



Inês Carvalho Mota
BSc in Biology

UNRAVELLING *NR5A2* TRANSCRIPTIONAL REGULATION IN PANCREATIC CANCER

MASTER IN MOLECULAR GENETICS AND BIOMEDICINE
NOVA University Lisbon
September, 2024

Unravelling *NR5A2* Transcriptional Regulation in Pancreatic Cancer

Inês Carvalho Mota

BSc in Biology

Adviser: Joana Teixeira

Researcher, Vertebrate Development and Regeneration Group, i3S, University of Porto

Co-advisers: Ana Rita Grosso

Assistant Professor, NOVA School of Science and Technology

Examination Committee:

Chair: Pedro M. Costa,

Assistant Professor, NOVA School of Science and Technology

Rapporteurs: Renata Freitas,

Assistant Professor, NOVA School of Science and Technology

Adviser: Joana Teixeira,

Researcher, Vertebrate Development and Regeneration Group, i3S, University of Porto

Unravelling *NR5A2* Transcriptional Regulation in Pancreatic Cancer

Copyright © Inês Mota, NOVA School of Science and Technology, NOVA University Lisbon.

The NOVA School of Science and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

ACKNOWLEDGMENTS

To my supervisor, Joana Teixeira, for her enormous dedication to always help me. I'm grateful for her guidance and all the indispensable advice. Much of what I have learned is thanks to the knowledge she has shared with me, which I will carry forward throughout my future career.

I must also thank Dr. José Bessa for allowing me to join his research group at i3S and to contribute for the development of this amazing project. His passion for science and research is evident in his speech and in his gaze, which obviously didn't go unnoticed and inspired me during this year of discoveries and hard work.

To my co-supervisor, Professor Ana Rita, I'm grateful for her constant support and kindness, even from afar. Her guidance and advice have been invaluable from the very beginning of this journey and will undoubtedly stay with me.

To the entire VDR lab for their help and support throughout this year, thank you Ana Eufrásio, Joana Marques, Isabel, Fábio, Francsico, Sofia, Marta Duque e João Nogueira. I have to say a special thank you to my companions, Diogo Pimenta, Marta Bao and Miguel, everything was easier with you and your smiles.

To one of my forever roommates, Nélia Fonseca, we did it, it's done and well done! Every day was brighter with you by my side!

To my family, thank you for being there for me, not only through these challenging years, but throughout my all life.

To my boyfriend, no words can truly express how grateful I am for everything you've done for me. Your love and support mean the world. Cheers to the beginning of our new journey together.

“Just keep swimming, swimming, swimming!”

ABSTRACT

Pancreatic cancer (PC) represents a significant health challenge with limited early detection methods and effective treatments. Pancreatic ductal adenocarcinoma accounts for most PC cases and is typically linked to duct cells due to its morphology and markers found. There is evidence that acinar cells can acquire ductal features, sparking debates regarding the origin of PC tumours. The identification of protein-coding genes like *NR5A2*, highly expressed in the pancreas, has been essential to understand PC development. Genome-wide association studies have uncovered some variants associated with PC in non-coding regions nearby *NR5A2* gene. Particular attention has been given to *cis*-regulatory elements (CREs), specifically enhancers, which interact with gene promoters, increasing their expression. However, the impact of enhancer disruption on PC remains unclear and poorly explored. In this project, exocrine pancreas enhancers of *NR5A2/nr5a2* using zebrafish were explored, to better understand their role in PC development. We validated *in vivo* enhancer activity of a human acinar putative enhancer sequence (PES) that overlap with a PC SNP rs3790844 (seq44). We found that this sequence has both pancreatic exocrine and endocrine enhancer activity, suggesting that the disruption of seq44 could have far-reaching effects on gene regulation within pancreas. Then, ATAC-seq and ChIP-seq data for epigenetic marks associated with enhancer activity were used to analyse the zebrafish genomic landscape of *nr5a2* in pancreatic acinar cells, identifying two PES: Region 1 and Region 2. Enhancer activity of these sequences was also tested *in vivo* and showed that Region 1 has no pancreatic enhancer activity, while Region 2 has acinar enhancer activity. CRISPR-Cas9-induced genomic deletions targeting those zebrafish PES were performed, however, we only successfully validated the genomic deletion for Region 1. Further studies are necessary to evaluate the impact of these sequences of *NR5A2/nr5a2* expression, particularly in the context of the human biology of PC.

Keywords: Pancreatic cancer, *nr5a2*, pancreas, enhancers, zebrafish

RESUMO

O cancro pancreático (CP) representa um desafio significativo para a saúde, com métodos de deteção e tratamentos limitados. O adenocarcinoma ductal pancreático é responsável pela maioria dos casos de CP, relacionado com células ductais pela sua morfologia e marcadores moleculares apresentados. Estudos recentes indicam que células pancreáticas acinares podem adquirir características ductais, questionando a origem do CP. A identificação de genes codificantes como o *NR5A2*, altamente expresso no pâncreas, tem sido essencial para compreender o desenvolvimento do CP. Estudos de associação de todo o genoma revelaram variantes associadas a CP em regiões não codificantes próximas do gene *NR5A2*. Especial atenção foi dada aos elementos *cis*-reguladores, especificamente aos potenciadores, que interagem com promotores, potenciando a expressão genica. Contudo, o impacto da disrupção destes elementos no CP permanece indefinido. Neste projeto, foram explorados possíveis potenciadores do gene *NR5A2/nr5a2* no pâncreas exócrino utilizando o peixe-zebra, para melhor compreender o seu papel no desenvolvimento do CP. Validámos *in vivo* a atividade de uma possível sequência potenciadora contendo um polimorfismo rs3790844 (seq44), associada com CP, na paisagem genómica humana do gene *NR5A2*. Verificámos que esta sequência tem atividade potenciadora no pâncreas exócrino e endócrino do peixe-zebra, sugerindo que a sua perturbação pode afetar amplamente a regulação génica pancreática. Dados de ATAC-seq e ChIP-seq para marcas epigenéticas associadas a atividade potenciadora foram utilizados para analisar a paisagem genómica do gene *nr5a2* em células acinares, identificando duas possíveis sequências com atividade potenciadora: Região 1 e Região 2. Testes *in vivo* revelaram que a Região 1 não tem atividade potenciadora, enquanto a Região 2 demonstrou atividade potenciadora em células acinares. Validamos com sucesso deleções genómicas induzidas por CRISPR-Cas9 visando a Região 1. Mais estudos são necessários para avaliar o impacto destas sequências na expressão do gene *NR5A2/nr5a2* no contexto da biologia humana do CP.

Palavras chave: Cancro pancreático, *nr5a2*, pâncreas, potenciador, peixe-zebra

CONTENTS

1	INTRODUCTION	1
1.1	Pancreas	1
1.1.1	Vertebrate Pancreas Development and Transcriptional Network.....	2
1.2	Transcriptional Regulation of Gene Expression	4
1.2.1	<i>Cis</i> -Regulatory Elements - Enhancers	5
1.2.2	Genome-Wide Assays to Identify and Predict Enhancer Regions	6
1.2.3	Pancreatic Cancer and <i>Cis</i> -Regulation	7
1.3	Zebrafish as a Model to Study Pancreatic Cancer	9
1.3.1	Zebrafish Pancreas Organization and Development.....	10
1.4	Genetic Tools to Study CREs <i>in vivo</i>	11
1.4.1	Reporter Assays to Study Enhancer Activity	11
1.4.2	Loss of Function Assays using CRISPR-Cas9	12
1.5	Main Objectives	12
2	MATERIALS AND METHODS.....	13
2.1	Zebrafish Stocks, Husbandry, Breeding, and Embryo Culture	13
2.1.1	Zebrafish Husbandry	13
2.1.2	Breeding and Egg Collection.....	13
2.1.3	Microinjection in 1-cell-stage Zebrafish Embryos	14
2.2	DNA Extraction from Zebrafish Embryos	14
2.3	SNPs and Putative Enhancers Sequences Selection.....	15
2.4	Generation of Plasmids with the Selected Sequences	15
2.4.1	Primers Design and PCR Amplification.....	15
2.4.2	Cloning of Fragments into pCR8™/GW/TOPO® Vector and Chemically Competent Bacteria Transformation.....	16

2.4.3	Z48 based Vector Recombination of the Selected Sequences.....	18
2.4.4	Phenol/Chloroform Purification of the Plasmid for Microinjection.....	19
2.5	Transgenesis.....	19
2.5.1	Microinjection of Plasmids and mRNA into Zebrafish Embryos.....	19
2.6	Screening for Enhancer Activity.....	20
2.6.1	Fixation and DAPI Staining (11-12dpf)	20
2.6.2	Fixation and Immunohistochemistry (48 hpf)	21
2.6.3	Leica SP5 Confocal Observation.....	22
2.7	Mutagenesis using CRISPR-Cas9 System.....	22
2.7.1	Cas9 Target Design and sgRNA Synthesis	22
2.7.2	Microinjection of sgRNAs into Zebrafish Embryos.....	25
2.7.3	Genotyping	26
2.8	Predicting Putative TFBS – JASPAR Analysis	26
3	RESULTS AND DISCUSSION.....	27
3.1	Prediction of CREs	27
3.1.1	Human <i>NR5A2</i> Regulatory Landscape Analysis and Selection of Putative Enhancer Sequences Overlapping with SNPs	27
3.1.2	Zebrafish <i>nr5a2</i> Regulatory Landscape Analysis and Selection of Putative Enhancer Sequences.....	28
3.2	PCR Amplification of the Putative Enhancer Sequences	29
3.3	Cloning of the PCR Products into pCR8™/GW/TOPO® Vector.....	30
3.4	Recombination of the Putative Enhancer Sequences into Z48 vector for Enhancer Reporter Assays	30
3.5	<i>In vivo</i> Transgenesis Reporter Assays	31
3.5.1	Assessment of Pancreatic Domains.....	31
3.5.2	Test Enhancer Activity of a Human <i>NR5A2</i> Putative Enhancer Sequence in Pancreatic Domains	33
3.5.3	Test Enhancer Activity of Zebrafish <i>nr5a2</i> Putative enhancers Sequences in Pancreatic Acinar Cells.....	38
3.6	<i>In vivo</i> Mutagenesis Assays	44
4	CONCLUSIONS AND FUTURE PERSPECTIVES	47

LIST OF FIGURES

Figure 1.1 Human pancreas anatomy and structure.	2
Figure 1.2. Overview of vertebrate pancreas development and differentiation.	3
Figure 1.3. Mechanisms of gene regulation and chromatin accessibility.....	6
Figure 1.4 Structural comparison between the human and zebrafish adult pancreas.	10
Figure 2.1 Graphical representation of pCR8™/GW/TOPO® vector (2817 bp) with the PES flanked by two attL and two Gateway® recombination sites.....	17
Figure 2.2 Graphic representation of the recombination of the PES between the pCR8™/GW/TOPO® and the Z48 vector (7890 bp).	18
Figure 2.3 Graphical representation of the pDR274 vector (2147 bp) with the scaffold where the sgRNAs were inserted.	24
Figure 3.1 The human <i>NR5A2</i> genomic landscape (hg38) containing a PES that overlaps with the PC SNP and epigenetic marks for enhancer activity.....	28
Figure 3.2 The zebrafish <i>nr5a2</i> genomic landscape (danRer10) containing two PES that overlap with regions of accessible chromatin and epigenetic marks for enhancer activity.....	29
Figure 3.3 PCR amplification products of Region 1 (470 bp) and Region 2 (798 bp) from zebrafish gDNA.	29
Figure 3.4 Analysis of pCR8™/GW/TOPO® vector digestion by EcoRI enzyme.....	30
Figure 3.5 Analysis of Z48 vector digestion by EcoRI enzyme.....	31
Figure 3.6 Representative images of zebrafish tg(zP.E3:EGFP, <i>ela</i> :mCherry) larvae at 11-12 dpf, tested to be used as positive control in enhancer reporter assays and to define the pancreatic acinar domain.	32
Figure 3.7 Representative images of the zebrafish tg(zP.E3:EGFP, <i>sst</i> :mCherry) embryos 48hpf tested to be used as positive control in enhancer reporter assays and to define the pancreatic progenitor and endocrine domain.	33
Figure 3.8 Representative images of a F1 zebrafish embryo midbrain at 48 hpf.....	34
Figure 3.9 Representative confocal images of the enhancer reporter assay for human seq44 in zebrafish larvae at 11-12 dpf.....	35

Figure 3.10 Representative confocal images of the enhancer reporter assays for human seq44 in zebrafish embryos at 48hpf.	37
Figure 3.11 Representative images of F1 zebrafish larvae at 11-12 dpf.	38
Figure 3.12 Representative images of the enhancer reporter assays for Region 1 in 11-12 dpf zebrafish larvae after outcrossing with tg(<i>ela</i> :mCherry).	39
Figure 3.13 The H3K4me3 HiChIP data from adult whole pancreas (red bar) shows chromatin putative interactions between Region 1 (orange) and the <i>nr5a2</i> promoter in the zebrafish genomic landscape (danRer10).	40
Figure 3.14 Representative confocal images of the enhancer reporter assays for Region 2 in zebrafish larvae at 11-12 dpf after outcrossing with tg(<i>ela</i> :mCherry).	41
Figure 3.15 The H3K4me3 HiChIP data from adult whole pancreas (red bar) shows chromatin putative interactions between Region 2 (orange) and the <i>nr5a2</i> promoter in the zebrafish genomic landscape (danRer10).	41
Figure 3.16 Vertebrate TFBS predicted by the JASPAR software within Region 2 (798 bp).	43
Figure 3.17 Design of sgRNAs targeting the Region 1.	45
Figure 3.18 Confirmation of the efficacy of two sgRNA pairs targeting Region 1 using the CRISPR-Cas9 system.	45
Figure 3.19 Zebrafish <i>nr5a2</i> genomic landscape (danRer10) containing the Region 2 (orange), flanked by repeating elements (grey bars).	46
Figure 3.20 Design of sgRNAs targeting the Region 2.	46

LIST OF TABLES

Table 2.1 Primers designed to amplify the PES selected in the zebrafish <i>nr5a2</i> regulatory landscape (Region 1 and Region 2).	16
Table 2.2 Primers designed to amplify the PES selected in the <i>NR5A2</i> regulatory landscape of human genome (Seq44).....	16
Table 2.3 Oligonucleotides designed for each sgRNA upstream and downstream targeting the PES in the zebrafish genome of <i>nr5a2</i> gene.	23
Table 3.1 JASPAR analysis of the top 5 vertebrate TFs predicted in bind to Region 2.....	43

ACRONYMS

ADM	Acinar-to-ductal metaplasia
ATAC-seq	Assay for Transposase-Accessible Chromatin with sequencing
Bp	Base pair
BSA	Bovine Serum Albumin
ChIP-seq	Chromatin immunoprecipitation with sequencing
CREs	<i>Cis</i> -Regulatory Elements
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphates
Dpf	Days post-fertilization
DSB	Double-strand brakes
DTT	Dithiothreitol
EDTA	Ethylenediamine Tetraacetic Acid

Ela	Elastase
EtOH	Ethyl Alcohol
Fw	Forward
gDNA	Genomic DNA
GFP	Green Fluorescent Protein
GWAS	Genome-Wide Association Studies
H3K27ac	Histone H3 lysine 27 acetylation
H3K4me3	Histone H3 lysine 4 tri-methylation
H3K27me3	Histone H3 lysine 27 tri-methylation
H3K4me1	Histone H3 lysine 4 mono-methylation
HiChIP	Highly Integrative Chromatin Immunoprecipitation
Hr(s)	Hour (s)
hpf	Hours post-fertilization
IN	Intestine
KRAS	Kirsten Ras homolog
LB	Lysogeny Broth
LRH-1	Liver receptor homologue-1
LRL	Liver Right Lobe
LT	Left Testis

min	Minute (s)
MPCs	Multipotent Progenitor Cells
mRNA	Messenger RNA
NHEJ	Non-homologous end joining
NKX6.1	Homeobox protein 6.1
NR5A2	Nuclear Receptor Subfamily 5 group A member 2
NTPs	Nucleoside Triphosphates
O.N.	Overnight
PanIN	Pancreatic intraepithelial neoplasia
PC	Pancreatic Cancer
PCR	Polymerase Chain Reaction
PDAC	Pancreatic Ductal Adenocarcinoma
PES	Putative enhancer sequence
pH	Potential of Hydrogen
PI	Principal Islet
pS	Pico Siemens
PTF1A	Pancreas-specific Transcription Factor 1A
PTMs	Post-Transcriptional Modifications
PTU	1-phenyl-2-thiourea
Rev	Reverse

RNA	Ribonucleic Acid
Rpm	Rotations per minute
SB	Swim Bladder
Sec	Second(s)
sgRNA	Single-guide RNA
SI	Secondary Islet
SNPs	Single-Nucleotide Polymorphism(s)
sst	Somatostatin
TE	Tris-EDTA
TFBS	TF-binding sites
TFs	Transcription Factors
UV	Ultraviolet
WT	Wild-Type

INTRODUCTION

1.1 Pancreas

The pancreas is a crucial organ located in the abdominal cavity within the digestive system (Anon 2010; Larsen and Grapin-Botton 2017) being responsible for glucose regulation, nutrition, metabolism, and digestion (Anon 2010; Atkinson et al. 2020; Shih et al. 2013). This retroperitoneal organ is divided in two major glandular systems: the endocrine and exocrine pancreas (Anon 2010; Atkinson et al. 2020; [Fig.1.1](#)). The endocrine pancreas is composed by five different types of cells: glucagon-producing α -cells, insulin-producing β -cells, somatostatin-producing δ -cells, ghrelin-producing ϵ -cells and pancreatic polypeptide-producing PP cells (Atkinson et al. 2020; Bakhti, Böttcher, and Lickert 2019). These cells are grouped in clusters, called islets of Langerhans, which secrete key hormones into the bloodstream, playing a crucial role in maintaining glucose homeostasis (Anon 2010; Bakhti et al. 2019; Shih et al. 2013). The exocrine gland comprises acini, clusters of pancreatic acinar and duct cells (Anon 2010; Atkinson et al. 2020; Low et al. 2010). Almost 80% of the exocrine pancreas is composed by acinar cells that secrete digestive enzymes, such as elastase, protease, lipase, and fluid into the duodenum through a network of ducts (Anon 2010; Atkinson et al. 2020; Low et al. 2010; Pan and Brissova 2014). These enzymes play a critical role in breaking down proteins, fat and carbohydrates, facilitating digestion (Anon 2010; Atkinson et al. 2020; Shih et al. 2013).

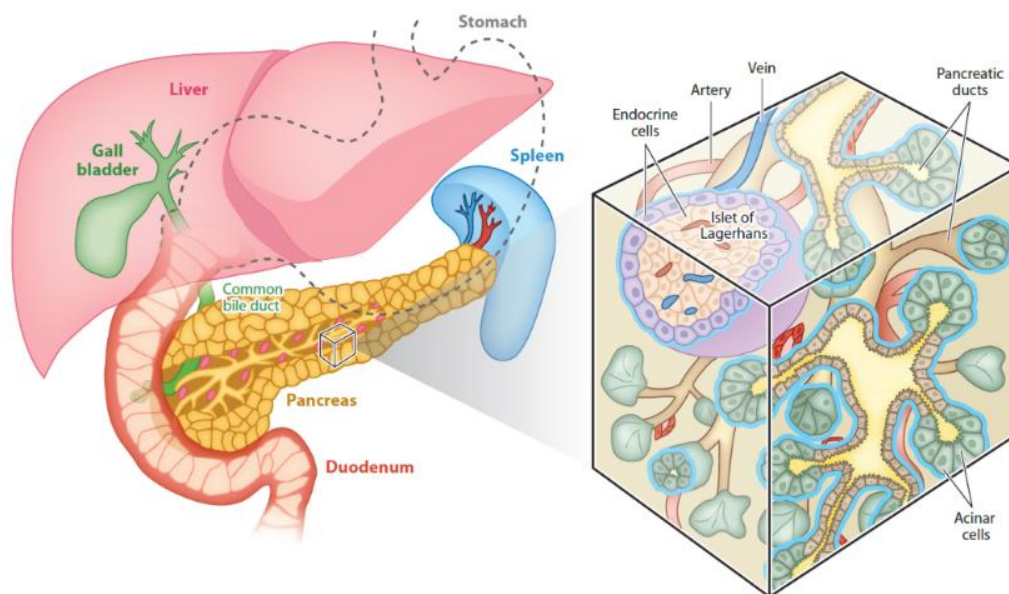


Figure 1.1 Human pancreas anatomy and structure. Schematic representation of the pancreas' anatomical localization in relation to the adjacent organs, including the liver, gall bladder, stomach, spleen and duodenum. View of the exocrine and endocrine pancreas, with detailed network of ducts connected with acinar cells, and a cluster of endocrine cells (islet of Langerhans) (Adapted from Shih et al. 2013).

1.1.1 Vertebrate Pancreas Development and Transcriptional Network

Pancreas development is a complex process tightly regulated by a network of transcription factors (TFs) that determine the differentiation of pancreatic progenitor cells into distinct cell types (Edlund 2001). The evagination of the embryonic endoderm in vertebrates is the beginning of the pancreas organogenesis, which forms the ventral and dorsal buds that ultimately fuse to form the pancreas (Arda, et al. 2013; Bastidas-Ponce et al. 2017; Pan and Brissova 2014; Fig.1.2 – A). Both of them are composed by multipotent progenitor cells (MPCs) that can develop all pancreatic cell types from tip and trunk cells with two major fate decisions: acinar and duct cell lineage or endocrine cell lineage (Aimeé Bastidas-Ponce et al. 2017; Pan and Brissova 2014; Fig.1.2 – B and C).

Pancreas-specific Transcription Factor 1A (PTF1A) promotes the formation of the tip domain, being responsible for the acinar cell differentiation (Aimeé Bastidas-Ponce et al. 2017; Duque et al. 2022; Shih et al. 2013). *Homeobox protein 6.1 (NKX6.1)* on the contrary drives the development of the trunk domain, being responsible for the differentiation into duct or endocrine cells (Aimeé Bastidas-Ponce et al. 2017; Duque et al. 2022; Schaffer et al. 2010; Shih et al. 2013). These two TFs are part of a positive

feedback loop of reciprocal inhibition essential for cell fate decisions that form the whole pancreas in vertebrates, such as humans and mice (Duque et al. 2022; Schaffer et al. 2010; Shih et al. 2013; Fig1.2 – C).

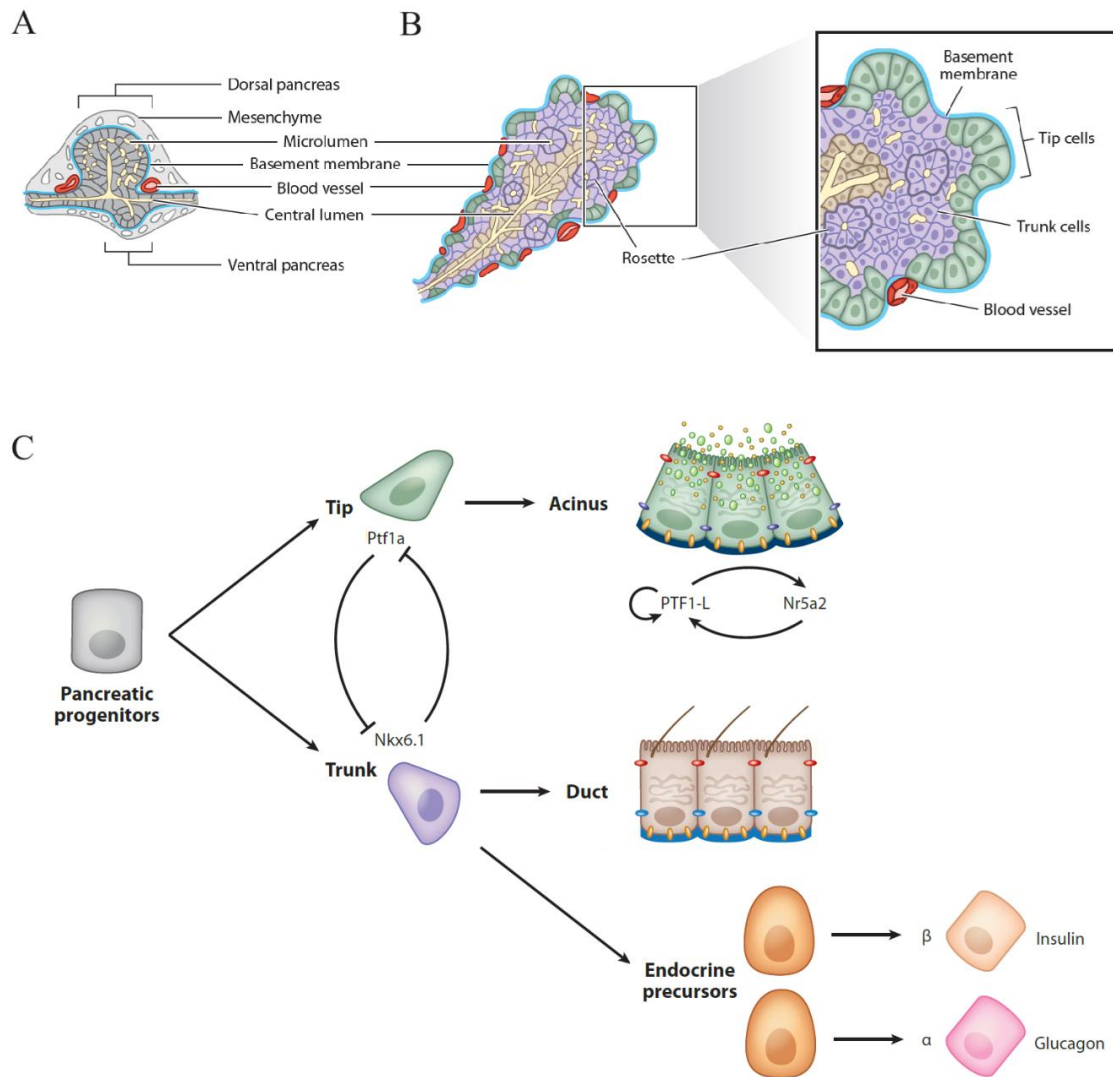


Figure 1.2. Overview of vertebrate pancreas development and differentiation. **A-** During the early stages of pancreatic development, the endoderm is divided into the dorsal and ventral pancreas which undergo evagination into the surrounding mesenchyme and become encapsulated by the basement membrane. Blood vessels (red) surrounding the epithelial buds with the microlumen (yellow) situated between them. **B-** In a later stage of the pancreatic development, a layer of tip cells (green) forms branching protrusions, while the trunk cells (purple) organize in rosettes around the lumen. **C-** Pancreatic progenitors differentiate into either tip or trunk cells, with a positive feedback loop of reciprocal inhibition between *Nkx6.1* and *Ptf1a* driving tip/trunk separation. Tip cells adopt an acinar phenotype maintained through the mutual positive regulation between *Ptf1a* and *Nr5a2*. Trunk cells are bipotent cells and therefore can differentiate into duct or endocrine cells. Endocrine precursors can even differentiate into five distinct endocrine cell types, although only α and β cells are represented in this image (Adapted from Shih et al. 2013).

This feedback loop is not only responsible for the early commitment of pancreatic progenitors, but also for maintaining acinar cell identity through the continuous expression of *PTF1A*, which is supported by the expression of *NR5A2* (*nuclear receptor subfamily 5 group A member 2*) (Azevedo-Pouly et al. 2023; Bastidas-Ponce et al. 2017; Hale et al. 2014; Larsen and Grapin-Botton 2017; Sandhu, et al. 2021).

The gene *NR5A2*, also known as *LRH-1* (*liver receptor homologue-1*), belongs to the hormone family of orphan nuclear receptors of ligand-activated TFs (Holmstrom et al. 2011; Luo et al. 2017; Sandhu et al. 2021). In humans, this gene is in chromosome 1q32.1 and is highly expressed in the pancreas, liver, intestine and ovary, where it contributes to many physiological processes, including pancreatic development and homeostasis (Azevedo-Pouly et al. 2023; Larsen and Grapin-Botton 2017; Sandhu et al. 2021). This gene plays an essential role in early pancreatic development as it drives the expansion of pancreatic progenitors (Hale et al. 2014; Sandhu et al. 2021). As development progresses *NR5A2* expression becomes restricted to pancreatic epithelial cells leading to terminal acinar differentiation and specialization (Hale et al. 2014; Larsen and Grapin-Botton 2017). In the adult pancreas, it is known that the expression of this gene continues to be essential to maintain acinar cell identity, being crucial in fetal organogenesis (Aimée Bastidas-Ponce et al. 2017; Hale et al. 2014; Sandhu et al. 2021). Many pathologies, such as gastrointestinal malignancies and chronic pancreatitis, have been associated to *Nr5a2* haploinsufficiency, highlighting this gene's significance in preserving healthy cellular homeostasis and averting illness (Hale et al. 2014; Sandhu et al. 2021).

1.2 Transcriptional Regulation of Gene Expression

In vertebrates, the intricate process of transcription regulation of gene expression is significantly influenced by the dynamic structure of chromatin and the interaction of different genetic elements (Elkon and Agami 2017; Maston, et al. 2006). The basic structural units of chromatin are nucleosomes. They are composed of DNA wrapped around histone proteins, and their modifications or positioning on DNA, play a key role in regulating gene expression (Schones et al. 2008). Chromatin can be found in two distinct states: condensed (closed) or relaxed (open) which directly affects the access to genetic information. In its open state, chromatin allows TFs and RNA polymerase II to bind to DNA and regulate gene transcription, while closed chromatin inhibits this interaction (Gowthaman et al. 2020; Maston et al. 2006; Shlyueva, Stampfel, and Stark 2014; [Fig. 1.3](#)).

These processes involve histones, proteins with N-terminal tails that can suffer post-translational modifications (PTM), such as acetylation or methylation (Hernandez-Garcia and Finer 2014; Maston et al. 2006). These histone modifications alter the interaction between histones and DNA, influencing how

tight the DNA is wrapped around the histones, therefore regulating the transition between open and closed chromatin states. By controlling DNA accessibility, these modifications ensure that genes are expressed in the right place at the right time, which is essential for the proper development and cellular function (Lodish et al. 2003; Schones et al. 2008).

1.2.1 *Cis*-Regulatory Elements - Enhancers

Cis-regulatory elements (CREs) are crucial non-coding DNA sequences responsible for the regulation of gene expression (Mathelier et al. 2015). The regulatory landscape of a gene is extremely complex, frequently involving more than one CREs that cooperate to define distinct patterns of gene expression in both temporal and spatial contexts (Lodish et al. 2003; Maston et al. 2006). CREs include promoters, enhancers, silencers and insulators, each with distinct roles in gene regulation (Maston et al. 2006; Mathelier et al. 2015; Riethoven 2010).

Enhancers have proven to be especially important in transcriptional regulation (Hernandez-Garcia and Finer 2014; Mathelier et al. 2015). These specific elements are sequences with 200-500 base pairs (bp) characterised by their ability to control the expression of one or more genes, independently from the distance or orientation relative to the particular gene promoter (Cho 2012; García-González et al. 2016; Natoli and Andrau 2012). Active enhancers are found in open chromatin regions, where DNA is free of nucleosomes and accessible for TFs to bind. This availability allows TFs to form DNA loops, bringing enhancers and core promoters of the target gene into close proximity facilitating and increasing transcription initiation (Maston et al. 2006; Natoli and Andrau 2012; Shlyueva et al. 2014; [Fig. 1.3](#)). Histones modifications act as markers for the different chromatin states, which can be detected using techniques that can help to predict and identify sequences with enhancer activity (Gowthaman et al. 2020). Specific histone marks are associated with enhancers activity (Shlyueva et al. 2014). Discovering the enhancers that regulate the transcription of specific target genes is essential for understanding their role in the mechanisms behind human development and disease (Shlyueva et al. 2014).

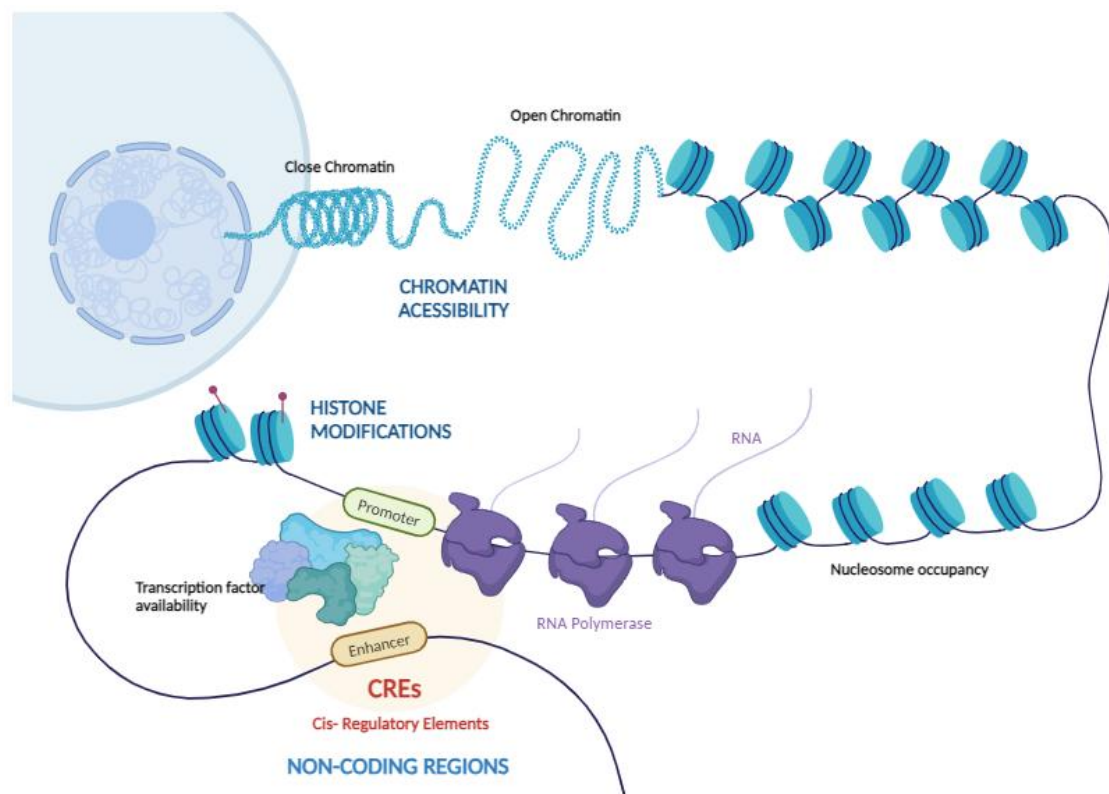


Figure 1.3. Mechanisms of gene regulation and chromatin accessibility. Chromatin can present itself in a condensed (closed) or relaxed (open) form, which directly influences the accessibility for transcription factors (TFs) and RNA polymerase to bind to DNA and consequently regulate gene transcription. The transition between chromatin states is often regulated by histone modifications. *Cis*-regulatory elements (CREs), such as enhancers, located in non-coding regions, play a crucial role in increasing gene transcription. This process occurs through the formation of DNA loops, facilitated by TFs, that bring enhancers into close proximity with their target promoters (Created with ©2024 BioRender).

1.2.2 Genome-Wide Assays to Identify and Predict Enhancer Regions

The identification and prediction of enhancers remains complex, however, advances in molecular and computational biology have improved the ability to detect them (Pennacchio et al. 2013; Wang et al. 2013). Genome-wide tools that profile epigenetic marks have become highly effective for the detection of candidate enhancers. Since the activity of CREs is regulated by DNA-binding proteins and histone modifications, methodologies that identify DNA accessibility and specific histone marks can provide a comprehensive overview of the regulatory landscapes across the genome, facilitating the identification and prediction of enhancers (Elkon and Agami 2017).

1.2.2.1 Chromatin Epigenetic Marks associated to Enhancers

Histone modifications can be useful markers to identify and predict enhancers. Histone H3 lysine 4 mono-methylation (H3K4me1) and the histone H3 lysine 27 acetylation (H3K27ac) marks together are typically associated with active enhancers (Shlyueva et al. 2014). Conversely, Histone H3 lysine 27 trimethylation (H3K27me3) mark is associated with poised enhancers, and Histone H3 lysine 4 trimethylation (H3K4me3) marks gene promoter regions (Shlyueva et al. 2014). The presence and combination of these histone modifications offer important information about the functional identity of CREs (Elkon and Agami 2017; Shlyueva et al. 2014).

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is one of the most used methods to identify and predict enhancers because it allows us to predict the localization of histone modifications associated with enhancer activity (Elkon and Agami 2017; Mundade et al. 2014; Shlyueva et al. 2014). ChIP-seq is a powerful tool to detect the binding of a specific TF to the DNA sequence, by using antibodies that target specific proteins or histone modifications of interest, such as H3K27ac, H3K4me1, H3K4me3, and H3K27me3 (Elkon and Agami 2017). After immunoprecipitation of the DNA-protein complexes obtained, they are sequenced, enabling genome-wide identification of enhancer activity (Elkon and Agami 2017; Mundade et al. 2014). With a low rate of false negatives, ChIP-seq has proven to be a robust method to predict and identify enhancer regions (Shlyueva et al. 2014).

1.2.2.2 Assays for Chromatin Accessibility

Active enhancers are usually localised in regions of open chromatin, which allows DNA to be accessible for TFs to bind (Shlyueva et al. 2014). Several genome-wide techniques have been developed to identify chromatin accessibility, with the Assay for Transposable Accessible Chromatin using Sequencing (ATAC-seq) standing out as a powerful method for locating active regulatory areas of open chromatin (Grandi et al. 2022). This approach uses the hyperactive Tn5 transposase that cuts DNA in zones of open chromatin, inserts adapters into these cut sites and the resulting DNA fragments are amplified and sequenced. Data analysis of the sequenced reads provides a comprehensive view of the chromatin accessibility profile, enabling the identification and prediction of enhancers (Buenrostro et al. 2013; Elkon and Agami 2017; Grandi et al. 2022; Shashikant and Ettensohn 2019).

1.2.3 Pancreatic Cancer and Cis-Regulation

More than 10% of people worldwide are thought to suffer from pancreatic diseases, with pancreatic ductal adenocarcinomas (PDAC) accounting for more than 90% of all cases (Arda et al. 2018; Grant et

al. 2016; Hawksworth et al. 2019). For this reason, throughout this thesis, PDAC will be referred to as pancreatic cancer (PC).

The origin of the PC has been a matter of discussion (Backx et al. 2022; Strobel et al. 2007). Historically, PC was thought to arise from duct cells based on the morphology of the tumour and molecular markers found (Backx et al. 2022; Grant et al. 2016). However, some studies have shown that PC can also originate from acinar cells through acinar-to-ductal metaplasia (ADM), a process in which acinar cells de-differentiate into cells with duct-like features (Baldan et al. 2024; Marstrand-Daucé et al. 2023; Storz 2017). This transformation often results in pancreatic intraepithelial neoplasia (PanIN), the most common precursor lesions of PC (Marstrand-Daucé et al. 2023; Pinho et al. 2011; Storz 2017). This inconclusiveness regarding the specific PC origin suggests that multiple pathways can be involved and that both duct and acinar cells can contribute to the development of this disease.

Due to its aggressive nature, PC is considered a high-lethal malignancy (Grant et al. 2016; Zhao and Liu 2020). Some advances have been made in the last few years, however, the 5-year survival rate for this specific type of cancer remains lower than 5% (Grant et al. 2016; Pinho et al. 2011; Vincent et al. 2011). This high mortality rate is largely due to PC resistance to chemotherapy and its tendency to late-stage diagnosis, as it often remains asymptomatic until it has grown to advanced stages (Aguirre and Collisson 2017; Zhao and Liu 2020). In light of these challenges, it is important to understand the molecular and genetic pathways underlying this disease.

Although recent studies have focused in understanding the impact of modifications in protein-coding genes related to PC, such as those found in the *KRAS* gene, these discoveries have not yet translated into effective therapies or tools for diagnosis (Aguirre and Collisson 2017; Venkat, et al. 2021; Vincent et al. 2011). This lack of understanding emphasises how important is to seek for other genetic factors, especially those present in DNA non-coding regions. Genome-wide association studies (GWAS) have been highlighting the importance of non-coding regions, revealing a significant number of single nucleotide polymorphisms (SNPs) associated with PC in these regions. A notable case is the landscape of *NR5A2* gene, where several SNPs have been detected in non-coding regions (Petersen et al. 2010). According to these findings, the transcription regulation of *NR5A2* is not only crucial for pancreatic development but can be also important for tumorigenesis (Chen et al. 2018).

Transcription regulation can be disturbed by alterations in CREs, including SNPs (Cho 2012). This genetic variation is the most common one, occurring almost once in every 1000 bp in the human genome, mostly in non-coding regions (Elkon and Agami 2017; Maston et al. 2006). While coding SNPs can directly impact protein function and structure, SNPs localised in non-coding regions can influence gene

expression by affecting CREs (Elkon and Agami 2017). Disruption of these elements can lead to dysregulation of key genes, such as *NR5A2*, involved in pancreatic function which could result in disease development (Chen et al. 2018; Zhang et al. 2014)

Although *NR5A2* is considered a tumour suppressor gene, its role in PC it is not fully understood yet (Sandhu et al. 2021). The downregulation of this specific gene leads to acinar cell instability and subsequently to ADM and PanIN lesions, both commonly associated with pancreatic inflammation (pancreatitis), causing irreversible damage to acinar cells, potentially contributing to the development of PC (Marstrand-Daucé et al. 2023; Sandhu et al. 2021; Cobo et al. 2018). Controversially, the overexpression of *NR5A2* is linked to tumour growth and poor prognosis in the later stages of PC (Luo et al. 2017; Sandhu et al. 2021). According to some studies using mice, heterozygous deletion of *Nr5a2* increases the susceptibility to pancreatic damage and the likelihood of developing *Kras*-driven tumours (Flandez et al. 2014).

As evidenced by the lack of information regarding the role of *NR5A2* in the development of PC and the doubts concerning the specific origin of this pathology, a more in-depth study of the *NR5A2* gene transcription regulation in the context of PC is essential.

So far, little has been done to explore the impact of the disruption of CREs regulating the expression of *NR5A2* in pancreatic acinar cells, probably due to the scarcity of acinar cell lines. This underscores the importance of identifying enhancers that control *NR5A2* gene expression, specifically in acinar cells, using alternative research models, such as zebrafish.

1.3 Zebrafish as a Model to Study Pancreatic Cancer

In the study of PC, animal models have become essential because they enable the investigation of tumour genetics in terms of development, progression and response to therapies in a whole organism context (Lorberbaum et al. 2020).

Zebrafish (*Danio rerio*) emerged as a powerful animal model in biomedicine and development biology (Choi et al. 2021). The use of these fish offers many advantages over traditional rodent models due to their rapid development (~3 months to adulthood) and high fecundity, which facilitate high-throughput genomic screenings (Aleström, et al. 2006; Hwang and Goessling 2016). Their embryos are transparent and develop outside the mother allowing the manipulation and detailed observation of development processes and tumour formation *in vivo* (Howe et al. 2013; Tehrani and Lin 2011).

These small tropical fish have a partially duplicated genome with 25 chromosomes and 26 206 protein-coding genes (Howe et al. 2013; Kwiatkowska et al. 2022; Tiso et al. 2009). Sharing high genetic homology with humans, approximately 70% of human genes have at least one ortholog in the zebrafish genome (Kwiatkowska et al. 2022). Genomic similarity makes zebrafish useful for studying the conserved pathways and histone modifications involved in vertebrate development and disease mechanisms, including PC (Mudbhary and Sadler 2011; Park and Leach 2018; Tiso et al. 2009).

1.3.1 Zebrafish Pancreas Organization and Development

The zebrafish pancreas resembles the mammalian pancreas in many morphological and functional aspects (Chen et al. 2007; Tehrani and Lin 2011). Such as the human pancreas, the zebrafish pancreas is divided in two compartments. The exocrine pancreas, that comprises acinar cells and duct cells which form a network capable of conducting digestive enzymes to the gut. The endocrine pancreas that is also composed by islets of α , β , δ and PP cells that secrete hormones into the bloodstream (Kimmel and Meyer 2010; c et al. 2024; Tiso et al. 2009; Van De Venter et al. 2021; Fig.1.4).

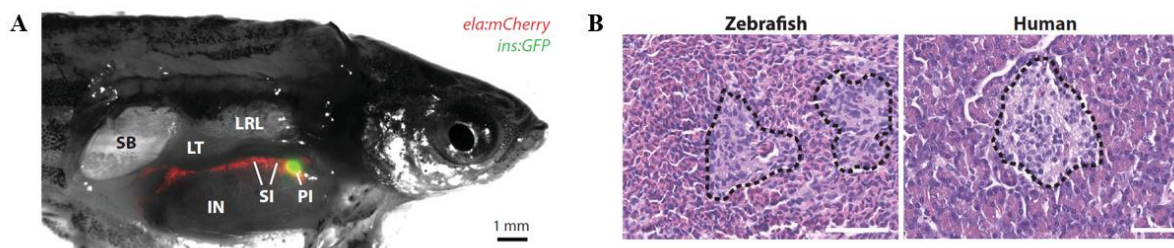


Figure 1.4 Structural comparison between the human and zebrafish adult pancreas. **A-** Zebrafish adult male from the transgenic line *tg(ins:GFP; ela:mCherry)*. GFP (Green Fluorescent Protein) highlights β -cells (green) under the control of the *insulin* promoter, while mCherry labels acinar cells (red) driven by the *elastase* promoter. Notable anatomical marks within the zebrafish pancreas include the secondary (SI) and principal islets (PI). Surrounding the pancreas is also noticeable the intestine (IN), the liver right lobe (LRL), the left testis (LT) and the swim bladder (SB). **B-** Histological cross-sections of the pancreas showcasing the islet of Langerhans (black dashed lines) surrounded by the exocrine tissue in both zebrafish (left) and human (right) samples. Scale bar: 1mm. Abbreviations: ela, elastase; ins, insulin (Adapted from Bordeira-Carriço et al. 2022).

Contrary to what happens in the development of the human pancreas, in the zebrafish, the dorsal bud gives rise only to the endocrine pancreas, while the ventral bud generates both endocrine and exocrine pancreas (Duque et al. 2022; Tehrani and Lin 2011). At 24 hours post-fertilization (hpf) the principle endocrine islet begins to develop and is almost totally developed at 48 hpf (Tehrani and Lin 2011). The

whole pancreas is only fully developed at 11-12 days post-fertilization (dpf) (Van De Venter et al. 2021). The genetic regulation of these development processes involves many genes and signalling pathways, such as *ptf1a*, *nkx6.1* and *nr5a2*, that are conserved between vertebrates (Duque et al. 2022; Mi et al. 2024; Yee, Lorent, and Pack 2005).

The conservation of the pancreatic anatomy, physiology and molecular pathways allows the use of this specific animal as a model to better understand pancreatic tumorigenesis.

1.4 Genetic Tools to Study CREs *in vivo*

1.4.1 Reporter Assays to Study Enhancer Activity

To better understand the role of CREs in human disease development using animal models, it is essential to find sequences that are functional equivalents. Conservation of the sequence alone does not indicate functional equivalency, which means that functional equivalent sequences could not share high sequence similarity (Bordeira-Carriço et al. 2022). This challenge can be mitigated using markers associated with CREs activity alongside enhancer reporter assays, one of the most effective methods to access enhancer activity *in vivo* (Bordeira-Carriço et al. 2022; Eufrásio et al. 2020; Narlikar and Ovcharenko 2009).

This technique is based on the insertion of PES in a vector, upstream of a reporter gene (Narlikar and Ovcharenko 2009). Green fluorescent protein (GFP), which enables the visualization and quantification of gene expression, is typically used as the reporter gene (Narlikar and Ovcharenko 2009; Soboleski, et al. 2005).

The vector also contains the TOL2 transposon flanking the PES. Its microinjection in zebrafish embryos is done alongside with messenger RNA (mRNA) encoding the TOL2 transposase that is in charge of the excision of the transposon and its random insertion in the zebrafish genome (Kemmler et al. 2023; Suster et al. 2009). This TOL2 transposon-based transgenic system guarantees the integration of transgenes into the zebrafish genome with more than 50% of efficiency (Chan et al. 2019; Kwan et al. 2007; Taminato et al. 2016). The flexibility of this system can be increased by the Gateway technology, which allows the easy recombination of DNA fragments between transgenic constructs in order to test multiple putative enhancers with minimal efforts (Chen, et al. 2016; Kwan et al. 2007).

Minimal promoters, such as the zebrafish *gata2a* promoter, are used in these type assays to provide a neutral background for the enhancer activity test, ensuring that the reporter gene expression is caused

by enhancer activity and not by the position effect or the genomic context in which the construct is inserted (Bessa et al. 2009; Kemmler et al. 2023).

1.4.2 Loss of Function Assays using CRISPR-Cas9

In the context of enhancer study, CRISPR-Cas9 system, derived from bacterial adaptative immune system, has been used as a loss of function assay with the capacity to target enhancer sequences (Choi et al. 2021; Gagnon et al. 2014). With high efficacy mutagenesis, this system can target almost any sequence in the genome, which makes it ideal for targeted mutagenesis in zebrafish (Albadri et al. 2017; Amorim et al. 2020; Li et al. 2016).

This revolutionary genome editing tool utilises single-guide RNAs (sgRNA) complementary to the target DNA sequence that direct Cas9 nuclease to make double-strand breaks (DSB) at this region. This breaks in the genome are repaired by a process called non-homologous end joining (NHEJ), which introduces small insertions or deletions disrupting the sequence and its function (Albadri et al. 2017; Gagnon et al. 2014). This approach not only provides insights into the role of these CREs in the regulation of gene expression but also enables the generation of stable transgenic zebrafish lines with the specific enhancer mutations facilitating the study over multiple generations (Auer et al. 2014; Li et al. 2016).

1.5 Main Objectives

The main goal of this master thesis is to validate and explore exocrine pancreas enhancers of *NR5A2/nr5a2* gene using an animal model to better understand the role of this gene in PC development. In order to achieve this goal, the project is divided into three main tasks:

1. Selection of pancreatic exocrine putative enhancers located in the human and zebrafish *NR5A2/nr5a2* genomic landscapes based on specific epigenetic marks and chromatin accessibility;
2. Perform enhancer reporter assays to validate the enhancer activity of the selected sequences;
3. Determine if the disruption of *nr5a2* putative enhancers results in PC.

These findings will deepen our understanding in how non-coding regions affect *NR5A2* transcription regulation, which is important to better understand early PC development and its genetic susceptibility.

MATERIALS AND METHODS

2.1 Zebrafish Stocks, Husbandry, Breeding, and Embryo Culture

The i3S animal facility and the project are licensed by the *Direcção Geral de Alimentação e Veterinária* (DGAV). All experimental protocols were approved by the i3S Animal Welfare and Ethics Review Body and were conducted by me, following completion of the Introductory Course in Laboratory Animal Science. This course was focused on Functions A and D of the European Directive 2010/63/EU, with specialisation in zebrafish, equivalent to FELASA Course Category B.

2.1.1 Zebrafish Husbandry

Adult zebrafish (*Danio rerio*) from AB WT and transgenic line strains were maintained in the i3S main facility with a light cycle of 14 hours (hrs) of light and 10 hrs of darkness, as well as daily monitoring of behaviour, clinical signs and feeding. Optimal water parameters were maintained at a temperature of around 27°C, pH (potential of hydrogen) of 7, and conductivity of 700 pS (picosiemens). All the tanks were connected by a recirculating water system with mechanical filters for solid residuals, biological filters for ammonia nitrogen removal, and chemical filters with activated carbon for organic substances. Additionally, there's a life support system including a recirculating pump, heaters, reverse osmosis, and UV (ultraviolet) sterilizer to eliminate harmful microorganisms.

2.1.2 Breeding and Egg Collection

The breeding method used was natural fertilization and a static tank strategy. In this method, sex identification is carried out in the late afternoon, at least 12 hrs before mating, and the fish are then moved to a breeding tank with a male:female ratio of 2:3, with a partition to separate the genders. In the next morning, the partition was removed, and the tank was tilted to create a less-deep zone over a photo of

gravel substrate. The eggs were moved to a labelled petri dish in E3 medium (NaCl, KCl, CaCl₂·2H₂O, MgCl₂·6H₂O and methylene blue; Applichem) or E3 supplemented with PTU (1-phenyl-2-thiourea, Sigma-Aldrich) and then incubated at 28°C.

Before getting into the nursery, the embryos were disinfected 10-28 hpf by bleaching with Sodium hypochlorite (Sigma-Aldrich). This technique consists in soaking the eggs in 5 containers for 5 minutes (min) each: 2 bleach solutions (0.0036% of Sodium hypochlorite in system water) and 3 tap water solutions.

2.1.3 Microinjection in 1-cell-stage Zebrafish Embryos

All microinjections were performed in 1-cell stage zebrafish embryos, which means that the eggs were collected within the first 20 min following fertilization, as described above. The glass microneedles (0.58x1.00x100mm; GB100F-10, SCIENCE PRODUCTS GmbH) used to do the injection were made by glass capillaries using Narishige PN-31 Horizontal Needle Puller (Tritech Research). Each embryo was injected with around 2-5 nanolitres (nL) of solution in the yolk or the cell using a Narishige injector (IM 300; Tritech Research) and a Leica SE6 stereomicroscope (KL300 LED, Motic SMZ-171, Leica microsystems). After, the eggs are kept in E3 medium or E3 with PTU 1x, at 28°C. The PTU is used to avoid the formation of pigmentation, which allows the selection of positive transgenic embryos (Antonucci and Hindges 2016).

2.2 DNA Extraction from Zebrafish Embryos

Zebrafish genomic DNA (gDNA) was extracted from 3 batches of 8 whole embryos 24 hpf, in a mixture of 70 µL of 5% Chelex100 sodium form resin (Sigma-Aldrich) diluted in TE buffer (Tris base, Nzytech; EDTA, Sigma-Aldrich), and 5 µL of Proteinase K (10mg/mL, ThermoFisher Scientific). Subsequently was incubated for 3 hrs at 56°C, followed by 10 min at 95°C. The gDNA was stored at -20°C and before being used was briefly centrifuged.

2.3 SNPs and Putative Enhancers Sequences Selection

In the human *NR5A2* regulatory landscape of pancreatic acinar and duct cell types, sequence 44 (containing the non-risk variant of the SNP; seq44) was previously selected for other studies in the lab taking into consideration the position of the PC risk SNP rs3790844 and the epigenetic profile of H3K27Ac, H3K4me1 and H3K4me3 ChIP-seq data (Arda et al. 2018; Petersen et al. 2010).

In the zebrafish genome, two PES were selected in the non-coding regions of the *nr5a2* genomic landscape, taking into consideration the simultaneous signal of ATAC-seq data from acinar cells and H3K27Ac ChIP-seq data from whole pancreatic tissue (unpublished data from the lab, not obtained in this thesis). The H3K4me3 HiChIP-seq from whole pancreas (unpublished data from the lab, not obtained in this thesis) was also used to predict the putative promoter region interaction with the putative pancreatic enhancers.

Both human (GRCh38/hg38) and zebrafish (GRCz10/danRer10) genomic landscapes were analysed using the UCSC Genome Browser (<https://genome.ucsc.edu/>).

2.4 Generation of Plasmids with the Selected Sequences

2.4.1 Primers Design and PCR Amplification

Primers for each PES were designed with Primer3web (version 4.1.0; <https://primer3.ut.ee/>) (Table 2.1). The Oligonucleotide Properties Calculator (Oligo Calc; version 3.27) was used to check primers self-complementary and each primer melting temperature.

PCR amplifications from zebrafish gDNA were performed using the proof-reading I-MAX II™ DNA Polymerase (iNtRON Biotechnology) in a final volume of 20 µL. The reaction was composed by 2 µL of 10X PCR respective buffer, 0.75 µL of Forward (Fw) and Reverse (Rev) primers (10ng/µL), 2 µL of dNTPs (10mM of each NTP; iNtRON Biotechnology), 13 µL of H₂O HyPure (HyClone™ Water), 0.50 µL of DNA Polymerase and 1 µL of zebrafish gDNA. In a thermocycler, first a 3 min cycle was performed at 94°C and then 35 cycles at 94°C for 30 seconds (sec), a range of temperatures (depending on melting temperature of each pair of primers, Table 2.1.) for 40 sec, 72°C for 1 min and 30 sec, and a final extension time of 10 min at 72°C. The PCR product was then confirmed in a 1% agarose gel stained with SYBR Safe (NZYtech) and the bands were excised and purified with the NZYGelpure kit (NZYTech) following the manufacturer's instructions.

Table 2.1 Primers designed to amplify the PES selected in the zebrafish *nr5a2* regulatory landscape (Region 1 and Region 2).

Sequence Name	Primer Orientation	Primer Sequence	Tm (°C)	Sequence Size
Region 1	Fw	TCACACAACTGCCCAAACTT	58	470 bp
	Rev	TGACTCCGCGCCAATTAAAA	58	
Region 2	Fw	ACCATATTGCCCAACACCAA	57.6	798 bp
	Rev	CCTTAGGGCTCATATGATTG	57	

The PCR amplification of seq44 from human gDNA was previously done in the lab using the primers described in Table 2.2. The I-MAX II™ DNA Polymerase was also used to amplify this sequence following the manufacturer's protocol instructions and the corresponding primer conditions.

Table 2.2 Primers designed to amplify the PES selected in the human *NR5A2* regulatory landscape (Seq44).

Sequence Name	Primer Orientation	Primer Sequence	Tm (°C)	Sequence Size
Seq44	Fw	GCTTGTCTCACTAGGTCAGCC	61	1697 bp
	Rev	GGCTCAGGCTCCAGTCCC	61	

2.4.2 Cloning of Fragments into pCR8™/GW/TOPO® Vector and Chemically Competent Bacteria Transformation

PCR products from Region 1 and Region 2 were cloned individually into the entry vector pCR8™/GW/TOPO® (Invitrogen, ThermoFisher Scientific) that is Gateway-adapted providing the recombination of the insert into any Gateway® destination vector (Fig. 2.1). The ligation reaction was done with 4 µL of insert, 1 µL of salt solution and 1 µL of pCR8™/GW/TOPO® vector diluted in 1/4 of dilution buffer that was then incubated at room temperature overnight (O.N).

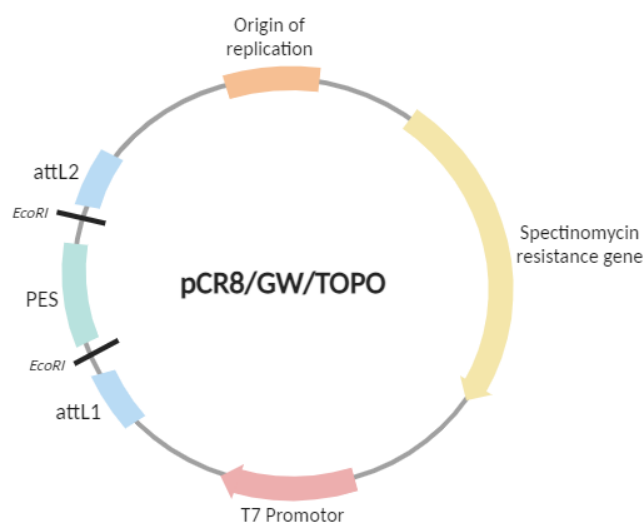


Figure 2.1 Graphical representation of pCR8™/GW/TOPO® vector (2817 bp) with the PES flanked by two attL and two Gateway® recombination sites. Those are both necessary for cloning and recombination with a destination vector, respectively. The vector also includes EcoRI enzyme restriction sites and spectinomycin resistance gene (Created with ©2024 BioRender).

Standard bacterial transformation was performed using 5 µL of the reaction in 50 µL of MultiShot™ FlexPLate Mach1 T1R competent bacteria (Invitrogen, ThermoFisher Scientific) and incubated on ice for 30 min. The cells were subjected to heat shock at 42°C for 30 sec and were incubated again on ice for only 2 min. Afterwards, 800 µL of Lysogeny Broth (LB) liquid medium was added, and competent cells were incubated at 37°C for 60 min in a shaking incubator. Transformed cells were finally plated into LB agar plates with 100 µg/ml Spectomycin (Sigma-Aldich) and grown at 37°C O.N. Then, colonies from each plate were pitched and inoculated in 3 ml of LB liquid medium with 100 µg/ml of the respective antibiotic and grown at 37°C shaking O.N. To extract plasmid DNA from each colony, an NZYMiniprep kit (NZYTech) was used following the manufacturer's instructions, and samples were then stored at -20°C.

The insertion of each sequence in the plasmid was confirmed using the EcoRI restriction enzyme (Anza, Invitrogen, ThermoFisher Scientific) in a 2 hrs reaction at 37°C: 3 µL of the miniprep product, 0.30 µL of enzyme, 2 µL of the respective buffer, and H₂O HyPure until a final volume of 20 µL. The digestion product was analysed in a 1% agarose gel to analyse the insertion of the sequence. Sanger Sequencing was used to confirm the correct nucleotide sequence, and all results were aligned and analysed using Benchling platform (<https://www.benchling.com/>).

2.4.3 Z48 based Vector Recombination of the Selected Sequences

Each PES was inserted in the pCR8™/GW/TOPO® vector and then recombined into the destination vector Z48 through the Gateway® system, in order to test the enhancer activity (Bessa et al. 2009 ; [Fig. 2.2](#)).

The Z48 vector has a *Gata2* minimal promoter that controls a GFP reporter gene. Upstream this promoter is a Z48 enhancer that drives GFP expression in the zebrafish midbrain, functioning as an internal control of transgenesis. If the sequence present in the cloning site has enhancer activity in a specific tissue, then GFP expression will also be observed in that particular tissue (Bessa et al. 2009, 2014; De La Calle-Mustienes et al. 2005; [Fig. 2.2](#)).

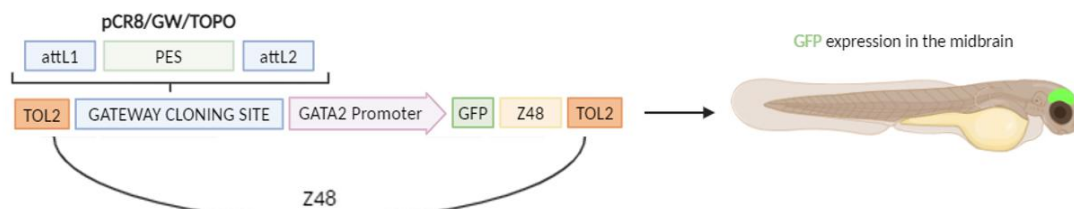


Figure 2.2 Graphic representation of the recombination of the PES between the pCR8™/GW/TOPO® and the Z48 vector (7890 bp). This happens through the attL sites within the Gateway® system localized upstream the *Gata2a* minimal promoter. Z48 enhancer drives GFP expression in zebrafish midbrain as an internal control of transgenesis. The TOL2 transposase recognition sites, also flanking the PES, make this vector a transposable element (Created with ©2024 BioRender).

The recombination reaction was performed using 0.50 µL of Gateway® LR Clonase® II Enzyme (Invitrogen, ThermoFisher Scientific), along with 1 µL of Z48 empty destination vector (50 ng/µL) and 1 µL of pCR8™/GW/TOPO® vector with the insert (50 ng/µL) in a final volume of 2.50 µL. After 30 min of incubation at 25°C in a thermocycler, 0.25 µL of Proteinase K (10mg/mL) were added, and the mixture was incubated for 10 min at 37°C.

Standard bacterial transformation was also performed as described above. Positive colony selection was done with 100 µg/ml of Ampicilin (NZYTech), which is the antibiotic resistance of the Z48 vector. Plasmids were purified using an NZYMiniprep kit (NZYTech) following the manufacturer's instructions, and the recombination was confirmed with the EcoRI restriction enzyme in a reaction similar to the one described in [section 2.4.2](#). The digestion product was analysed in a 1% agarose gel, and the

plasmid content was confirmed by Sanger sequencing using sequence-specific primers. Analysis of sequencing results was performed using Benchling (<https://www.benchling.com/>).

2.4.4 Phenol/Chloroform Purification of the Plasmid for Microinjection

A phenol/chloroform purification protocol was used to purify the final plasmids before microinjection. First, 100 μ L of Phenol:Chloroform:Isoamyl Alcohol (UltraPure™, Invitrogen, ThermoFisher Scientific) was added to the samples diluted with H₂O HyPure (to a final volume of 100 μ L). The mixture was vortexed and centrifuged for 5 min (13000 rpm) for phase separation. The upper aqueous phase was transferred to an RNase-free 1.5mL Eppendorf along with 100 μ L of chloroform. Following vortexing, the mixture was centrifuged under previously defined conditions. Then, the aqueous phase was transferred to a new RNase-free 1.50 mL Eppendorf. To precipitate the sample, 10 μ L of sodium acetate (3M) and 200 μ L of cold 100% EtOH (ethyl alcohol) were added to each 100 μ L of aqueous phase and incubated O.N. at -20°C. The mixture was then centrifuged for 15 min at 4°C (13000 rpm), and the ethanol left was removed. The pellet was diluted in 15 μ L of H₂O HyPure and stored at -20 °C. The concentration of purified plasmid DNA was determined using a spectrophotometer (NanoDrop One, ThermoFisher Scientific).

2.5 Transgenesis

2.5.1 Microinjection of Plasmids and mRNA into Zebrafish Embryos

The solution injected in this assay was composed by Z48 vector carrying the sequence of interest (50 ng/ μ L) along with the TOL2 mRNA transposase (50 ng/ μ L) and 0.05% phenol-red (Sigma-Aldrich) to obtain a final concentration of 25 ng/nL. These mixtures were microinjected into one-cell stage embryos according to the conditions described above. Then, the internal control of transgenesis, GFP expression in the midbrain, allowed selection of microinjected embryos at 48 hpf. To establish stable transgenic zebrafish lines, those embryos were then disinfected by bleaching (as described in [section 2.1.2.](#)) and grew until adulthood (~ 3 months).

2.6 Screening for Enhancer Activity

Adult fish injected with the human and zebrafish sequences were outcrossed with the transgenic line *tg(ela:mCherry)*, which has an *ela* (*elastase*) promoter that drives the expression of the mCherry marker in acinar cells, to assess the enhancer activity of each sequence in the exocrine pancreas (Bordeira-Carriço et al. 2022).

The offspring (F1) was screened for the internal control of transgenesis at 48 hpf, GFP expression in the midbrain, meaning that germline transmission occurred. When the exocrine pancreas is fully developed in zebrafish larvae, at 11-12 dpf, they were screened for mCherry expression in acinar cells, labelling the exocrine pancreas and a possible co-localization with GFP expression in the case of an acinar enhancer sequence (Wilson and Matsudaira 2004).

Later on, the fish injected with the human sequence (seq44) were also outcrossed with another transgenic line: *tg(sst:mCherry)*, which has a *sst* (*somatostatin*) promoter that drives the expression of the mCherry marker in δ -cells, to assess endocrine pancreas enhancer activity (Eufrásio et al. 2020). The offspring, in this case, was only screened at 48 hpf once the endocrine domain is already formed at this timepoint in the zebrafish (Tiso et al. 2009). In order to analyse the enhancer activity of this sequence in pancreatic progenitor cells, mouse anti-Nkx6.1 antibody (F55A10; Hybridoma bank) was used in the immunohistochemistry as a molecular marker for these cells (Binot et al. 2010).

As a positive control of these enhancer reporter assays, we used a stable zebrafish line that contains a distal *ptf1a* enhancer (zP.E3) that shows a strong GFP expression in progenitor cells and differentiated acinar cells (Bordeira-Carriço et al. 2022). The transgenic line *tg(zP.E3:EGFP, ela:mCherry)* was used to assess acinar cell enhancer activity at 11-12dpf. The transgenic lines *tg(zP.E3:EGFP)* and *tg(sst:mCherry)* were also crossed, the offspring was screened at 48 hpf for GFP expression, and they passed through immunohistochemistry with anti-Nkx6.1 to mark the progenitor cells (Binot et al. 2010).

2.6.1 Fixation and DAPI Staining (11-12dpf)

F0 carrying each PES were crossed with the line *tg(ela:mCherry)* to assess enhancer activity in pancreatic acinar cells (Bordeira-Carriço et al. 2022). To observe the cell nucleus ahead in fluorescence microscopy, DAPI staining was performed.

At 11-12 dpf larvae were anaesthetised by adding MS-222 (Tricaine; ethyl 3-aminobenzoate methane-sulphonate salt, Sigma-Aldrich) into the E3 medium for selection into an RNase-free 1.5 mL Eppendorf, via mCherry expression in the pancreas. Afterwards, E3 medium was discarded, and larvae were

fixed in 600 μ L of Formaldehyde solution 4% (Sigma-Aldrich) in PBS-1x (Phosphate-buffered Saline; Gibco™, ThermoFisher Scientific) for 1 hr. Then, formaldehyde was wasted, 500 μ L of PBS-1x were added and larvae were stored at 4°C until staining. This step was followed by a wash with 500 μ L of 0.1% Triton X-100 (Sigma-Aldrich) in PBS-1x for 5 min. Permeabilization was performed with 500 μ L of PBS-T (0.5% Triton X-100 in PBS-1x) for at least 30 min, followed by another 5 min wash with 500 μ L of PBS-T 0.1%. Then, the mixture was incubated O.N. with PureBlu™ DAPI nuclear staining dye (1:1000; Bio-Rad) diluted in PBS-T 0.1% at 4°C shaking. The next day, larvae were washed six times with 600 μ L of PBS-T 0.1% for 10 min (last wash of 30 min) and 500 μ L of 50% glycerol (VWR Chemicals) in PBS 1x were added. All incubation times were performed at room temperature with shaking (except the ones that were specified). This step was also performed for the positive control: tg(zP.E3:EGFP, *ela*:mCherry) at 11-12 dpf.

2.6.2 Fixation and Immunohistochemistry (48 hpf)

F0 adult fish injected with the Z48 vector carrying the human seq44 were crossed with the transgenic line tg(*sst*:mCherry) to evaluate the enhancer activity of these sequences in endocrine pancreatic cells. Immunohistochemistry was performed using mouse anti-Nkx6.1 antibody (F55A10; Hybridoma bank) to assess enhancer activity in progenitor cells and with DAPI for nuclear staining.

At 48 hpf, embryos were anaesthetized (as described in [section 2.6.1.](#)), selected for GFP expression in the midbrain, and dechorionated. After being fixed O.N. in 600 μ L of Formaldehyde 4% (in PBS-1x) at 4°C shaking, embryos were incubated in 500 μ L of PBS-1x at 4°C until the immunostaining was performed. Fixed embryos were washed in 500 μ L of PBS-T 0.1% for 5 min, followed by permeabilization for at least 2 hrs in 500 μ L of PBS-T (1% Triton in PBS-1x). Another 5 min wash was done with 500 μ L of PBS-T 0.1% and blockage was performed with 200 μ L of 5% BSA (bovine serum albumin, NZYTech) in PBS-T 0.1% for 1 hr. The embryos were incubated O.N. at 4 °C along with 200 μ L of mouse anti-Nkx6.1 (F55A10; Hybridoma bank) primary antibody (1:75) diluted in BSA in PBS-T (5%). Subsequently, six 10 min washes were performed with 500 μ L of PBS-T 0.1% each, and the embryos were O.N. incubated again, but this time with the goat anti-mouse AlexaFlour 647 (1:1800, Invitrogen, ThermoFisher Scientific) secondary antibody along with DAPI (1:1000, Bio-Rad), both diluted in PBS-T 0.1%. To finish, six washes of 10 min each with 600 μ L of PBS-T 0.1% (last wash of 30 min) were performed, and the embryos were imbibed in 500 μ L of Glycerol 50% (VWR Chemicals) in PBS 1x. All incubation times were performed at room temperature with shaking (except the ones that are specified). This step was also performed for the positive control: tg(zP.E3:EGFP) crossed with tg(*sst*:mCherry) at 48 hpf.

2.6.3 Leica SP5 Confocal Observation

Either the yolk or the eyes were removed from each embryo/larvae for confocal imaging, depending on the developmental stage, 48 hpf or 11-12 dpf, respectively.

Slides for microscopy were prepared using 50% glycerol in PBS 1x. Leica TCS SP5II confocal microscope (Leica microsystems) along with the Leica Application Suite Advanced Fluorescence software (Leica LAS AF software, version 2.6.3.8173) were used to capture the confocal images, which were then processed using ImageJ software (version 1.54i).

2.7 Mutagenesis using CRISPR-Cas9 System

2.7.1 Cas9 Target Design and sgRNA Synthesis

Single-guide RNAs (sgRNAs) targeting the chosen Region 1 and Region 2, in the zebrafish *nr5a2* genomic landscape were predicted and selected using the Crisprscan programme (<https://www.crisprscan.org/>; Moreno-Mateos et al. 2015; Amorim et al. 2020; [Table 2.3](#)).

To anneal the oligonucleotides of each sgRNA, a mix of the following reagents was heated at 95°C for 5 min: 1.5 µL of each oligo 1 and 2 (100 µM, Sigma-Aldrich), 24 µL of the T4 ligase annealing buffer (ThermoFisher Scientific), and H₂O HyPure to make a final volume of 50 µL. The vector used to transcribe the sgRNA was pDR274 (Addgene; Bordeira-Carriço et al. 2022; [Fig. 2.3](#)). To clone the annealed oligonucleotides first the vector was linearized by BsaI digestion (Anza™ 36 Eco31, Invitrogen, ThermoFisher Scientific) in a mix of 2 µL of restriction enzyme (10 ng/µL), 3 µL of pDR274, 2 µL of the corresponding restriction buffer (10x) and H₂O HyPure up to 20 µL of final volume.

Table 2.3 Oligonucleotides designed for each sgRNA upstream and downstream targeting the PES in the zebrafish genome of *nr5a2* gene.

	sgRNA	Oligonucleotide Orientation	Sequence (5'-3')	TM (°C)
Region 1	sgRNA 1	oligo 1	TAGGTCACGGGGGCTGAGCGCA	69.5
		oligo 2	AAACTGCGCTCAGCCCCCGTGA	67.9
	sgRNA 2	oligo 1	TAGGGCAGTCCGTGGAGCCGCT	69.5
		oligo 2	AAACACCGGCTCCACGGACTGC	67.9
	sgRNA 3	oligo 1	TAGGCTCACGGGGGCTGAGCGC	71.6
		oligo 2	AAACGCGCTCAGCCCCCGTGAG	69.5
	sgRNA 4	oligo 1	TAGGCAGTCCGTGGAGCCGCT	67.3
		oligo 2	AAACACCGGCTCCACGGACTG	65.3
	sgRNA 5	oligo 1	TAGGCACTTTTATAGATGTCCA	56.4
		oligo 2	AAACTGGACATCTATAAAAGTG	54.7
	sgRNA 6	oligo 1	TAGGGCACTGTCCCGGGACATT	65.8
		oligo 2	AAACAATGTCCCGGGACAGTGC	64.2
	sgRNA 7	oligo 1	TAGGTTTCTTTCTGCGGTGACC	62.1
		oligo 2	AAACGGTCACCGCAGAAAGAAA	60.1
	sgRNA 8	oligo 1	TAGGATTGATCAGGGAGCCGCT	64.2
		oligo 2	AAACAGCGGCTCCCTGATCAAT	62.1
Region 2	sgRNA 1	oligo 1	TAGGTCCCGGCCGAATCCCCAA	67.9
		oligo 2	AAACTTGGGGATTTCGGCCGGA	65.8
	sgRNA 2	oligo 1	TAGGTTCTAGATCACAAGTGA	58.4
		oligo 2	AAACCTCAGTTGTGATCTAGAA	56.4
	sgRNA 3	oligo 1	TAGGAGACAGGCTTGTAGAGGC	64.2
		oligo 2	AAACGCCTCTACAAGCCTGTCT	62.1
	sgRNA 4	oligo 1	TAGGCAGGGGTATGAAGCAAT	59.5
		oligo 2	AAACATTGCTTCATACCCCTG	57.5

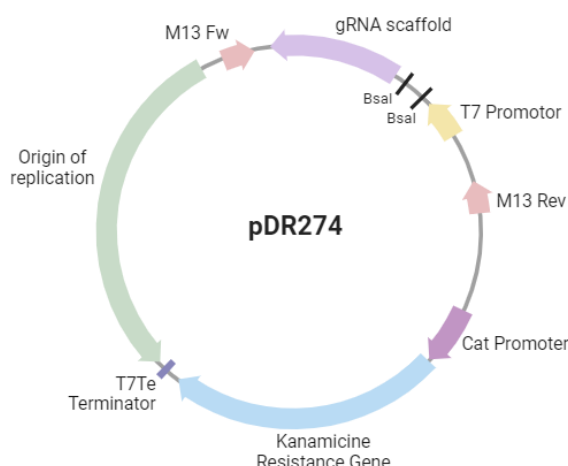


Figure 2.3 Graphical representation of the pDR274 vector (2147 bp) with the scaffold where the sgRNAs were inserted. The vector has two sites for the BsaI restriction enzyme used to linearize the plasmid before the sgRNA cloning. The primers M13 Forward and M13 Reverse were used to confirm the cloning (Created with ©2024 BioRender).

The vector linearization was confirmed in a 1% agarose gel stained with SYBR Safe (NZYtech) along with NZYDNA Loading Dye (NZYtech). Using the NZY GelPure gel extraction kit (NZYTech) following the manufacturer's instructions, the linearized plasmid DNA was extracted for the gel bands.

Next, to join the annealed oligos to the linearised plasmid, a 10 μ L ligation mix was prepared with: 2.50 μ L of linearised vector, 1 μ L of T4 ligase annealing buffer (ThermoFisher Scientific), 0.50 μ L of T4 ligase enzyme (ThermoFisher Scientific), and 1.50 μ L of annealed oligos (100 ng/ μ L). Moreover, incubated at room temperature for 3 hr.

MultiShot™ FlexPLate Mach1 T1R competent bacteria (Invitrogen, ThermoFisher Scientific) were used to do a transformation and insert each sgRNA oligo into the pDR274 plasmid. Vector selection was performed using 50 μ g/ml kanamycin A (NZYTech) in the LB culture medium. DNA from each colony was purified using an NZYMiniprep kit (NZYTech) following the manufacturer's instructions and was confirmed by Sanger sequencing using sequence-specific primers (Table 2.1.).

To select bacterial colonies containing the plasmid with the cloned sgRNA oligos, colony PCR was performed using 0.25 μ L of M13 Forward and sgRNA oligo1 as primers, 5 μ L of NZYTaq II 2 $\times$ Green Master Mix (NZYtech), and H₂O HyPure up to 10 μ L of final volume. Finally, the following amplification protocol was used: 95°C for 3 min, 94°C for 20 sec, 30 sec at melting temperature, 72°C for 20 sec, and finally 5 min at 72°C. The PCR products were loaded onto a 2% agarose gel with SYBR Safe (NZYtech) to confirm the insertion of each sgRNA into the plasmid.

At least two positive colonies from each were inoculated in 3 mL of LB medium along with kanamycin A (50 µg/mL) and incubated at 37°C O.N shaking. The DNA of each colony was extracted using an NZY Miniprep kit (NZYTech), and the cloning was confirmed using Sanger sequencing. The plasmid containing each sgRNA oligo was linearised using the restriction enzyme HindIII (Anza™ 16, Invitrogen, ThermoFisher Scientific), and the product was purified by Phenol/Chloroform (as described in [section 2.4.4.](#)).

To synthesise the sgRNAs, an *in vitro* RNA transcription reaction was performed in a mix of 15 µL of H₂O HyPure, 10 µL of Transcription buffer 5x (ThermoFisher Scientific), 5 µL of DTT (dithiothreitol, 50mM, Invitrogen, ThermoFisher Scientific) and 5 µL of NTP mix (10 mM of each: A, U, C, G; New England Biolabs), incubated for 5 min at 37°C. Then 12 µL of the DNA was added, incubated at 37°C for 1 min, 1 µL of RNase inhibitor (NZYTech) was again incubated at 37°C for 1 min, and then 2 µL of T7 RNA polymerase (20 units/µL, ThermoFisher Scientific) were added, and the mix was incubated for 1 hr. Next, 1 more µL of polymerase was added, and the mix was incubated for another hr. Before 1 µL of DNase I (ThermoFisher Scientific) and finally incubated for 30 min at 37°C. Then, the RNA product was Phenol/Chloroform purified again and stored at -80°C.

2.7.2 Microinjection of sgRNAs into Zebrafish Embryos

The solution injected for the loss of function assays was composed by the sgRNAs pair (100 ng/µL each) designed for the two selected zebrafish sequences (Region 1 and Region 2) along with cas9 mRNA (600 ng/µL; PNA Bio, Inc) and 0.05% phenol-red (Sigma-Aldrich). As negative controls, a batch of embryos were injected with Cas9 protein only (600 ng/µL) to control any possible toxicity and another batch of not-injected embryos was used to control embryo viability.

To evaluate the efficacy of the sgRNAs deleting each sequence, each combination of the sgRNAs designed was tested by microinjection in one-cell stage embryos, followed by DNA extraction ([section 2.2.](#)) and conforming by PCR ([section 2.7.3.](#)).

To establish a stable transgenic line, the effective sgRNAs pairs were injected in embryos from tg(zP.E3:EGFP, *ela*:mCherry). This transgenic line express mCherry in pancreatic acinar cells, labelling the exocrine domain.

2.7.3 Genotyping

In order to confirm the genomic deletions a PCR amplification was done with DNA extracted from the injected embryos (as described in 2.2.) and primers flanking the Region 1 and Region 2 (Table 2.1.).

The following components were present in the solution: 1 µL of DNA, 0.25 µL of each Fw and Rev primer, 5 µL of NZYTaQ II 2× Green Master Mix (NZYtech) and H₂O HyPure to a final volume of 10 µL. PCR conditions were as follows: 3 min denaturation step at 95°C, 30 cycles of 30 sec at 94°C, 30 sec at 55°C-60°C (primer specific melting temperature) and 1 min at 72°C, followed by a 5 min final step at 72°C for extension. Next, the PCR amplification was confirmed in a 2% agarose gel stained with SYBR Safe (NZYtech) by electrophoresis. NZYGelpure kit (NYZtech) was used to remove the DNA from the desired bands and purify them.

To confirm the presence of the genomic deletions caused by the designed sgRNA pairs, the genotyping PCR bands were all confirmed by Sanger Sequencing using specific primers and the results were analysed in Benchling (<https://www.benchling.com/>).

2.8 Predicting Putative TFBS – JASPAR Analysis

The JASPAR software (<https://jaspar.elixir.no/>) utilizes a set of frequently updated annotated position weight matrices to predict putative TF binding sites (TFBS) in DNA sequences. This software is based on the homology with the DNA-binding domains of vertebrate TFs that have already been characterized (Rauluseviciute et al. 2024). These predictions are based on a relative score, a more accurate metric than the total score, since it is calculated using the formula: $(\text{score} - \text{min_score}) / (\text{max_score} - \text{min_score})$. The relative score ranges between 0 and 1, providing a threshold to evaluate the likelihood of a TF binding to a specific site. Sequences with a higher relative score indicate stronger potential for TF binding (Sandelin et al. 2004).

RESULTS AND DISCUSSION

3.1 Prediction of CREs

3.1.1 Human *NR5A2* Regulatory Landscape Analysis and Selection of Putative Enhancer Sequences Overlapping with SNPs

In previous studies conducted in the lab (unpublished data), a PES located in the human *NR5A2* regulatory landscape, denominated seq44, was selected as a putative human pancreatic enhancer. This region overlaps with a PC SNP (rs3790844) and with the enriched signal of H3K27ac (active enhancers) and H3K4me1 (active enhancers) ChIP-seq data from human pancreatic acinar and duct cells, the most common source of pancreatic adenocarcinomas (Backx et al. 2022; Petersen et al. 2010; [Fig. 3.1](#)).

Luciferase reporter assays were previously performed in the lab in a human pancreatic duct cell line (hTERT-HPNE) to test the enhancer activity of this sequence. The results showed significant luciferase activity in this seq44 compared with the control sequence, demonstrating that this sequence is a pancreatic duct enhancer in this pancreatic cell line (unpublished data, and not performed in this thesis). Therefore, our objective is to investigate whether this sequence can also have enhancer activity in pancreatic acinar cells. However, human pancreatic acinar cell lines are scarce. Therefore, we decided to test the acinar enhancer activity of human seq44 using zebrafish as animal model, which offers many advantages, such as the conservation of pancreatic anatomy, physiology and molecular pathways that allow us to better understand transcription regulation and pancreatic tumorigenesis (Bordeira-Carriço et al. 2022; Kimmel and Meyer 2010; Kwiatkowska et al. 2022; Mi et al. 2024).

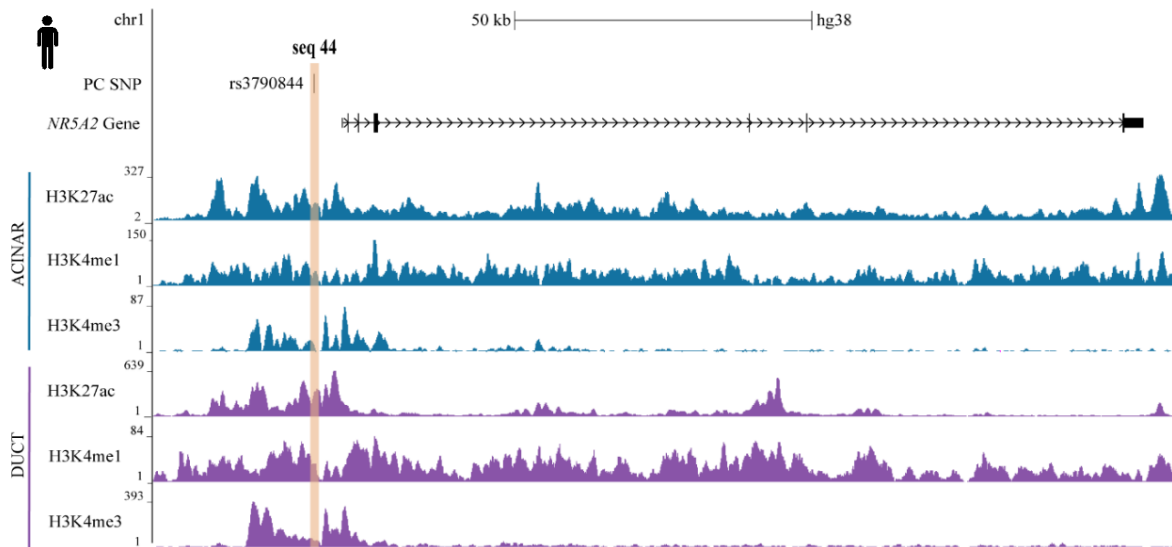


Figure 3.1 The human *NR5A2* genomic landscape (hg38) containing a PES that overlaps with the PC SNP and epigenetic marks for enhancer activity. The epigenetic profile of HK27Ac (active enhancers), H3K4me1 (active enhancers) and H3K4me3 (active promoters) ChIP-seq data from pancreatic acinar (dark blue) and duct cells (purple) can be observed in the regulatory landscape. In orange are labelled the PES selected to test for enhancer activity (seq44) and the PC risk SNP that overlap this sequence is rs3790844 (Unpublished data from the lab).

3.1.2 Zebrafish *nr5a2* Regulatory Landscape Analysis and Selection of Putative Enhancer Sequences

In order to find a functional equivalent pancreatic enhancer of human seq44 in the zebrafish *nr5a2* genomic landscape, we combined information from accessible chromatin, ATAC-seq datasets from zebrafish pancreatic acinar cell lines and whole pancreas, with H3K27ac (active enhancers) ChIP-seq data from zebrafish whole pancreas (Unpublished data from the lab). From this analysis, we selected two PES, Region 1 and Region 2, based on the overlap with ATAC-seq signal enrichment with the epigenetic mark for enhancer activity (Fig. 3.2).

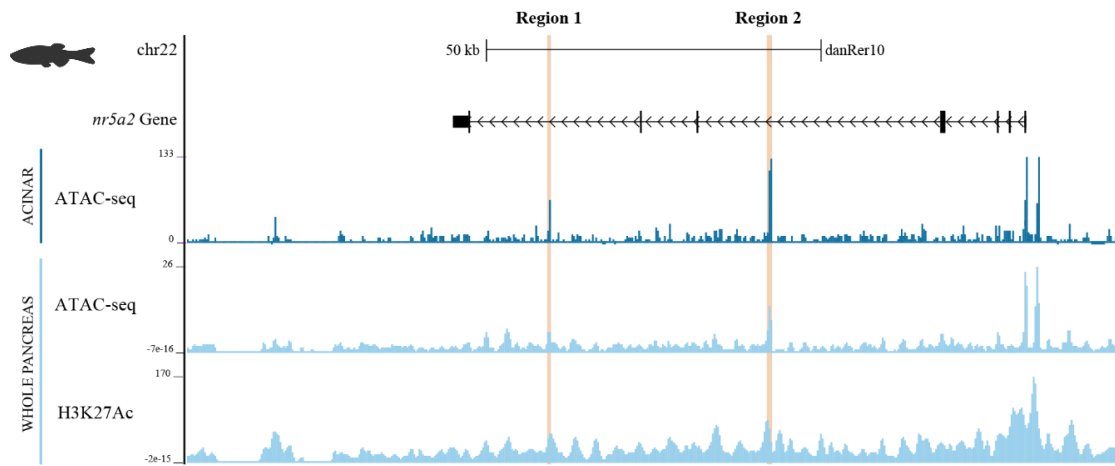


Figure 3.2 The zebrafish *nr5a2* genomic landscape (danRer10) containing two PES that overlap with regions of accessible chromatin and epigenetic marks for enhancer activity. ATAC-seq data shows elevated signal in both zebrafish pancreatic acinar cells (dark blue) and whole pancreas (light blue), indicating that the chromatin is putatively accessible in those regions. H3K27ac ChIP-seq data indicative of active enhancers is also represented in light blue. In orange are labelled the two selected regions to test for enhancer activity (Region 1 and Region 2) (Unpublished data from the lab).

3.2 PCR Amplification of the Putative Enhancer Sequences

To amplify PES, a PCR reaction was conducted using genomic DNA (gDNA) extracted from zebrafish embryos, specific primers for Region 1 (470 bp) and Region 2 (798 bp) (Table 2.1), and a high-fidelity DNA polymerase, as detailed in section 2.4.1. The PCR products were confirmed on a 1% agarose gel (Fig.3.3).

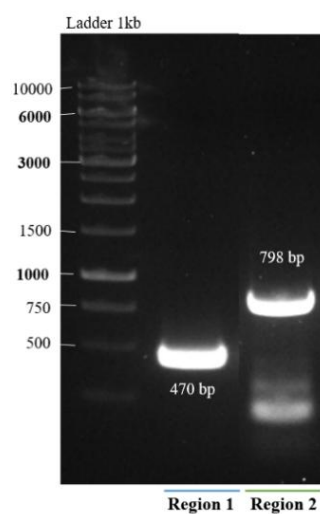


Figure 3.3 PCR amplification products of Region 1 (470 bp) and Region 2 (798 bp) from zebrafish gDNA. Ladder Gene Ruler 1kb (ThermoFisher Scientific).

3.3 Cloning of the PCR Products into pCR8TM/GW/TOPO[®] Vector

Cloning of each PES into the pCR8TM/GW/TOPO[®] vector was performed using chemically competent bacteria and following the protocols outlined in [section 2.4.2](#). After DNA purification from some selected colonies, the positive colonies, containing the PES, was confirmed by a digestion with the EcoRI restriction enzyme, and the products were run on a 1% agarose gel. EcoRI restriction enzyme has only two restriction sites within the pCR8TM/GW/TOPO[®] vector that flank the cloning site, meaning that only two bands were expected, in case of a positive colony, the vector backbone (2817 bp) and a second band corresponding to the PES ([Fig. 3.4](#)).

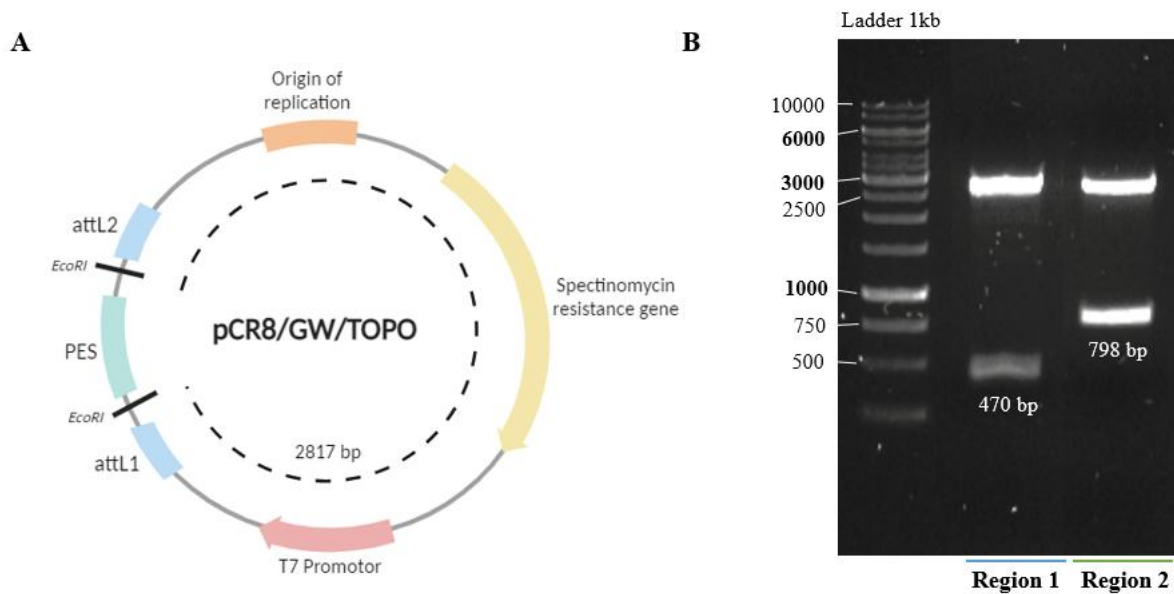


Figure 3.4 Analysis of pCR8TM/GW/TOPO[®] vector digestion by EcoRI enzyme. A- Graphical representation of pCR8TM/GW/TOPO[®] vector indicating the EcoRI restriction sites (black lines) (Created with ©2024 BioRender). B- Gel electrophoresis showing the EcoRI digestion product of pCR8TM/GW/TOPO[®] vector containing the cloned PES. The upper band corresponds to the vector backbone (2617 bp) and the lowest band corresponds to the PES. (Region 1- 470 bp; Region 2- 798 bp). Ladder Gene Ruler 1kb (ThermoFisher Scientific).

3.4 Recombination of the Putative Enhancer Sequences into Z48 vector for Enhancer Reporter Assays

Each PES was cloned into the Z48 vector via the Gateway recombination system from the pCR8TM/GW/TOPO[®] vector. To determine whether sequence insertion within the destination vector correctly occurred, a digestion reaction was performed using the EcoRI restriction enzyme. The digested product was confirmed using a 1% agarose gel. The enzyme used has three restriction binding sites

within the Z48 vector, which means that in case of a successfully recombination three bands are expected: two from the plasmid backbone with 4279 bp and 1835 bp and another with the PES (Fig.3.5).

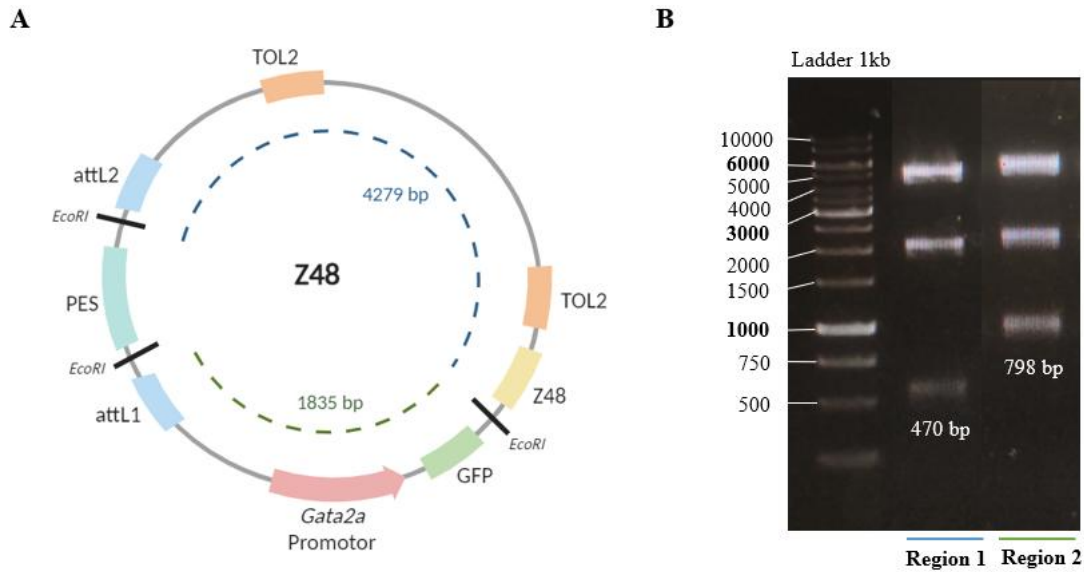


Figure 3.5 Analysis of Z48 vector digestion by EcoRI enzyme. **A-** Graphical representation of Z48 vector indicating the EcoRI restriction sites (black lines) (Created with ©2024 BioRender). **B-** Gel electrophoresis showing the EcoRI enzyme digestion product of the Z48 vector (7890 bp) containing the cloned PES. In each lane, three bands can be visualized. Two of them correspond to the digested vector backbone with approximate sizes of 4279 bp and 1835 bp. The third lowest band corresponds to the size of the PES. (Region 1- 470 bp; Region 2- 798 bp) Ladder Gene Ruler 1kb (ThermoFisher Scientific).

3.5 *In vivo* Transgenesis Reporter Assays

3.5.1 Assessment of Pancreatic Domains

To perform proper enhancer reporter assays, it is essential to include a validated positive control sequence to confirm accurate enhancer signal and also define the proper domains for analysis. With this purpose, a transgenic line tg(zP.E3:EGFP), containing a strong distal enhancer of the *ptf1a* gene (zP.E3), was used as a positive control. This line was selected because the zP.E3 enhancer drives expression in both progenitor and exocrine pancreatic cells, specifically in differentiated acinar cells (Bordeira-Carriço et al. 2022).

Since one of the main objectives of this thesis is to test the enhancer activity *in vivo* of zebrafish PES specifically in pancreatic acinar cells, it was important to use a marker of these type of cells to outline the exocrine domain of the zebrafish pancreas. In order to proper localise the zebrafish pancreatic acinar cells, a transgenic line tg(*ela*:mCherry) was also used. This zebrafish reporter line drives mCherry expression in acinar cells, the producers of elastase (Bordeira-Carriço et al. 2022).

Therefore, the transgenic line *tg(zP.E3:EGFP, ela:mCherry)* was crossed and the progeny grew until 11-12 dpf, timepoint where the zebrafish pancreas is already fully differentiated (Van De Venter et al. 2021). As described in the bibliography the distal enhancer showed strong GFP expression in differentiated pancreatic acinar cells co-localising with mCherry marker (n=16), confirming that this sequence is a good positive control for enhancer reporter assays in pancreatic acinar cells (Bordeira-Carriço et al. 2022; Fig. 3.6).

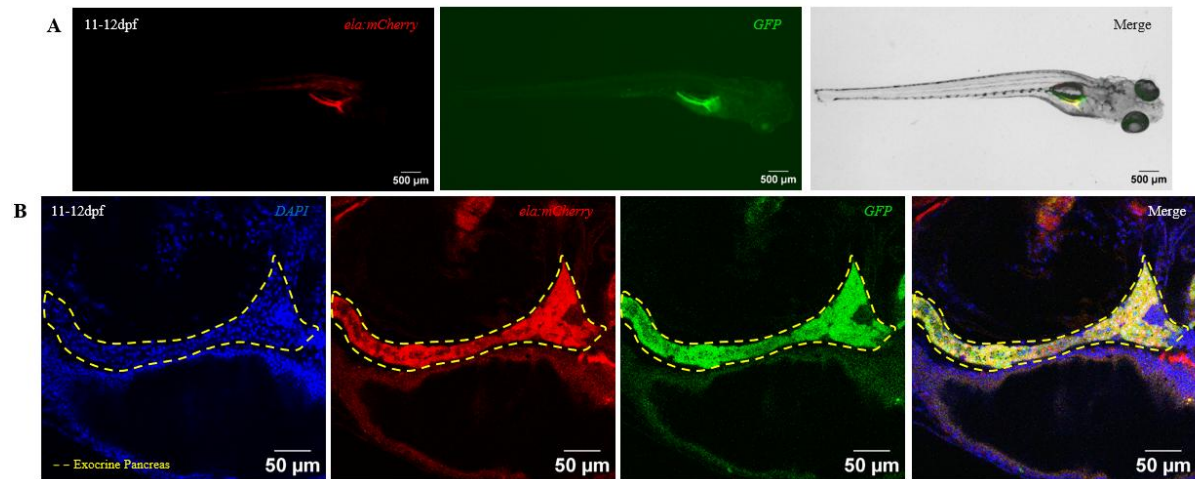


Figure 3.6 Representative images of zebrafish *tg(zP.E3:EGFP, ela:mCherry)* larvae at 11-12 dpf, tested to be used as positive control in enhancer reporter assays and to define the pancreatic acinar domain. **A-** Brightfield and epifluorescence representative full-body image of zebrafish larvae, side oriented, anterior right, dorsal top, showing GFP (green) expression in the exocrine pancreas. Images were captured with Leica M205 Fluorescence microscope; Zoom 3,00 x. Scale bar: 500µm. **B-** Representative confocal image of the zebrafish larvae (n=16) showing expression of GFP (green) in the pancreatic exocrine domain (yellow dashed line). Elastase promoter is driving mCherry expression in acinar cells (red) and the nuclei are labelled in blue. Images were captured with Leica confocal SP5II; Magnification 40x. Scale bar: 50µm. Abbreviations: *ela*, elastase.

Although the main objective is to evaluate the enhancer activity of PES in acinar cells, we also examined the pancreatic progenitor and endocrine domains to investigate whether these sequences can act as enhancers in other pancreatic developmental time and in other pancreatic cell types. Thus, to define the pancreatic endocrine domain, the zebrafish reporter line *tg(sst:mCherry)* was used since the somatostatin (*sst*) promoter drives mCherry expression in δ -cells, one of the pancreatic cell types present in the endocrine pancreas. The pancreatic progenitor's domain, on the other hand, was accessed by immunohistochemistry using an anti-Nkx6.1 antibody. The embryos were analysed at 48 hpf, the timepoint where endocrine and progenitor cells are already formed (Fig. 3.7). Then, the *tg(zP.E3:EGFP)* line was crossed

with tg(*ss:t:mCherry*) line, followed by immunohistochemistry using an anti-Nkx6.1 antibody to determine whether the positive control exhibits enhancer activity in progenitor and endocrine domains.

As outlined in the literature, embryos at 48hpf, showed co-localisation of GFP expression with Nkx6.1 staining in the progenitor domain but not in endocrine domain (n=11), confirming that this sequence is a good positive control for enhancer reporter assays for pancreatic progenitor cells but not for endocrine cells (Fig. 3.7).

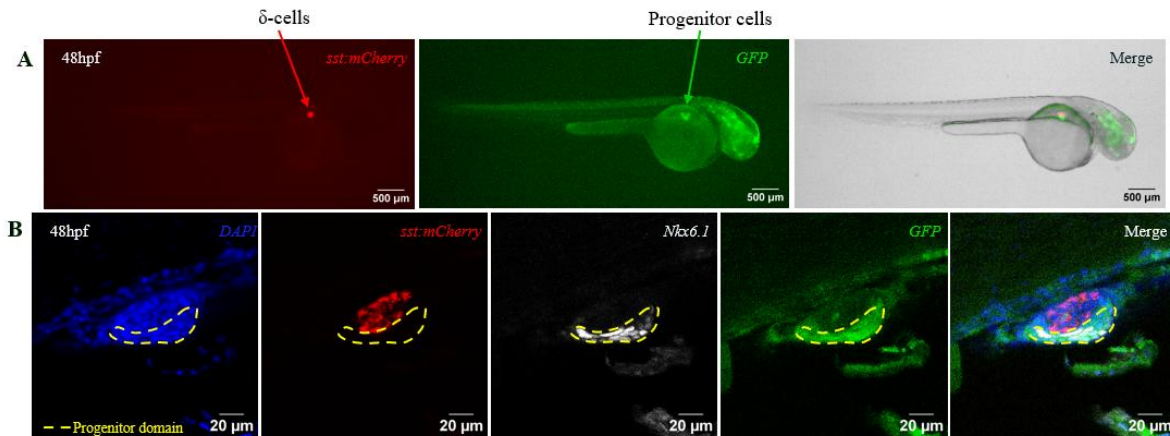


Figure 3.7 Representative images of the zebrafish tg(zP.E3:EGFP, *ss:t:mCherry*) embryos 48hpf tested to be used as positive control in enhancer reporter assays and to define the pancreatic progenitor and endocrine domain. **A-** Representative full-body brightfield and epifluorescence image of a side-oriented, anterior right, dorsal top zebrafish embryo demonstrating GFP expression (green arrow) in the pancreatic progenitor domain and mCherry expression in endocrine domain. Images were captured with Leica M205 Fluorescence microscope; Zoom 3,00 x. Scale bar: 500μm. **B-** Representative confocal image of the zebrafish embryo (n=11) showing expression of GFP (green) in the progenitor domain labelled with Nkx6.1 (white). The yellow dashed line delineates the progenitor domain, while the nuclei are stained with DAPI (blue) Images were captured with Leica confocal SP5II; Zoom 2,10 x; Magnification 40x. Scale bar: 20μm. Abbreviations: ela, elastase; sst, somatostatin.

3.5.2 Test Enhancer Activity of a Human NR5A2 Putative Enhancer Sequence in Pancreatic Domains

To test the enhancer activity of the human sequence that overlaps with the PC risk SNP rs3790844, the Z48 vector containing seq44 was injected into one-cell stage embryos along with TOL2 mRNA. At 48 hpf, the zebrafish embryos injected were selected by the internal control, the GFP expression in the midbrain. When the twenty-five fish reached sexual maturity (~3 months), they were outcrossed with the tg(*ela:mCherry*) line to access pancreatic enhancer activity specifically in acinar cells (Bordeira-Carriço et al. 2022).

The offspring (F1) was also screened at 48 hpf for the internal control of transgenesis to confirm transgenesis and germline transmission of seq44 (Fig. 3.8), revealing that only eight out of the twenty-five (8/25) injected fish were classified as founders. Over time, only four out of the eight founders survived throughout the screening period when maintained in individual breeding tanks.

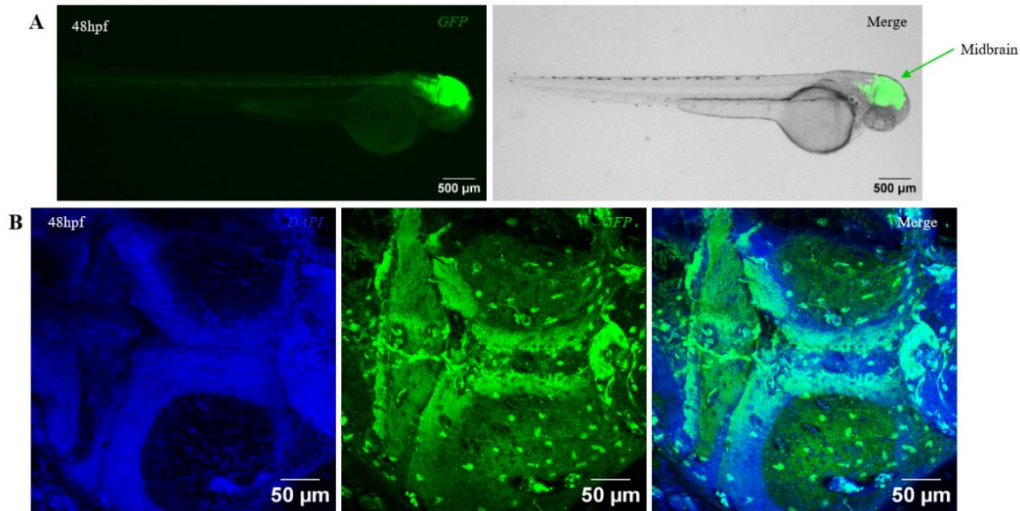


Figure 3.8 Representative images of a F1 zebrafish embryo midbrain at 48 hpf. **A-** Zebrafish embryo full-body brightfield and epifluorescence representative image, side-oriented, anterior right, dorsal top, showing Z48-mediated GFP expression (green arrow) in the midbrain as a positive control of transgenesis and germline transmission of seq44. Images were captured with Leica M205 Fluorescence microscope; Zoom 3,00 x. Scale bar: 500μm. **B-** Representative confocal image of the zebrafish embryo showing GFP expression (green) with more detail in the midbrain. The nuclei are stained with DAPI (blue). Images were captured with Leica confocal SP5II; Zoom 2,10 x; Magnification 40x. Scale bar: 50μm.

Founders I (n=8) and II (n=12) did not drive GFP expression in the exocrine pancreas of the analysed larvae (Fig.3.9 - I and II). Founder III (n=5) did not exhibit GFP expression in the exocrine domain. However, Founder III unexpectedly showed GFP expression in some pancreatic endocrine cells of 40% (2/5) of the larvae, even without a marker of the pancreatic endocrine domain (Fig.3.9). For Founder IV (n=6), in contrast, 33% (2/6) of the analysed larvae appeared to express GFP in acinar cells (Fig.3.9 - IV).

These results suggest that different genomic integration sites may have influenced enhancer activity, leading to distinct expression patterns in different pancreatic tissues (Cho 2012; Ishibashi et al. 2013). Since *nr5a2* gene is known to be broadly expressed in pancreatic acinar cells, the same injected sequence was expected to drive constant expression in the exocrine pancreas across all progeny of the founders (Bordeira-Carriço et al. 2022; Gou et al. 2020; Larsen and Grapin-Botton 2017; Sandhu et al. 2021). However, we observed a variable and mild GFP expression in the zebrafish pancreas, indicating multiple insertion events due to the use of TOL2 transposase for transgenesis (Ishibashi et al. 2013). Additionally,

we unexpectedly observed GFP expression in the endocrine domain, indicating that this sequence could have some activity in the pancreatic domain, probably in a different timepoint of pancreatic development (Bordeira-Carriço et al. 2022; Galang et al. 2020; Uttley et al. 2023).

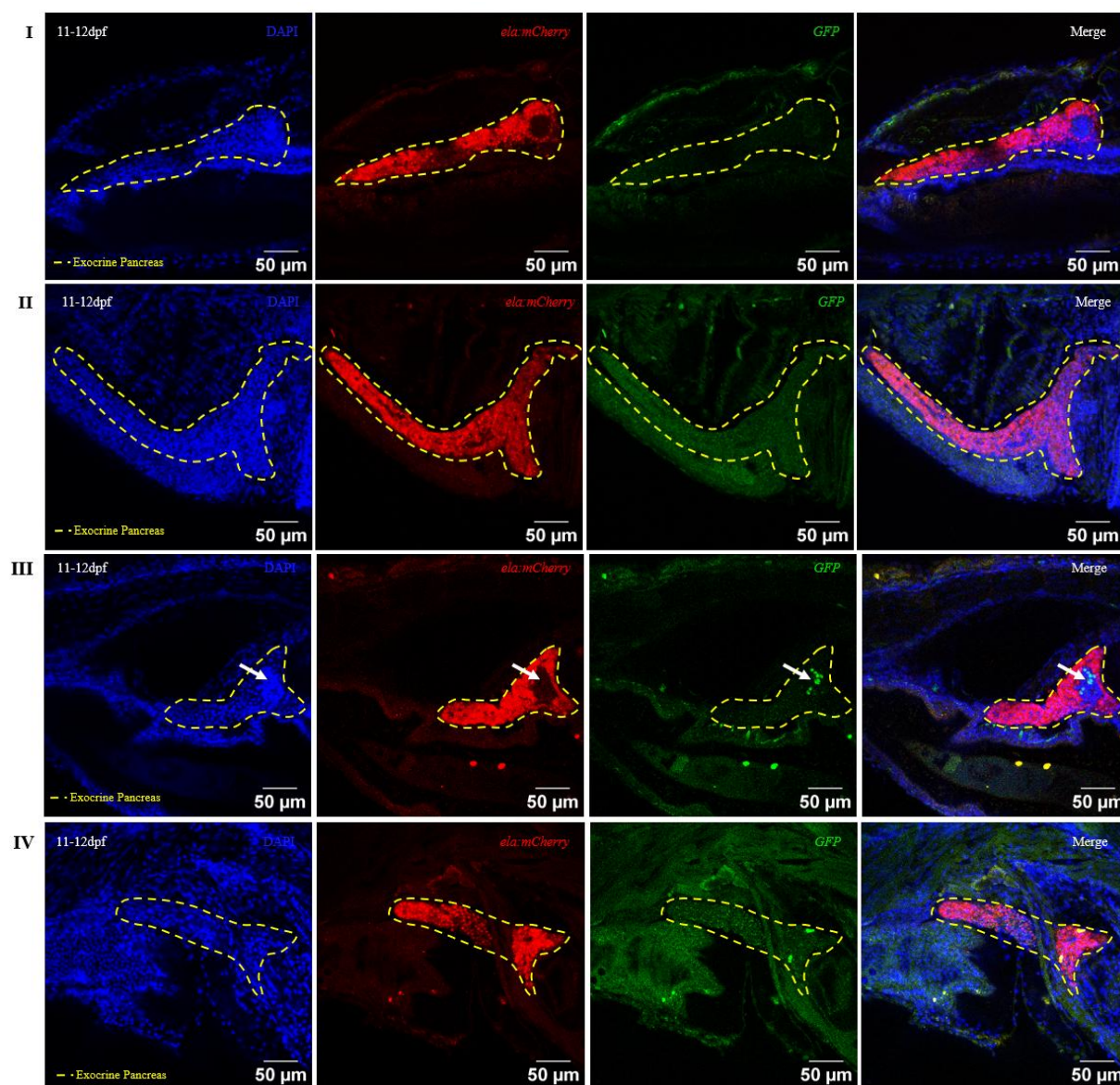


Figure 3.9 Representative confocal images of the enhancer reporter assay for human seq44 in zebrafish larvae at 11-12 dpf. Founder I (n=8) and II (n=12) showed no GFP expression (green) in zebrafish exocrine pancreas (yellow dashed line). Founder III (n= 5) also didn't show GFP expression in the exocrine domain, but even without endocrine pancreas markers some GFP-positive cells can be observed in the endocrine domain (white arrows). Founder IV (n=6) is the only one showing GFP expression co-localizing with pancreatic acinar cells (red). The yellow dashed line delineates the pancreatic exocrine domain, while the nuclei are stained with DAPI (blue). Images were captured with Leica confocal SP5II microscope; Magnification 40x. Scale bar: 50µm. Abbreviations: ela, elastase.

As discussed above, the fact that seq44 can drive a mild and variable GFP expression in exocrine and endocrine pancreas, points towards the possibility that this enhancer may have the ability to act in other pancreatic cell types or in a different pancreatic development stage. This hypothesis will be explored in the following results, where the same founders were crossed with the tg(*sst:mCherry*) transgenic line to access pancreatic enhancer activity specifically in endocrine δ -cells, a specific cell type of the endocrine pancreas (Eufrásio et al. 2020). Additionally, to analyse if this sequence can have an earlier enhancer activity, the progeny of the described outcross, was immunoassayed with anti-Nkx6.1 at 48 hpf, to observe the pancreatic progenitor cells (Binot et al. 2010).

In Founder I (n=15), GFP expression was observed in the endocrine domain of 53% (8/15) of the embryos. Surprisingly, only 12.5% (1/8) of the embryos showed GFP expression co-localising with the endocrine δ -cells (red) (Fig.3.10 - I.1). In the remaining 87.5% (7/8) of the embryos, GFP expression was detected in the endocrine domain but did not overlap with either Nkx6.1 staining or mCherry expression. (Fig.3.10 - I.2). Founders II (n=13) and IV (n=10), remained consistent with the previous results, showing no GFP expression in the endocrine and neither in progenitor domains (Fig.3.10 - II and IV). Founder III (n=7), as previously described, showed GFP expression in the endocrine domain in 43% (3/7) of the embryos, however, this GFP expression did not resemble either endocrine δ -cells (red) or progenitor domains (Fig.3.10 - III). These results are in line with the results from Founder I, both suggest enhancer activity of seq44 in other endocrine cell types.

Although seq44 is not able to drive expression in the progenitor domain, its unexpected ability to drive GFP expression in the endocrine pancreas supports the hypothesis that the seq44 may have a broader regulatory role within the pancreas than previously thought. This is also indicated by the Islet Regulome Browser (Mularoni et al. 2017) data, which shows that the *NR5A2* gene, although in low levels, is also expressed in pancreatic endocrine cells. These facts raise the possibility that seq44 could regulate *NR5A2* expression in both endocrine and exocrine tissues, acting as a dual-function enhancer (Halfon 2020; Ishibashi et al. 2013).

This dual functionality coupled with the PC SNP rs3790844 located within this sequence can imply that the disruption of seq44 would affect multiple pancreatic cell types, beyond just pancreatic acinar cells. This would increase the complexity of how this SNP might influence *nr5a2* gene expression in pancreas and thereby contribute to developmental anomalies or protection (Cho 2012; Elkon and Agami 2017; Eufrásio et al. 2020; Kleinjan and Coutinho 2009; Lettice et al. 2003; Roman et al. 2017). This highlights the importance of finding which specific endocrine cell type is regulated by seq44, which could be achieved by using other endocrine cell type-specific markers.

Even with these varying patterns, all observed expression was still confined to the pancreas, indicating that the human seq44 maintains some degree of pancreas specificity, which underscores the potential regulatory significance of the sequence within this organ.

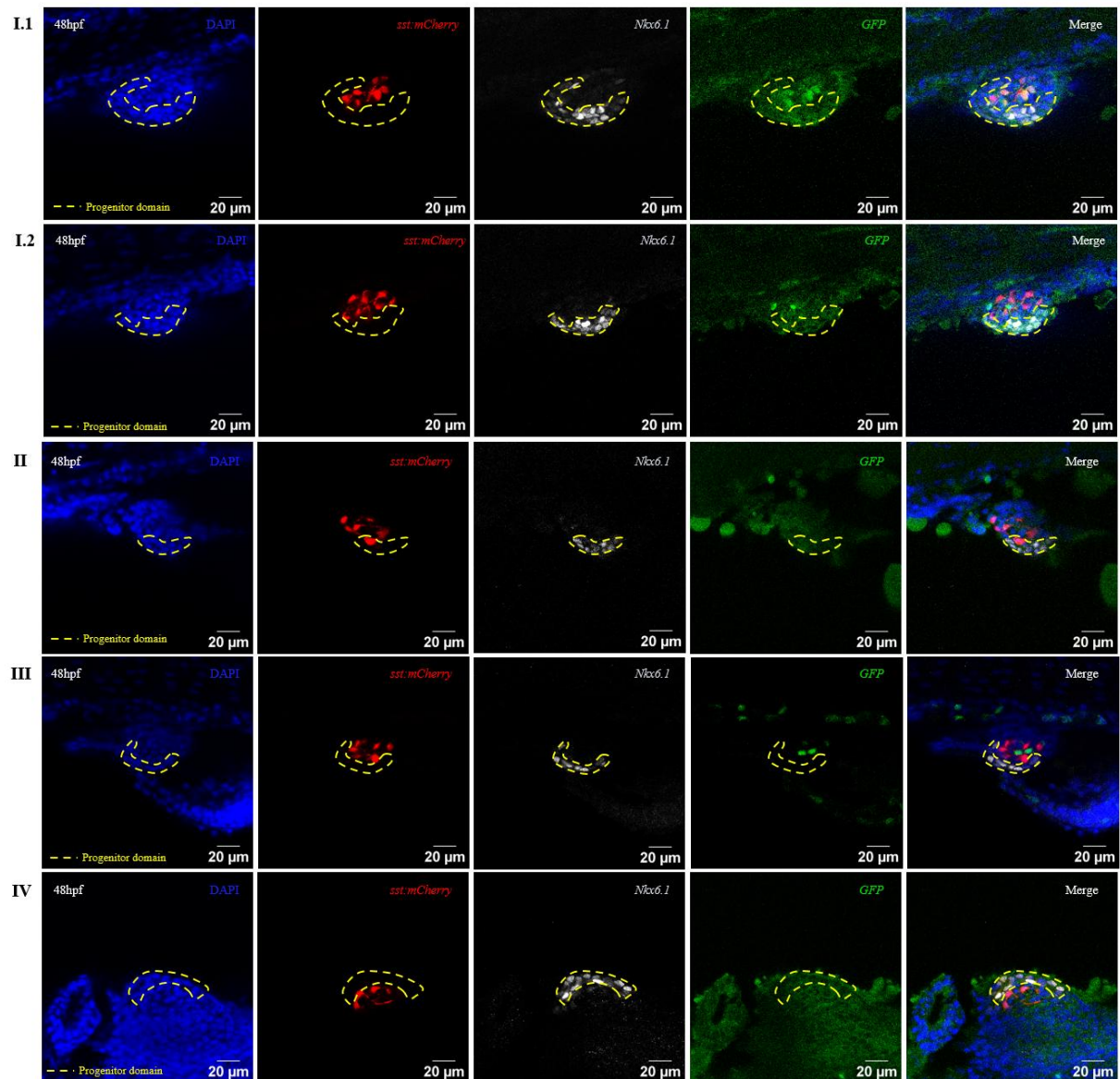


Figure 3.10 Representative confocal images of the enhancer reporter assays for human seq44 in zebrafish embryos at 48hpf. The yellow dashed line delineates the progenitor domain, stained by Nkx6.1 (white) while the nuclei are stained with DAPI (blue). Additionally, Somatostatin-producing cells (δ -cells) express mCherry (red). Founder I (n=15) has two different patterns of expression where some embryos show GFP expression (green) co-localising with the δ -cells (Founder I.1) and others show GFP expression but not co-localizing either with the progenitor domain and δ -cells (Founder I.2). In Founder III (n=7) the same can be observed, although GFP positive cells can be observed they don't co-localize either progenitor domain or δ -cells. Founders II (n=13) and IV (n=10) on the other hand, showed no GFP expression in the analysed embryos. Images were captured with Leica confocal SP5II; Zoom 2,10 x; Magnification 40x. Scale bar: 20 μ m. Abbreviations: ela, elastase; sst, somatostatin.

3.5.3 Test Enhancer Activity of Zebrafish *nr5a2* Putative enhancers Sequences in Pancreatic Acinar Cells

To test the enhancer activity of the zebrafish Region 1 and Region 2, the Z48 vectors carrying each PES were microinjected, along with TOL2 transposase mRNA, in one-cell stage WT zebrafish embryos. The injected fish were screened for GFP expression in the midbrain at 48 hpf to confirm genome integration of the sequence in the zebrafish genome. Three months later, those fish were outcrossed with the transgenic line tg(*ela:mCherry*) to test enhancer sequence activity in pancreatic acinar cells (Bordeira-Carriço et al. 2022). At 11-12 dpf, F1 larvae were also screened for the internal control of transgenesis in the midbrain and mCherry expression in the exocrine pancreas, marking acinar-producing cells (Fig.3.11).

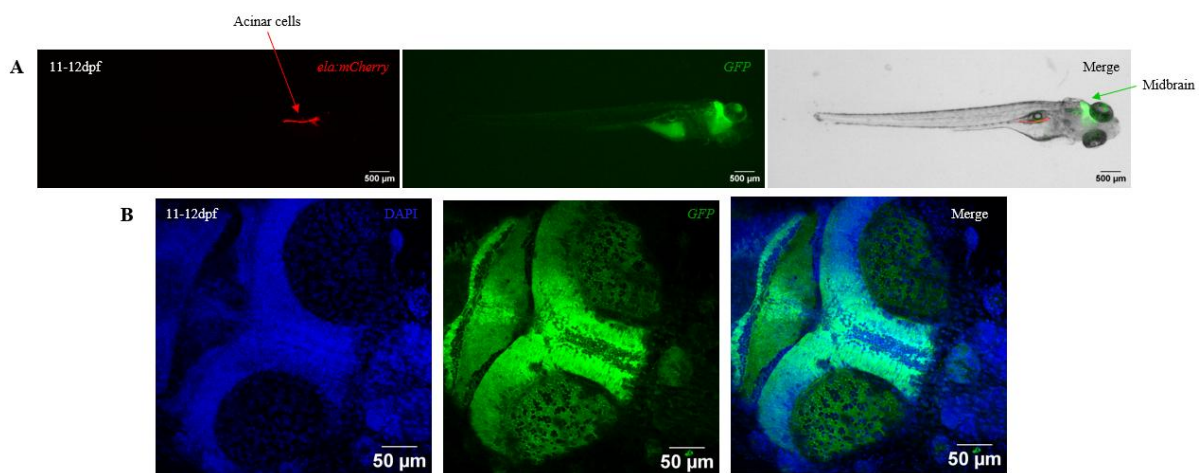


Figure 3.11 Representative images of F1 zebrafish larvae at 11-12 dpf. **A-** Full-body representative brightfield and epifluorescence image of a side-oriented, anterior right, dorsal top zebrafish embryo demonstrating GFP expression in the midbrain (green arrow) and mCherry expression in elastase-producing cells (acinar cells, red arrow). Images were captured with Leica M205 Fluorescence microscope; Zoom 3,00 x. Scale bar: 500µm. **B-** Representative confocal image of the zebrafish embryo showing GFP (green) expression in the midbrain as a positive control of transgenesis and germline transmission of zebrafish PES. The nuclei are stained with DAPI (blue). Images were captured with Leica confocal SP5II; Magnification 40x. Scale bar: 50µm.

Regarding the analysis of Region 1, the screening of the F1 progeny for GFP expression in the midbrain revealed that from the twenty-five injected fish (F0) only two (8%) can be classified as founders.

However, neither of them (Founder I – n=6; Founder II – n=12) showed GFP expression in the exocrine pancreas (Fig.3.12). Additionally, although no endocrine marker was used in this screening, no GFP expression was detected in the endocrine domain either, suggesting that Region 1 sequences do not exhibit enhancer activity either in pancreatic acinar or endocrine cells at this specific development stage of zebrafish.

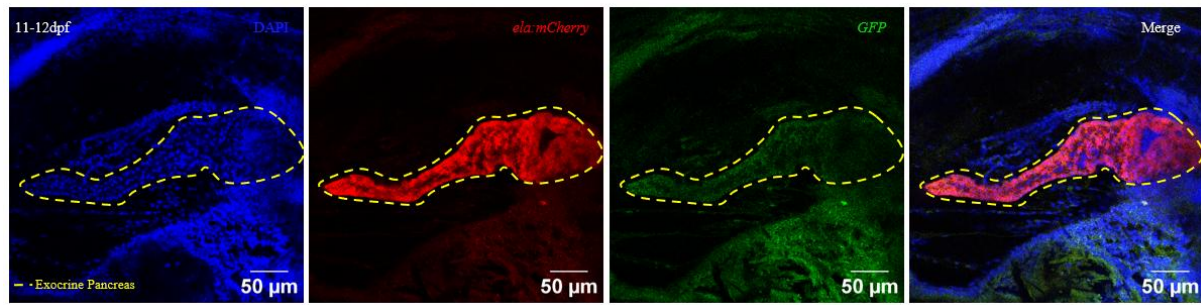


Figure 3.12 Representative images of the enhancer reporter assays for Region 1 in 11-12 dpf zebrafish larvae after outcrossing with *tg(ela:mCherry)*. Showing no GFP expression (green). The marker of the exocrine domain (yellow dashed line) is mCherry expression (red). All samples were stained with DAPI (blue). Images captured with *Leica confocal SP5II*; Zoom 2,10 x; Magnification 40x. Scale bar: 50μm. Abbreviations: ela, elastase.

The lack of detectable enhancer activity when we tested this PES can be explained by the position effect of the transgene, caused by the use of a transient reporter assay as Z48 vector (Bessa et al. 2009; Kemmler et al. 2023). Region 1 is randomly inserted in the zebrafish genome, which means that the genomic context in which this sequence is inserted can include other CREs that can modulate its enhancer activity (Bessa et al. 2009; Kemmler et al. 2023). Silencers, for example, can recruit TFs or structural proteins that can repress gene expression, impairing the activity of this PES (Maston et al. 2006; Riethoven 2010). Or even, insulators, elements that block communication between the PES and its target gene by disrupting the enhancer-promoter interaction or inhibiting chromatin looping, preventing enhancer activation in this context (Ali, et al. 2016; Maston et al. 2006; Riethoven 2010).

On the other hand, despite the absence of enhancer activity for this sequence at the specific timepoint analysed, this result does not exclude the possibility that this PES could be active at an earlier developmental stage. Given that enhancers often exhibit transient activity during development, it is possible that Region 1 was active during a different developmental window, but its activity was not detected at the selected timepoint (Ishibashi et al. 2013).

HiChIP (highly integrative chromatin immunoprecipitation) data for H3K4me3 from zebrafish whole pancreas was analysed in order to identify putative chromatin interactions between a PES and a specific promoter. Notably, an interaction was observed between Region 1 and the promoter of *nr5a2* gene (Bordeira-Carriço et al. 2022; Fig. 3.13). This observation could be an indication that Region 1 could be a functional enhancer that regulates *nr5a2* expression in one or more pancreatic cell types in zebrafish.

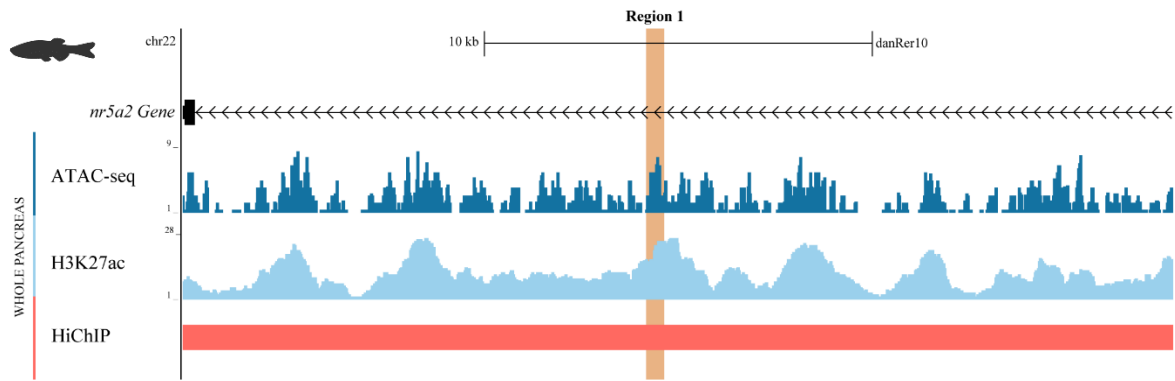


Figure 3.13 The H3K4me3 HiChIP data from adult whole pancreas (red bar) shows chromatin putative interactions between Region 1 (orange) and the *nr5a2* promoter in the zebrafish genomic landscape (danRer10). ATAC-seq (dark blue) data and H3K27ac ChIP-seq (light blue) data from adult whole pancreas shows an overlap between signal enrichment and the analysed sequence (Bordeira-Carriço et al. 2022).

However, we must consider the possibility that Region 1 may not possess enhancer activity in pancreatic acinar cells (Bordeira-Carriço et al. 2022). It is important to emphasise that the presence of epigenetic marks associated with enhancer activity and the chromatin accessibility observed in this region (section 3.1.2.), which made us choose Region 1 as a PES, are only predictions rather than definitive proof of enhancer function (Eufrásio et al. 2020).

For Region 2, the screening of the F1 larvae revealed that out of twenty-two injected fish only three (14%) can be classified as founders. Two of them, Founder I (n=44) and II (n=9), showed GFP expression in pancreatic acinar cells, co-localising with pancreatic acinar cells, in 30% (13/44) and 56% (5/9) of the larvae analysed, respectively (Fig. 3.14 – I and II). This co-localisation between GFP and mCherry expression is a strong indication of acinar-specific enhancer activity, as elastase is a well-established marker of exocrine tissue (Wan et al. 2006). In contrast, the other founder (III – n=5) lacked GFP expression in the exocrine pancreas (Fig. 3.14 - III), which may reflect variability in enhancer activity, potentially due to more than one insertion events (Ishibashi et al. 2013).

These findings strongly suggest that this sequence acts as an enhancer in pancreatic acinar cells, implying that Region 2 plays a role in regulating gene expression in the exocrine pancreas. However, this is not enough to prove that Region 2 is directly regulating *nr5a2* gene expression. Interestingly, when we analysed H3K4me3 HiChIP data from adult zebrafish whole pancreas, a technique that allows to identity all the putative interactions with putative active promoters, we observed a putative interaction between Region 2 and *nr5a2* gene promoter, indicating that maybe this enhancer sequence has a role in the regulation in the expression of this gene (Bordeira-Carriço et al. 2022; Fig. 3.15).

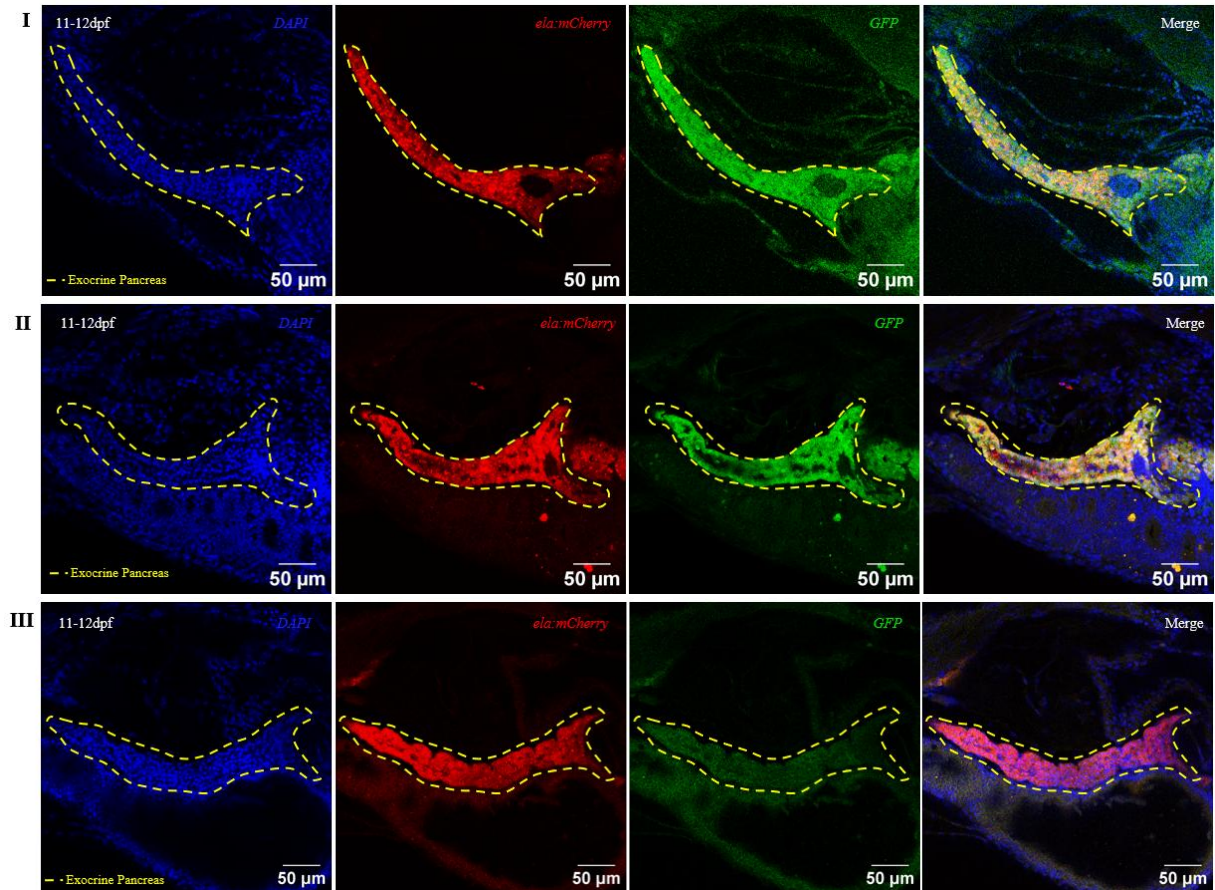


Figure 3.14 Representative confocal images of the enhancer reporter assays for Region 2 in zebrafish larvae at 11-12 dpf after outcrossing with *tg(ela:mCherry)*. The yellow dashed line delineates the exocrine domain, where elastase promoter is driving mCherry expression in acinar cells (red). The nuclei are stained with DAPI (blue). Founders I and II show GFP expression (green) in the pancreatic exocrine domain (yellow dashed line). Founder III shows no GFP expression in the exocrine pancreas. Images were captured with Leica confocal SP5II; Zoom 2,10 x; Magnification 40x. Scale bar: 50µm. Abbreviations: ela, elastase.

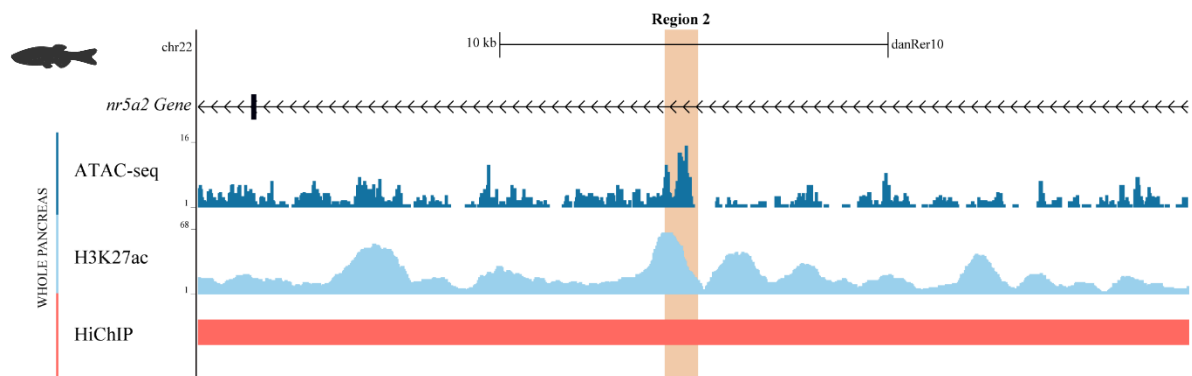


Figure 3.15 The H3K4me3 HiChIP data from adult whole pancreas (red bar) shows chromatin putative interactions between Region 2 (orange) and the *nr5a2* promoter in the zebrafish genomic landscape (danRer10). ATAC-seq (dark blue) data and ChIP-seq H3K27ac mark (light blue) from adult whole pancreas showing signal enrichment overlapping with the selected sequence (Region 2) (Bordeira-Carriço et al. 2022).

To further understand the regulatory function of Region 2, the JASPAR software (<http://jaspar.genereg.net>) (Rauluseviciute et al. 2024) was used to predict the TFs that have the ability to bind to this sequence (as described in [section 2.8](#)). By identifying the specific TFs that interact with this region provides a clearer understanding of which TFs are most likely to influence its enhancer activity.

Using JASPAR with a relative profile score threshold of 80%, 9607 putative vertebrate TFBS were predicted to bind to Region 2. However, as JASPAR includes multiple matrices for each TF, we identified a total of 718 different TFs that have the ability to bind to the analysed sequence.

Since Region 2 showed enhancer activity in pancreatic acinar cells, we decided to overlap this list of JASPAR predicted TFs with a list of TFs known to be associated with exocrine functions (Sox9, Hnf1b, Hnf6, Gata6, Gata4, Ptf1a, Nr5A2, Rpbj, Mist1, ERRgamma) (Bordeira-Carriço et al. 2022). We identified 3 exocrine TFs (Ptf1a, Gata4, Nr5A2) that overlapped with our predicted list of TFs. As previously described in [section 1.1.1](#), Ptf1a and Nr5a2 TFs are key drivers of acinar cell differentiation and stability, such as Gata4/6 factors that are essential for proper formation of the exocrine pancreatic tissue (Azevedo-Pouly et al. 2023; Aimeé Bastidas-Ponce et al. 2017; Duque et al. 2022; Hale et al. 2014; Larsen and Grapin-Botton 2017; Sandhu et al. 2021). The binding of these TFs to Region 2 suggests that this enhancer may influence pancreatic exocrine development processes and help maintain acinar cell identity.

Although there was no evidences of Region 2 endocrine expression at the timepoint analysed, we decided to also overlap our list of JASPAR predicted TFs with a list of TFs known to be associated with endocrine functions (Pax4, MafA, MafB, Nkx2.2, Nkx6.1, Isl1, NeuroD1, Pax6, Hb9, Pdx1, Npas4, Acl1, FoxA2, Ngn3, PGR) (Bordeira-Carriço et al. 2022), in order to analyse if the sequence might have endocrine activity at another timepoint not analysed. In the endocrine lineage, Pdx1 and Foxa2 point to potential roles in early pancreatic development and β -cell differentiation (Azevedo-Pouly et al. 2023; Aimeé Bastidas-Ponce et al. 2017; Tarifeño-Saldivia et al. 2017). These findings suggest that Region 2 could have endocrine enhancer activity possibly in an earlier development stage.

The relatively small overlap observed between the predicted TFs that bind in our sequence and those associated with pancreatic function, may be due to the reduced number of TFs present in our exocrine and endocrine lists. Yet, we were able to identify important TFs with relevant functions in pancreatic acinar cells.

JASPAR software contains an extensive vertebrate TFBS database, however, TFs specific to the zebrafish genome are poorly annotated (Rauluseviciute et al. 2024). This limitation could be contributing to the lack of enrichment observed in this analysis.

Furthermore, after ranking our list of predicted TFs from the highest relative score to the lowest, we obtained a top five of TFs (Table 3.1 and Fig. 3.16). These are the ones that, according to JASPAR, are most likely to bind to Region 2, providing insights into their regulatory potential and exploring their role on pancreatic biology could be relevant to better understand their potential impact in PC development.

Table 3.1 JASPAR analysis of the top 5 vertebrate TFs predicted in bind to Region 2. The relative score is filtered from the highest score, which is 1, to the lowest, which is 0.

Name	Score	Relative score	Start	End	Strand	Predicted sequence
FOS::JUN	10.663977	1.000000015	669	675	-	tgactca
Ahr::Arnt	8.844031	1.000000014	446	450	+	gcgtg
Ahr::Arnt	8.844031	1.000000014	501	505	-	gcgtg
ISL2	10.031854	1.000000013	269	274	-	cactta
ISL2	10.031854	1.000000013	688	693	-	cactta
SOX13	12.543894	1.000000012	71	77	+	acaatgg
Msx3	11.105006	1.000000011	33	38	-	caatta



Figure 3.16 Vertebrate TFBS predicted by the JASPAR software within Region 2 (798 bp). The TFBS represented are the top five more prone vertebrate TF to bind to this specific sequence. Msx3 TFBS is represented in red. SOX13 TFBS is represented in pink. ISL2 TFBS is represented in blue. Ahr::Arnt complex TFBS is represented in green. FOX::JUN complex TFBS is represented in yellow.

This TFs predicted to interact with Region 2 suggest that this sequence may have regulatory roles in acinar cell function and PC development. The first identified TF complex, FOS::JUN, part of the AP-1 (activator protein-1) family, is involved in inflammatory responses, indicating that Region 2 may mediate pro-inflammatory pathways in pancreatic acinar cells, increasing susceptibility to damage and contributing to PC development (Baldan et al. 2024; Cobo et al. 2018; Grabinski et al. 2023; Kubra et al. 2022).

The binding of Ahr::Arnt (Aryl hydrocarbon receptor :: AHR nuclear translocator) complex, known for responding to environmental factors and to be elevated in PC (Masoudi et al. 2019; Ren et al. 2020; Sondermann et al. 2023). This suggests that Region 2 could be influenced by external factors, which could potentially lead to acinar cell dysfunction, a key factor in the early stages of PC development (Fletterick 2017; Storz 2017).

Isl2 (ISL LIM homeobox 2a) TFs, linked to exocrine pancreas development, binding may support Region 2 in the regulation of *nr5a2* expression during acinar differentiation (Azevedo-Pouly et al. 2023; Larsen and Grapin-Botton 2017; Li et al. 2007; Sandhu et al. 2021; Wilfinger, Arkhipova, and Meyer 2013).

SOX13 (SRY-box TF 13) has been shown to be upregulated in PC, where it can promote invasion and colony formation in cancer cells (Nelson et al. 2020). The interaction of this TF with Region 2 could indicate a role in tumorigenesis, particularly in acinar cells. Given the emerging evidences that *NR5A2* plays a protective role against PC, the interaction between SOX13 and this sequence could be relevant in understanding how the dysregulation of this enhancer can contribute to PC development (Sandhu et al. 2021).

Finally, Msx3 (muscle segment homeobox 3) TFs, known for transcription repression, may regulate chromatin accessibility, affecting how Region 2 controls acinar cell-specific gene expression (Liu et al. 2004).

These interactions collectively support the hypothesis that Region 2 acts as an enhancer in pancreatic acinar cells, playing a role in regulating acinar cell identity and pancreatic homeostasis, key functions associated with *nr5a2* gene (Hale et al. 2014; Larsen and Grapin-Botton 2017; Sandhu et al. 2021). Therefore, suggesting that Region 2 controls the expression of this gene specifically.

3.6 *In vivo* Mutagenesis Assays

To further investigate the potential regulatory functions of Region 1 and Region 2 sequences in controlling the expression of the *nr5a2* gene in pancreatic acinar cells, *in vivo* mutagenesis of these sequences was performed using the CRISPR-Cas9 system (as described in [section 2.7.](#); Amorim et al. 2020). These loss-of-function assays aimed to delete part of the sequences to assess their impact on *nr5a2* gene expression.

Several sgRNA pairs were designed to target different parts of Region 1 and Region 2 in the zebrafish genome. Each sgRNA combination was tested within the pDR274 vector by microinjecting in one-cell stage WT embryos. As control for the Cas9-mRNA toxicity, embryos were injected only with cas9 protein. At 48 hpf, DNA was extracted from the injected embryos in order to assess if the sgRNA combinations successfully deleted the sequence fragments. A PCR reaction was performed with the specific primers of each sequence and the PCR product was run in a 2% agarose gel. DNA extracted from non-injected embryos was used as negative controls to ensure deletion specificity of the assay.

For Region 1 (470 bp), after testing multiple sgRNA combinations ([Table 2.3](#) and [Fig.3.17](#)) in zebrafish WT embryos, two effective sgRNA pairs were confirmed via PCR amplification and sequencing.



Figure 3.17 Design of sgRNAs targeting the Region 1. The total length of the sequence is 470 bp and in red are labelled the sgRNAs selected to test.

The first sgRNA pair (sgRNA 1 and sgRNA 2) covers 420 bp of the sequence. After performing the mutagenesis assay and a PCR with specific primers, the PCR product showed a band with 265 bp, confirming the successful deletion of part of Region 1 (Fig. 3.18 – A). The second sgRNA pair (sgRNA 1 and sgRNA 8) covers 211 bp, and the PCR product ran in the agarose gel showed a 259 bp band, confirming the successful deletion (Fig. 3.18 – B). Injected animals are mosaics for the deletion, so the PCR products ran in the agarose gel, showed two bands, one corresponding to the full sequence of Region 1 and another band with a lower size corresponding to the deleted sequence (Fig. 3.18).

Then, these genomic deletions were confirmed by Sanger sequencing, each sgRNA pair was injected into the transgenic zebrafish line tg(zP.E3:EGFP, *ela*:mCherry) contains an enhancer sequence that drives GFP expression in acinar cells (Bordeira-Carriço et al. 2022). The use of this zebrafish transgenic line will further help understanding the upcoming phenotypes in pancreatic acinar cells that may appear in response to the genomic deletions.

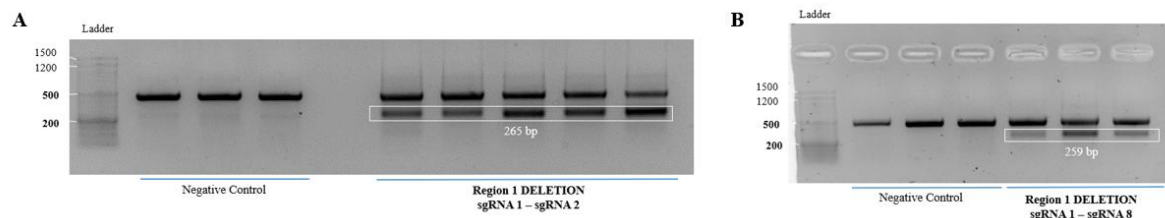


Figure 3.18 Confirmation of the efficacy of two sgRNA pairs targeting Region 1 using the CRISPR-Cas9 system. **A-** PCR result of Region 1 deletion using sgRNA 1 and sgRNA 2. The negative control shows the full-length of the sequence (470 bp) which is replicated in our results. The successful deletion results is confirmed by a reduced band with 265 bp, indicating the excision of 240 bp. **B-** PCR result of Region 1 deletion using sgRNA 1 and sgRNA 8. The negative control also shows the full-length of the sequence (470 bp) which is replicated in our results. The successful deletion results in a reduced band with 259 bp, confirming the excision of 211 bp. Ladder IV Gene Ruler 50-1500 bp (NZYTech).

The mosaicism observed in F0 embryos presents challenges for direct phenotype analysis, as injected animals consist in a mixture of mutant and WT cells. This mixture may dilute any observed effects of the mutation, which highlights the importance of establishing stable mutant lines (Amorim et al. 2020;

Maston et al. 2006). After crossing F0 fish with WT adults, mutations can be isolated in the F1 generation and phenotypes can only be directly evaluated in the F2 embryos that are homozygous for the deletion. This process is crucial because it could be possible to access to the functional impact of Region 1 and Region 2 in *nr5a2* gene expression.

Until now, with the enhancer reporter assays performed, it was not possible to show that the Region 1 is an enhancer in zebrafish pancreatic acinar cells, since other timepoints can be tested and analysed. However, depending on the genomic context or proximity to other regulatory elements in which this region is inserted, it could influence *nr5a2* expression (Bordeira-Carriço et al. 2022; Kemmler et al. 2023). Therefore, the loss-of-function assay can still yield informative results about this sequence regulatory functions.

The designing of sgRNAs targeting Region 2 (798 bp) proved to be particularly challenging due to the presence of repetitive elements flanking the sequence, which also complicated the design of specific primers for the screening (Fig. 3.19). Despite the test of several sgRNA combinations (Table 2.3 and Fig. 3.20), none of them was effective in creating genomic deletions at the target sequence. This can be explained by the genomic complexity of Region 2 (Luo et al. 2024).

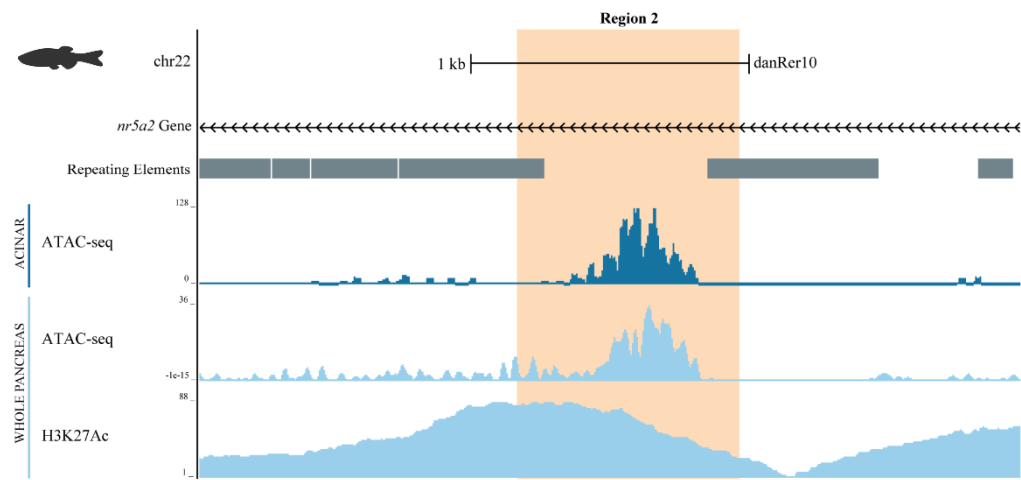


Figure 3.19 Zebrafish *nr5a2* genomic landscape (danRer10) containing the Region 2 (orange), flanked by repeating elements (grey bars). This sequence overlaps with the signal enrichment of H3K27Ac ChIP-seq from adult whole pancreas (dark blue) and the ATAC-seq data from both pancreatic acinar cells (light blue) and adult whole pancreas.



Figure 3.20 Design of sgRNAs targeting the Region 2. The total length of the sequence is 798 bp and in red are labelled the sgRNAs selected to test.

CONCLUSIONS AND FUTURE PERSPECTIVES

In this work, we wanted to explore the acinar enhancer activity of a human PES that overlaps with a PC SNP rs3790844 (also termed as seq44) using zebrafish as a model to access its role in the regulation of the *nr5a2* gene. Through enhancer reporter assays in pancreatic exocrine, endocrine and progenitor domains, we were able to demonstrate that seq44 exhibits mild and variable enhancer activity in the zebrafish pancreas, with a surprising dual role in both pancreatic endocrine and exocrine tissues ([section 3.5.2.](#)). These findings suggest that the disruption of this sequence could have widespread effects on gene regulation in the pancreas. However, the impact of this sequence on *nr5a2* gene expression in pancreatic acinar cells remains unclear. In future research, we could perform reporter assays in various timepoints in pancreatic acinar cells to determine whether this sequence functions as an acinar-specific enhancer (Bordeira-Carriço et al. 2022). Then, if we identify a specific timepoint at which the sequence is active, we could then delete it in this specific cell type and assess its impact on *nr5a2* expression through qPCR analysis (Bordeira-Carriço et al. 2022; Eufrásio et al. 2020).

While the zebrafish model has provided valuable insights into the seq44 activity, additional studies in human pancreatic cells are necessary to comprehensively assess the functional role of this sequence as an enhancer in its physiological context. To investigate this further, we could conduct *in vitro* enhancer reporter assays using human acinar cell lines (Bordeira-Carriço et al. 2022; Ferreira et al. 2023). Given the limited availability of human acinar cell lines, pancreatic acinar organoids could be established as a model system to perform these enhancer assays (Huang et al. 2021).

Regarding the zebrafish *nr5a2* genomic landscape, two PES, Region 1 and Region 2, were selected ([section 3.2.](#)) to test their acinar enhancer activity and to further investigate their potential roles in regulating *nr5a2* gene transcription and further impact in the development of PC. Findings from the enhancer reporter assays indicate that these two zebrafish sequences function differently, however, both appear to have significant roles in the regulation of pancreatic cell function.

Region 1 showed no pancreatic acinar enhancer activity under the tested conditions (section 3.5.3.). This result could be explained by the fact that enhancers could be active in different timepoints (Cho 2012). This hypothesis could be explored by testing if this sequence is active in earlier stages of pancreatic development, by performing immunoassays with other cell-specific markers, such as *Nkx6.1*, a marker of the pancreatic progenitor cells (Binot et al. 2010). Other zebrafish transgenic lines could also be used to access enhancer activity in other pancreatic cell types (Eufrásio et al. 2020).

If any enhancer activity is detected in early zebrafish pancreatic development, the genomic deletions generated by the two effective sgRNA pairs, designed and validated in this thesis using CRISPR-Cas9 assays (section 3.6), can be useful to investigate the impact of these PES on *nr5a2* expression through qPCR (Bordeira-Carriço et al. 2022; Eufrásio et al. 2020). As this gene is known to play an important role in the development of the pancreas (Hale et al. 2014; Larsen and Grapin-Botton 2017; Sandhu et al. 2021), would be relevant to understand if Region 1 influences pancreatic development and contributes to the onset of an inflammatory state, a known factor in the progression of PC (Flandez et al. 2014).

Region 2 showed a clear enhancer activity in pancreatic acinar cells. The predicted results of H3K4me3 HiChIP from adult zebrafish pancreas (section 3.5.3.) indicate a putative interaction between this region and the *nr5a2* promoter. Additionally, JASPAR analysis (section 3.5.3) suggests that key pancreatic transcription factors may bind to this particular sequence. Based on these findings we can speculate that this sequence could play an important role in the transcription regulation of the *nr5a2* gene. However, in order to refine these regulatory predictions, further research should be made in order to expand the list of endocrine and exocrine TFs and to improve the existing zebrafish-specific TF matrices.

The loss-of-function assays performed for Region 2 encountered significant technical difficulties because of repeating elements flanking this sequence complicating the design of sgRNAs and primers (section 3.6.). Some efforts must be made to overcome these technical issues. A possible solution could be the redesign of sgRNAs avoiding the repetitive elements followed by the optimization of the CRISPR-Cas9 mutagenesis assay (Amorim et al. 2020; Gagnon et al. 2014; Luo et al. 2024).

Once mutagenesis is optimized with efficient sgRNA pairs, the impact of this acinar enhancer sequence specifically on *nr5a2* expression could be accessed through qPCR (Bordeira-Carriço et al. 2022; Eufrásio et al. 2020). If the qPCR showed a clear negative impact in the transcription of this gene, stable mutant lines carrying Region 2 genomic deletions could be developed in a genomic background more sensitive to tumour formation, such as those with oncogenic *Kras* mutations (Flandez et al. 2014). Previous studies have shown that pancreatic-specific partial loss of *Nr5a2* gene alone in mice is enough to trigger pancreatic inflammation, leading to pancreatitis, but is not sufficient to cause tumours (Cobo et al. 2018; Flandez et al. 2014). However, *Nr5a2* mutations in combination with oncogenic *Kras* mutations can drive the progression of PC (Flandez et al. 2014). Additionally, using an acinar cell marker to

track cell-specific behaviour and organisation within pancreatic tissue, would be useful to evaluate tumorigenesis progression.

Overall, this work shed light on the regulatory roles of specific PES located in the landscape of *NR5A2/nr5a2* gene, particularly in pancreatic acinar cells. Despite considerable advances in elucidating the transcriptional regulation of this gene, translating these discoveries into human biology represents the next important challenge, particularly in the context PC. By exploring these future research avenues, we will improve our understanding of how *nr5a2* enhancer activity can contribute PC progression, bringing us one step closer to potential therapeutic interventions targeting gene regulation at an enhancer level.

BIBLIOGRAPHY

- Aguirre, Andrew J., and Eric A. Collisson. 2017. *Advances in the Genetics and Biology of Pancreatic Cancer*.
- Albadri, Shahad, Filippo Del Bene, and Céline Revenu. 2017. "Genome Editing Using CRISPR/Cas9-Based Knock-in Approaches in Zebrafish." *Methods* 121–122:77–85.
- Aleström, Peter, Jethro L. Holter, and Rasoul Nourizadeh-Lillabadi. 2006. "Zebrafish in Functional Genomics and Aquatic Biomedicine." *Trends in Biotechnology* 24(1):15–21.
- Ali, Tamer, Rainer Renkawitz, and Marek Bartkuhn. 2016. "Insulators and Domains of Gene Expression." *Current Opinion in Genetics and Development* 37:17–26.
- Amorim, João Pedro, Renata Bordeira-Carriço, Ana Gali-Macedo, Chiara Perrod, and José Bessa. 2020. "CRISPR-Cas9-Mediated Genomic Deletions Protocol in Zebrafish." *STAR Protocols* 1(3). doi: 10.1016/j.xpro.2020.100208.
- Anon. 2010. "Overview of the Pancreas." *Advances in Experimental Medicine and Biology* 690:3–12.
- Antinucci, Paride, and Robert Hindges. 2016. "A Crystal-Clear Zebrafish for in Vivo Imaging." *Scientific Reports* 6. doi: 10.1038/srep29490.
- Arda, H. Efsun, Cecil M. Benitez, and Seung K. Kim. 2013. "Gene Regulatory Networks Governing Pancreas Development." *Developmental Cell* 25(1):5–13.
- Arda, H. Efsun, Jennifer Tsai, Yenny R. Rosli, Paul Giresi, Rita Bottino, William J. Greenleaf, Howard Y. Chang, and Seung K. Kim. 2018. "A Chromatin Basis for Cell Lineage and Disease Risk in the Human Pancreas." *Cell Systems* 7(3):310–322.e4. doi: 10.1016/j.cels.2018.07.007.
- Atkinson, Mark A., Martha Campbell-Thompson, Irina Kusmartseva, and Klaus H. Kaestner. 2020. "Organisation of the Human Pancreas in Health and in Diabetes." *Diabetologia* 63(10):1966–73.

- Auer, Thomas O., Karine Durore, Anne De Cian, Jean Paul Concordet, and Filippo Del Bene. 2014. "Highly Efficient CRISPR/Cas9-Mediated Knock-in in Zebrafish by Homology-Independent DNA Repair." *Genome Research* 24(1):142–53. doi: 10.1101/gr.161638.113.
- Azevedo-Pouly, Ana, Michael A. Hale, Galvin H. Swift, Chinh Q. Hoang, Tye G. Deering, Jumin Xue, Thomas M. Wilkie, L. Charles Murtaugh, and Raymond J. MacDonald. 2023. "Key Transcriptional Effectors of the Pancreatic Acinar Phenotype and Oncogenic Transformation." *PLoS ONE* 18(10 October). doi: 10.1371/journal.pone.0291512.
- Backx, Elyne, Katarina Coolens, Jan Lars Van den Bossche, Isabelle Houbracken, Elisa Espinet, and Ilse Rooman. 2022. "On the Origin of Pancreatic Cancer: Molecular Tumor Subtypes in Perspective of Exocrine Cell Plasticity." *CMGH* 13(4):1243–53.
- Bakhti, Mostafa, Anika Böttcher, and Heiko Lickert. 2019. "Modelling the Endocrine Pancreas in Health and Disease." *Nature Reviews Endocrinology* 15(3):155–71.
- Baldan, Jonathan, Juan Camacho-Roda, Marta Ballester, Kristina Høj, Anita Kurilla, H. Carlo Maurer, Sebastian Arcila-Barrera, Xinyi Lin, Zhaolong Pan, Joana Leitão Castro, Alejandro Enrique Mayorca-Guiliani, Charlotte Vestrup Rift, Jane Hasselby, Luc Bouwens, Véronique Lefebvre, Charles J. David, Oren Parnas, Kathleen E. Del-Giorno, Janine Terra Erler, Ilse Rooman, and Luis Arnes. 2024. "Resolution of Acinar Dedifferentiation Regulates Tissue Remodeling in Pancreatic Injury and Cancer Initiation." *Gastroenterology*. doi: 10.1053/j.gastro.2024.04.031.
- Bastidas-Ponce, Aimée, Sara S. Roscioni, Ingo Burtscher, Erik Bader, Michael Sterr, Mostafa Bakhti, and Heiko Lickert. 2017. "Foxa2 and Pdx1 Cooperatively Regulate Postnatal Maturation of Pancreatic β -Cells." *Molecular Metabolism* 6(6):524–34. doi: 10.1016/j.molmet.2017.03.007.
- Bastidas-Ponce, Aimeé, Katharina Scheibner, Heiko Lickert, and Mostafa Bakhti. 2017. "Cellular and Molecular Mechanisms Coordinating Pancreas Development." *Development (Cambridge)* 144(16):2873–88.
- Bessa, José, Mario Luengo, Solangel Rivero-Gil, Ana Ariza-Cosano, António H. F. Maia, Francisco J. Ruiz-Ruano, Pablo Caballero, Silvia Naranjo, Jaime J. Carvajal, and José Luis Gómez-Skarmeta. 2014. "A Mobile Insulator System to Detect and Disrupt Cis-Regulatory Landscapes in Vertebrates." *Genome Research* 24(3):487–95. doi: 10.1101/gr.165654.113.
- Bessa, José, Juan J. Tena, Elisa De La Calle-Mustienes, Ana Fernández-Miñán, Silvia Naranjo, Almudena Fernández, Lluís Montoliu, Altuna Akalin, Boris Lenhard, Fernando Casares, and José Luis Gómez-Skarmeta. 2009. "Zebrafish Enhancer Detection (ZED) Vector: A New Tool to Facilitate Transgenesis and the Functional Analysis of Cis-Regulatory Regions in Zebrafish." *Developmental Dynamics* 238(9):2409–17. doi: 10.1002/dvdy.22051.

- Binot, A. C., I. Manfroid, L. Flasse, M. Winandy, P. Motte, J. A. Martial, B. Peers, and M. L. Voz. 2010. "Nkx6.1 and Nkx6.2 Regulate α - and β -Cell Formation in Zebrafish by Acting on Pancreatic Endocrine Progenitor Cells." *Developmental Biology* 340(2):397–407. doi: 10.1016/j.ydbio.2010.01.025.
- Bordeira-Carriço, Renata, Joana Teixeira, Marta Duque, Mafalda Galhardo, Diogo Ribeiro, Rafael D. Acemel, Panos N. Firbas, Juan J. Tena, Ana Eufrásio, Joana Marques, Fábio J. Ferreira, Telmo Freitas, Fátima Carneiro, José Luís Gómez-Skarmeta, and José Bessa. 2022. "Multidimensional Chromatin Profiling of Zebrafish Pancreas to Uncover and Investigate Disease-Relevant Enhancers." *Nature Communications* 13(1). doi: 10.1038/s41467-022-29551-7.
- Buenrostro, Jason D., Paul G. Giresi, Lisa C. Zaba, Howard Y. Chang, and William J. Greenleaf. 2013. "Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position." *Nature Methods* 10(12):1213–18. doi: 10.1038/nmeth.2688.
- Chan, Shuheng, Dan Shen, Yatong Sang, Saisai Wang, Yali Wang, Cai Chen, Bo Gao, and Chengyi Song. 2019. "Development of Enhancer-Trapping and-Detection Vectors Mediated by the Tol2 Transposon in Zebrafish." *PeerJ* 2019(4). doi: 10.7717/peerj.6862.
- Chen, Qun, Hao Yuan, Guo Dong Shi, Yang Wu, Dong Fang Liu, Yu Ting Lin, Lei Chen, Wan Li Ge, Kuirong Jiang, and Yi Miao. 2018. "Association between NR5A2 and the Risk of Pancreatic Cancer, Especially among Caucasians: A Meta-Analysis of Case-Control Studies." *OncoTargets and Therapy* 11:2709–23. doi: 10.2147/OTT.S157759.
- Chen, Shen, Chaohong Li, Guangming Yuan, and Fukang Xie. 2007 *Anatomical and Histological Observation on the Pancreas in Adult Zebrafish*.
- Chen, Xiaowen, Dafne Gays, and Massimo M. Santoro. 2016. "Transgenic Zebrafish." Pp. 107–14 in *Methods in Molecular Biology*. Vol. 1464. Humana Press Inc.
- Cho, Ken W. Y. 2012. "Enhancers." *Wiley Interdisciplinary Reviews: Developmental Biology* 1(4):469–78. doi: 10.1002/wdev.53.
- Choi, Tae Young, Tae Ik Choi, Yu Ri Lee, Seong Kyu Choe, and Cheol Hee Kim. 2021. "Zebrafish as an Animal Model for Biomedical Research." *Experimental and Molecular Medicine* 53(3):310–17.
- Cobo, Isidoro, Paola Martinelli, Marta Flández, Latifa Bakiri, Mingfeng Zhang, Enrique Carrillo-De-Santa-Pau, Jinping Jia, Víctor J. Sánchez Arévalo Lobo, Diego Megías, Irene Felipe, Natalia Del Pozo, Irene Millán, Liv Thommesen, Torunn Bruland, Sara H. Olson, Jill Smith, Kristina Schoonjans, William R. Bamlet, Gloria M. Petersen, Núria Malats, Laufey T. Amundadottir, Erwin F. Wagner, and Francisco X. Real. 2018. "Transcriptional Regulation by NR5A2 Links Differentiation and Inflammation in the Pancreas." *Nature* 554(7693):533–37. doi: 10.1038/nature25751.

- Duque, Marta, João Pedro Amorim, and José Bessa. 2022. "Ptf1a Function and Transcriptional Cis-Regulation, a Cornerstone in Vertebrate Pancreas Development." *FEBS Journal* 289(17):5121–36.
- Edlund, Helena. 2001. *Section 1:-Cell Differentiation and Growth Developmental Biology of the Pancreas*. Vol. 50.
- Elkon, Ran, and Reuven Agami. 2017. "Characterization of Noncoding Regulatory DNA in the Human Genome." *Nature Biotechnology* 35(8):732–46.
- Eufrásio, Ana, Chiara Perrod, Fábio J. Ferreira, Marta Duque, Mafalda Galhardo, and José Bessa. 2020. "In Vivo Reporter Assays Uncover Changes in Enhancer Activity Caused by Type 2 Diabetes–Associated Single Nucleotide Polymorphisms." *Diabetes* 69(12):2794–2805. doi: 10.2337/db19-1049.
- Ferreira, Fábio J., Mafalda Galhardo, Joana Teixeira, Elsa Logarinho, and Abel Salazar. n.d. "FOXM1 Expression Reverts Aging Chromatin Profiles through Repression of the Senescence-Associated Pioneer Factor AP-1 2 3 4 5." 16. doi: 10.1101/2023.05.04.539315.
- Flandez, Marta, Jaroslaw Cendrowski, Marta Cañamero, Antonio Salas, Natalia Del Pozo, Kristina Schoonjans, and Francisco X. Real. 2014. "Nr5a2 Heterozygosity Sensitises to, and Cooperates with, Inflammation in KRasG12V-Driven Pancreatic Tumorigenesis." *Gut* 63(4):647–55. doi: 10.1136/gutjnl-2012-304381.
- Fletcher, Robert. 2017. "NR5A2 Discovering Compounds That Block Tumor Growth in PDAC." *Journal of Surgical Oncology* 116(1):89–93.
- Gagnon, James A., Eivind Valen, Summer B. Thyme, Peng Huang, Laila Ahkmetova, Andrea Pauli, Tessa G. Montague, Steven Zimmerman, Constance Richter, and Alexander F. Schier. 2014. "Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs." *PLoS ONE* 9(5). doi: 10.1371/journal.pone.0098186.
- Galang, Giselle, Ravi Mandla, Hongmei Ruan, Catherine Jung, Tanvi Sinha, Nicole R. Stone, Roland S. Wu, Brandon J. Mannion, Prasanna K. R. Allu, Kevin Chang, Ashwin Rammohan, Marie B. Shi, Len A. Pennacchio, Brian L. Black, and Vasanth Vedantham. 2020. "ATAC-Seq Reveals an Isl1 Enhancer That Regulates Sinoatrial Node Development and Function." *Circulation Research* 127(12):1502–18. doi: 10.1161/CIRCRESAHA.120.317145.
- García-González, Estela, Martín Escamilla-Del-Arenal, Rodrigo Arzate-Mejía, and Félix Recillas-Targa. 2016. "Chromatin Remodeling Effects on Enhancer Activity." *Cellular and Molecular Life Sciences* 73(15):2897–2910.
- Gou, Yawei, Wei Sun, Lingling Liu, Mingming Zhang, Jianan Du, Ruonan Wang, and Xuesong Xu. 2020. "Construction of Irf4a Transgenic Zebrafish Using Tol2 System and Its Potential Application." *Dose-Response* 18(2). doi: 10.1177/1559325820926733.

- Gowthaman, Uthra, Desiré García-Pichardo, Yu Jin, Isabel Schwarz, and Sebastian Marquardt. 2020. "DNA Processing in the Context of Noncoding Transcription." *Trends in Biochemical Sciences* 45(12):1009–21.
- Grabinski, Sarah E., Dhvani Parsana, and Brian D. Perkins. 2023. "Comparative Analysis of Transcriptional Changes in Zebrafish Cep290 and Bbs2 Mutants by RNA-Seq Reveals Upregulation of Inflammatory and Stress-Related Pathways." *Frontiers in Molecular Neuroscience* 16. doi: 10.3389/fnmol.2023.1148840.
- Grandi, Fiorella C., Hailey Modi, Lucas Kampman, and M. Ryan Corces. 2022. "Chromatin Accessibility Profiling by ATAC-Seq." *Nature Protocols* 17(6):1518–52.
- Grant, T. J., K. Hua, and A. Singh. 2016. "Molecular Pathogenesis of Pancreatic Cancer." Pp. 241–75 in *Progress in Molecular Biology and Translational Science*. Vol. 144. Elsevier B.V.
- Hale, Michael A., Galvin H. Swift, Chinh Q. Hoang, Tye G. Deering, Toshi Masui, Youn Kyoung Lee, Jumin Xue, and Raymond J. MacDonald. 2014. "The Nuclear Hormone Receptor Family Member NR5A2 Controls Aspects of Multipotent Progenitor Cell Formation and Acinar Differentiation during Pancreatic Organogenesis." *Development (Cambridge)* 141(16):3123–33. doi: 10.1242/dev.109405.
- Halfon, Marc S. 2020. "Silencers, Enhancers, and the Multifunctional Regulatory Genome." *Trends in Genetics* 36(3):149–51.
- Hawksworth, Graeme, Jake Hales, Fransico Martinez, Allen Hynes, Angela Hamilton, and Veronica Fernandez. 2019. "Pancreatic Cancer Trends in Europe: Epidemiology and Risk Factors." *Medical Studies* 35(2):164–71. doi: 10.5114/ms.2019.86336.
- Heinz, Sven, Casey E. Romanoski, Christopher Benner, and Christopher K. Glass. 2015. "The Selection and Function of Cell Type-Specific Enhancers." *Nature Reviews Molecular Cell Biology* 16(3):144–54.
- Hernandez-Garcia, Carlos M., and John J. Finer. 2014. "Identification and Validation of Promoters and Cis-Acting Regulatory Elements." *Plant Science* 217–218:109–19.
- Holmstrom, Sam R., Tye Deering, Galvin H. Swift, Frank J. Poelwijk, David J. Mangelsdorf, Steven A. Kliewer, and Raymond J. Macdonald. 2011. "LRH-1 and PTF1-L Coregulate an Exocrine Pancreas-Specific Transcriptional Network for Digestive Function." *Genes and Development* 25(16):1674–79. doi: 10.1101/gad.16860911.
- Howe, Kerstin, Matthew D. Clark, Carlos F. Torroja, James Torrance, Camille Berthelot, Matthieu Muffato, John E. Collins, Sean Humphray, Karen McLaren, Lucy Matthews, Stuart McLaren, Ian Sealy, Mario Caccamo, Carol Churcher, Carol Scott, Jeffrey C. Barrett, Romke Koch, Gerd Jörg Rauch, Simon White, William Chow, Britt Kilian, Leonor T. Quintais, José A. Guerra-Assunção, Yi Zhou, Yong Gu, Jennifer Yen, Jan Hinnerk Vogel, Tina Eyre, Seth Redmond, Ruby Banerjee, Jianxiang Chi, Beiyuan Fu, Elizabeth Langley, Sean F. Maguire, Gavin K. Laird, David Lloyd,

Emma Kenyon, Sarah Donaldson, Harminder Sehra, Jeff Almeida-King, Jane Loveland, Stephen Trevanion, Matt Jones, Mike Quail, Dave Willey, Adrienne Hunt, John Burton, Sarah Sims, Kirsten McLay, Bob Plumb, Joy Davis, Chris Clee, Karen Oliver, Richard Clark, Clare Riddle, David Elliott, Glen Threadgold, Glenn Harden, Darren Ware, Beverly Mortimer, Giselle Kerry, Paul Heath, Benjamin Phillimore, Alan Tracey, Nicole Corby, Matthew Dunn, Christopher Johnson, Jonathan Wood, Susan Clark, Sarah Pelan, Guy Griffiths, Michelle Smith, Rebecca Glithero, Philip Howden, Nicholas Barker, Christopher Stevens, Joanna Harley, Karen Holt, Georgios Panagiotidis, Jamieson Lovell, Helen Beasley, Carl Henderson, Daria Gordon, Katherine Auger, Deborah Wright, Joanna Collins, Claire Raisen, Lauren Dyer, Kenric Leung, Lauren Robertson, Kirsty Ambridge, Daniel Leongamornlert, Sarah McGuire, Ruth Gilderthorpe, Coline Griffiths, Deepa Manthravadi, Sarah Nichol, Gary Barker, Siobhan Whitehead, Michael Kay, Jacqueline Brown, Clare Murnane, Emma Gray, Matthew Humphries, Neil Sycamore, Darren Barker, David Saunders, Justene Wallis, Anne Babbage, Sian Hammond, Maryam Mashreghi-Mohammadi, Lucy Barr, Sancha Martin, Paul Wray, Andrew Ellington, Nicholas Matthews, Matthew Ellwood, Rebecca Woodmansey, Graham Clark, James Cooper, Anthony Tromans, Darren Grafham, Carl Skuce, Richard Pandian, Robert Andrews, Elliot Harrison, Andrew Kimberley, Jane Garnett, Nigel Fosker, Rebekah Hall, Patrick Garner, Daniel Kelly, Christine Bird, Sophie Palmer, Ines Gehring, Andrea Berger, Christopher M. Dooley, Zübeyde Ersan-Ürün, Cigdem Eser, Horst Geiger, Maria Geisler, Lena Karotki, Anette Kirn, Judith Konantz, Martina Konantz, Martina Oberländer, Silke Rudolph-Geiger, Mathias Teucke, Kazutoyo Osoegawa, Baoli Zhu, Amanda Rapp, Sara Widaa, Cordelia Langford, Fengtang Yang, Nigel P. Carter, Jennifer Harrow, Zemin Ning, Javier Herrero, Steve M. J. Searle, Anton Enright, Robert Geisler, Ronald H. A. Plasterk, Charles Lee, Monte Westerfield, Pieter J. De Jong, Leonard I. Zon, John H. Postlethwait, Christiane Nüsslein-Volhard, Tim J. P. Hubbard, Hugues Roest Crollius, Jane Rogers, and Derek L. Stemple. 2013. "The Zebrafish Reference Genome Sequence and Its Relationship to the Human Genome." *Nature* 496(7446):498–503. doi: 10.1038/nature12111.

Huang, Ling, Ridhdhi Desai, Daniel N. Conrad, Nayara C. Leite, Dipikaa Akshinthala, Christine Maria Lim, Raul Gonzalez, Lakshmi B. Muthuswamy, Zev Gartner, and Senthil K. Muthuswamy. 2021. "Commitment and Oncogene-Induced Plasticity of Human Stem Cell-Derived Pancreatic Acinar and Ductal Organoids." *Cell Stem Cell* 28(6):1090-1104.e6. doi: 10.1016/j.stem.2021.03.022.

Hwang, Katie L., and Wolfram Goessling. 2016. "Baiting for Cancer: Using the Zebrafish as a Model in Liver and Pancreatic Cancer." Pp. 391–410 in *Advances in Experimental Medicine and Biology*. Vol. 916. Springer New York LLC.

Ishibashi, Minaka, Alejandro S. Mechaly, Thomas S. Becker, and Silke Rinkwitz. 2013. "Using Zebrafish Transgenesis to Test Human Genomic Sequences for Specific Enhancer Activity." *Methods* 62(3):216–25. doi: 10.1016/j.ymeth.2013.03.018.

- Jennings, Rachel E., Andrew A. Berry, James P. Strutt, David T. Gerrard, and Neil A. Hanley. 2015. "Human Pancreas Development." *Development (Cambridge)* 142(18):3126–37.
- Kemmler, Cassie L., Hannah R. Moran, Brooke F. Murray, Aaron Scoresby, John R. Klem, Rachel L. Eckert, Elizabeth Lepovsky, Sylvain Bertho, Susan Nieuwenhuize, Sibylle Burger, Gianluca D'Agati, Charles Betz, Ann Christin Puller, Anastasia Felker, Karolina Ditrychova, Seraina Bötschi, Markus Affolter, Nicolas Rohner, C. Ben Lovely, Kristen M. Kwan, Alexa Burger, and Christian Mosimann. 2023. "Next-Generation Plasmids for Transgenesis in Zebrafish and Beyond." *Development (Cambridge)* 150(8). doi: 10.1242/DEV.201531.
- Kimmel, Robin A., and Dirk Meyer. 2010. *Molecular Regulation of Pancreas Development in Zebrafish*. Vol. 100.
- Kleinjan, Dirk Jan, and Pedro Coutinho. 2009. "Cis-Ruption Mechanisms: Disruption of Cis-Regulatory Control as a Cause of Human Genetic Disease." *Briefings in Functional Genomics and Proteomics* 8(4):317–32.
- Kubra, Khadizatul, Gurveer K. Gaddu, Clifford Liongue, Somayyeh Heidary, Alister C. Ward, Amardeep S. Dhillon, and Faiza Basheer. 2022. "Phylogenetic and Expression Analysis of Fos Transcription Factors in Zebrafish." *International Journal of Molecular Sciences* 23(17). doi: 10.3390/ijms231710098.
- Kwan, Kristen M., Esther Fujimoto, Clemens Grabher, Benjamin D. Mangum, Melissa E. Hardy, Douglas S. Campbell, John M. Parant, H. Joseph Yost, John P. Kanki, and Chi Bin Chien. 2007. "The Tol2kit: A Multisite Gateway-Based Construction Kit for Tol2 Transposon Transgenesis Constructs." *Developmental Dynamics* 236(11):3088–99. doi: 10.1002/dvdy.21343.
- Kwiatkowska, Iwona, Justyna Magdalena Hermanowicz, Zaneta Iwinska, Krystyna Kowalczyk, Jolanta Iwanowska, and Dariusz Pawlak. 2022. "Zebrafish – An Optimal Model in Experimental Oncology." *Molecules* 27(13).
- De La Calle-Mustienes, Elisa, Cármen Gloria Feijóo, Miguel Manzanares, Juan J. Tena, Elisa Rodríguez-Seguel, Annalisa Letizia, Miguel L. Allende, and José Luis Gómez-Skarmeta. 2005. "A Functional Survey of the Enhancer Activity of Conserved Non-Coding Sequences from Vertebrate Iroquois Cluster Gene Deserts." *Genome Research* 15(8):1061–72. doi: 10.1101/gr.4004805.
- Larsen, Hjalte List, and Anne Grapin-Botton. 2017. "The Molecular and Morphogenetic Basis of Pancreas Organogenesis." *Seminars in Cell and Developmental Biology* 66:51–68.
- Lettice, Laura A., Simon J. H. Heaney, Lorna A. Purdie, Li Li, Philippe de Beer, B. A. Oostra, Debbie Goode, Greg Elgar, Robert E. Hill, and Esther de Graaff. 2003. "A Long-Range Shh Enhancer Regulates Expression in the Developing Limb and Fin

- and Is Associated with Preaxial Polydactyly." *Human Molecular Genetics* 12(14):1725–35. doi: 10.1093/hmg/ddg180.
- Li, Mingyu, Liyuan Zhao, Patrick S. Page-McCaw, and Wenbiao Chen. 2016. "Zebrafish Genome Engineering Using the CRISPR-Cas9 System." *Trends in Genetics* 32(12):815–27.
- Li, Yongqing, Yuejuan Zhang, Bei He, Yuequn Wang, Zengjin Yuan, Wuzhou Yuan, Peng Liao, Yun Deng, Jing Xiao, Chuanbing Zhu, Ying Wang, Xiushan Wu, and Mingyao Liu. 2007. "Cloning and Expression of a Novel Human Gene, Isl-2, Encoded a LIM-Homeodomain Protein." *Molecular Biology Reports* 34(1):19–26. doi: 10.1007/s11033-006-9003-0.
- Liu, Ying, Amy W. Helms, and Jane E. Johnson. 2004. "Distinct Activities of Msx1 and Msx3 in Dorsal Neural Tube Development." *Development* 131(5):1017–28.
- Lodish, Harvey, Arnold Berk, Paul Matsudaira, Chris A. Kaiser, Monty Krieger, Matthew P. Scott, Lawrence Zipursky, and James Darnell. 2003. *Molecular Cell Biology Fifth Edition*.
- Lorberbaum, David S., Fiona M. Docherty, and Lori Sussel. 2020. "Animal Models of Pancreas Development, Developmental Disorders, and Disease." Pp. 65–85 in *Advances in Experimental Medicine and Biology*. Vol. 1236. Springer.
- Low, Jiun T., Alka Shukla, and Peter Thorn. 2010. "Pancreatic Acinar Cell: New Insights into the Control of Secretion." *International Journal of Biochemistry and Cell Biology* 42(10):1586–89.
- Luo, Juanjuan, Chunjiao Lu, and Xiaojun Yang. 2024. "Protocol for Tissue-Specific Mutagenesis with Fluorescent Labeling in Zebrafish." *STAR Protocols* 5(3). doi: 10.1016/j.xpro.2024.103207.
- Luo, Zhaofan, Yanan Li, Mingxin Zuo, Chang Liu, Dong Yan, Huamin Wang, and Donghui Li. 2017. "Effect of NR5A2 Inhibition on Pancreatic Cancer Stem Cell (CSC) Properties and Epithelial-Mesenchymal Transition (EMT) Markers." *Molecular Carcinogenesis* 56(5):1438–48. doi: 10.1002/mc.22604.
- Marstrand-Daucé, Louis, Diane Lorenzo, Anaïs Chassac, Pascal Nicole, Anne Couvelard, and Cécile Haumaitre. 2023. "Acinar-to-Ductal Metaplasia (ADM): On the Road to Pancreatic Intraepithelial Neoplasia (PanIN) and Pancreatic Cancer." *International Journal of Molecular Sciences* 24(12).
- Masoudi, Sahar, Amin Hassanzadeh Nemati, Hamid Reza Fazli, Samira Beygi, Maliheh Moradzadeh, Akram Pourshams, and Ashraf Mohamadkhani. 2019. "An Increased Level of Aryl Hydrocarbon Receptor in Patients with Pancreatic Cancer." *Middle East Journal of Digestive Diseases* 11(1):38–44. doi: 10.15171/mejdd.2018.126.

- Maston, Glenn A., Sara K. Evans, and Michael R. Green. 2006. "Transcriptional Regulatory Elements in the Human Genome." *Annual Review of Genomics and Human Genetics* 7:29–59.
- Mathelier, Anthony, Wenqiang Shi, and Wyeth W. Wasserman. 2015. "Identification of Altered Cis-Regulatory Elements in Human Disease." *Trends in Genetics* 31(2):67–76.
- Mi, Jiarui, Lipeng Ren, and Olov Andersson. 2024. "Leveraging Zebrafish to Investigate Pancreatic Development, Regeneration, and Diabetes." *Trends in Molecular Medicine*.
- Moreno-Mateos, Miguel A., Charles E. Vejnár, Jean Denis Beaudoin, Juan P. Fernandez, Emily K. Mis, Mustafa K. Khokha, and Antonio J. Giraldez. 2015. "CRISPRscan: Designing Highly Efficient SgRNAs for CRISPR-Cas9 Targeting in Vivo." *Nature Methods* 12(10):982–88. doi: 10.1038/nmeth.3543.
- Mudbhary, Raksha, and Kirsten C. Sadler. 2011. "Epigenetics, Development, and Cancer: Zebrafish Make Their ARK." *Birth Defects Research Part C - Embryo Today: Reviews* 93(2):194–203.
- Mularoni, Loris, Mireia Ramos-Rodríguez, and Lorenzo Pasquali. 2017. "The Pancreatic Islet Regulome Browser." *Frontiers in Genetics* 8(FEB). doi: 10.3389/fgene.2017.00013.
- Mundade, Rasika, Hatice Gulcin Ozer, Han Wei, Lakshmi Prabhu, and Tao Lu. 2014. "Role of ChIP-Seq in the Discovery of Transcription Factor Binding Sites, Differential Gene Regulation Mechanism, Epigenetic Marks and Beyond." *Cell Cycle* 13(18):2847–52.
- Narlikar, Leelavati, and Ivan Ovcharenko. 2009. "Identifying Regulatory Elements in Eukaryotic Genomes." *Briefings in Functional Genomics and Proteomics* 8(4):215–30.
- Natoli, Gioacchino, and Jean Christophe Andrau. 2012. "Noncoding Transcription at Enhancers: General Principles and Functional Models." *Annual Review of Genetics* 46:1–19.
- Nelson, Shannon R., Sandra Roche, Maura Cotter, Pablo Anton Garcia, Daniela Reitmeier, Elisabeth Zollbrecht, Fiona O'Neill, Martin Clynes, Padraig Doolan, Jai P. Medha, Niall Swan, Anne Marie Larkin, and Naomi Walsh. 2020. "Genomic Profiling and Functional Analysis of Let-7c miRNA-mRNA Interactions Identify SOX13 to Be Involved in Invasion and Progression of Pancreatic Cancer." *Journal of Oncology* 2020. doi: 10.1155/2020/2951921.
- Pan, Fong Cheng, and Marcela Brissova. 2014. "Pancreas Development in Humans." *Current Opinion in Endocrinology, Diabetes and Obesity* 21(2):77–82.
- Park, Joon Tae, and Steven D. Leach. 2018. "Zebrafish Model of KRAS-Initiated Pancreatic Cancer." *Animal Cells and Systems* 22(6):353–59. doi: 10.1080/19768354.2018.1530301.
- Pennacchio, Len A., Wendy Bickmore, Ann Dean, Marcelo A. Nobrega, and Gill Bejerano. 2013. "Enhancers: Five Essential Questions." *Nature Reviews Genetics* 14(4):288–95.

- Petersen, Gloria M., Laufey Amundadottir, Charles S. Fuchs, Peter Kraft, Rachael Z. Stolzenberg-Solomon, Kevin B. Jacobs, Alan A. Arslan, H. Bas Bueno-De-Mesquita, Steven Gallinger, Myron Gross, Kathy Helzlsouer, Elizabeth A. Holly, Eric J. Jacobs, Alison P. Klein, Andrea Lacroix, Donghui Li, Margaret T. Mandelson, Sara H. Olson, Harvey A. Risch, Wei Zheng, Demetrius Albanes, William R. Bamlet, Christine D. Berg, Marie Christine Boutron-Ruault, Julie E. Buring, Paige M. Bracci, Federico Canzian, Sandra Clipp, Michelle Cotterchio, Mariza De Andrade, Eric J. Duell, J. Michael Gaziano, Edward L. Giovannucci, Michael Goggins, Göran Hallmans, Susan E. Hankinson, Manal Hassan, Barbara Howard, David J. Hunter, Amy Hutchinson, Mazda Jenab, Rudolf Kaaks, Charles Kooperberg, Vittorio Krogh, Robert C. Kurtz, Shannon M. Lynch, Robert R. McWilliams, Julie B. Mendelsohn, Dominique S. Michaud, Hemang Parikh, Alpa V. Patel, Petra H. M. Peeters, Aleksandar Rajkovic, Elio Riboli, Laudina Rodriguez, Daniela Seminara, Xiao Ou Shu, Gilles Thomas, Anne Tjønneland, Geoffrey S. Tobias, Dimitrios Trichopoulos, Stephen K. Van Den Eeden, Jarmo Virtamo, Jean Wactawski-Wende, Zhaoming Wang, Brian M. Wolpin, Herbert Yu, Kai Yu, Anne Zeleniuch-Jacquotte, Joseph F. Fraumeni, Robert N. Hoover, Patricia Hartge, and Stephen J. Chanock. 2010. "A Genome-Wide Association Study Identifies Pancreatic Cancer Susceptibility Loci on Chromosomes 13q22.1, 1q32.1 and 5p15.33." *Nature Genetics* 42(3):224–28. doi: 10.1038/ng.522.
- Pinho, Andreia V., Ilse Rooman, Maximilian Reichert, Nele De Medts, Luc Bouwens, Anil K. Rustgi, and Francisco X. Real. 2011. "Adult Pancreatic Acinar Cells Dedifferentiate to an Embryonic Progenitor Phenotype with Concomitant Activation of a Senescence Programme That Is Present in Chronic Pancreatitis." *Gut* 60(7):958–66. doi: 10.1136/gut.2010.225920.
- Rauluseviciute, Ieva, Rafael Riudavets-Puig, Romain Blanc-Mathieu, Jaime A. Castro-Mondragon, Katalin Ferenc, Vipin Kumar, Roza Berhanu Lemma, Jérémy Lucas, Jeanne Chèneby, Damir Baranasic, Aziz Khan, Oriol Fornes, Sveinung Gundersen, Morten Johansen, Eivind Hovig, Boris Lenhard, Albin Sandelin, Wyeth W. Wasserman, François Parcy, and Anthony Mathelier. 2024. "JASPAR 2024: 20th anniversary of the Open-Access Database of Transcription Factor Binding Profiles." *Nucleic Acids Research* 52(D1):D174–82. doi: 10.1093/nar/gkad1059.
- Ren, Fei, Cheng Ji, Yujie Huang, Stanley Aniagu, Yan Jiang, and Tao Chen. 2020. "AHR-Mediated ROS Production Contributes to the Cardiac Developmental Toxicity of PM2.5 in Zebrafish Embryos." *Science of the Total Environment* 719. doi: 10.1016/j.scitotenv.2019.135097.
- Riethoven, Jean Jack M. 2010. "Regulatory Regions in DNA: Promoters, Enhancers, Silencers, and Insulators." *Methods in Molecular Biology (Clifton, N.J.)* 674:33–42.
- Roman, Tamara S., Maren E. Cannon, Swarooparani Vadlamudi, Martin L. Buchkovich, Brooke N. Wolford, Ryan P. Welch, Mario A. Morken, Grace J. Kwon, Arushi Varsh-

- ney, Romy Kursawe, Ying Wu, Anne U. Jackson, Michael R. Erdos, Johanna Kuusisto, Markku Laakso, Laura J. Scott, Michael Boehnke, Francis S. Collins, Stephen C. J. Parker, Michael L. Stitzel, and Karen L. Mohlke. 2017. "A Type 2 Diabetes-Associated Functional Regulatory Variant in a Pancreatic Islet Enhancer at the ADCY5 Locus." Pp. 2521–30 in *Diabetes*. Vol. 66. American Diabetes Association Inc.
- Sandelin, Albin, Wynand Alkema, Pär Engström, Wyeth W. Wasserman, and Boris Lenhard. 2004. "JASPAR: An Open-Access Database for Eukaryotic Transcription Factor Binding Profiles." *Nucleic Acids Research* 32(DATABASE ISS.). doi: 10.1093/nar/gkh012.
- Sandhu, Nikita, Satyavati Rana, and Kiran Meena. 2021. "Nuclear Receptor Subfamily 5 Group A Member 2 (NR5A2): Role in Health and Diseases." *Molecular Biology Reports* 48(12):8155–70.
- Schaffer, Ashleigh E., Kristine K. Freude, Shelley B. Nelson, and Maïke Sander. 2010. "Nkx6 Transcription Factors and Ptf1a Function as Antagonistic Lineage Determinants in Multipotent Pancreatic Progenitors." *Developmental Cell* 18(6):1022–29. doi: 10.1016/j.devcel.2010.05.015.
- Schones, Dustin E., Kairong Cui, Suresh Cuddapah, Tae Young Roh, Artem Barski, Zhibin Wang, Gang Wei, and Keji Zhao. 2008. "Dynamic Regulation of Nucleosome Positioning in the Human Genome." *Cell* 132(5):887–98. doi: 10.1016/j.cell.2008.02.022.
- Shashikant, Tanvi, and Charles A. Ettensohn. 2019. "Genome-Wide Analysis of Chromatin Accessibility Using ATAC-Seq." Pp. 219–35 in *Methods in Cell Biology*. Vol. 151. Academic Press Inc.
- Shih, Hung Ping, Allen Wang, and Maïke Sander. 2013. "Pancreas Organogenesis: From Lineage Determination to Morphogenesis." *Annual Review of Cell and Developmental Biology* 29:81–105.
- Shlyueva, Daria, Gerald Stampfel, and Alexander Stark. 2014. "Transcriptional Enhancers: From Properties to Genome-Wide Predictions." *Nature Reviews Genetics* 15(4):272–86.
- Soboleski, Mark R., Jason Oaks, and William P. Halford. 2005. *Green Fluorescent Protein Is a Quantitative Reporter of Gene Expression in Individual Eukaryotic Cells*.
- Sondermann, Natalie C., Sonja Faßbender, Frederick Hartung, Anna M. Hätälä, Katharina M. Rolfes, Christoph F. A. Vogel, and Thomas Haarmann-Stemmann. 2023. "Functions of the Aryl Hydrocarbon Receptor (AHR) beyond the Canonical AHR/ARNT Signaling Pathway." *Biochemical Pharmacology* 208.
- Storz, Peter. 2017. "Acinar Cell Plasticity and Development of Pancreatic Ductal Adenocarcinoma." *Nature Reviews Gastroenterology and Hepatology* 14(5):296–304.

- Strobel, Oliver, Yuval Dor, Janivette Alsina, Amy Stirman, Gregory Lauwers, Amanda Trainor, Carlos Fernández Del Castillo, Andrew L. Warshaw, and Sarah P. Thayer. 2007. "In Vivo Lineage Tracing Defines the Role of Acinar-to-Ductal Transdifferentiation in Inflammatory Ductal Metaplasia." *Gastroenterology* 133(6):1999–2009. doi: 10.1053/j.gastro.2007.09.009.
- Suster, Maximiliano L., Hiroshi Kikuta, Akihiro Urasaki, Kazuhide Asakawa, and Koichi Kawakami. 2009. "Transgenesis in Zebrafish with the Tol2 Transposon System." *Methods in Molecular Biology* 561:41–63. doi: 10.1007/978-1-60327-019-9_3.
- Taminato, Tomohito, Daisuke Yokota, Soh Araki, Hiroki Ovara, Kyo Yamasu, and Akinori Kawamura. 2016. "Enhancer Activity-Based Identification of Functional Enhancers Using Zebrafish Embryos." *Genomics* 108(2):102–7. doi: 10.1016/j.ygeno.2016.05.005.
- Tarifeño-Saldivia, Estefania, Arnaud Lavergne, Alice Bernard, Keerthana Padamata, David Bergemann, Marianne L. Voz, Isabelle Manfroid, and Bernard Peers. 2017. "Transcriptome Analysis of Pancreatic Cells across Distant Species Highlights Novel Important Regulator Genes." *BMC Biology* 15(1). doi: 10.1186/s12915-017-0362-x.
- Tehrani, Zahra, and Shuo Lin. 2011. "Endocrine Pancreas Development in Zebrafish." *Cell Cycle* 10(20):3466–72.
- Tiso, Natascia, Enrico Moro, and Francesco Argenton. 2009. "Zebrafish Pancreas Development." *Molecular and Cellular Endocrinology* 312(1–2):24–30.
- Uttley, Kirsty, Andrew S. Papanastasiou, Manuela Lahne, Jennifer M. Brisbane, Ryan B. MacDonald, Wendy A. Bickmore, and Shipra Bhatia. 2023. "Unique Activities of Two Overlapping PAX6 Retinal Enhancers." *Life Science Alliance* 6(11). doi: 10.26508/lsa.202302126.
- Venkat, Swati, Abdulrahman A. Alahmari, and Michael E. Feigin. 2021. "Drivers of Gene Expression Dysregulation in Pancreatic Cancer." *Trends in Cancer* 7(7):594–605.
- Van De Venter, Maryna, Jenske Didloff, Shanika Reddy, Bresler Swanepoel, Sharlene Govender, Ntokozo Shirley Dambuza, Saralene Williams, Trevor Craig Koekemoer, and Luanne Venables. 2021. "Wild-Type Zebrafish (Danio Rerio) Larvae as a Vertebrate Model for Diabetes and Comorbidities: A Review." *Animals* 11(1):1–21.
- Vincent, Audrey, Joseph Herman, Rich Schulick, Ralph H. Hruban, and Michael Goggins. 2011. "Pancreatic Cancer." Pp. 607–20 in *The Lancet*. Vol. 378. Elsevier B.V.
- Wan, Haiyan, Svitlana Korzh, Zhen Li, Sudha Puttur Mudumana, Vladimir Korzh, Yun Jin Jiang, Shuo Lin, and Zhiyuan Gong. 2006. "Analyses of Pancreas Development by Generation of Gfp Transgenic Zebrafish Using an Exocrine Pancreas-Specific ElastaseA Gene Promoter." *Experimental Cell Research* 312(9):1526–39. doi: 10.1016/j.yexcr.2006.01.016.

- Wang, Chengqi, Michael Q. Zhang, and Zhihua Zhang. 2013. "Computational Identification of Active Enhancers in Model Organisms." *Genomics, Proteomics and Bioinformatics* 11(3):142–50.
- Wilfinger, Armin, Valeriya Arkhipova, and Dirk Meyer. 2013. "Cell Type and Tissue Specific Function of Islet Genes in Zebrafish Pancreas Development." *Developmental Biology* 378(1):25–37. doi: 10.1016/j.ydbio.2013.03.009.
- Wilson, Leslie, and Paul Matsudaira. n.d. *Series Editors*.
- Yee, Nelson S., Kristin Lorent, and Michael Pack. 2005. "Exocrine Pancreas Development in Zebrafish." *Developmental Biology* 284(1):84–101. doi: 10.1016/j.ydbio.2005.04.035.
- Zhang, Xunlei, Dongying Gu, Mulong Du, Meilin Wang, Chunxiang Cao, Lili Shen, Meng Kuang, Yongfei Tan, Xinying Huo, Weida Gong, Zhi Xu, Jinfei Chen, Zhengdong Zhang, and Cuiju Tang. 2014. "Associations of NR5A2 Gene Polymorphisms with the Clinicopathological Characteristics and Survival of Gastric Cancer." *International Journal of Molecular Sciences* 15(12):22902–17. doi: 10.3390/ijms151222902.
- Zhao, Zhi Yu, and Wei Liu. 2020. "Pancreatic Cancer: A Review of Risk Factors, Diagnosis, and Treatment." *Technology in Cancer Research and Treatment* 19.



2024

Ines Mota

UNRAVELLING NR5A2 TRANSCRIPTIONAL REGULATION IN PANCREATIC CANCER