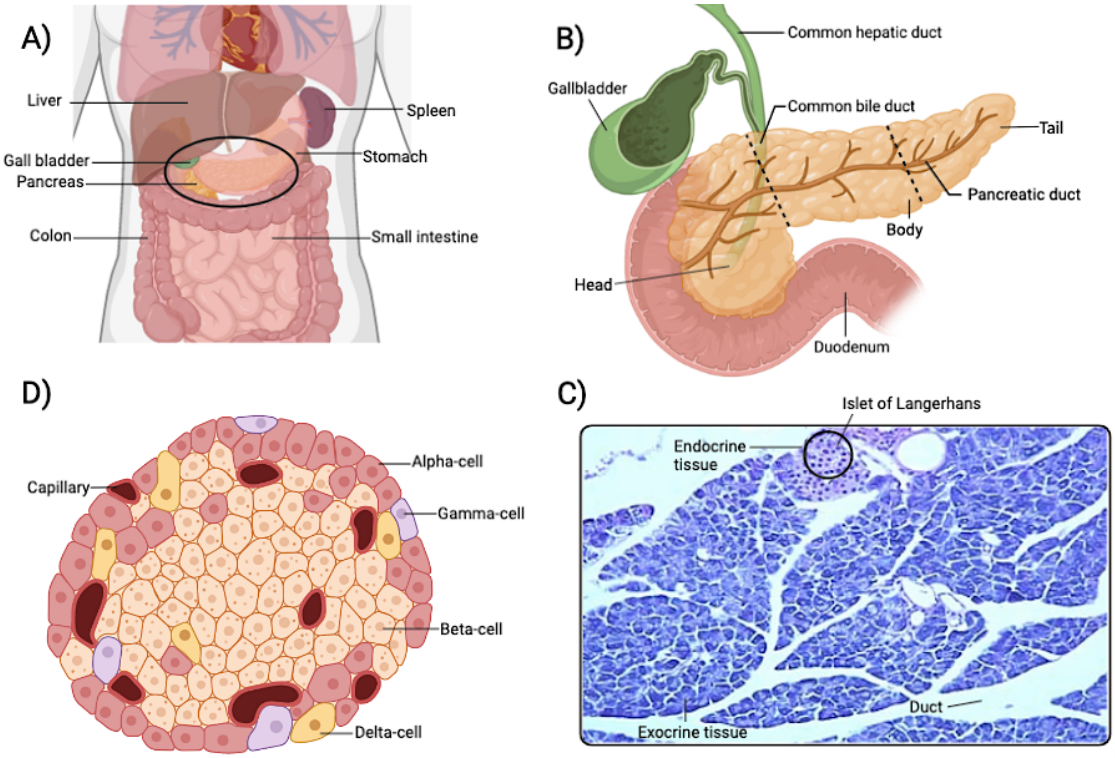

Figure 1.1: Anatomical organisation of the pancreas. A) The pancreas (circled) is located in the abdominal cavity behind the stomach. It is surrounded by multiple organs, including the spleen, the gallbladder and the small intestine. **B)** The pancreas is anatomically divided into a head, body and tail. The pancreatic duct merges with the common bile duct into the duodenum, transporting digestive enzymes and bicarbonate ions into the digestive system. **C)** A longitudinal section of pancreatic tissue stained with haematoxylin displays the pancreatic duct, exocrine and endocrine tissues (circled). The endocrine tissue is organised in cell clusters called islets of Langerhans. The image is adapted from Lise Gundersen. **D)** The islets of Langerhans comprise at least four cell types: alpha-, beta-, gamma- and delta-cells. Alpha-cells are typically located in the periphery, while beta-cells are located in the middle of the islet. The figure is made in BioRender.

The pancreas plays an essential role in controlling blood glucose levels by secreting the antagonistic hormones insulin and glucagon.[10,12] When blood glucose levels are low, glucagon released from alpha-cells signals the liver to break down glycogen into glucose. Glucose is then released into the bloodstream and transported to cells/tissues for energy production. This

process is known as glycogenolysis.[12] Additionally, glucagon drives hepatic and renal gluconeogenesis, a process in which glucose is synthesised from non-carbohydrate sources, such as amino acids and glycerol.[13] On the other hand, when glucose levels are high, insulin is secreted from beta-cells stimulating glucose uptake by cells/tissues. In addition, insulin promotes glycogenesis, the process where glucose is converted into glycogen for energy storage. This entire process is known as glucose homeostasis and is essential for maintaining cellular respiration and biosynthesis.[14] Understanding the intricate balance of glucose homeostasis is crucial for developing effective treatments for metabolic disorders such as diabetes.[12]

1.1.3 Classification of diabetes

Diabetes is a complex disease that presents in various forms. While all subtypes exhibit hyperglycemia, they differ in underlying causes, disease mechanisms, phenotypes and response to treatment.[15] The two most common types are type 1 and type 2 diabetes.[1] Other forms also exist, including monogenic diabetes, gestational diabetes and other diabetic-associated syndromic diseases (Figure 1.2). Monogenic diabetes results from changes in a single gene and includes maturity-onset diabetes of the young (MODY) and neonatal diabetes. Unfortunately, these subtypes are often misdiagnosed because of their low prevalence and because they share symptoms with the more common types of diabetes.[16] Hence, expanding our knowledge of diabetes subtypes can increase our opportunities for correct diagnosis and optimal treatment.

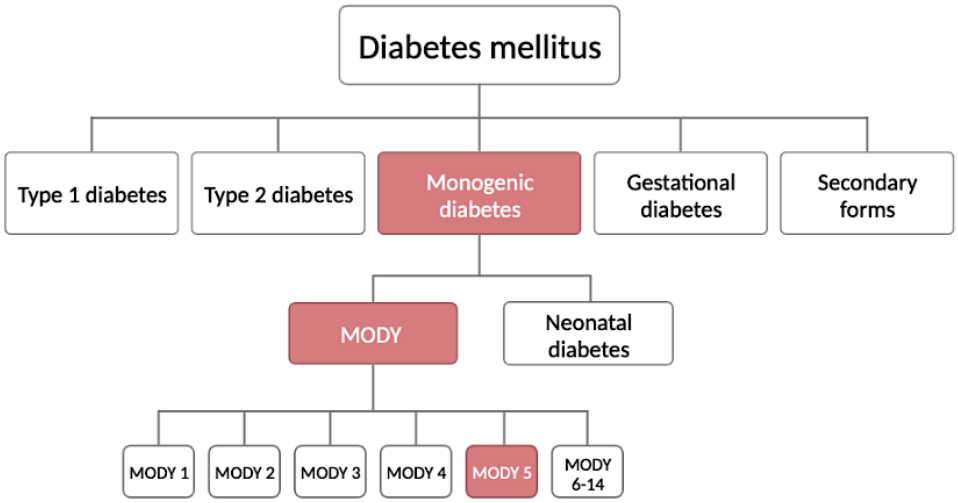

Figure 1.2: Classification of diabetes mellitus. Diabetes can be categorised into type 1 and type 2 diabetes, monogenic diabetes, gestational diabetes and other diabetic-associated syndromic diseases (secondary forms). Monogenic diabetes includes maturity-onset diabetes of the young (MODY) and neonatal diabetes. MODY can be divided into MODY1-14. The subtypes 6-14 are represented in one box to illustrate their low prevalence compared to subtypes 1-5. This study focuses on MODY5. The figure is created in BioRender.

1.1.4 Type 1 diabetes

Type 1 diabetes (T1D) is an autoimmune disease in which the body's immune system attacks beta-cells in the pancreas, leading to insulin deficiency and hyperglycemia.[17] Individuals with T1D depend on insulin injections for proper glucose regulation, making access to affordable treatment vital for survival.[1] T1D is typically diagnosed in children and young adults <20 years of age and accounts for more than 85 % of all diabetes cases in that population.[17] However, the disease can develop at any age, and 5-10% of adult cases of diabetes are T1D.[1] Common symptoms include excessive thirst, fatigue, constant hunger, weight loss and blurred vision.[1,4] The cause of the disease is not fully understood, but complex interactions between environmental and genetic factors have been associated with the development of the disease. For instance, vitamin D deficiency, decreased gut-microbiome diversity, and certain virus infections have been associated with T1D development.[18] The genetic risk is mainly associated with HLA class II genes, which play a crucial role in controlling immune responses.[19] Despite ongoing research, there is currently no cure or prevention for T1D.[4]

1.1.5 Type 2 diabetes

Type 2 diabetes (T2D) is characterised by insulin resistance or beta-cell dysfunction. It occurs when the body is unable to respond adequately to insulin, or when the islets fail to keep pace with the increasing insulin resistance.[19] T2D typically develops in adulthood (>40 years of age) and is mainly associated with obesity, physical inactivity, unhealthy diets and stress. However, there is also a strong genetic component to the development of T2D, and multiple risk genes have been identified.[19] The symptoms of T2D are similar to those of T1D but are generally less pronounced. Diagnosing disease onset is therefore challenging, increasing the risk of developing late diabetic complications such as heart disease and visual impairment.[1] Today, T2D accounts for over 90% of diabetes cases worldwide, and its prevalence is rising across all regions and age groups. This is primarily due to economic development and increasing urbanisation leading to unhealthy diets and more sedentary lifestyles.[1] While there is currently no cure for T2D, it can be managed through lifestyle changes such as regular exercise and healthy diets, or by initiating oral medication such as metformin.[20]

1.2 Maturity-onset diabetes of the young (MODY)

MODY is an autosomal dominantly inherited form of monogenic diabetes.[21] The disease was first described by Tattersall in 1974,[22] where it was defined as "fasting hyperglycemia diagnosed under age 25 which could be treated without insulin for more than two years". Today,

the term MODY describes a heterogeneous group of monogenic diabetes associated with beta-cell dysfunction.[23] The most common clinical presentation is mild, asymptomatic hyperglycemia in nonobese children and young adults (<25 years of age).[21,23] MODY is also characterised by measurable serum C-peptide levels and insulin resistance.[24] Additionally, most affected patients have a prominent family history of diabetes.[21]

1.2.1 Misdiagnosis of MODY

MODY is expected to account for 1-2% of all diabetes cases worldwide.[24] It is a rare condition, and awareness of its clinical features is likely to vary.[23] Furthermore, the clinical features of MODY often overlap with phenotypes of the more common types of diabetes.[23,25] As a result, many cases are misdiagnosed as either T1D or T2D.[23,24] According to a UK report from 2010, more than 80% of patients with MODY are incorrectly diagnosed with T1D or T2D.[23] The misdiagnosing is especially high among young adults diagnosed with T2D, and studies have shown that 25% of individuals diagnosed with T2D before 30 years of age had MODY.[24] Thus, the true prevalence of MODY is considered to be higher than previously predicted. Correct diabetes diagnosis is essential for identifying an optimal treatment strategy.[25,26] For instance, for MODY patients (MODY1/MODY3) who have been receiving insulin injections based on incorrect T1D diagnosis, switching to oral agents such as sulfonylureas can, in some cases, improve their glycaemic control and hence life quality.[26]

1.2.2 MODY-associated genes

MODY is caused by heterozygous pathogenic variants in genes essential for normal beta-cell development and function.[21] MODY-related mutations have been reported in at least 14 different genes: *HNF4A, GCK, HNF1A, PDX1, HNF1B, NEUROD1, KLF11, CEL, PAX4, INS, BLK, ABCC8, KCNJ11* and *APPL1,* which accordingly are associated with MODY1-14.[*] Although all fourteen subtypes have impaired beta-cell function, they vary in clinical presentation, disease course and response to treatment.[27,28] MODY2 and MODY3 are the most common subtypes, accounting for 30-70% of all cases.[29] MODY2 is caused by mutations in the glucokinase (*GCK*) gene and is characterised by glucose-sensing defects leading to an increase

[*] Hepatocyte nuclear factor-4 alpha (*HNF4A*), Glucokinase (*GCK*), Hepatocyte nuclear factor-1 alpha (*HNF1A*), Pancreatic and duodenal homeobox 1 (*PDX1*), hepatocyte nuclear factor-1 beta (*HNF1B*), neuronal differentiation 1 (*NEUROD1*), krueppel-like factor 11 (*KLF11*), carboxyl ester lipase (*CEL*), paired box 4 (*PAX4*), insulin (*INS*), BLK proto-oncogene (*BLK*), ATP-binding cassette transporter sub-family C member 8 (*ABCC8*), potassium inwardly rectifying channel subfamily j member 11(*KCNJ11*), adaptor protein phosphotyrosine interacting with ph domain and leucine zipper 1 (*APPL1*)

in the insulin secretion threshold, as illustrated in Figure 1.3. MODY3 is caused by mutations in the hepatocyte nuclear factor-1 alpha (*HNF1A*) gene, which encodes a nuclear transcription factor expressed in the liver, kidney, intestine and pancreatic beta-cells.[29] The identification of MODY genes has expanded our knowledge of factors affecting the insulin signalling pathway and glucose homeostasis, providing potential targets for treating MODY diabetes.[30]

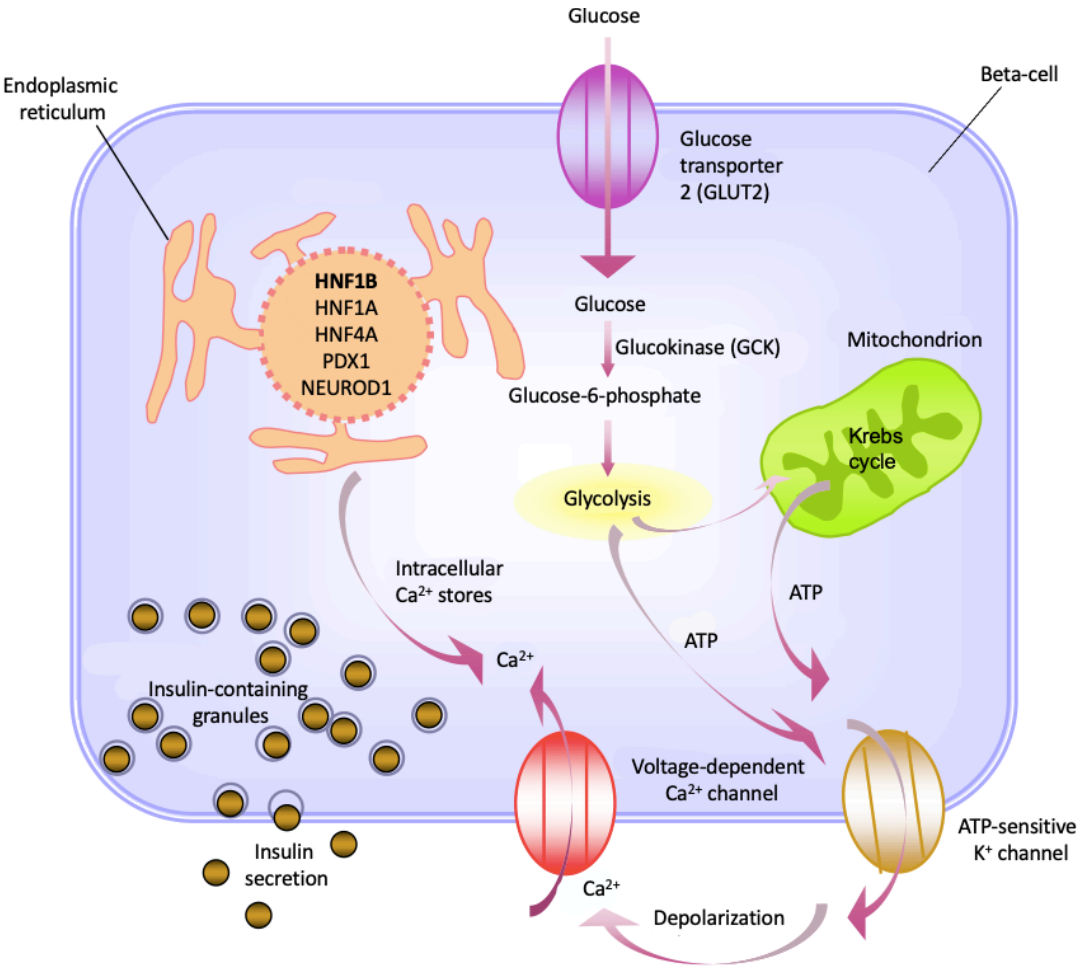

Figure 1.3: The pancreatic beta-cell and some of the proteins involved in maturity-onset diabetes of the young (MODY). Glucose is transported into the beta-cell by the glucose transporter 2 (GLUT2) and is phosphorylated by the glycolytic enzyme glucokinase (GCK), forming glucose-6-phosphate. This molecule is further involved in glycolysis and the Krebs cycle, generating ATP. The increase of ATP closes the ATP-sensitive potassium channels, leading to depolarization of the plasma membrane. This opens the voltage-dependent calcium channels resulting in an influx of extracellular calcium ions and a mobilization of intracellular calcium. This leads to the release of insulin-containing granules into the bloodstream, stimulating the uptake and metabolism of glucose. Mutations in the *GCK* gene results in decreased glucose phosphorylation and metabolism. Mutations in the transcription factors located in the nucleus of the beta-cell, such as *HNF1B, HNF1A, HNF4A, PDX1* and *NEUROD1*, affect the regulation of multiple genes essential for normal beta-cell function. This can result in the development of MODY.[21] The figure is adapted and modified from Fajans et al.[21]

1.3 Hepatocyte nuclear factor-1 beta (HNF1B)

1.3.1 *HNF1B* gene and protein isoforms

The hepatocyte nuclear factor-1 beta (*HNF1B*) gene is located on chromosome 17q12 and encodes the 557 amino acids (aa) long HNF1B protein. The gene, also known as transcription factor 2 (*TCF-2*), comprises nine coding exons spanning around 59 kilobases (kb).[31] Three alternative splicing variants of *HNF1B* have been discovered, all identical in their N-terminal ends but differ in their C-terminus (Figure 1.4). Transcript variant A is the canonical transcript and encodes the longest isoform. Transcript variant B lacks an internal segment compared to isoform A, while transcript variant C lacks exons 8 and 9. As a result, the HNF1B isoforms differ in function and target sequences.[32] Variants A and B are both transcriptional activators, while variant C has lost its transactivational activity and acts as an inhibitor of the other HNF1B isoforms.[33] The multiple isoforms add complexity to the gene's potential to regulate different cellular and tissue functions.[34]

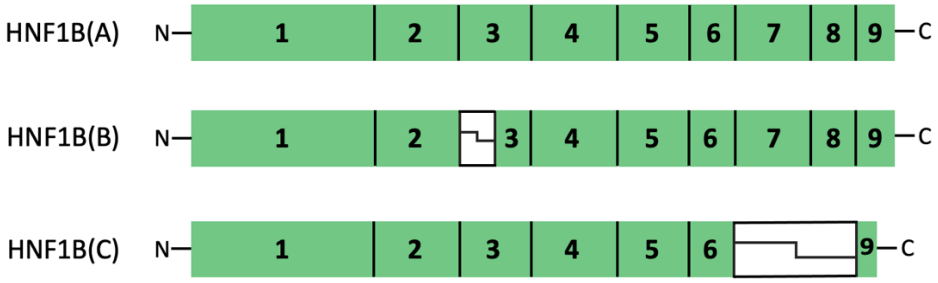

Figure 1.4: *HNF1B* mRNA transcript variants in humans. The variants are generated by alternative splicing and the differential use of polyadenylation sites. HNF1B(A) is the canonical transcript. HNF1B(B) lacks 26 aa at the 5' end of exon 3, and HNF1B(C) lacks the entire exons 7 and 8, which results in a frameshift in exon 9 and an early stop codon.[32] The white boxes illustrate the missing sequences compared to the canonical transcript. The lengths of the exons are proportional.

1.3.2 Structure and function of HNF1B

HNF1B is part of the homeobox gene family, and consists of four functional domains: a dimerization domain, a DNA binding domain composed of an atypical POU-specific domain (POU_S) and a POU-homeodomain (POU_H), and a transactivation domain.[35] The HNF1B protein structure is illustrated in Figure 1.5. The DNA binding domain enables HNF1B to recognise and bind to the promoter regions of target genes, thereby regulating their transcription. The cooperation between the POU_S and POU_H domains enhances the binding affinity and specificity of DNA binding.[36] The protein's transactivation domain facilitates the recruitment of co-activators and other transcription factors to the promotor region of its target genes.

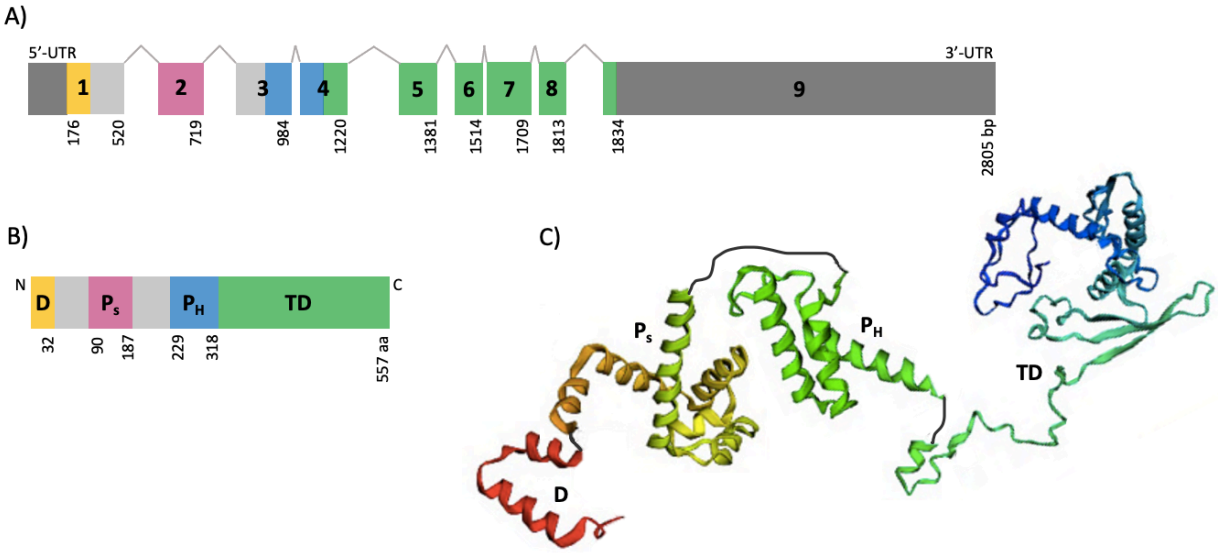

Figure 1.5: *HNF1B* mRNA sequence and HNF1B protein structure. The human *HNF1B* gene encodes a 9-exon mRNA transcript, where the conical transcript consists of 2805 bp **(A)**. The transcript is translated into the 557 aa long HNF1B protein **(B** and **C)**. HNF1B consists of a dimerization domain (D), a POU-specific domain (P_S), a POU-homeodomain (P_H) and a transactivation domain (TD).[35] The numbers represent the exons, and the colours distinguish between the functional domains of the protein. The dark grey boxes indicate untranslated regions of the exons. The size of the exons and introns are proportional. The protein structure in C is adapted and modified from Çubuk & Çaban.[37]

The HNF1B protein is a widely distributed transcription factor, highly conserved in structure and function within vertebrates.[38] The protein is mainly expressed in the primitive endoderm during early embryonic development and throughout organogenesis of the liver, kidney, pancreas and digestive tract. Its expression is also maintained in these organs in the adult stage.[39]

HNF1B binds to DNA sequences as a homodimer or as a heterodimer with the HNF1A protein.[36] HNF1A is a transcription factor structurally related to HNF1B, sharing more than 80% sequence homology.[40] Like HNF1B, HNF1A is also abundantly expressed in the pancreas, liver, kidney and digestive tract.[41] Despite their considerable homology and shared binding site, HNF1A and HNF1B do not compensate for each other.[31,40] The transcription factors differ in their expression levels and timing of expression during embryonic development, with HNF1B being expressed earlier than HNF1A.[40,42] Additionally, the diverse transactivation domains are believed to facilitate interactions with different sets of proteins.[36]

1.3.3 Pancreatic requirements of HNF1B

HNF1B is required for the differentiation of the primitive endoderm. The activation of HNF1B directs the differentiation of endodermal cells into pancreatic multipotent progenitor cells (PMC).[38] Additionally, the protein is a key member of a network of transcription factors that regulate the differentiation of pancreatic PMC into acinar, ductal and endocrine cells (Figure 1.6).[38] Mutations in *HNF1B* can disrupt this intricate network, leading to the absence of endocrine cells and abnormal beta-cell development.[43] However, it is difficult to distinguish the role of HNF1B in the pancreas from its role in the differentiation of the primitive endoderm.[43] Much is still unknown about its precise function, particularly regarding the complex gene regulatory network it is a part of.

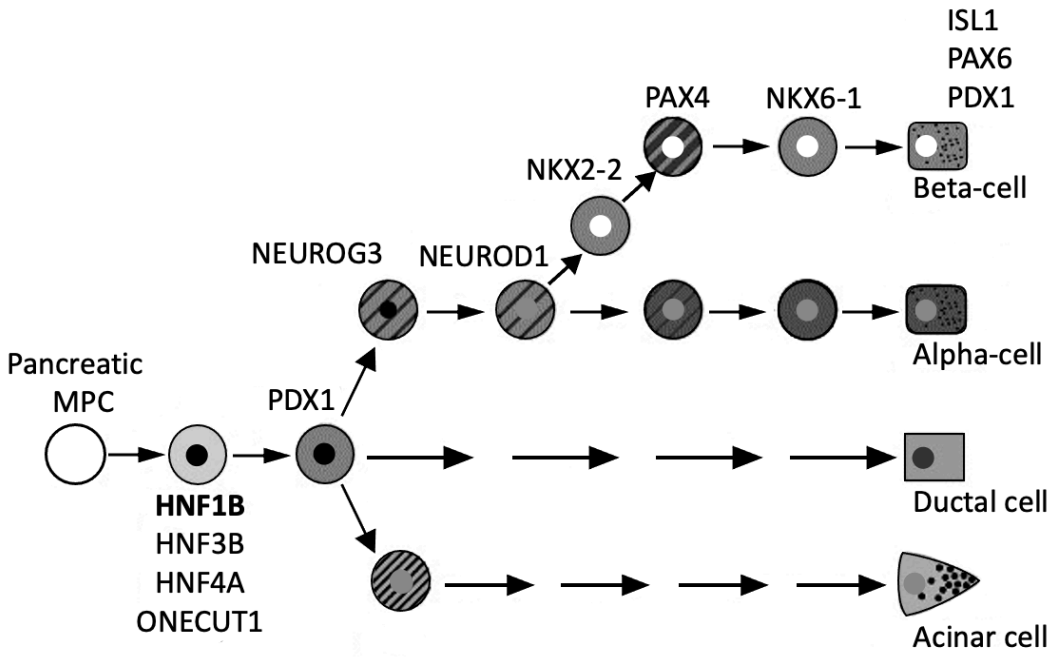

Figure 1.6: A simplified schematic presentation of the role of HNF1B in endocrine differentiation in the developing pancreas. The protein is part of a network regulating the differentiation of pancreatic multipotent progenitor cells (MPC) into endocrine and exocrine cells. The position of each transcription factor is based on its predominant functional role, time of expression, or both, focusing on genes relevant to this thesis (except *foxa3,* as it is regulated earlier in the foregut endoderm). The lineage is based on mouse data. The figure is adapted and modified from Wilson et al.[44]

1.4 MODY5

Heterozygous mutations in the *HNF1B* gene can cause MODY5, which accounts for 5-10% of all MODY cases.[45,46] The disease is characterised by reduced insulin secretion, pancreas atrophy, impaired renal function and genital malformations. Progressive nondiabetic nephropathy and liver test abnormalities have also been reported, making MODY5 a complex

multi-organ disease.[46,47] MODY5 was first described by Horikawa in 1997,[45] and since then, various aspects of the disease have been studied. However, most studies consist of case reports, which limits our understanding of the disease mechanisms.

A recent systematic study of MODY5-literature by Ge et al.[48] revealed that renal cysts and other renal abnormalities are the most common clinical feature of the disease, affecting 72% of all patients. This feature is considered a hallmark of MODY5 and distinguishes it from other MODY subtypes.[46] In addition, most MODY5 patients have pancreatic dysplasia, hypomagnesemia and are normal weight (BMI from 18.5 to 25 kg/m^2). Clinical management is critical, and many patients require insulin therapy.[46,49]

Whole-gene deletion and large genomic rearrangements are the most frequent molecular alteration observed in MODY5 patients.[46] Point mutations are detected in the majority of remaining cases, with most mutations clustering in the first four exons,[50,51] as illustrated in Figure 1.7. Although most mutations are de novo, some inherited mutations have also been reported to cause MODY5.[48,50] MODY5-associated mutations in *HNF1B* are shown to reduce the overall transcriptional activity, as a loss-of-function mechanism, resulting in impaired tissue/organ development.[36]

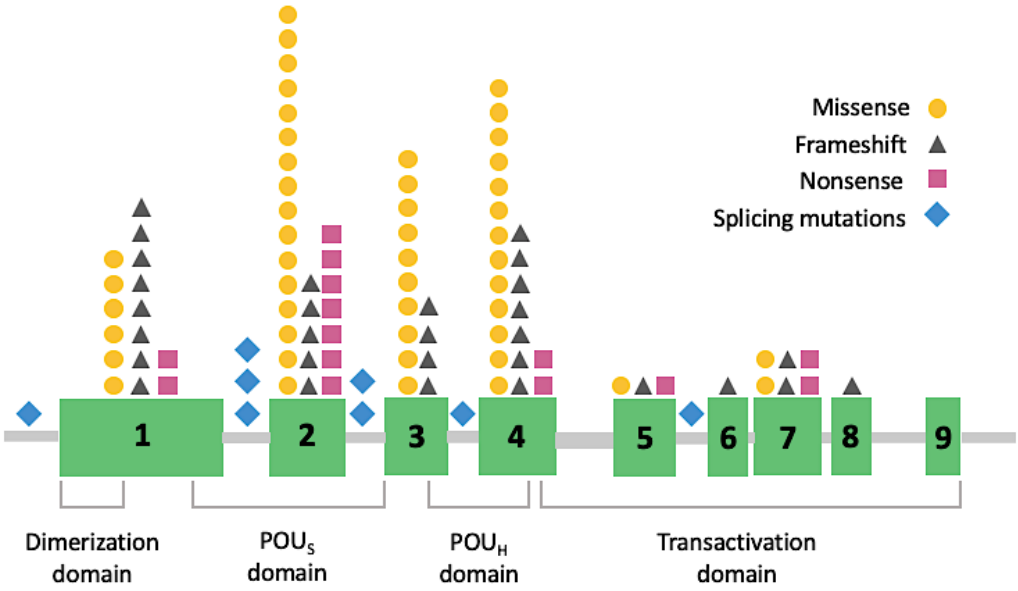

Figure 1.7: Representative distribution of pathogenic mutations within the *HNF1B* gene. The numbered boxes refer to the exons, and the symbols refer to unique mutations. The domains of the HNF1B protein are shown beneath the gene transcript. POU$_S$ is the POU-specific domain, and POU$_H$ is the POU homeodomain. The figure is adapted and moderated from Clissold et al.[52]

1.5 Zebrafish (*Danio rerio*)

The zebrafish (*Danio rerio*) is a tropical freshwater fish in the minnow family. It is native to South Asia and grows to a length of 6-7.5 cm as adults. Zebrafish attain sexual maturity around 3 months of age and usually have a lifespan of 2-3 years under laboratory conditions.[53] The zebrafish genome has been sequenced and assembled, facilitating a high-quality reference genome. The genome possesses more than 26,200 protein-coding genes.[54]

1.5.1 Zebrafish as a model organism

The zebrafish was introduced as a genetically tractable organism by Streisinger et al.[55] in the 1980s. Since then, the zebrafish has emerged as an important model organism for biomedical research.[56] It offers several advantages compared to other vertebrate models. Due to their small size, zebrafish can be housed at higher densities than rodents, making maintenance and breeding easier and more cost-effective. The fish have a rapid development cycle and produce hundreds of embryos per mating, enabling larger sample sizes and faster breeding processes. Moreover, zebrafish develop externally and are optically transparent at embryonic and early larvae stages, allowing imaging of live embryos and larvae during early life stages.[56]

The last common ancestor of humans and zebrafish is estimated to have lived around 450 million years ago.[57] Since the divergence, zebrafish and other teleost have undergone a whole-genome duplication event named teleost-specific genome duplication (TGD). This event resulted in a temporary increase in the number of genes.[58] In most cases, there was no selective advantage to maintain the duplicated gene, and the duplicated genes were gradually lost. However, some genes acquired new functions or conserved only specialised functions of the primary gene.[58] The early divergence of humans and teleost complicates the use of zebrafish as a model organism compared to other mammalian models, such as rodents, as the species are genetically less similar, and there are often two homologues for a single copy human gene.[59] However, more than 70% of zebrafish genes have a human orthologue, and 82% of all human disease-related genes have an orthologue in zebrafish.[60] Furthermore, the rapid advancements in CRISPR and next-generation sequencing technology make it easy to create various mutations in the zebrafish genome. When these mutations are bred to homozygosity, they can generate defects in multiple organ systems, resulting in pathologies similar to human disease.[60] Genes critical for pancreas development are highly conserved between zebrafish and higher vertebrates, making zebrafish an ideal model organism for investigating pancreas formation.[42,61]

1.5.2 CRISPR-Cas9 methodology

CRISPR-Cas9 is a revolutionary genome editing technology that allows for precise and site-specific changes in DNA sequences. The tool has transformed biomedical research, providing new opportunities for understanding the underlying mechanisms of a broad range of human diseases.[62] The system is based on an adaptive immune system found in bacteria, consisting of CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and Cas9 (CRISPR-associated endonuclease) proteins.[62] Here, the bacterial cell protects itself against viral infections by editing its genome, allowing it to recognise and destroy viral RNA. As an experimental tool, the DNA endonuclease activity of the CAS protein cause double-stranded DNA breaks in the genome at the targeted location determined by the guide RNA (gRNA). The gRNA facilitates RNA-DNA complementary binding, allowing for accurate identification of the target site. Once the DNA is cut, the DNA sequence is repaired by homology-directed repair (HDR) or non-homologous end joining (NHEJ). NHEJ introduces insertions or deletions that can cause frameshifts and knock out the gene, whilst HDR relies on a template strand to correct and elongate the interrupted sequence. Scientists can manipulate this template strand to contain a sequence of interest, thereby altering the DNA sequence.

In diabetic research, the CRISPR-Cas9 system has been used primarily for precise gene editing purposes.[63] It has shown great promise in studying monogenic diabetes, such as MODY. For example, Song et al.[64] used CRISPR-Cas9 to delete short fragments of the *GCK* gene in rabbits, providing a new model for preclinical studies and drug screening of MODY2. Overall, the CRISPR-Cas9 system has revolutionised biomedical research and continues to expand our possibilities in understanding and treating complex diseases.

1.5.3 The HNF1B ortholog

The hepatocyte nuclear factor 1-beta a (*hnf1ba*) gene is the zebrafish ortholog of the human MODY5-associated *HNF1B* gene. Due to the smaller introns, the size of the zebrafish gene (15.7 kb) is about one-fourth of the human *HNF1B* gene in size.[41] The genes share 86% similarity, with most differences located in the interdomain space.[41] As for the human ortholog, the zebrafish *hnf1ba* gene comprises 9 exons (Figure 1.8). The protein exhibits the same functional domains, with an almost identical homeodomain.[42] As for the human ortholog, Hnf1ba is abundantly expressed in the liver, pancreas, gut and kidney. Hnf1ba is also strongly expressed in the swim bladder and weakly expressed in the brain, muscles and eyes.[41] The first

mRNA expression is detectable between the 256-cell and 1000-cell stages, approximately 3 hours post fertilisation (hpf).[42]

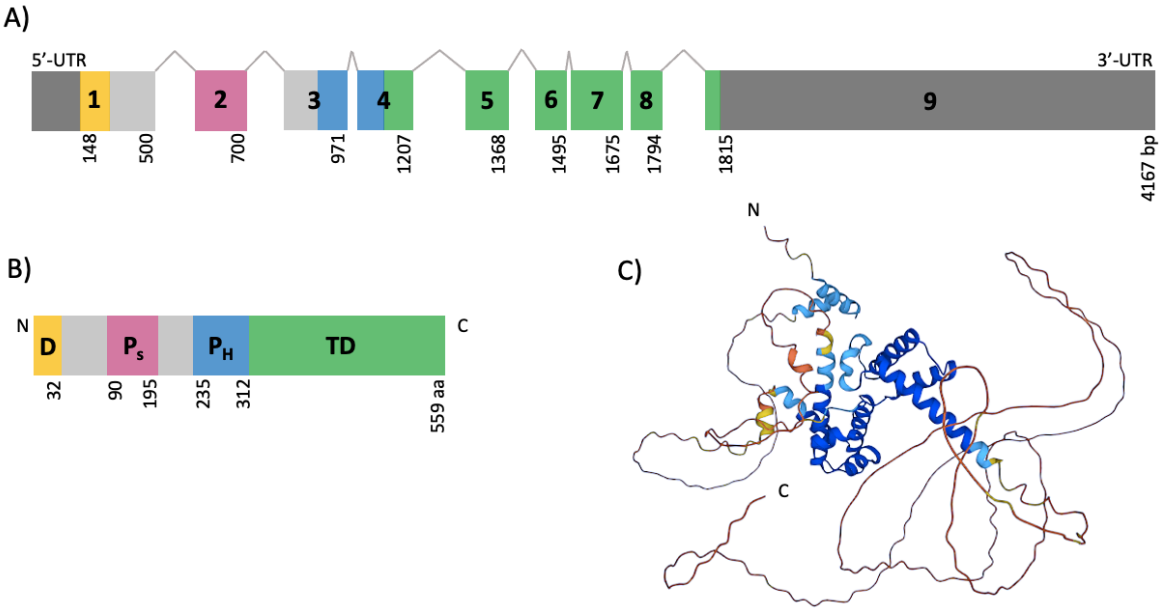

Figure 1.8: *hnf1ba* mRNA sequence and Hnf1ba protein structure. The zebrafish *hnf1ba* gene **(A)** encodes a 9-exon mRNA transcript which is translated into the 559 aa Hnf1ba protein **(B** and **C)**. Non-coding sequences are coloured dark grey. The size of the exons and introns are proportional. Hnf1ba consists of a dimerization domain (D), a POU-specific domain (P_S), a POU-homeodomain (P_H) and a transactivation domain (TD). The protein structure in C is predicted by Alphafold.[65,66]

1.5.4 Hnf1ba and the pancreatic gene regulatory network

Like the human orthologue, the zebrafish Hnf1ba protein is involved in a gene regulatory network controlling the expression of multiple genes essential for pancreatic development and function.[42,67] Some of these genes include *foxa3, gcga*, *gcgb*, *hnf1bb*, *insa, insb, pdx1* and *pax6b. Foxa3* is the zebrafish ortholog of the human *FOXA3* gene, a transcription factor required for the differentiation of endoderm-derived tissues such as the pancreas.[34] The gene is positively regulated by HNF1B and is important for maintaining the function of mature beta-cells by regulating the expression of the glucose transporter GLUT2 protein.[34,68] The role of GLUT2 was visualised in Figure 1.3. *Gcga* and *gcgb* are the duplicated orthologs of the human glucagon gene (GCG) and are believed to result from the teleost-specific genome duplication event. Why zebrafish have maintained multiple glucagon genes is currently unclear. However, studies suggest that *gcga* is sub-functionalised to produce GLP-2, a protein that plays a role in maintaining intestinal function.[69] The zebrafish *hnf1ba* gene has a paralog named *hnf1bb,* which encodes a shorter protein (532 aa) and is expressed later in embryonic development, starting at the segmentation stage at 10 hpf.[70] *Hnf1ba* is suggested as the primary functional ortholog of

the mammalian *HNF1B* during embryonic development.[67] However, much is still unknown about the role and distinct function of *hnf1bb*.

Zebrafish possess two insulin paralogs, *insa* and *insb*. Both genes have insulin activity, but their expression patterns differ.[59,61] *Insa* is believed to be expressed in pancreatic islets and functions like the mammalian ortholog in blood glucose regulation. *Insb* is suggested to be expressed earlier, in extrapancreatic tissues, playing a more important role in development.[59] However, there is limited research on the precise role of the paralogs, and many studies do not distinguish between them. *Pdx1* is one of the earliest markers of pancreatic development, and its proper regulation is essential for pancreas development due to its role during the proliferation and differentiation of the pancreatic buds.[42,71] The onset of *pdx1* expression is initiated by one cut homeobox 1 (ONECUT1), a transcription factor activated by HNF1B.[34] The role of *pdx1* in pancreas development and beta-cell differentiation is conserved in zebrafish, and the gene is not known to have a paralog.[71] Hnf1ba also regulates the expression of paired box 6b (Pax6b), a transcription factor important for the development of the endocrine pancreas, eyes and brain in vertebrates.[72] Pax6b-deleted zebrafish embryos have been shown to have almost no beta-cells, while partial reduction of pax6b-function has been shown to result in a substantial rise in the number of alpha-cells. These findings highlight the importance of fine-tuning pax6b levels for its biological activity.[72]

1.5.5 The *hnf1ba* mutant

The zebrafish mutant utilised in this project was modified by a previous master student, Rosemary Hoff, using the CRISPR-Cas9 system.[73] 11 base pairs (bp) were deleted in exon 1, resulting in a frameshift generating an early stop-codon (Figure 1.9). As a consequence, the dimerization domain of the Hf1ba protein was truncated. The deleted bp was located at position 46-56 in the coding region of the *hnf1ba* gene, corresponding to the L16fs mutation in the human *HNF1B* gene. We are not aware of any animal models representative for MODY5 that investigate this mutation.

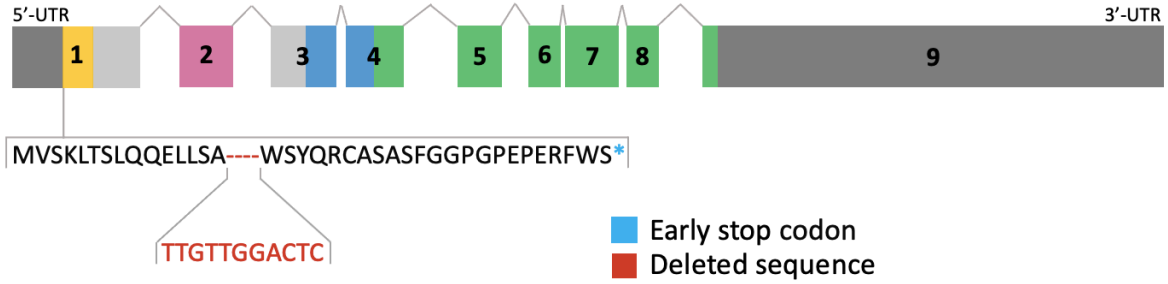

Figure 1.9: CRISPR-Cas9 mediated knockout of *hnf1ba* in zebrafish. 11 base pairs (marked in red) were deleted in the first exon of the zebrafish *hnf1ba* gene, generating a frameshift and an early stop codon (marked in blue). This mutation resulted in a truncated Hnf1ba protein. The numbers indicate the exons and the colours distinguish between the functional domains of the protein. Yellow = dimerization domain, pink = POU-specific domain, blue = POU-homeodomain, green = transactivation domain. The non-coding regions of exon 1 and exon 9 are coloured dark grey.

1.6 Aim of the study

Main aim:

To investigate whether the *hnf1ba*(-/-) zebrafish larvae could be a model for human MODY5 disease. We wanted to study how the mutation affected the pancreatic gene regulatory network, in addition to pancreas structure and function.

Sub aims:

Comparing wild-type (wt) and *hnf1ba*(-/-) zebrafish larvae by:

- examining the transcript levels of genes important for pancreas development.
- investigating pancreas structure and function by immunohistochemical analysis of the organization of alpha-cells and beta-cells.
- observing the morphological phenotypes to better understand the effect of the mutation.

2 Materials and methods

Adult zebrafish were genotyped and heterozygous *hnf1ba*(+/-) individuals were crossed to produce homozygous larvae (*hnf1ba*(-/-)). At 4 days post fertilisation (dpf), larvae displaying abnormal phenotypes were genotyped. Homozygous individuals were selected for immunohistochemistry (IHC) or reverse transcription quantitative real-time PCR (RT-qPCR). The phenotype of the homozygous larvae was examined through imaging. Figure 2.1 shows the general workflow.

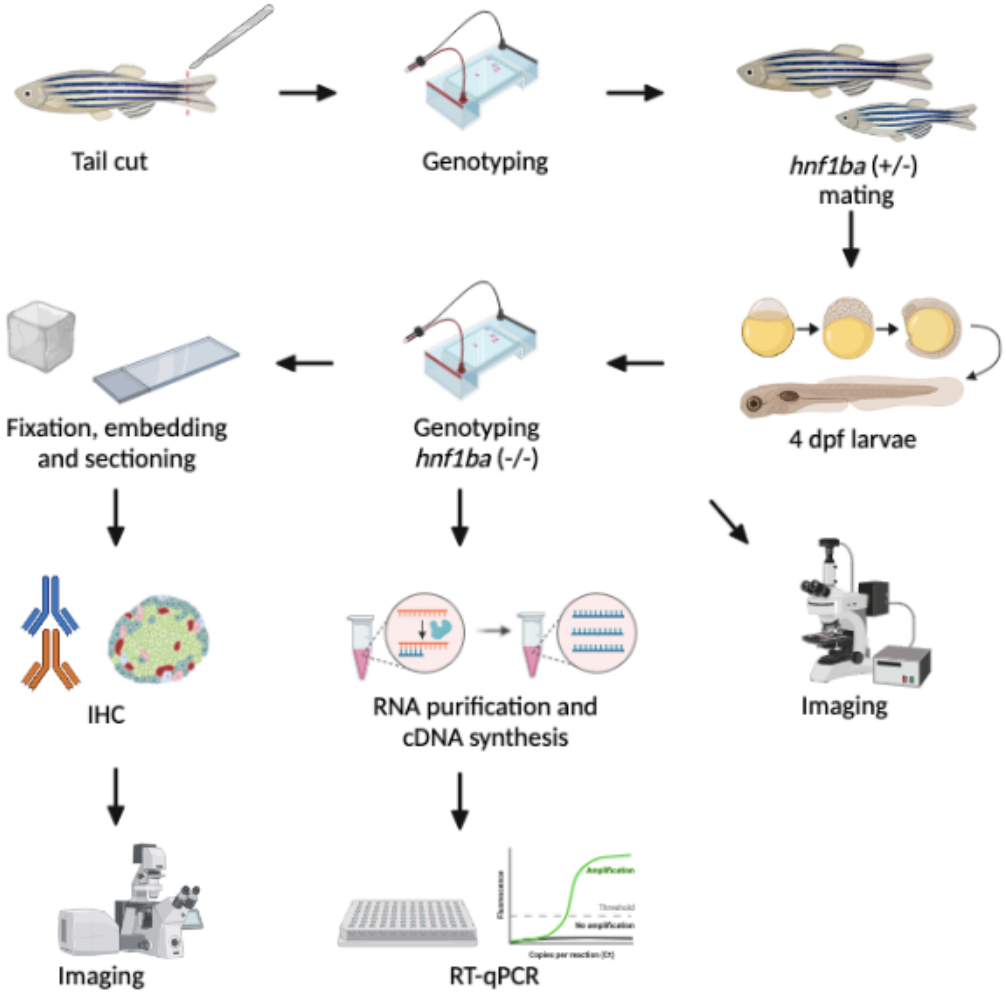

Figure 2.1: General workflow for studying *hnf1ba*(-/-) mutant zebrafish. *Hnf1ba*(+/-) adult zebrafish were genotyped and crossed, and offspring presenting an abnormal phenotype were further analysed. The genotype of the larvae was determined, and the phenotype of the 4 dpf *hnf1ba*(-/-) was characterised by microscope imaging. The pancreas of 4 dpf *hnf1ba*(-/-) and wt larvae was studied using IHC, and the expression of specific pancreatic genes was studied using RT-qPCR. The figure is made in BioRender.

2.1 Zebrafish husbandry and breeding

2.1.1 Husbandry

Zebrafish were maintained at the zebrafish facility at the Department of Biological Sciences at the University of Bergen, following guidelines described by Aleström et al.[74] These guidelines include housing adult fish in tanks filled with recirculating water of 28°C, feedings twice a day and maintaining a standardised light-dark cycle of 10 hours (h) dark and 14 h light. The fish were housed at densities of 2 to 15 fish per 3-litre tank, and behavioural patterns were monitored daily.

2.1.2 Breeding and sampling

Zebrafish heterozygous for *hnf1ba*(+/-), as well as wt zebrafish, were mated. The day before fertilisation, crossings were set up in static water with a perforated container separating males and females within the same tank. Right after the lights were turned on the following morning, the fish were placed in clean tanks, no longer separated by the perforated container. Green plastic strings were added to the breeding tanks to optimise breeding conditions, providing artificial spawning sites and places of refuge for the fish.[75] Eggs were collected after 30-40 minutes (min). Fertilised eggs were kept in Petri dishes filled with blank E3 medium at 28°C for 4 days. The Petri dishes had a maximum density of 50 embryos per dish, and dead embryos were removed daily.

At 4 dpf, we counted the embryos and examined the morphology of offspring from the *hnf1ba*(+/-) fish using a Nikon SMZ645 stereomicroscope. Larvae with an abnormal phenotype were genotyped, while the remaining larvae were disposed of as they were unsuitable for additional experiments. All larvae were euthanised by being placed in ice water for 20 min. The larvae that were confirmed to be homozygous, as well as wt larvae, were either directly fixated for immunohistochemistry or stored at -20°C in RNAlater™ for RT-qPCR analysis. However, after finding that most samples in RNAlater™ got degraded, the larvae were immediately processed for RNA extraction instead.

The females spawned between 100 to 300 eggs per mating, and multiple rounds of crossings were needed to generate sufficient amounts of genetic material for the RT-qPCR analyses. We ensured a recovery period of at least one week between each spawning to allow sufficient regeneration and minimise stress.

2.2 Genotyping

2.2.1 Individuals of interest

A total of 23 adult F1 zebrafish were genotyped. These fish originated from a cross between a CRISPR-Cas9 modified F0 founder male and a wt female. The lineage is visualised in Figure 2.2. We were interested in the F1 individuals heterozygous for *hnf1ba*(+/-) as they could be used to generate homozygous larvae carrying the mutation of interest. Genotyping was performed on F2 larvae, and individuals homozygous for *hnf1ba*(-/-) were selected for further analysis.

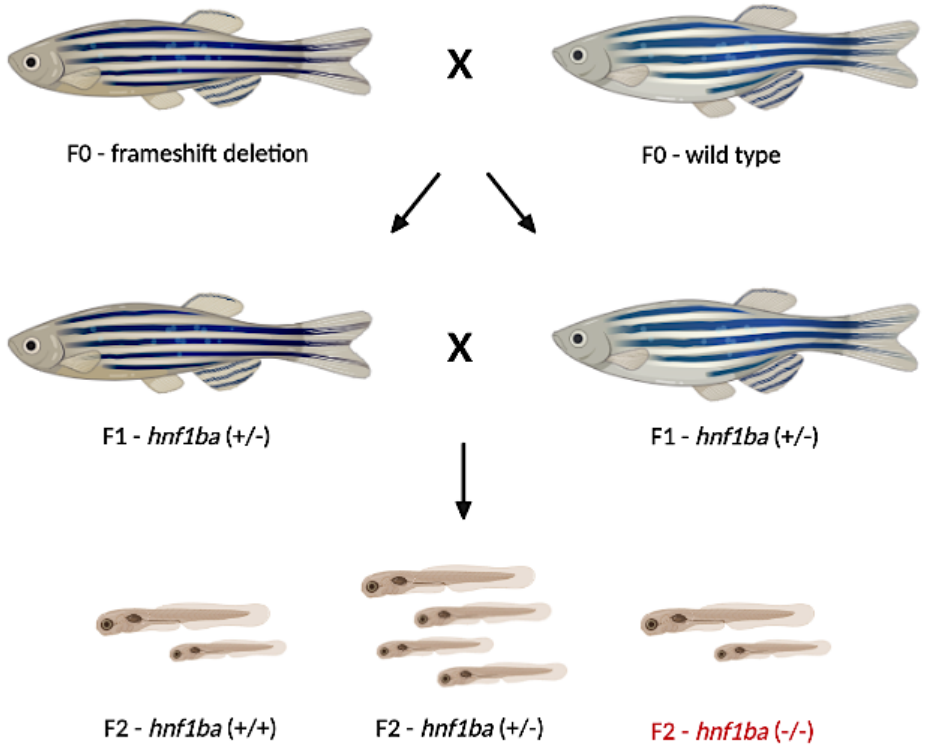

Figure 2.2: Lineage of modified zebrafish. A CRISPR-Cas9 modified F0 founder male carrying a frameshift deletion mutation in the *hnf1ba* gene was crossed with a wild-type female. F1 offspring were genotyped, and heterozygous individuals were crossed. Larvae showing an abnormal phenotype were genotyped, and homozygous *hnf1ba*(-/-) larvae (typed in red) were used in RT-qPCR analysis or IHC. The figure was made in BioRender.

2.2.2 DNA extraction – adult fish

Tail cuts were performed to extract genomic DNA from adult zebrafish. The fish were anaesthetised in 4 % tricaine for 4 min or until no gill movement could be observed. Approximately 2/3 of the tail fin was cut and placed in 100 µl lysis buffer:

Reagent	Volume
1M Tris-HCl (pH 8.0)	10 µl
5M NaCl	4 µl
0.5M EDTA (pH 8.0)	1 µl
Proteinase K (fresh) (20 mg/mL)	1 µl
ddH_2O	84 µl
Total volume per sample	100 µl

After incubating the samples for 20 min at 65°C, they were vortexed for 1 min and incubated overnight at 65°C. The next day, the samples were vortexed for 1 min and centrifuged (Eppendorf, 5424R) for 2 min at room temperature (RT) and max speed (13300 x g). The supernatant was transferred to new tubes, and 210 µl absolute ethanol was added to each sample. After inverting the samples several times, any precipitated organic material was removed, and the samples were incubated on ice for 15 min. The samples were centrifuged at 4°C at 15 000 x g for 20 min, and the supernatant was discarded. Next, the samples were washed with 210 µl 70% ethanol and centrifuged at 15 000 x g at 4°C for 10 min. The supernatant was again discarded, and the pellets were airdried for 20 min at RT in a fume hood. Finally, 15 µl nuclease-free water was added to dissolve the pellet, and the samples were incubated at 55°C for 60 min. The DNA concentrations were measured using Nanodrop One spectrophotometer (Thermo Scientific). The samples were diluted to DNA concentrations <300 ng/µl as higher concentrations could inhibit the PCR reaction.

2.2.3 DNA extraction – larvae

DNA was extracted from 4 dpf larvae using the Qiagen DNeasy Blood & Tissue Kit. For larvae used in RT-qPCR analysis, genomic DNA was amplified from an RNA pool. For larvae used in IHC, tail cut was performed on all individuals. Each tail cut was transferred to a 1.5 ml microcentrifuge tube on ice. After removing all E3 medium, 180 µl of ATL buffer and 20 µl Proteinase K were added to each tube. The tubes were vortexed and incubated at 56 °C until complete lysis of the larvae tails, which took approximately 90 min. The samples were then vortexed for 15 seconds (s). To extract the DNA, a mix of 200 µl AL buffer and 200 µl absolute ethanol was added to each sample. The samples were vortexed and transferred to DNeasy Mini spin columns placed in a 2 ml collection tube. The samples were centrifuged at 6000 x g for 1 min, and the flow-through was discarded. After placing the columns in new collection tubes, the samples were washed with 500 µl Buffer AW1 and centrifuged for 1 min at 6000 x g. The

columns were placed in new tubes, 500 µl Buffer AW2 was added, and the tubes were centrifuged for 3 min at 20 000 x g to dry the DNeasy membrane. The flow-through was discarded, and the samples were again transferred to new collection tubes. To elute the DNA, 50 µl Buffer AE was added directly onto the DNeasy membrane. The samples were incubated at RT for 1 min and centrifuged for 1 min at 6000 x g. This last step was repeated once to increase the overall DNA yield. The DNA concentrations were measured using Nanodrop One spectrophotometer.

2.2.4 PCR gene amplification

To amplify the genomic sequences surrounding the *hnf1ba* mutation (Figure 1.9), a PCR reaction master mix containing a sequence-specific primer pair was prepared:

Reagent	Volume
2X DreamTaq Green PCR Master Mix	12.5 µl
10 µM *hnf1ba* forward primer	1 µl
10 µM *hnf1ba* reverse primer	1 µl
ddH_2O	8.5 µl
Total volume per sample	23 µl

The forward and reverse primer sequences were 5'-CGCATTTCCCTCTCTAGACC-3' and 5'-GCTCTTTGAGAATTGGCGGT-3', respectively. 2 µl of DNA (400-600 ng/µl) was added to each tube filled with 23 µl PCR reaction mix. The content was mixed with a pipette and spun down before running the following PCR program using the GeneAmp® PCR System 2700 (Applied Biosystems):

Pre-denaturation	95°C	2 min	
Denaturation	95°C	30 s	x 40 cycles
Annealing	60°C	30 s	
Elongation	72°C	30 s	
Final elongation	72°C	30 s	

2.2.5 Purifying PCR products

Before proceeding with the restriction digestion, the DNA fragments were purified using the Qiagen MinElute PCR purification kit. 5 volumes of Buffer PB were added to 1 volume of PCR

product. The samples were transferred to MinElute columns placed in 2 ml collection tubes, and centrifuged at 17900 x g for 1 min at RT. The flow-through was discarded, and the samples were washed with 750 µl Buffer PE and centrifuged using the same program. The flow-through was again discarded and the columns were centrifuged for an additional 1 min. After placing the columns in clean tubes, the DNA was eluted by adding 10 µl nuclease-free water to the centre of the membrane, waiting 1 min and then centrifuging for 1 min. The DNA concentrations were measured using the Nanodrop spectrophotometer.

2.2.6 Restriction digestion

To detect the *hnf1ba* mutation, a restriction digestion assay was conducted using the BstNI restriction enzyme. The restriction site of BstNI is 5'-CCTGG-3'. The restriction digestion assay was already verified by the PhD student running the project, by the use of Sanger sequencing. The following digestion reaction mix was prepared, and the samples were incubated at 60°C for 15 min:

Reagents	Volume
10X NEBuffer r3.1	2.5 µl
Purified DNA (90 ng/µl)	8
BstNI (10,000units/ml)	0.5 µl
ddH_2O	14 µl
Total volume per sample	25 µl

2.2.7 Agarose gel electrophoresis

The digestion products were run on a 2% agarose gel. 6X loading dye was added to each sample before loading them onto the gel alongside a 100 bp DNA ladder, a positive and negative control. The electrophoresis was run at 80 V for 40 min, and the gel was imaged using the G:Box (Syngene).

2.3 Imaging phenotypes

The phenotype of wt and *hnf1ba*(-/-) mutant larvae were compared at different developmental stages by imaging. Tricaine buffer was added drop by drop onto the Petri dishes to sedate the larvae. Approximately ten drops were needed until no movement could be detected. The larvae were assembled in 3% methylcellulose and imaged at 16x and 20x magnification using the

Leica M420 photomacroscope with an Infinity camera (Lumera) and the Image-Pro® Premier software (Version 9.1.4). The larvae were imaged at 24 hpf, 48 hpf, 72 hpf and 96 hpf.

2.4 RT-qPCR

To investigate the impact of *hnf1ba* knockout on the expression of genes related to pancreas development, we conducted RT-qPCR on RNA extracts from 4 dpf old wt and *hnf1ba*(-/-) larvae.

2.4.1 RNA extraction

RNA extraction was performed using the Qiagen RNeasy® Mini Kit. To ensure sufficient amounts of RNA for analysis, 30 larvae were pooled together in a single tube. 40 µl of 1M dithiothreitol (DTT) was added to 1 ml Buffer RLT. The tissues were disrupted and homogenised in 350 µl Buffer RLT + DTT using a mechanical homogeniser. The lysate was centrifuged at 15,000 x g at RT for 3 min, and the supernatant was collected and mixed with 350 µl 70% ethanol. Next, 700 µl of each sample was transferred to an RNeasy Mini spin column placed in a 2 ml collection tube. The columns were centrifuged for 15 s at 8000 x g, and the flow-through was discarded. 700 µl Buffer RW1 was added, the columns were centrifuged for 15 s at 8000 x g, and the flow-through was discarded. The columns were then washed by adding 500 µl Buffer RPE, centrifuging using the same program, and discarding the flow-through. This washing step was repeated once. To dry the spin column membrane, the spin columns were centrifuged at 15,000 x g for 1 min. The columns were placed in new 1.5 ml tubes, and 30 µl RNase-free water was added directly to the membrane to elute the RNA. The columns were then centrifuged for 1 min at 8000 x g. The RNA concentrations were measured using the NanoDrop One spectrophotometer, and the samples were standardised to the lowest concentration by diluting in nuclease-free water and adjusting to a volume of 11 µl. The samples were stored at -80°C to limit RNA degradation.

2.4.2 Assessing RNA integrity

The RNA integrity number (RIN) was determined using the Agilent 2100 Bioanalyzer (G2939BA) and the Agilent RNA 6000 Nano Kit. 550 µl of gel matrix were placed onto the top receptacle of a spin filter and centrifuged for 10 min at 1500 x g at RT. The RNA 6000 Nano dye concentrate was vortexed for 10 s and spun down. The gel-dye mix was prepared by adding 1 µl of the Nano dye concentrate to 65 µl aliquot of the filtered gel already prepared.

The solution was vortexed thoroughly and spun down for 10 min at RT at 13 000 x g. 9 µl of the gel-dye mix was pipetted to the bottom of the well on the disposable RNA Nano Chip. A plunger positioned at 1 ml was closed on the chip priming station for 30 s. 9 µl of gel-dye mix was pipetted into two wells, and 5 µl of the RNA 6000 Nano Marker was loaded into the ladder-well and the 12 sample-wells. The samples were heat denatured at 70°C for 2 min to minimise the RNA's secondary structure, and 1 µl of each sample was loaded onto the chip. 1 µl of the DNA ladder was pipetted into the ladder-well and the chip was vortexed at 2400 rpm for 60 s. Finally, the chip was inserted into the 2100 Bioanalyzer and run using the 2100 Expert Software (Version B.02.08.SI648).

2.4.3 cDNA synthesis

Complementary DNA (cDNA) was synthesised using the Invitrogen SuperScript III Reverse Transcriptase kit. Each sample was added 2 µl of Mix 1 (table below) and incubated at 65°C for 5 min. The samples were put on ice for 1 min and briefly centrifugated before 7 µl of Mix 2 (table below) was added to the samples. A no-reverse transcriptase (NRT) control was prepared by running the cDNA synthesis as mentioned but replacing SuperScript™ III RT in Mix 2 with nuclease-free water. This negative control allowed us to assess the primers' ability to amplify the traces of genomic DNA left in the original RNA samples. All samples were incubated at 50°C for 60 min, and the reaction was finally inactivated by heating at 70°C for 15 min. The samples were stored at -20°C.

Mix 1:

Reagents	Volume
50 µM Oligo(dt)$_{20}$ Primer	1 µl
10 mM dNTP Mix	1 µl
Total volume per sample	2 µl

Mix 2:

Reagents	Volume
5X First-Strand Buffer	4 µl
0.1 M DTT	1 µl
RNaseOUT (40U/µl)	1 µl
SuperScript™ III RT (200 U/µl)	1 µl
Total volume per sample	7 µl

2.4.4 Primer design

The primer pairs utilised in the RT-qPCR analysis were already designed by a PhD student in this project, using BioRad's qPCR Assay Design and Optimalization page to select target sequences and primer design.[76] Zebrafish genes with human orthologs associated with early pancreas development were selected, as explained in the introduction. Three housekeeping

genes (HKG) were also selected: *hprt1l, tmem50a* and *tuba1c*. The table below presents all tested genes and their corresponding primer sequences. Teleost-specific paralogous genes are distinguished by "a" and "b".

Table 2.1: Primer sequences and characteristics. The hashtag (#) distinguishes between primer pairs tested for the same gene. Tm is the melting temperature of the primer, GC is the %GC-content of the primer, and size is the length of the PCR product. The asterisk (*) indicates the junction between an exon and an intron. The primer sequences of genes written in bold green were considered optimal and selected for the RT-qPCR sample analysis. *Insb* primer pair 2 (written in orange) was not considered optimal but was still included in the analysis to get an insight into the relationship between the insulin paralogs.

Gene	**#**	**Primer Sequence (5'→3')**		**T_m (°C)**	**GC (%)**	**Size (bp)**
actb1	-	Forward	GAGAAGATGACACAGATCATGTTC	61.5	41.7	118
		Reverse	CAGAGTCCATCACAATACCAG	59.6	47.6	
foxa3	1	Forward	CTACAGTGAGGCAAATGAG*ATG	61.3	45.5	142
		Reverse	GCTGCTGTTGAGACCTGC	62.6	61.1	
foxa3	2	Forward	GATGTTGAGCTCCGTGAAGATG	65.9	50.0	193
		Reverse	GCTGTTGAGACCTGCGCTG	67.5	63.2	
gcga	1	Forward	GACAGCACAAG*CACAGAG	58.1	55.6	130
		Reverse	CTTGTGCTCTTCTGGTCTC	57.6	52.6	
gcga	2	Forward	CGCTAAGAGAAATGG*TGG	58.4	50.0	124
		Reverse	GGGTTGTCCGGATTTTAG	59.4	50.0	
gcgb	1	Forward	GAAAACGGCAG*CCTTATG	60.8	50.0	131
		Reverse	CTTGTGCTCTCCTGGTCTC	60.2	57.9	
gcgb	2	Forward	CTCTAAGAGGAGTGG*AGTCCC	60.8	57.1	78
		Reverse	GGTCTGAAGAATCCTCATCG	61.2	50.0	
hnf1ba	-	Forward	CAAGATGCAAG*GTGTCCG	63.1	55.6	128
		Reverse	CCTGGAGAGACCTGCTGTAAC	62.8	57.1	
hnf1bb	-	Forward	CATGGGCTTGGATCTAAC	58.0	50.0	190
		Reverse	GTTGACTGTACCTGAGCT*TTG	59.0	47.6	
hprt1l	1	Forward	GCCTCATTAAAGACAG*GACGG	64.6	52.4	75
		Reverse	GAGCCACGATATGATGACCG	65.4	50.0	
hprt1l	2	Forward	GTGAAAGTGGTGAGTCTGCTG	62.6	52.4	174
		Reverse	CTTCTCCCTTGCTCTGTCG	62.8	57.9	

insa	-	Forward	ATGGCAGTGTGGCTTCAGG	66.4	57.9	230
		Reverse	AATGCAAAGTCAGCCACCTCAG	67.5	50.0	
insb	1	Forward	ATCTGTGTGGTTCAAGCCTG	62.8	50.0	112
		Reverse	GAGAGCAAAG*CCAGCAAG	61.5	55.6	
insb	2	Forward	GAGACCTTGCTGG*CTTTGC	65.1	57.9	91
		Reverse	CCCTCTTCATCTTCATCACCTTC	64.9	47.8	
nkx6.1	1	Forward	TTCTGCTCGACAAAGATGGG	65.3	50.0	168
		Reverse	ACCACACCTTGA*CTTGACTC	59.2	50.0	
nkx6.1	2	Forward	GAAAAGGAAACACACGCGAC	64.5	50.0	147
		Reverse	CACACCTTGA*CTTGACTCTCTG	61.6	50.0	
nkx6.2	-	Forward	GGACGCCTCATGGAATAAGC	65.6	55.0	250
		Reverse	ATCATGTTTGCTTGGG*CCG	68.1	52.6	
onecut1	1	Forward	GGAAATGGTTGCAGGAGC	63.3	55.6	183
		Reverse	ACGGACGCTTGTTCTCTTTG	64.3	50.0	
onecut1	2	Forward	CTCAAGTCTGGTCGGGAAAC	63.6	55.0	102
		Reverse	CTTTCGTTTGCACG*CTGC	65.8	55.6	
pax6b	-	Forward	AAGCCCAGAGTAGCGACG	63.2	61.1	150
		Reverse	GGTTTATCGACGACAC*GCTC	64.5	55.0	
pdx1	-	Forward	ACTCTCTGGACCTCTGCGG	64.7	63.2	115
		Reverse	ACCATATAAGGGC*CTGTCCAC	63.7	52.4	
tmem50a	-	Forward	CTTGGGACAAACAG*GTGC	61.8	55.6	79
		Reverse	CCACATAGAGGCGATGAGAG	62.4	55.0	
tuba1c	-	Forward	TGCCTCAATCTTGGACAGTG	64.0	50.0	193
		Reverse	TGGATGCCATGCTCAAGAC	65.1	52.6	

2.4.5 Primer validation

Dilution and melting curves were generated for all primer pairs to determine primer efficiencies and specificities, and to determine the optimal cDNA concentration for the RT-qPCR analysis. A 2-fold dilution series from a cDNA pool of wt and *hnf1ba*(-/-) mutant larvae was prepared, testing the concentrations 1:10, 1:20, 1:40, 1:80 and 1:160. The following PCR reaction mix was prepared for each primer pair:

Reagent	Volume
2X iTaq™ Universal SYBR® green Supermix	12.5 µl
5 µM each forward and reverse primer mix	1 µl
ddH_2O	6.5 µl
Total volume per primer pair	20 µl

5 µl of diluted cDNA and 20 µl of PCR reaction mix were added to the wells of a 96-well plate in technical triplicates, alongside a no-template control (NTC) and an NRT control. The NTC contained nuclease-free water instead of cDNA and was aimed to test for contamination or primer-dimer formation. The plate was sealed with plastic adhesive film, spun down and run using the following PCR program using the C1000 Touch ™ Thermal Cycler (BioRad, CFX96™ Real-Time System).

Pre-denaturation	95°C	2 min	
Denaturation	95°C	15 s	x 37 cycles, plate read
Annealing + elongation	* °C	25 s	
Pre-melt denaturation	95°C	10 s	
Annealing	55°C	5 s	
Melting curve analysis	55-95°C	30 s	x 80 cycles, 0.5°C/cycle, plate read

The asterisk (*) indicates the temperature that was to be optimised. This temperature was tested in the range of 58°C-62°C. Ct values generated from the RT-qPCR run were plotted against the negative logarithmic dilution, and a linear regression line was fitted. The slope of the regression line was used to calculate the primer efficiency by equations 1 and 2 adapted from Pfaffl.[77]

$$E = 10^{-\left(\frac{-1}{slope}\right)} \quad (1)$$

$$\%E = (E - 1) * 100 \quad (2)$$

E = primer efficiency

Primer efficiencies between 85-105% were considered optimal. Efficiencies above 105% could result from high Ct value variances between the dilutions, while efficiencies below 85% were considered too low. Ct values under 30 were considered optimal as Ct values above 30 could be high enough for other sequences to be amplified and give false positive results.

The primer specificity was studied by analysing melting curves generated from the RT-qPCR-runs. The curve should display only one peak indicating the melting point of the sequence. A second peak would indicate that the primers were non-specific, that primer dimers were formed, or that the genomic version of the gene could be amplified. Neither the NTC nor NRT controls were accepted with Ct values less than ten cycles higher than the experimental samples.

2.4.6 Sample analysis

The RT-qPCR analysis was performed on those genes that possessed primer pairs with an effective design (Table 2.1), utilising the optimal temperature and dilution as determined from the dilution curves. PCR reaction mixes were prepared as described earlier, except that half the volumes were used. 2.5 µl of cDNA (1:10 dilution) and 10 µl PCR reaction mix were added to the wells of a 96-well plate in technical duplicates, alongside a NTC. The plate was sealed with plastic adhesive film, spun down and run using the same PCR program as earlier, with optimised amplification temperatures.

2.4.7 Relative quantification and statistical analysis

The relative gene expression (RGE) of the genes of interest (GOI) was calculated using the following equation:

$$RGE = \frac{(E_{GOI})^{\Delta Ct_{meanGOI}}}{(E_{Ref})^{\Delta Ct_{meanref}}} \qquad (3)$$

E_{GOI} = Efficiency of gene of interest

E_{ref} = Efficiency of reference gene (*hprt1l* and *tuba1c*)

$\Delta CT_{meanGOI}$ = the mean Ct value of 2 sample replicates

$\Delta CT_{meanref}$ = the mean Ct value of 2 sample replicates of reference genes

Statistical analysis was performed in RStudio (Version 2022.07.2). The normality of the data was tested using the Shapiro-Wilk normality test ($p < 0.05$). An independent-sample Mann-Whitney U test was conducted to determine whether there was a significant difference in the mRNA expression levels of wt and *hnf1ba*(-/-) zebrafish. The significance threshold was set at $p < 0.05$ (indicated *) and $p < 0.005$ (indicated **). The RGE was presented as boxplots using the same software.

2.5 Immunohistochemistry

Immunohistochemistry was performed on sections of 4 dpf wt and *hnf1ba*(-/-) mutant zebrafish larvae. The sections were stained with glucagon and insulin antibodies as markers for alpha- and beta-cells to visualise the pancreas' structure and developmental state.

2.5.1 Fixation and cryopreservation

Individuals carrying the *hnf1ba*(-/-) mutation and wt larvae were selected for IHC. At 4 dpf, the larvae were euthanised in ice water for 20 min. The larvae were put in 4% paraformaldehyde (PFA) at 4°C for 48 h to fixate the structure. After fixation, the larvae were washed with PBS for 3x5 min using a nutating mixer (VWR, 82007-202). The larvae were cryopreserved in 25% sucrose in PBS at 4°C for 24 h.

2.5.2 Embedding and cryosectioning

The larvae were either stored in 25% sucrose + 25% Tissue-Tek® in PBS at -20°C until usage or directly embedded in Tissue-Tek®. To embed the larvae, 1 cm x 1.5 cm moulds of aluminium foil were prepared and filled with Tissue-Tek®. Several individuals were aligned in the medium, with the dorsal side facing upwards and anterior towards a pre-determined side of the mould. The mould was transferred to dry ice for hardening, wrapped in parafilm and stored at -80°C for at least one night. Cryosectioning was performed with a Leica CM 3050 cryostat. The blocks were mounted to the cryostat holder, and 10 µm sections were generated and collected sequentially onto a series of 4 room-tempered microscope slides. The slides were incubated at 65°C for 10 min to fix the sections to the glass slides, and stored at -80°C until staining.

2.5.3 Antibody staining

The slides were thawed at RT for 1 h and incubated at 65°C for 10 min. To rehydrate and permeabilise the sections, the slides were washed in PBS containing 0.05% EcoSurfEH-9 for 10 min (referred to as PBS 0.05% ES from here on). The back of the slides was dried with a Kimtech wipe, and the outer sections were dried with Whatman paper. The slides were then placed in a dark, humified chamber and covered with 200 µl of 1:100 rabbit-anti-insulin and 1:100 mouse-anti-glucagon antibody solution. The antibody solution was made in PBS containing 0.05% EcoSurfEH-9 and 1X Casein buffer (referred to as PBS 0.05% ES-Cas from here on). Two slides were left unstained and were used as controls. After incubating the slides at RT overnight, the primary antibody solution was carefully flushed away with PBS 0.05%

ES. The slides were washed 3x15 min in PBS 0.05% ES and dried as previously described. The slides were then placed in a dark, humified chamber and covered with 200 µl of secondary antibody solution (1:400 goat anti-rabbit Alexa Fluor Plus 555 and goat anti-mouse Alexa Fluor Plus 647 diluted in PBS 0.05% ES-Cas). After 45 min, the secondary antibody solution was carefully flushed away with PBS 0.05% ES. The slides were washed 2x5 min in PBS 0.05% ES and then 3x5 min in PBS without ES to prevent background fluorescence during imaging. Slides were dried as previously described, mounted with 2 drops of ProLong™ Glass Antifade Mountant with NucBlue™ Stain and covered with glass coverslips. The slides were sealed with nail polish to prevent dehydration, and kept at 4°C overnight for curing.

2.5.4 Imaging

The slides were scanned using a Zeiss Axio Scan Z.1 slide scanner to locate the most relevant sections. This selection was made to minimise the impact of irrelevant slides, such as those showing only the ends of the pancreas instead of the main part. The selected sections were then imaged using the Olympus FV3000 confocal microscope and the FluoView31S software (Version 2.4.1.198). Sections were imaged using z-sections, 2.20 zoom and 60x magnification. Images were then edited using ImageJ (Version 1.53).

2.6 Alpha- and beta-cell quantification

2.6.1 Cell count

The number of glucagon- and insulin-producing cells were counted on confocal images of cryosections containing the pancreas of 4 dpf *hnf1ba*(-/-) mutant and wt larvae. These cells were used as markers for alpha- and beta-cells. We analysed the pancreas of 6 mutants and 6 wt individuals and counted the cells on the z-sections containing the most cells per individual.

2.6.2 Statistical analysis

Statistical analysis was performed in RStudio (Version 2022.07.2). The normality of the data was tested using the Shapiro-Wilk normality test ($p < 0.05$), and the Breusch-Pagan test was used to test for heteroscedasticity ($p < 0.05$). An independent t-test was conducted to determine whether there was a significant difference in the number of insulin and glucagon cells between wt and *hnf1ba*(-/-) zebrafish. The significance threshold was set at $p < 0.05$ (indicated *). The number of cells was presented as boxplots.

2.7 Materials

Reagents and kits

Name	Manufacturer	Catalogue #
100 bp DNA ladder	BioLabs	N3231S
Agarose	Sigma	A9539-100G
Agilent RNA 6000 Nano Kit	Agilent Technologies	5067-1511
BstNI	BioLabs	R0168S
ddH_2O	MilliPore	A10 Toc Monitor
DL-Dithiotreitol solution (DTT)	Sigma	BCBW7450
DNeasy Blood & Tissue Kit	Qiagen	69506
dNTPmix (10mM ea)	Invitrogen	10534823
DreamTaq Green PCR Master Mix (2X)	Thermo Scientific	K1081
ECOSURF™ EH-9 solution	Sigma	STS0012-100ML
Ethanol Absolute	VWR	20821.330
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich	BCBV6404
Gel Loading Dye, Purple (6X)	BioLabs	B7024S
GelRed® Nucleic Acid Gel Stain	Biotium	41003
Goat anti-Mouse Alexa Fluor™ Plus 647	Invitrogen	A32728
Goat anti-Rabbit Alexa Fluor™ Plus 555	Invitrogen	A32732
Insulin antibody	GeneTex	GTX128490
iTaq™ Universal SYBR® Green Supermix	BioRad	172-5121
Methylcellulose	Merck	M0512
MinElute® PCR purification kit	Qiagen	28006
Monoclonal Anti-Glucagon antibody	Sigma-Aldrich	G2654-.2ML
Nuclease-Free Water	Qiagen	129112
$Oligo(dT)_{20}$ Primer	Invitrogen	11685581
Paraformaldehyde (PFA)	Electron Microscopy Sciences	15714-S
PBS Tablets	Merck	524650-1EA
ProLong™ Glass Antifade Mountant with NucBlue™ Stain	Invitrogen	P36981
Proteinase K	Qiagen	19131

RNAlater™	Sigma	R0901-100ML
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Invitrogen	10154652
RNeasy® Mini kit	Qiagen	74104
Sodium chloride	Merck	STBG9158
Sucrose	Sigma	SLCH3216
SuperScript® III Reverse Transcriptase Kit	Invitrogen	10432122
TAE Buffer (50X) (Tris-Acetate-EDTA buffer)	Apollo Scientific	BI6050
Tissue-Tek®	Sakura	4583
Tricaine methanesulfonate	VWR	KS413
Tris(hydroxymethyl)aminomethane (Tris-HCl)	Sigma-Aldrich	BCBT8082

Buffers and solutions

Name	**Protocol**
1X E3 medium	60X E3 diluted in ddH_2O
1X PBS (pH 7.4)	1 PBS tablet in 1 L ddH_2O
1X PBS + 0.05% EcoSurf EH-9	250 µl EcoSurf EH-9 in 500 mL 1X PBS
1X PBS + 0.05% EcoSurf EH-9 + 1X casein buffer	25µl EcoSurf EH-9 and 5 ml 10X casein buffer in 45 ml 1X PBS
4% tricaine	4.2 ml tricaine solution diluted in 100 ml tank water
4% PFA in PBS	32% PFA diluted in 1x PBS
Tricaine solution	400 mg tricaine powder diluted in 97.9 ml ddH_2O and 2.1 ml 1M Tris (pH 9)

Instruments

Name	**Manufacturer**	**Catalogue #**
2100 Bioanalyzer	Agilent	G2939BA
C1000 Touch™ Thermal Cycler	BioRad	CFX96™ Real-Time System
Centrifuge	Eppendorf	5424R/5804R

Centrifuge	Thermo Scientific	Heraeus Pico 17
Confocal Microscope	Olympus	FV3000
G:Box	Syngene	F3 GelDoc
GeneAmp® PCR System 2700	Applied Biosystems	
Leica Cryostat	Leica	CM3050 S
Microscopy Camera	Lumenera	Infinity 3-3UR
MiniStar Silverline Centrifuge	VWR	
Nanodrop Spectrophotometer	Thermo Scientific	ND-ONE-W
Nutating Mixer	VWR	82007-202
Photomacroscope	Leica	M420
PowerPac™ Basic	BioRad	
Slide Scanner	Zeiss	Axio Scan Z.1
Stereomicroscope	Nikon	SMZ645

Equipment

Name	**Manufacturer**	**Catalogue #**
Adhesive Foil for Microplates	VWR	60941-076
Hard-Shell PCR Plates 96-well	BioRad	HSP9601
Immuno Stain Moisture Chamber	EMS	62010-37
KIMBLE® Pellet Pestle® Cordless Motor	Merck	Z359971
Menzel-Gläser Cover glasses	VWR	#1,5
Microseal® 'B' seal	BioRad	MSB1001
Superfrost™ Plus Adhesion Microscope Slide White Tab	Epredia	J1800AMNZ

Software

Name	**Supplier**	**Version**
2100 Expert	Agilent Technologies	B.02.08.SI648
FluoView31S	Olympus	2.4.1.198
Image-Pro® Premier	MediaCybernetics	9.1.4
ImageJ	National Institute of Health	1.53
RStudio	Posit	2022.07.2

3 Results

3.1 Genotyping of *hnf1ba* mutants

Larvae and adult zebrafish were genotyped to identify larvae homozygous for the *hnf1ba* mutation. Genomic sequences surrounding the mutation were amplified by PCR, and the resulting amplification products were digested with the BstNI restriction enzyme. The target sequence of BstNI (CCTGG) is present once in wt sequences and twice in sequences carrying the frameshift deletion mutation. This is due to the deletion of 11 base pairs, which generates a target sequence at the deletion site in the *hnf1ba* mutant. By using the location of the restriction site to predict the size of the DNA fragments, we could identify wt, homozygous and heterozygous individuals. Wt larvae showed bands at 143 and 195 bp, while homozygous mutants exhibited bands at 195, 91 and 41 bp (Figure 3.1). Heterozygous individuals showed all four band sizes.

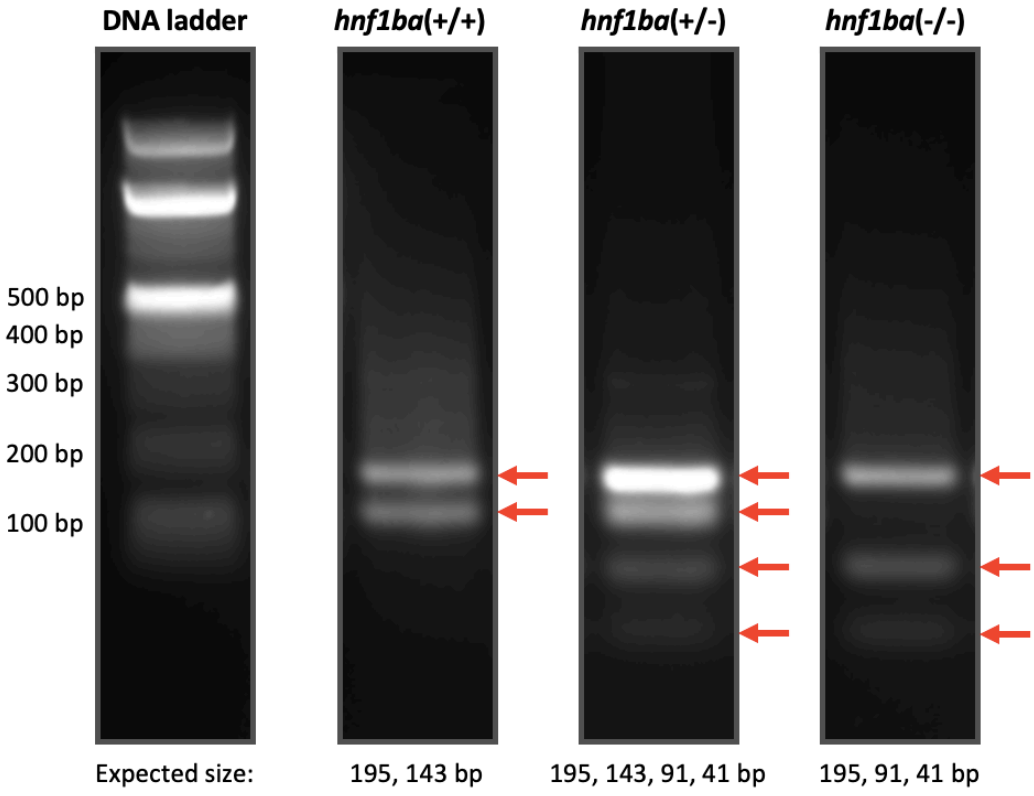

Figure 3.1: Genotyping of *hnf1ba* mutant zebrafish by restriction enzyme digestion and agarose gel electrophoresis. DNA was extracted from zebrafish larvae and adults, and genomic sequences surrounding the *hnf1ba* mutation were amplified by PCR. The PCR products were cut with the BstNI restriction enzyme and run on a 2% agarose gel to verify the *hnf1ba* mutation. A 100 bp DNA ladder (left) was used to analyse the size of the DNA fragments.

3.2 Phenotype analysis

We compared the phenotype of *hnf1ba*(-/-) larvae with that of wt larvae at 24hpf, 48hpf, 72hpf and 96hpf (Figures 3.2 and 3.3). Homozygous mutants were indistinguishable from wt larvae at the 24h and 48h stages. At 72h, the differences between the groups were more noticeable, enabling us to identify homozygous individuals. Abnormal characteristics observed in affected individuals at 72hpf include a curved tail, an enlarged yolk sac, blood accumulation in the pericardium, pericardial edemas, and a slow heart rate (Figure 3.2). Furthermore, most examined homozygous larvae had grey spots on their eyes and showed variation in the protrusion of their cornea. At 96hpf, morphological differences were more evident, making it easier to distinguish homozygous mutants from wt larvae (Figure 3.3). The differences in the yolk sac size were more prominent as wt larvae almost had no yolk sac remaining. In addition to the mentioned features, homozygous larvae had shorter body lengths and underdeveloped swim bladders and pectoral fins compared to wt larvae. Mutant larvae had little or no movement, even when provoked with a needle tip. However, the extent of the features varied greatly between the mutant individuals. Some individuals exhibited only one or two of the mentioned features, such as an enlarged yolk sac or a reduced heart rate. Others exhibited all of the mentioned features. None of the differences were quantified.

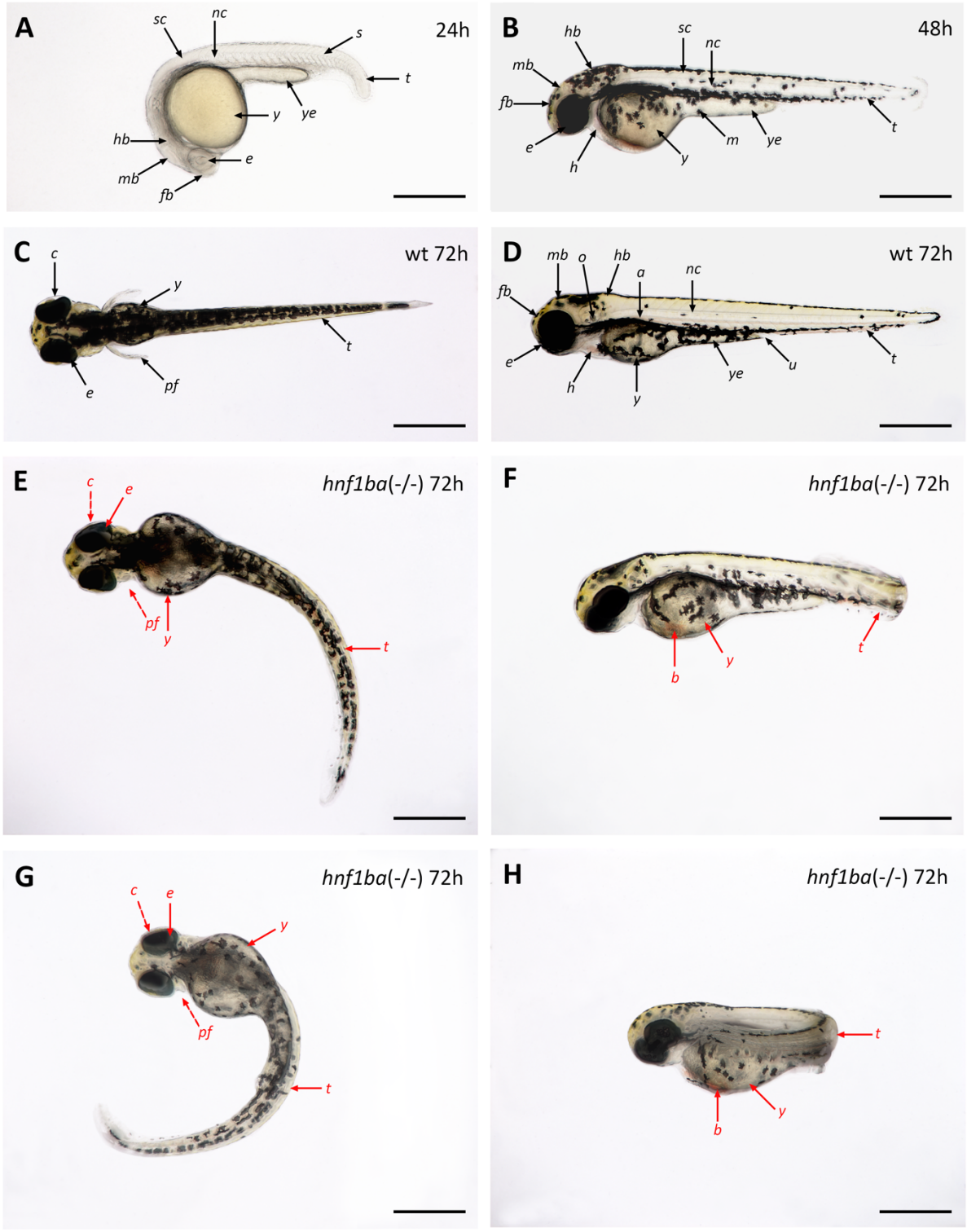

Figure 3.2: Morphological comparison of *hnf1ba*(-/-) and wild-type (wt) zebrafish larvae at different developmental stages. After removing the chorion, the larvae were sedated and positioned in methylcellulose for imaging at 20x magnification at 24h, 48h and 72h post fertilisation. The images C, E and G were captured top-down, while A, B, D, F and H show a side view of the larvae. The two genotype groups could not be distinguished at 24h (A) and 48h (B) as no abnormal morphological features were present. At 72h (C-H), mutant *hnf1ba*(-/-) individuals could be identified by showing specialised features such as curved tails and enlarged yolk sacs. Two representative mutant individuals were imaged at 72h to illustrate the varying phenotype (E, G and F, H). Black

arrows indicate general anatomical structures, while red arrows indicate distinctive features separating mutants from wt larvae. The dotted arrows indicate underdeveloped structures in the *hnf1ba* mutant compared to wt larvae. Explanations: a = dorsal aorta, b = blood accumulation, c = cornea, e = eyes, fb = forebrain, h = heart, hb = hindbrain, m = melanocytes, mb = midbrain, n = notochord, o = otolith, pf = pectoral fin, s = somites, t = tail, u = urogenital opening, y = yolk sac, ye = yolk extension. Scale bar = 5 mm.

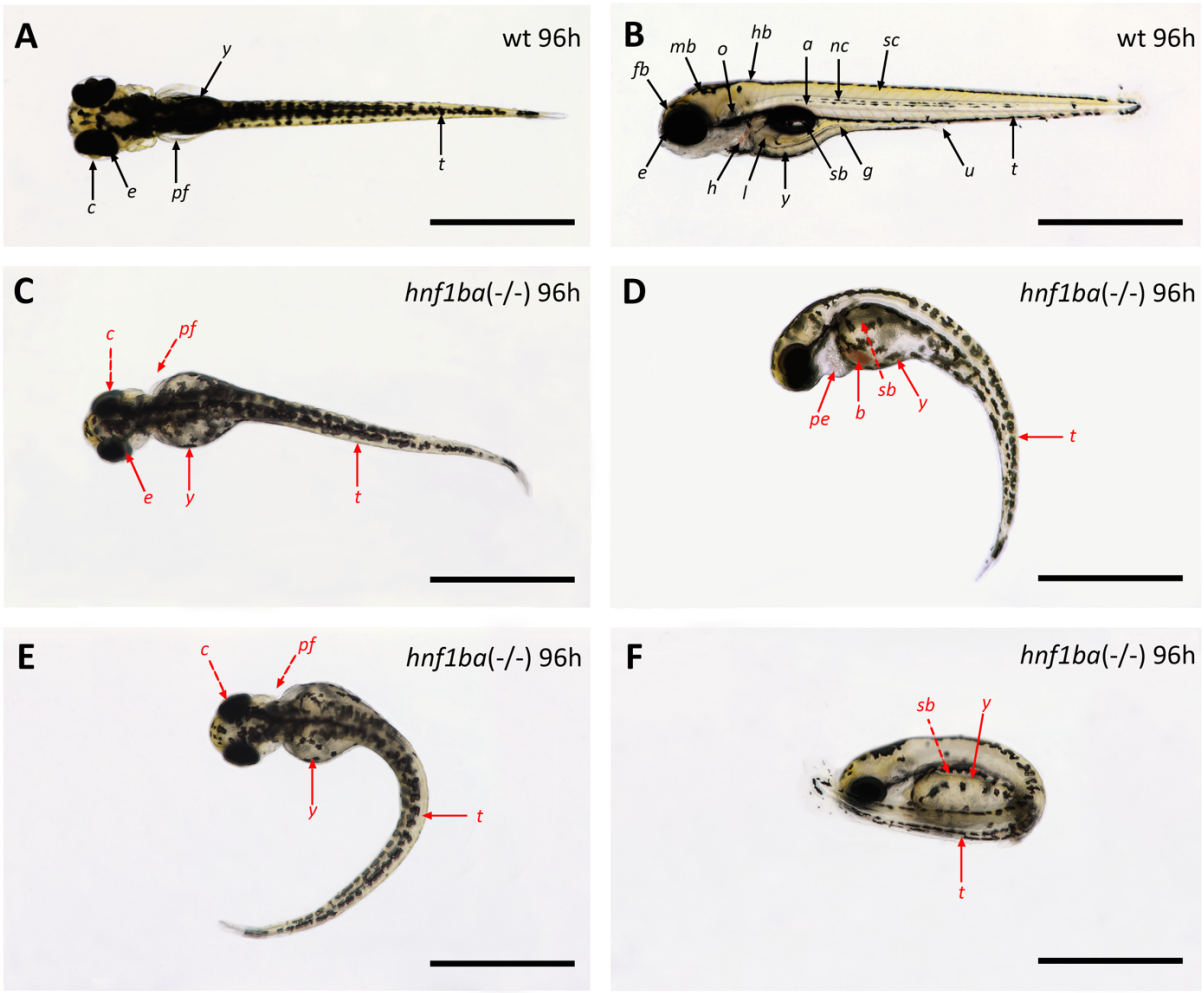

Figure 3.3: Morphological comparison of *hnf1ba*(-/-) mutants and wild-type (wt) zebrafish larvae at 96 hours post fertilisation. The larvae were sedated, organised in methylcellulose and imaged using 16x magnification. Image A, C and E were captured top-down, while B, D and F show a side view of the larvae. Two representative mutant individuals were imaged to illustrate the varying phenotype (C, E and D, F). Black arrows indicate general anatomical structures, while red arrows indicate distinctive features separating mutants from wt larvae. The dotted arrows indicate missing structures in the *hnf1ba* mutant compared to wt larvae. Explanations: a = dorsal aorta, b = blood accumulation, c = cornea, e = eyes, fb = forebrain, h = heart, hb = hindbrain, g = gut, l = liver, mb = midbrain, nc = notochord, o = otolith, pe = pericardial edema, pf = pectoral fin, s = somites, sb = swim bladder, sc = spinal cord, t = tail, u = urogenital opening, y = yolk sac. Scale bar = 1 cm.

3.3 The pancreatic gene regulatory network

The transcription level of genes important for early pancreas development was examined to investigate the effect of the *hnf1ba* knockout. These genes included *foxa3, hnf1ba, hnf1bb, insa, insb, gcga, gcgb, pax6b,* and *pdx1*. RNA was extracted from 4 dpf homozygous and wt larvae, and used as a template for cDNA synthesis. RT-qPCR was performed, and the relative gene expression was determined by normalising to HKGs. This allowed us to examine potential changes in pancreas development at the transcriptional level in the *hnf1ba*(-/-) larvae.

3.3.1 Determination of primer efficiencies

The efficiency of the designed primer pairs was tested by preparing dilution curves. Ct-values were plotted against the logarithm of five cDNA dilutions, and linear regression was performed (Figure 3.4). The slope of the regression line was used to calculate primer efficiencies and the coefficient of determination to determine the linearity of the amplification. All dilution curves showed a linear decrease in Ct-values with an increase in log10 dilution, and 1:10 was selected as the optimal cDNA dilution. After several rounds of optimisation, an efficiency percentage between 85-105% and Ct-values below 30 at 1:10 dilution were considered acceptable. Eight primer pairs were found optimal, in addition to the primer pairs of the HKGs *hprt1l* and *tuba1c*, as shown in Table 3.1. The primer pair of *insb* was also included in the analysis despite having efficiency (78%) and Ct values (>32) just outside the optimal range. Neither the NTC nor the NRT controls had values less than ten cycles higher than the experimental samples, indicating no contamination, primer dimer formation, or unspecific amplification of genomic DNA in the samples. The melting curves showed one single peak, confirming the findings.

Before proceeding with RT-qPCR analysis, the RNA integrity number (RIN) was determined for each sample. Initial RIN values ranged from 2.80 to 5.10. After finding that the RNA was degraded when storing the larvae in RNAlater™, we removed the step and directly purified the RNA at 96 hpf. This step drastically improved the RNA integrity number, and values now ranged from 9.10 to 9.70.

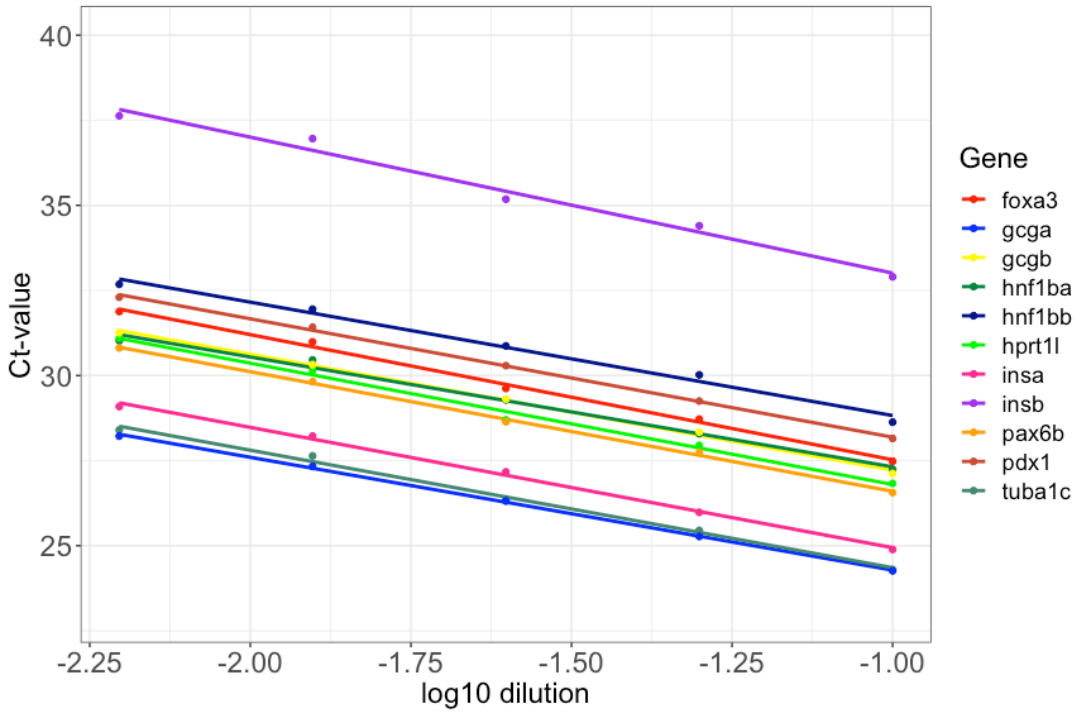

Figure 3.4: Dilution curves of primer pairs designed for genes important for pancreas development. The efficiency of the primers was studied by running RT-qPCR on a cDNA dilution series of 1:10, 1:20, 1:40, 1:80 and 1:160. Linear regression was performed, and the slope and the coefficient of determination were read. Only verified primer pairs and *insb* are listed. The figure is made in RStudio.

Table 3.1: Optimised amplification temperatures (T), slopes, coefficients of determination (R^2) and efficiencies for utilised primer pairs. Values were obtained by performing linear regression on RT-qPCR dilution curves. The slopes were used to calculate primer efficiencies.

Gene	T (°C)	Slope	R^2	Efficiency	Efficiency (%)
foxa3	60.0	-3.6774	0.9959	1.8704	87.04
gcga	59.5	-3.3286	0.9992	1.9972	99.72
gcgb	60.0	-3.4083	0.9976	1.9652	96.52
hnf1ba	62.0	-3.2256	0.9905	2.0418	104.18
hnf1bb	60.0	-3.3319	0.9887	1.9959	99.59
hprt1l	62.0	-3.5644	0.9929	1.9079	90.79
insa	62.0	-3.5345	0.9971	1.9183	91.83
insb	62.0	-3.9950	0.9826	1.7801	78.01
pax6b	60.0	-3.5146	0.9985	1.9253	92.53
pdx1	62.0	-3.4781	0.9992	1.9387	93.87
tuba1c	60.0	-3.4515	0.9953	1.9486	94.86

3.3.2 Altered gene expressions in *hnf1ba*(-/-) mutants

To assess the gene expression in *hnf1ba*(-/-) mutants, we performed RT-qPCR and calculated the relative gene expression (RGE) using Ct-values and primer efficiencies. Statistical analysis was performed, and the RGE was plotted for each gene (Figures 3.5 and 3.6).

We employed the non-parametric Mann-Whitney U test since the data did not exhibit a normal distribution. The test revealed significant changes in transcript levels for all tested genes (Figures 3.5 and 3.6). However, the extent and direction of the changes varied widely. The expression level of both *hnf1ba* and *hnf1bb* mRNA had a threefold increase in the *hnf1ba*(-/-) mutant compared to wt larvae. Similarly, *foxa3* and *pax6b* transcript levels were over twofold higher in the mutants than in wt larvae. In contrast, the *pdx1* mRNA expression was sevenfold lower in the mutants than in wt larvae.

The change in transcript levels varied among the paralogs for both glucagon and insulin. *Gcga* and *insb* transcript levels increased in the mutants, whereas *gcgb* and *insa* levels decreased. The transcript level of *insa* was almost undetectable in the mutant. As previously mentioned, the primer pair of *insb* was not considered optimal. Outliers were present in most box plots, indicating variation in the expression levels within the genotype groups.

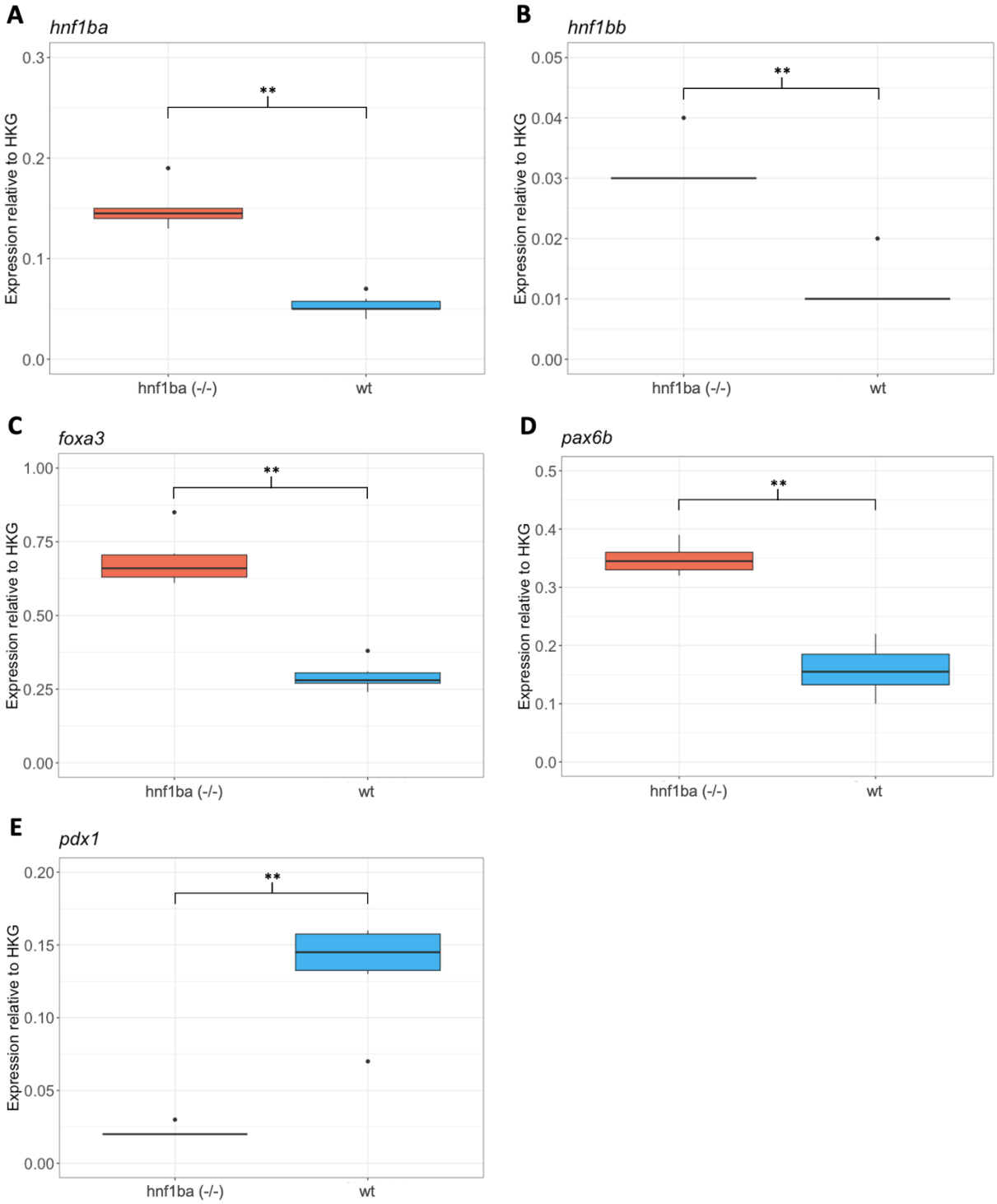

Figure 3.5: Relative mRNA expression of *hnf1ba* (A), *hnf1bb* (B), *foxa3* (C), *pax6b* (D) and *pdx1* (E) in 4 dpf *hnf1ba*(-/-) and wild-type (wt) zebrafish larvae. The mRNA levels were obtained from RT-qPCR analysis of homozygous and wt zebrafish (n = 6 per group). The mRNA levels were measured for multiple genes important for pancreatic development and normalised to the housekeeping genes (HKG) *tuba1c* and *hprt1l*. P-values were calculated using the Mann-Whitney U test. Statistical significance is indicated with an asterisk (**p < 0.005). The boxplots are made in RStudio.

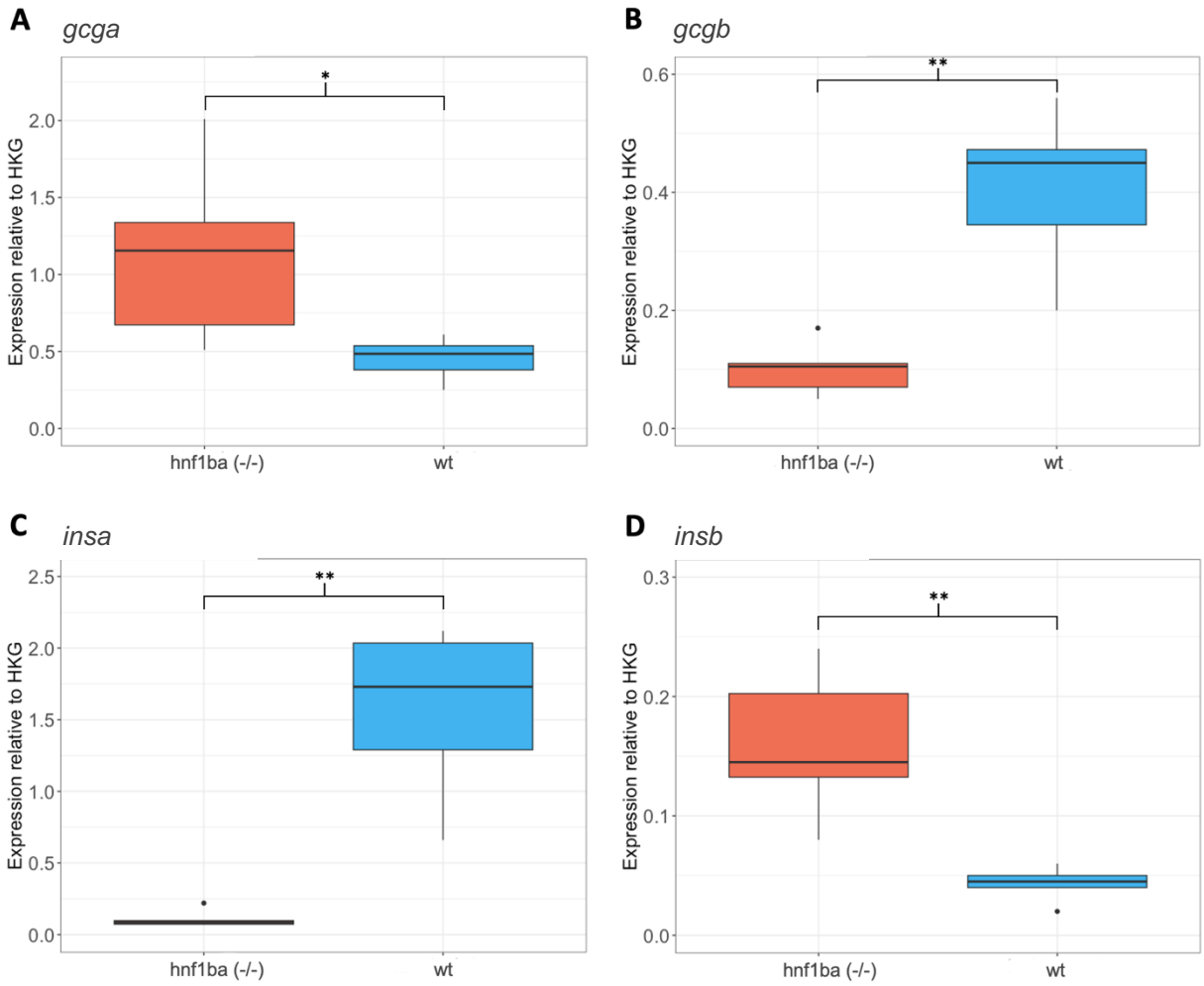

Figure 3.6: Relative mRNA expression of *gcga* (A), *gcgb* (B), *insa* (C) and *insb* (D) in 4 dpf *hnf1ba*(-/-) and wild-type (wt) zebrafish larvae. The mRNA levels were obtained from RT-qPCR analysis of homozygous mutants and wt zebrafish (n = 6 per group). The mRNA levels were measured for the *insulin* and *glucagon* paralogs and normalised to the housekeeping genes (HKG) *tuba1c* and *hprt1l*. P-values were calculated using the Mann-Whitney U test, and statistical significance is indicated with an asterisk (*p < 0.05; **p < 0.005). The boxplots are made in RStudio.

3.4 Characterisation of pancreas structure

To investigate the endocrine pancreas structure of the *hnf1ba*(-/-) mutant, the upper abdomen of 4 dpf mutant and wt larvae was cross-sectioned to 10 μm. Immunostaining using anti-glucagon and anti-insulin antibodies allowed for the visualisation of alpha- and beta-cells, respectively.

3.4.1 Variations in pancreas structure in *hnf1ba*(-/-) zebrafish

The endocrine pancreas was observed in cryosections for both wt and mutant larvae (Figure 3.7). In all wt larvae (n = 6), the pancreas was located on the right lateral side, and both glucagon- and insulin-producing cells were present. Islets were oval-shaped, with glucagon-producing cells surrounding the insulin-producing cells (Figure 3.8, A-B). In mutant larvae (n = 6), the structure of the pancreas varied greatly. Some individuals (n = 3) displayed a pancreas structure similar to wt larvae, with insulin-producing and glucagon-producing cells arranged in the pattern described (Figures 3.7 and 3.8, C). In other homozygous individuals (n = 3), only a few insulin-producing cells were observed, and the location of these cells varied among the individuals (Figures 3.7 and 3.8, D). Glucagon-producing cells were not detectable in these individuals. Additionally, these mutants exhibited morphological abnormalities such as an overall body deformation, an enlarged yolk sac, and a smaller gut tube and swim bladder compared to wt larvae (Figure 3.7 D).

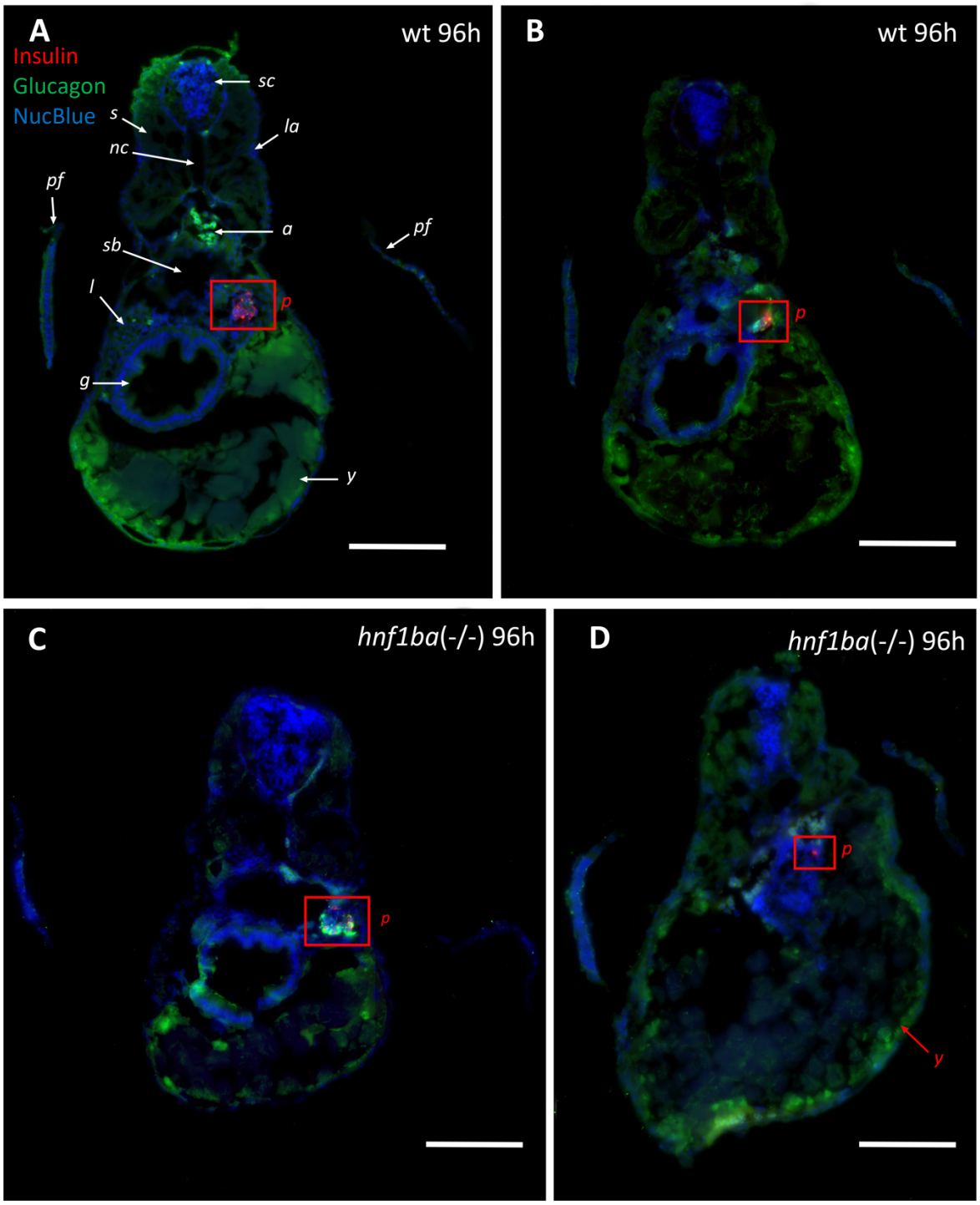

Figure 3.7: Location of the endocrine pancreas in wild-type (wt) and *hnf1ba*(-/-) zebrafish larvae. The upper abdomen of 4 dpf zebrafish larvae were cross-sectioned and immunostained using anti-glucagon (green) and anti-insulin (red) as alpha- and beta-cell markers. NucBlue (blue) stained the nuclei. The sections were imaged using a slide scanner and 40x magnification. Two individuals were imaged per genotype to illustrate the differences within the group. White arrows indicate general anatomical structures, while red arrows indicate distinctive features separating mutants from wt larvae. A red rectangle annotates the pancreas. The pancreas of each individual (A-D) is further studied in Figure 3.8. Explanations: a = dorsal aorta, g = gut, la = lateral line, l = liver, n = notochord, o = otolith, p = pancreas, pf = pectoral fin, s = somites, sb = swim bladder, sc = spinal cord, y = yolk sac. Scale bar = 100 µm.

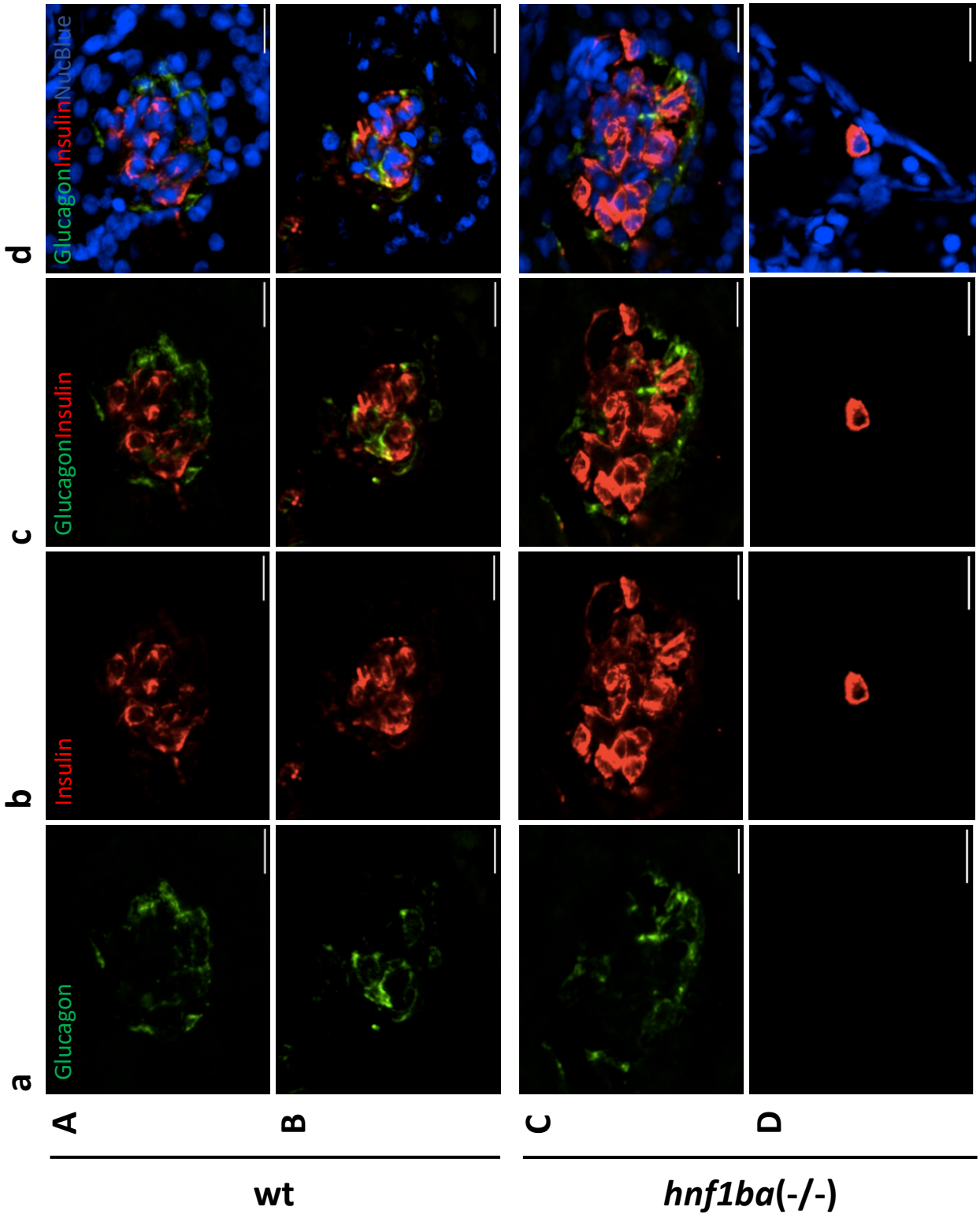

Figure 3.8: Characterisation of alpha- and beta-cells in the pancreas of wild-type (wt) and *hnf1ba*(-/-) mutant zebrafish larvae. The upper abdomen of 4 dpf zebrafish larvae were cross-sectioned and immunostained using anti-glucagon (green) and anti-insulin (red) as alpha- and beta-cell markers. NucBlue (blue) stained the nuclei. Sections containing the pancreas were imaged at 60x magnification using a confocal microscope at 2.2 zoom. Two individuals are included per genotype to illustrate the differences within the group. Scalebar: 10 µm.

3.4.2 Variations in alpha- and beta-cell numbers with a trend towards reduction

The number of glucagon- (alpha-cells) and insulin-producing cells (beta-cells) in the pancreas of 4 dpf *hnf1ba*(-/-) mutant and wt larvae was quantified in confocal images of cryosections. Wt larvae had a higher mean number of both alpha- and beta-cells compared to mutant larvae, with 9 and 6.3 alpha-cells and 27.5 and 14.2 beta-cells in wt and homozygous larvae, respectively (Figure 3.9). However, the independent samples t-test did not yield significant differences in the number of alpha-cells ($t(10) = 0.70$, $p = 0.4979$) or beta-cells ($t(10) = 1.87$, $p = 0.091$) between the two groups.

The number of alpha- and beta-cells varied greatly within each genotype group. In wt larvae, the number of counted alpha-cells varied from 3 to 15, and the number of beta-cells varied from 15 to 45. In mutant larvae, alpha-cells varied from 0 to 17, and the number of beta-cells varied from 1 to 30. Three individuals had no alpha-cells and 5 or fewer beta-cells, while the 3 other homozygous individuals had cell numbers that resembled those of wt larvae (Appendix 2). One homozygous individual had more glucagon-producing cells (17 cells) than any studied wt larvae (Figure 3.8, C).

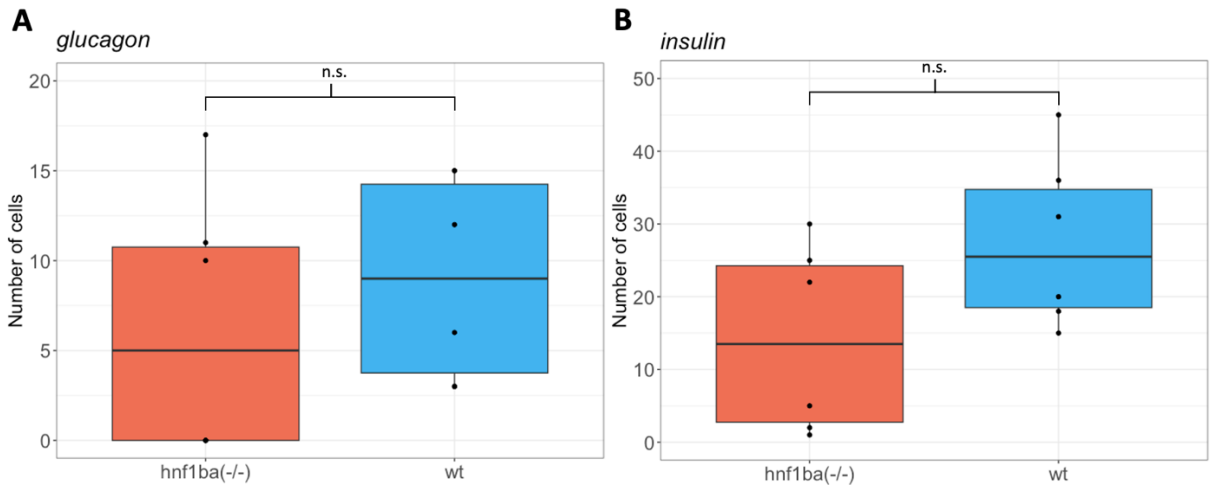

Figure 3.9: Number of glucagon-producing cells (A) and insulin-producing cells (B) in the pancreas of 4 dpf *hnf1ba*(-/-) mutants and wild-type (wt) zebrafish larvae. Counting was conducted on confocal images of 10 µm cryosections of the pancreas, immunostained with anti-glucagon and anti-insulin as alpha- and beta-cell markers. A total of 12 individuals were studied, with n = 6 per genotype. Each dot represents one individual, except for the cases where 3 *hnf1ba*(-/-) larvae had 0 glucagon-producing cells, 2 wt larvae had 3 glucagon-producing cells and 2 wt larvae had 15 glucagon-producing cells. The independent t-test indicated no significant (n.s.) difference in the means of the groups ($p > 0.05$). The boxplots are made in RStudio.

4 Discussion

MODY5 is a multisystemic disease resulting from heterozygous mutations in the *HNF1B* gene. The disease exhibits a wide range of clinical manifestations and is often misdiagnosed as T1D or T2D.[48] The precise role of HNF1B in diabetes and pancreas development is unclear, limited by the lack of a fully representative animal model.[67] Studies employing mice or zebrafish larvae with heterozygous *HNF1B* mutations do not display pancreas defects characteristic of MODY5, suggesting differences in the regulation of *HNF1B* among vertebrates.[67,78] Notably, a zebrafish model with a homozygous *hnf1ba* hypomorphic mutation has been utilised to study MODY5. The model exhibited fewer beta-cells, disorganised islets, and pancreas hypoplasia, resembling the pancreas pathologies observed in humans.[67]

In MODY5 patients, whole-gene deletion is the most observed molecular alteration.[46] Hence, our objective was to investigate whether our *hnf1ba*(-/-) mutant zebrafish could serve as a MODY5 model. The studied mutation involved an 11-base pair deletion in exon 1, similar to L16fs in humans, causing a frameshift leading to a premature stop-codon. Thus, the mutation is with high probability a null mutation (knockout) affecting all isoforms. Our study focused on pancreas structure and gene expression, given the crucial role of the pancreas in glucose homeostasis and the development of diabetes. We show that Hnf1ba is required to properly regulate the expression of genes essential for pancreas development in zebrafish. Given the great variability in pancreas structure observed among the mutants, we also propose variable penetrance of the *hnf1ba*(-/-) mutation. The similarity between *hnf1ba*-associated larval phenotypes and human symptoms suggests a correlation between the developmental role of *hnf1ba* and the molecular cause of MODY5.

The *hnf1ba*(-/-) mutant larvae exhibit an abnormal phenotype which becomes visible at 3 dpf. We chose to study 4 dpf larvae to capture a more pronounced manifestation of this phenotype. At this developmental stage, most organs are well-developed and important physiological systems have initiated their processes. The pancreas has acquired a relatively mature structure with functional endocrine and exocrine compartments.[79] Most mutant larvae do not survive beyond 5 dpf, and therefore, our decision to focus on 4 dpf larvae aligns with ethical guidelines aiming to minimise suffering. Also, studying larvae at this stage may help avoid confounding secondary effects that could arise in later stages.

4.1 Variable phenotype of *hnf1ba*(-/-) mutants

To gain insights into the effects of the *hnf1ba*(-/-) mutation, we conducted a morphological comparison between wt and mutant larvae at different developmental stages. As mentioned, abnormal features became apparent in the mutant at 3 dpf and were more pronounced at 4 dpf.

4.1.1 Variable penetrance of the *hnf1ba* mutation

The severity of the abnormal features varied considerably between the mutant larvae. Some individuals exhibited only one or two abnormal features, such as a slow heart rate or pericardial edemas, while others displayed all mentioned features. These findings suggest a variable penetrance of the *hnf1ba*(-/-) mutation. This observation is consistent with previous reports of varying phenotypes in MODY5 patients harbouring the same inherited mutation. Siblings carrying early whole deletion mutations in *HNF1B* have been found to differ in clinical features, including pancreatic hypoplasia and genital structural anomalies.[80] Furthermore, non-related patients carrying the same mutation also exhibit differences in features such as renal cysts, liver anomalies, hypomagnesemia and incidence of diabetes.[80,81]

4.1.2 Abnormal features observed in *hnf1ba*(-/-) larvae

One of the most prominent features observed in mutants at 4 dpf was a curved tail/body. The underlying cause of this phenotype is currently unknown, and possible explanations may include muscular atrophy due to hyperglycemia or hypomagnesemia, and the possibility of neurological issues. Hyperglycemia is known to cause muscle atrophy by altering metabolic pathways or through toxic effects leading to neuropathy, thus impairing muscle function.[82] Hypomagnesemia is commonly observed in individuals with MODY5 and is often used as a diagnostic clue.[83] The severity of hypomagnesemia varies depending on the mutation type, with studies indicating that patients with *HNF1B* deletions experience more severe hypomagnesemia than those with other mutations.[48] As magnesium is vital for nerve transmission and muscle contraction[84], hypomagnesemia may explain the curved body structure observed in many homozygous mutants. Muscular dystrophy has been reported in a MODY5 patient who also suffered from hypomagnesemia, further supporting this explanation.[85] Moreover, the curvature and lack of movement could be associated with developmental neurological issues, considering the expression of Hnf1ba in the hindbrain during the tailbud stage.[41] The protein is also weakly expressed in zebrafish muscles, but its specific function is unknown.[41]

Other abnormal features observed in zebrafish mutants at 4 dpf included an enlarged yolk sac, small swim bladder and pectoral fins, blood accumulation in the pericardium, pericardial edemas, and a slow heart rate. The yolk sac sustains metabolic function and growth by providing the larva with proteins, lipids and micronutrients during the first five days of life.[86] An enlarged yolk sac suggests that the larva is not effectively utilising all its nutritional resources, potentially contributing to a delay in development. The presence of a small swim bladder and pectoral fins further indicates this underdeveloped state. Hnf1ba has not been reported to be expressed in the yolk sac, but it is expressed in the swim bladder during embryonic development.[41] The findings indicate a profound role of *hnf1ba* in foregut patterning, consistent with previous studies.[42] The expression of Hnf1ba in the zebrafish heart may explain the observed slow heart rate, blood accumulation and pericardial edemas. Cardiac abnormalities are not commonly examined in MODY5 patients; however, a study involving this parameter reported no abnormalities in a MODY5 patient with *HNF1B* deletions.[85] It is noteworthy that our zebrafish model is homozygous compared to the heterozygous MODY5 condition in humans, which may explain the extreme phenotype observed. Additionally, the mentioned features may be due to secondary effects. Interestingly, we observed that homozygous larvae commonly had grey spots on their eyes and showed variation in the protrusion of their cornea. Eye disease is a typical feature of MODY and other forms of diabetes, as long-term hyperglycemia often leads to damage to the retina's small blood vessels.[1,87]

4.2 Altered gene expression of pancreatic genes

Genome-wide association studies, transcriptome profiling, and the study of human genetic conditions associated with diabetes have revealed candidate genes believed to affect pancreatic function.[79] To validate the significance of these genes, it is necessary to conduct experiments using in vivo models. We aimed to investigate the effect of *hnf1ba* knockout on the pancreas by analysing the expression levels of candidate genes using RT-qPCR. Our findings revealed significant alterations in the expression levels of all nine examined genes, including *hnf1ba, hnf1bb, gcga, gcgb, insa, insb, pax6b, pdx1* and *foxa3*.

4.2.1 Feedback loops and compensatory mechanisms of Hnf1bb

We observed a threefold increase in *hnf1ba* transcript levels in the mutant larvae, which may indicate the presence of a Hnf1ba-directed feedback loop. Such a feedback loop could lead to an upregulation of *hnf1ba* gene expression in the absence of Hnf1ba. Similarly, *hnf1bb*

transcript levels showed a threefold increase in the mutant larvae. This finding suggests that Hnf1bb may exhibit compensatory mechanisms, or have conserved domains that lead to the upregulation of Hnf1bb along with Hnf1ba. However, even with this potential compensatory role of Hnf1bb, considerable phenotypical differences were still observed between the mutant and wt larvae. Additionally, Hnf1bb is not expressed in the primitive endoderm, and its expression begins later in development (10 hpf).[67,70] These findings suggest that Hnf1bb is unable to fully substitute the function of its absent paralog.

4.2.2 Differences within insulin and glucagon paralogs

The relative expression of *insa* and *insb* was altered in the mutant, with *insa* showing a sevenfold decrease and *insb* showing a threefold increase in expression. These findings suggest that Hnf1ba plays a crucial role in regulating insulin gene expression and that the protein is essential for proper beta-cell function. Furthermore, the analysis suggests divergent functions of the insulin paralogs during development, supported by previous studies.[59,61] *Insa* is expressed in the islet of Langerhans and is a beta-cell marker, and decreased transcript levels might indicate underdeveloped or dysfunctional islets. Consequently, the *hnf1ba* inactivation could lead to imbalanced glucose homeostasis and potentially hyperglycemia in the zebrafish larvae. The increased expression of *insb* might be associated with its proposed involvement in regulating embryonic development in extrapancreatic tissues, although its precise role remains unclear.[59]

Similarly, the mutant larvae exhibited significant changes in the expression of glucagon paralogs. *Gcga* exhibited a twofold increase, while *gcgb* showed a fourfold decrease in expression. Both *gcga* and *gcgb* are markers for alpha-cells, but they differ in their functions.[88] While both produce glucagon and GLP-1, *gcga* has been subfunctionalised to also produce GLP-2, making the structure and function of *gcga* more similar to the mammalian glucagon gene.[88] It is unclear if there is an overall change in the amount of glucagon protein in the mutants compared to the wt. Conducting additional experiments, such as western blot or mass spectrometry, to quantify glucagon protein levels could provide further insights into these findings.

4.2.3 Altered gene expression of *pax6b, pdx1* and *foxa3*

Our findings concerning glucagon expression are consistent with a previous study by Lavergne et al.[89] demonstrating reduced *gcgb* expression and slightly increased *gcga* expression in Pax6b-

downregulated mutant zebrafish.[89] Yet, in our mutant, the *pax6b* transcript levels were more than twofold higher compared to wt larvae. Previous research has shown that the absence of Pax6b leads to a severe reduction in beta-cell numbers in zebrafish embryos. Hence the increased *pax6b* expression may indicate a compensatory mechanism to maintain beta-cell development and function.[72] However, the nearly undetectable expression levels of *insa* suggest impaired beta-cell functions in the mutant. It should be noted that the study of pax6b-mutants focuses on pancreatic buds in 27 hpf larvae, which differs from our RT-qPCR analysis on whole 4 dpf larvae. The relationship between *hnf1ba* and *pax6b* is not fully understood, and further investigations are needed to elucidate this mechanism. Exploring changes in the expression of transcription factors that interact with both Pax6b and Hnf1ba, such as Neurod1 and Nk2 homeobox 2 (Nkx2.2), would be valuable for future research.

The expression level of *pdx1* was barely detectable in the mutants, showing a sevenfold decrease compared to wt larvae. These findings are consistent with previous studies demonstrating the essential role of Hnf1ba in activating Pdx1 expression in the endoderm.[42] Pdx1 plays a critical role in the specification of pancreatic cell fate, and at later stages, the regulation of pancreatic endocrine differentiation.[90] Mutations in the *pdx1* gene are associated with the development of MODY4.[42] It is reasonable to speculate that the inactivation of Hnf1ba may disrupt the regulatory pathway controlling Pdx1 expression, ultimately leading to severe impairments in pancreas formation.

Surprisingly, we observed a twofold increase in *foxa3* transcript levels in the mutant larvae compared to wt larvae. Previous in vitro and in vivo mice studies suggest that HNF1B acts as an upstream activator of FOXA3 during development,[91] leading to the expectation of decreased *foxa3* levels in the absence of Hnf1ba. However, it should be noted that HNF1B primarily regulates FOXA3 in the early gut endoderm, which occurs at a developmental stage preceding our study.[34] Furthermore, FOXA3 is also positively regulated by HNF1A during liver development.[34] Given the underdeveloped state of the mutant larvae, the altered expression of other genes involved in regulating FOXA3 may also contribute to our observations. Further investigation is necessary to confirm and understand this observation, as it deviates from the established understanding of HNF1B's regulatory role on FOXA3 expression.

Collectively, the presented findings offer valuable insights into the regulatory role of Hnf1ba in pancreas development and function. While compensatory mechanisms likely exist, such as the upregulation of Hnf1bb in the absence of Hnf1ba, the results indicate that other proteins do

not fully compensate for the functions of Hnf1ba. The findings further highlight the intricate nature of the regulatory network Hnf1ba is a part of, and emphasise its importance in ensuring proper pancreas development and function. It should be noted that the data did not follow a normal distribution, and we therefore used the non-parametric Mann-Whitney U test to test for statistical significance. Non-parametric tests are less powerful than parametric tests, and choosing this test may therefore have introduced a higher degree of uncertainty in the study's findings.

4.3 Pancreas structure and function

In order to study the effect of *hnf1ba* knockout on the morphology of the pancreas and alpha- and beta-cells, the upper abdomen of 4 dpf zebrafish larvae were cross-sectioned and immunostained with anti-glucagon and anti-insulin antibody solutions.

4.3.1 Varying pancreas structure in *hnf1ba*(-/-) larvae

The slide scanner images of the immunostained cross-sections revealed notable variations in the structure of the endocrine pancreas within the mutant larvae. Half of the individuals exhibited a complete absence of alpha-cells and a limited number of beta-cells. In contrast, the remaining half displayed a cellular organisation similar to that observed in wt larvae. Moreover, a wide diversity in overall body structure was observed among the mutant larvae. While half of the individuals displayed a "normal" body morphology, the remaining half exhibited enlarged yolk sacs and distorted gut and swim bladders. These findings align with the morphological analysis and emphasise the considerable heterogeneity within the mutant larvae. Consequently, our observations further support the notion that the *hnf1ba* mutation has variable penetrance.

4.3.2 Unaltered mean alpha- and beta-cell numbers

The confocal images of the endocrine pancreas were analysed to determine the number of alpha- and beta-cells. Surprisingly, no significant difference between the mutant and wt larvae was found. The alpha-cell count ranged from 3 to 15 in wt larvae and from 0 to 17 in homozygous individuals, while the beta-cell count ranged from 15 to 45 in wt larvae and from 1 to 30 in homozygous individuals. These observations suggest heterogeneity in the alpha-cell and beta-cell populations independent of the genotype. However, all wt larvae exhibited both cell types, whereas half of the homozygotes lacked alpha-cells and had a limited number of beta-cells. Notably, the IHC experimental group consisted of only six individuals per genotype. Future

studies with larger sample sizes might provide a deeper understanding of the phenotypic variability of the endocrine pancreas associated with the *hnf1ba* mutation.

Furthermore, uneven sections during histological preparation could influence the observed differences in pancreas structure. In cases where larvae were not perfectly spatially orientated during sectioning, sections may provide a misleading impression of the pancreas' extent. This might be the case for individual C in Figure 3.7, where other organs, such as the spinal cord, appear larger in the mutant compared to the wt. This potential discrepancy may also apply to other individuals, impacting cell count measurements. The choice of 10 µm sections may also affect our results, as the pancreas structure appeared markedly different in two subsequent sections of the same individuals. Utilising 8 µm sections might provide more informative and representative data for a greater number of individuals. Furthermore, variations in antibody stability and binding affinities might explain some discrepancies. During confocal imaging, insulin fluorescent signals were consistently detectable, while glucagon signals were weaker and more prone to bleaching. The anti-insulin staining was designed for the Insa protein, and its binding affinity to Insb is unknown. Given the nearly identical binding area of the glucagon paralogs, the anti-glucagon staining likely binds to both Gcga and Gcgb. Overall, these additional factors should be considered when interpreting the results.

4.4 Model for MODY5

Our zebrafish *hnf1ba* knockout model exhibits several key characteristics that make it a promising tool for studying MODY5 disease mechanisms. Firstly, the phenotype analysis reveals abnormalities in multiple organs in the mutant larvae, resembling the multisystemic nature of MODY5. While our model shows more extreme abnormalities compared to the heterozygous state in humans, the reported absence of abnormal phenotypes in heterozygous larvae and mice suggests that our model better mimics the human condition. Secondly, the RT-qPCR analysis shows altered expression levels of genes crucial for early pancreas development. This finding suggests the significance of *hnf1ba* in the pancreatic gene regulatory network, consistent with findings in MODY5 patients with *HNF1B* mutations. Thus, our zebrafish model holds promise for investigating the molecular pathways underlying MODY5.

Additionally, the IHC analysis provides valuable insights into the impact of the *hnf1ba* knockout on the endocrine cell population. Although overall cell numbers are not significantly reduced, a distinct pattern emerges in some homozygous individuals with a lack of alpha-cells

and a limited number of beta-cells. These findings, together with the phenotype analysis, reflect the variable penetrance observed in MODY5 patients. The severity of pancreatic hypoplasia has been shown to vary greatly among MODY5 patients carrying the same mutation, even within the same family. Nevertheless, studies show that most MODY5 patients (~70%) exhibit some degree of pancreatic hypoplasia.[48] Likewise, the RT-qPCR results reveal almost undetectable *insa* levels in the mutants, indicating severe pancreatic impairment in most individuals.

HNF1B knockout models have previously faced criticism from Prince et al.,[79] who suggested that the inactivation of *HNF1B* leads to profound endoderm defects, resulting in an ineffective disease model. While this idea warrants further investigation, our IHC results contradict this assumption by demonstrating the presence of beta-cells in all mutant individuals. Additionally, half of the studied mutants had pancreas structures indistinguishable from those of wt individuals. These findings suggest a different outcome of *hnf1ba* knockout models than previously suggested. However, most mutants do not survive beyond 5 dpf, indicating the development of severe defects and restricting the model to the larval stage. Furthermore, by focusing on the pancreas in this study, the potential effects of the gene knockout on other endoderm-derived organs remains unknown. Renal abnormalities are common in MODY5 patients and is a natural focus area for further analysis.

Our model exhibits some similarities to the previously mentioned hypomorph zebrafish model.[67] In the hypomorph model, mutants displayed pancreas hypoplasia, disorganised islets, and significantly reduced beta-cell numbers. We did not observe significant differences in the number of alpha- and beta-cells between wt and mutants. Furthermore, half of our mutants had normally organised endocrine cells forming islets. Hence, our model might provide a better representation of the variable penetrance of the *hnf1ba* mutation as observed in humans. It is worth noting that the study of the hypomorph model performed *in situ* hybridization and IHC, but no RT-qPCR, which makes the comparison of the two studies challenging. Furthermore, in their IHC analysis, the hypomorph study did not include anti-glucagon staining, and the comparison is therefore limited to beta-cells.

5 Conclusion

As the prevalence of diabetes continues to rise, there is an urgent need for effective treatment options. Understanding the underlying disease mechanisms is necessary to meet this challenge. While mice have been extensively used as a model system for studying pancreatic development, it is important to recognise the distinct advantages offered by other model systems. MODY5 is associated with more than 200 mutations in the *HNF1B* gene, with over half of these mutations resulting in whole gene deletion. In this context, we investigated the homozygous *hnf1ba* knockout zebrafish larvae as a model for MODY5.

By comparing the morphology of 4 dpf wt and mutant larvae, we found that the *hnf1ba* inactivation resulted in severe multisystemic effects. The mutants exhibited abnormal phenotypes such as an enlarged yolk sac, curved tail, pericardial edemas, and underdeveloped swim bladder and pectoral fins. These findings align with the multisystemic nature of MODY5 and emphasise the important role of the widely expressed transcription factor *hnf1ba*.

Through RT-qPCR analysis, in the mutants compared to wt larvae, we observed altered expression levels of genes essential for pancreas development and function. These genes included *foxa3, hnf1ba, hnf1bb, gcga, gcgb, insa, insb, pax6b* and *pdx1,* underscoring the impact of *hnf1ba* inactivation on the regulatory network involved in pancreas development.

Immunostaining of alpha- and beta-cells revealed that *hnf1ba* homozygous larvae did not exhibit a significantly lower number of cells compared to wt, although a reduction was observed. Instead, a distinct pattern emerged. Half of the homozygous individuals lacked alpha-cells and showed a limited number of beta-cells, while the remaining half displayed a cellular organisation similar to wt larvae. These findings demonstrate the considerable variability among affected individuals, emphasising the variable penetrance of the *hnf1ba* mutation.

5.1 Future perspectives

To further investigate the *hnf1ba* knockout model for studying MODY5 disease mechanisms in pancreas development, we suggest incorporating a classification system for phenotypes, such as mild, moderate and severe. This approach would allow for better monitoring of the distribution and correlation of morphological features. A more comprehensive understanding of the model can be achieved by following individual samples and examining their gene

expression profiles and pancreas structure. Quantifying observed morphological features would provide insights into the extent of variable penetrance associated with the mutation.

Our study focused on a specific set of genes involved in pancreas development and function. Including further genes essential for pancreas development, such as *ONECUT1* and *NKX2.2*, would be valuable for exploring the effects of *hnf1ba* inactivation on the pancreas. Additionally, employing RNA sequencing (RNA-Seq) for transcriptome profiling in both wt and mutant pancreas, even considering single-cell sequencing, could potentially provide novel and more complete information about the underlying mechanisms of the phenotypes observed in the *hnf1ba* mutant.

As previously mentioned, variability was observed when analysing the mutants' alpha- and beta-cell counts. Future IHC studies should include bigger sample sizes to provide higher credibility. Moreover, including additional antibodies, such as for Islet-1 and Pax6b, would yield more information about the presence of endocrine cells that might not be actively secreting insulin or glucagon. To better understand the regulation of the insulin and glucagon paralogs, we propose to investigate protein levels over time by western blot analysis. Furthermore, it would be interesting to study a double mutant where both *hnf1ba* and *hnf1bb* are inactivated. This would prevent potential compensatory mechanisms by Hnf1bb in the *hnf1ba*(-/-) mutant.

Finally, conducting analyses on other organs commonly affected in MODY5 patients, such as the liver and kidneys, would be informative. Additionally, it would be beneficial to assess hyperglycemia, hypomagnesemia or other common symptoms of MODY5. These investigations have the potential to provide deeper insights into the *hnf1ba* knockout model and may ultimately contribute to the development of new treatments for pancreatic diseases such as MODY5.

6 References

1. International Diabetes Federation. IDF Diabetes Atlas 10th edition. 2021. https://www.diabetesatlas.org.
2. Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract*. 2022;183. doi:10.1016/j.diabres.2021.109119
3. Beagley J, Guariguata L, Weil C, Motala AA. Global estimates of undiagnosed diabetes in adults. *Diabetes Res Clin Pract*. 2014;103(2):150-160. doi:10.1016/j.diabres.2013.11.001
4. World Health Organization. Global report on diabetes World Health Organization. 2016. https://apps.who.int/iris/handle/10665/204871.
5. Wilcox G. Insulin and Insulin Resistance. *Clin Biochem Rev*. 2005;26(2):19-39. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1204764/.
6. Giri B, Dey S, Das T, Sarkar M, Banerjee J, Dash SK. Chronic hyperglycemia mediated physiological alteration and metabolic distortion leads to organ dysfunction, infection, cancer progression and other pathophysiological consequences: An update on glucose toxicity. *Biomed Pharmacother*. 2018;107:306-328. doi:10.1016/J.BIOPHA.2018.07.157
7. Campos C. Chronic Hyperglycemia and Glucose Toxicity: Pathology and Clinical Sequelae. *Postgrad Med*. 2015;124(6):90-97. doi:10.3810/PGM.2012.11.2615
8. Slack JM. Developmental biology of the pancreas. *Development*. 1995;121(6):1569-1580. doi:10.1242/dev.121.6.1569
9. M Das SL, C Kennedy JI, Murphy R, et al. Relationship between the exocrine and endocrine pancreas after acute pancreatitis. *World J Gastroenterol*. 2014;20(45):17196-17205. doi:10.3748/wjg.v20.i45.17196

10. A-Kader HH, Ghishan FK. The Pancreas. *Textb Clin Pediatr*. 2012:1925-1936. doi:10.1007/978-3-642-02202-9_198

11. Brissova M, Fowler MJ, Nicholson WE, et al. Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *J Histochem Cytochem*. 2005;53(9):1087-1097. doi:10.1369/jhc.5C6684.2005

12. Röder P V, Wu B, Liu Y, Han W. Pancreatic regulation of glucose homeostasis. *Exp Mol Med*. 2016;48(3):e219. doi:10.1038/emm.2016.6

13. Freychet L, Desplanque N, Zirinis P, et al. Effect on intranasal glucagon on blood glucose levels in healthy subjects and hypoglycaemic patients with insulin-dependent diabetes. *Lancet*. 1988;331(8599):1364-1366. doi:10.1016/S0140-6736(88)92181-2

14. Nakrani MN, Wineland RH, Anjum F. Physiology, Glucose Metabolism. *StatPearls*. July 2022. https://www.ncbi.nlm.nih.gov/books/NBK560599/.

15. Vaxillaire M, Bonnefond A, Froguel P. The lessons of early-onset monogenic diabetes for the understanding of diabetes pathogenesis. *Best Pract Res Clin Endocrinol Metab*. 2012;26(2):171-187. doi:10.1016/J.BEEM.2011.12.001

16. Kleinberger JW, Copeland KC, Gandica RG, et al. Monogenic diabetes in overweight and obese youth diagnosed with type 2 diabetes: the TODAY clinical trial. *Genet Med*. 2018;20(6):583-590. doi:10.1038/GIM.2017.150

17. Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of Type 1 Diabetes. *Endocrinol Metab Clin North Am*. 2010;39(3):481-497. doi:10.1016/J.ECL.2010.05.011

18. Rewers M, Ludvigsson J. Environmental risk factors for type 1 diabetes. *Lancet*. 2016;387(10035):2340-2348. doi:10.1016/S0140-6736(16)30507-4

19. Ali O. Genetics of type 2 diabetes. *World J Diabetes*. 2013;4(4):114-123. doi:10.4239/wjd.v4.i4.114

20. Borse SP, Chhipa AS, Sharma V, Singh DP, Nivsarkar M. Management of Type 2 Diabetes: Current Strategies, Unfocussed Aspects, Challenges, and Alternatives. *Med Princ Pract Int J Kuwait Univ Heal Sci Cent*. 2021;30(2):109-121. doi:10.1159/000511002

21. Fajans SS, Bell GI, Polonsky KS. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N Engl J Med*. 2001;345(13):971-980. doi:10.1056/NEJMRA002168

22. Tattersall RB. Mild familial diabetes with dominant inheritance. *Q J Med*. 1974;43(170):339-357.

23. Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT, Ellard S. Maturity-onset diabetes of the young (MODY): How many cases are we missing? *Diabetologia*. 2010;53(12):2504-2508. doi:10.1007/S00125-010-1799-4

24. Thanabalasingham G, Pal A, Selwood MP, et al. Systematic assessment of etiology in adults with a clinical diagnosis of young-onset type 2 diabetes is a successful strategy for identifying maturity-onset diabetes of the young. *Diabetes Care*. 2012;35(6):1206-1212. doi:10.2337/DC11-1243/-/DC1

25. Kim SH. Maturity-Onset Diabetes of the Young: What Do Clinicians Need to Know? *Diabetes Metab J*. 2015;39(6):468-477. doi:10.4093/DMJ.2015.39.6.468

26. Shepherd M, Shields B, Ellard S, Rubio-Cabezas O, Hattersley AT. A genetic diagnosis of HNF1A diabetes alters treatment and improves glycaemic control in the majority of insulin-treated patients. *Diabet Med*. 2009;26(4):437-441. doi:10.1111/J.1464-5491.2009.02690.X

27. Urakami T. Maturity-onset diabetes of the young (MODY): current perspectives on diagnosis and treatment. *Diabetes, Metab Syndr Obes*. 2019;12:1047-1056. doi:10.2147/DMSO.S179793

28. Gundersen LB, Aukrust I. Diabetes Mellitus - analyser av genvarianter assosiert med subtyper av Maturity-Onset Diabetes of the Young (MODY). *Bioingeniøren*. 2022;2. doi:10.1016/B978-0-323-67254-2.00255-2

29. Peixoto-Barbosa R, Reis AF, Giuffrida FMA. Update on clinical screening of maturity-onset diabetes of the young (MODY). *Diabetol Metab Syndr*. 2020;12(1):50. doi:10.1186/S13098-020-00557-9

30. Giuffrida FMA, Reis AF. Genetic and clinical characteristics of maturity-onset diabetes of the young. *Diabetes, Obes Metab*. 2005;7(4):318-326. doi:10.1111/J.1463-1326.2004.00399.X

31. Horikawa Y. Maturity-onset diabetes of the young as a model for elucidating the multifactorial origin of type 2 diabetes mellitus. *J Diabetes Investig*. 2018;9(4):704-712. doi:10.1111/JDI.12812

32. Bártů M, Dundr P, Němejcová K, et al. The Role of HNF1B in Tumorigenesis of Solid Tumours: a Review of Current Knowledge. *Folia Biol*. 2018;64(3):71-83.

33. Bach I, Yaniv M. More potent transcriptional activators or a transdominant inhibitor of the HNF1 homeoprotein family are generated by alternative RNA processing. *EMBO J*. 1993;12(11):4229-4242. doi:10.1002/J.1460-2075.1993.TB06107.X

34. Lau HH, Ng NHJ, Loo LSW, Jasmen JB, Teo AKK. The molecular functions of hepatocyte nuclear factors – In and beyond the liver. *J Hepatol*. 2018;68(5):1033-1048. doi:10.1016/J.JHEP.2017.11.026

35. Wu G, Bohn S, Ryffel GU. The HNF1β transcription factor has several domains involved in nephrogenesis and partially rescues Pax8/lim1-induced kidney malformations. *Eur J Biochem*. 2004;271(18):3715-3728. doi:10.1111/J.1432-1033.2004.04312.X

36. Lu P, Rha GB, Chi Y-I. Structural Basis of Disease-Causing Mutations in Hepatocyte

Nuclear Factor 1β. *Biochemistry*. 2007;46(43):12071-12080. doi:10.1021/bi7010527

37. Çubuk H, Yalçın Çapan Ö. A Review of Functional Characterization of Single Amino Acid Change Mutations in HNF Transcription Factors in MODY Pathogenesis. *Protein J*. 2021;40(3):348-360. doi:10.1007/s10930-021-09991-8

38. El-Khairi R, Vallier L. The role of hepatocyte nuclear factor 1β in disease and development. *Diabetes, Obes Metab*. 2016;18(1):23-32. doi:10.1111/dom.12715

39. Coffinier C, Thépot D, Babinet C, Yaniv M, Barra J. Essential role for the homeoprotein vHNF1/HNF1β in visceral endoderm differentiation. *Development*. 1999;126(21):4785-4794. doi:10.1242/DEV.126.21.4785

40. Pearson ER, Badman MK, Lockwood CR, et al. Contrasting diabetes phenotypes associated with hepatocyte nuclear factor-1alpha and-1beta Mutations. *Diabetes Care*. 2004;25(5):1102-1107. doi:10.2337/diacare.27.5.1102

41. Gong HY, Lin CJF, Chen MHC, et al. Two distinct teleost hepatocyte nuclear factor 1 genes, hnf1α/tcf1 and hnf1β/tcf2, abundantly expressed in liver, pancreas, gut and kidney of zebrafish. *Gene*. 2004;338(1):35-46. doi:10.1016/J.GENE.2004.05.003

42. Sun Z, Hopkins N. vhnf1, the MODY5 and familial GCKD-associated gene, regulates regional specification of the zebrafish gut, pronephros, and hindbrain. *Genes Dev*. 2001;15(23):3217-3229. https://doi.org/10.1101/gad946701.

43. De Vas MG, Kopp JL, Heliot C, Sander M, Cereghini S, Haumaitre C. Hnf1b controls pancreas morphogenesis and the generation of Ngn3+ endocrine progenitors. *Dev* . 2015;142(5):871-882. doi:10.1242/DEV.110759/-/DC1

44. Wilson ME, Scheel D, German MS. Gene expression cascades in pancreatic development. *Mech Dev*. 2003;120(1):65-80. doi:10.1016/s0925-4773(02)00333-7

45. Horikawa Y, Iwasaki N, Hara M, et al. Mutation in hepatocyte nuclear factor–1β gene (TCF2) associated with MODY. *Nat Genet 1997 174*. 1997;17(4):384-385.

doi:10.1038/ng1297-384

46. Bellanné-Chantelot C, Clauin S, Chauveau D, et al. Large Genomic Rearrangements in the Hepatocyte Nuclear Factor-1β (TCF2) Gene Are the Most Frequent Cause of Maturity-Onset Diabetes of the Young Type 5. *Diabetes*. 2005;54(11):3126-3132. doi:10.2337/DIABETES.54.11.3126

47. Delvecchio M, Pastore C, Giordano P. Treatment Options for MODY Patients: A Systematic Review of Literature. *Diabetes Ther*. 2020;11(8):1667-1685. doi:10.1007/S13300-020-00864-4/TABLES/4

48. Ge S, Yang M, Cui Y, et al. The Clinical Characteristics and Gene Mutations of Maturity-Onset Diabetes of the Young Type 5 in Sixty-One Patients. *Front Endocrinol*. 2022;13. doi:10.3389/fendo.2022.911526

49. Hattersley AT, Greeley SAW, Polak M, et al. ISPAD Clinical Practice Consensus Guidelines 2018: The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatr Diabetes*. 2018;19 Suppl 1:47-63. doi:10.1111/PEDI.12772

50. Edghill EL, Bingham C, Ellard S, Hattersley AT. Mutations in hepatocyte nuclear factor-1beta and their related phenotypes. *J Med Genet*. 2006;43(1):84-90. doi:10.1136/JMG.2005.032854

51. Quilichini E, Fabre M, Nord C, et al. Insights into the etiology and physiopathology of MODY5/HNF1B pancreatic phenotype with a mouse model of the human disease. *J Pathol*. 2021;254(1):31-45. doi:10.1002/PATH.5629

52. Clissold RL, Hamilton AJ, Hattersley AT, Ellard S, Bingham C. HNF1B-associated renal and extra-renal disease—an expanding clinical spectrum. *Nat Rev Nephrol 2014 112*. 2014;11(2):102-112. doi:10.1038/nrneph.2014.232

53. Westerfield M. *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio Rerio)*. 5th ed. Eugene; 2007.

54. Collins JE, White S, Searle SMJ, Stemple DL. Incorporating RNA-seq data into the zebrafish Ensembl genebuild. *Genome Res*. 2012;22(10):2067-2078. doi:10.1101/GR.137901.112

55. Streisinger G, Walker C, Dower N, Knauber D, Singer F. Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature*. 1981;291(5813):293-296. doi:10.1038/291293a0

56. Choi TY, Choi TI, Lee YR, Choe SK, Kim CH. Zebrafish as an animal model for biomedical research. *Exp Mol Med*. 2021;53(3):310-317. doi:10.1038/s12276-021-00571-5

57. Kumar S, Hedges SB. A molecular timescale for vertebrate evolution. *Nature*. 1998;392(6679):917-920. doi:10.1038/31927

58. Glasauer SMK, Neuhauss SCF. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics*. 2014;289(6):1045-1060. doi:10.1007/S00438-014-0889-2/METRICS

59. Irwin DM. A second insulin gene in fish genomes. *Gen Comp Endocrinol*. 2004;135(1):150-158. doi:10.1016/J.YGCEN.2003.08.004

60. Howe K, Clark MD, Torroja CF, et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 2013;496(7446):498-503. doi:10.1038/nature12111

61. Elo B, Villano CM, Govorko D, White LA. Larval zebrafish as a model for glucose metabolism: expression of phosphoenolpyruvate carboxykinase as a marker for exposure to anti-diabetic compounds. *J Mol Endocrinol*. 2007;38(4):433-440. doi:10.1677/JME-06-0037

62. Katti A, Diaz BJ, Caragine CM, Sanjana NE, Dow LE. CRISPR in cancer biology and therapy. *Nat Rev Cancer 2022 225*. 2022;22(5):259-279. doi:10.1038/s41568-022-

00441-w

63. Yi P, Morrow N. Applying CRISPR Screen in Diabetes Research. *Diabetes*. 2021;70(9):1962-1969. doi:10.2337/DBI20-0047

64. Song Y, Sui T, Zhang Y, et al. Genetic deletion of a short fragment of glucokinase in rabbit by CRISPR/Cas9 leading to hyperglycemia and other typical features seen in MODY-2. *Cell Mol Life Sci 2019 7716*. 2019;77(16):3265-3277. doi:10.1007/S00018-019-03354-4

65. Varadi M, Anyango S, Deshpande M, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*. 2022;50(D1):D439-D444. doi:10.1093/NAR/GKAB1061

66. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nat 2021 5967873*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

67. Lancman JJ, Zvenigorodsky N, Gates KP, et al. Specification of hepatopancreas progenitors in zebrafish by hnf1ba and wnt2bb. *Development*. 2013;140(13):2669-2679. doi:10.1242/DEV.090993

68. Shen W, Scearce LM, Brestelli JE, Sund NJ, Kaestner KH. Foxa3 (Hepatocyte Nuclear Factor 3γ) is Required for the Regulation of Hepatic GLUT2 Expression and the Maintenance of Glucose Homeostasis during a Prolonged Fast. *J Biol Chem*. 2001;276(46):42812-42817. doi:10.1074/jbc.M106344200

69. Irwin DM, Mojsov S. Diversification of the functions of proglucagon and glucagon receptor genes in fish. *Gen Comp Endocrinol*. 2018;261:148-165. doi:10.1016/J.YGCEN.2018.03.003

70. Choe SK, Hirsch N, Zhang X, Sagerström CG. hnf1b Genes in Zebrafish Hindbrain

Development. *Zebrafish*. 2008;5(3):179. doi:10.1089/ZEB.2008.0534

71. Milewski WM, Duguay SJ, Chan SJ, Steiner DF. Conservation of PDX-1 Structure, Function, and Expression in Zebrafish. *Endocrinology*. 1998;139:1440-1449. doi:10.1210/endo.139.3.5768

72. Verbruggen V, Ek O, Georlette D, et al. The Pax6b homeodomain is dispensable for pancreatic endocrine cell differentiation in Zebrafish. *J Biol Chem*. 2010;285(18):13863-13873. doi:10.1074/jbc.M110.108019

73. Hoff R. Gene-editing of the monogenic diabetes associated gene hnf1ba to model multisystemic MODY5 disease mechanisms in zebrafish (DR). 2021. https://hdl.handle.net/11250/2834970.

74. Aleström P, D'Angelo L, Midtlyng PJ, et al. Zebrafish: Housing and husbandry recommendations. *Lab Anim*. 2020;54(3):213-224. doi:10.1177/0023677219869037

75. Nasiadka A, Clark MD. Zebrafish Breeding in the Laboratory Environment. *ILAR J*. 2012;53(2):161-168. doi:10.1093/ILAR.53.2.161

76. BioRad. qPCR Assay Design and Optimization . https://www.bio-rad.com/en-no/applications-technologies/qpcr-assay-design-optimization?ID=LUSO7RIVK. Accessed May 26, 2023.

77. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29(9):e45. doi:10.1093/nar/29.9.e45

78. Barbacci E, Reber M, Ott MO, Breillat C, Huetz F, Cereghini S. Variant hepatocyte nuclear factor 1 is required for visceral endoderm specification. *Development*. 1999;126(21):4795-4805. doi:10.1242/DEV.126.21.4795

79. Prince VE, Anderson RM, Dalgin G. Zebrafish Pancreas Development and Regeneration: Fishing for Diabetes Therapies. *Curr Top Dev Biol*. 2017;124:235-276. doi:10.1016/BS.CTDB.2016.10.005

80. Madariaga L, García-Castaño A, Ariceta G, et al. Variable phenotype in HNF1B mutations: extrarenal manifestations distinguish affected individuals from the population with congenital anomalies of the kidney and urinary tract. *Clin Kidney J*. 2019;12(3):373-379. doi:10.1093/CKJ/SFY102

81. Lim SH, Kim JH, Han KH, et al. Genotype and Phenotype Analyses in Pediatric Patients with HNF1B Mutations. *J Clin Med*. 2020;9(7):1-11. doi:10.3390/JCM9072320

82. Yagihashi S, Mizukami H, Sugimoto K. Mechanism of diabetic neuropathy: Where are we now and where to go? *J Diabetes Investig*. 2011;2(1):32. doi:10.1111/J.2040-1124.2010.00070.X

83. Verhave JC, Bech AP, Wetzels JFM, Nijenhuis T. Hepatocyte Nuclear Factor 1β–Associated Kidney Disease: More than Renal Cysts and Diabetes. *J Am Soc Nephrol*. 2016;27(2):345-353. doi:10.1681/ASN.2015050544

84. Potter JD, Robertson SP, Johnson JD. Magnesium and the regulation of muscle contraction. *Fed Proc*. 1981;40(12):2653-2656.

85. Cheng Y, Zhong DP, Ren L, Yang H, Tian CF. Unusual manifestations of young woman with MODY5 based on 17q12 recurrent deletion syndrome. *BMC Endocr Disord*. 2022;22(1):1-6. doi:10.1186/S12902-022-00989-6

86. Quinlivan VH, Farber SA. Lipid Uptake, Metabolism, and Transport in the Larval Zebrafish. *Front Endocrinol (Lausanne)*. 2017;8(319). doi:10.3389/fendo.2017.00319

87. Wang W, Lo ACY. Diabetic Retinopathy: Pathophysiology and Treatments. *Int J Mol Sci*. 2018;19(6):1816-1830. doi:10.3390/ijms19061816

88. Zhou L, Irwin DM. Fish proglucagon genes have differing coding potential. *Comp Biochem Physiol Part B Biochem Mol Biol*. 2004;137(2):255-264. doi:10.1016/J.CBPC.2003.11.009

89. Lavergne A, Tarifeño-Saldivia E, Pirson J, et al. Pancreatic and intestinal endocrine cells

in zebrafish share common transcriptomic signatures and regulatory programmes. *BMC Biol*. 2020;18(1):109. doi:10.1186/S12915-020-00840-1

90. Sousa M, Rego T, Armas JB. Insights into the Genetics and Signaling Pathways in Maturity-Onset Diabetes of the Young. *Int J Mol Sci*. 2022;23(21). doi:10.3390/ijms232112910

91. Hiemisch H, Schütz G, Kaestner KH. Transcriptional regulation in endoderm development: characterization of an enhancer controlling Hnf3g expression by transgenesis and targeted mutagenesis. *EMBO J*. 1997;16(13):3995-4006. doi:10.1093/emboj/16.13.3995

92. National Center for Biotechnology Information. ClinVar; [VCV000012648.3]. https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000012648.3. Accessed May 27, 2023.

93. Rebouissou S, Vasiliu V, Thomas C, et al. Germline hepatocyte nuclear factor 1α and 1β mutations in renal cell carcinomas. *Hum Mol Genet*. 2005;14(5):603-614. doi:10.1093/HMG/DDI057

Appendix

A1 Ct-values from RT-qPCR analysis of pancreatic genes

		tuba1c (2 replicates)		*hprt1l* (2 replicates)		*hnf1ba* (2 replicates)		*hnf1bb* (2 replicates)		*foxa3* (2 replicates)	
wt	1	22.45	22.47	26.92	26.88	25.42	25.58	28.09	28.10	26.43	26.46
	2	22.85	22.78	26.05	25.95	27.75	27.67	29.14	28.94	27.15	27.27
	3	23.55	23.61	27.50	27.24	26.91	27.19	29.88	29.93	28.10	28.18
	4	23.93	23.88	26.63	26.63	27.01	26.97	29.87	30.90	28.97	28.58
	5	24.48	24.46	26.38	26.30	26.11	26.11	31.43	31.17	28.02	28.11
	6	24.74	24.77	28.05	28.08	27.95	28.11	31.35	31.41	29.25	29.37
hnf1ba (-/-)	1	23.20	23.19	26.55	26.57	25.12	25.10	28.07	28.16	26.05	26.16
	2	23.13	23.09	25.94	26.04	24.89	28.97	28.07	27.92	26.19	26.10
	3	23.54	23.56	26.34	26.39	25.38	25.36	28.22	28.28	26.36	26.46
	4	23.66	23.59	26.42	26.39	25.58	25.66	28.66	28.64	26.59	26.65
	5	23.38	23.30	26.18	26.16	25.13	25.28	28.20	28.18	26.36	26.34
	6	23.34	23.56	26.42	26.44	24.86	24.92	28.05	28.13	25.90	26.13

		gcga (2 replicates)		*gcgb* (2 replicates)		*insa* (2 replicates)		*insb* (2 replicates)		*pax6b* (2 replicates)		*pdx1* (2 replicates)	
wt	1	24.45	24.63	25.63	25.34	24.59	24.55	33.46	34.12	26.09	26.14	27.77	27.46
	2	24.32	24.20	25.00	25.02	24.01	23.75	32.71	32.73	26.61	26.51	26.85	26.92
	3	24.42	24.52	25.14	25.09	24.08	24.11	33.62	33.78	27.39	27.37	27.35	27.20
	4	24.80	24.58	25.36	25.51	24.28	24.17	33.51	33.51	28.01	27.96	27.58	27.52
	5	25.02	25.26	25.94	25.93	24.68	24.59	33.54	34.40	28.87	28.70	28.31	28.34
	6	25.31	25.36	26.02	26.10	25.04	24.92	34.34	35.12	29.45	29.47	28.55	28.37
hnf1ba (-/-)	1	23.14	23.08	26.30	26.32	26.76	26.97	31.97	32.25	26.18	26.13	29.70	29.55
	2	22.66	22.51	26.68	26.75	28.22	28.14	30.95	30.89	25.82	25.78	29.82	29.75
	3	22.86	23.01	27.40	27.39	28.57	28.52	31.20	31.49	26.11	26.13	30.47	30.45
	4	22.46	22.39	27.15	27.32	28.30	28.46	31.60	31.81	26.44	26.48	30.18	30.41
	5	24.15	24.15	27.84	27.86	28.76	28.84	30.50	30.06	26.06	26.03	30.20	30.10
	6	24.05	24.22	28.17	28.40	29.17	28.66	30.45	30.78	26.35	26.25	30.30	30.32

Table A1: Ct-values obtained from RT-qPCR for wild-type (wt) and *hnf1ba*(-/-) mutant zebrafish larvae. The values were used to calculate the relative gene expression of pancreas genes. Only genes with optimised primer pairs are included (and *insb* due to relevance). The numbers 1-6 illustrate the six different samples per genotype group.

A2 Number of counted cells for all individuals

	#	Insulin	Glucagon	Other cells
wt	1	15	6	2
	2	18	15	2
	3	20	3	4
	4	31	15	4
	5	36	12	4
	6	45	3	8
hnf1ba (-/-)	1	1	0	0
	2	2	0	0
	3	5	0	0
	4	25	17	4
	5	22	10	3
	6	30	11	1

Table A2: The number of insulin- and glucagon-producing cells counted per individual for wild-type (wt) and *hnf1ba*(-/-) mutants zebrafish larvae. The cells were counted on confocal images of cross-sections of the immunostained pancreas. The numbers 1-6 illustrate the six individuals per genotype group ."Other cells" indicates the number of cell nuclei found within the endocrine pancreas that did not have any detectable insulin- or glucagon-signal. These cells are assumed to be either gamma- or delta-cells, or part of insulin- or glucagon-producing cells located more anterior or posterior to the cross-section.