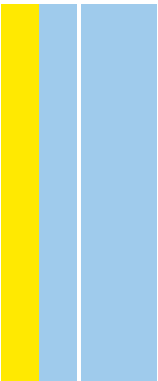


DOUTORAMENTO
BIOLOGIA MOLECULAR E CELULAR

Cis-regulatory elements controlling the identity of pancreatic progenitors

João Pedro Curado Agra Amorim

D
2024

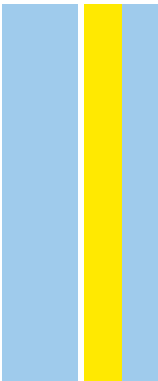


João Pedro Curado Agra Amorim
Cis-regulatory elements controlling the identity of pancreatic progenitors



D.ICBAS 2024

Cis-regulatory elements controlling the identity of pancreatic progenitors
João Pedro Curado Agra Amorim



João Pedro Curado Agra Amorim

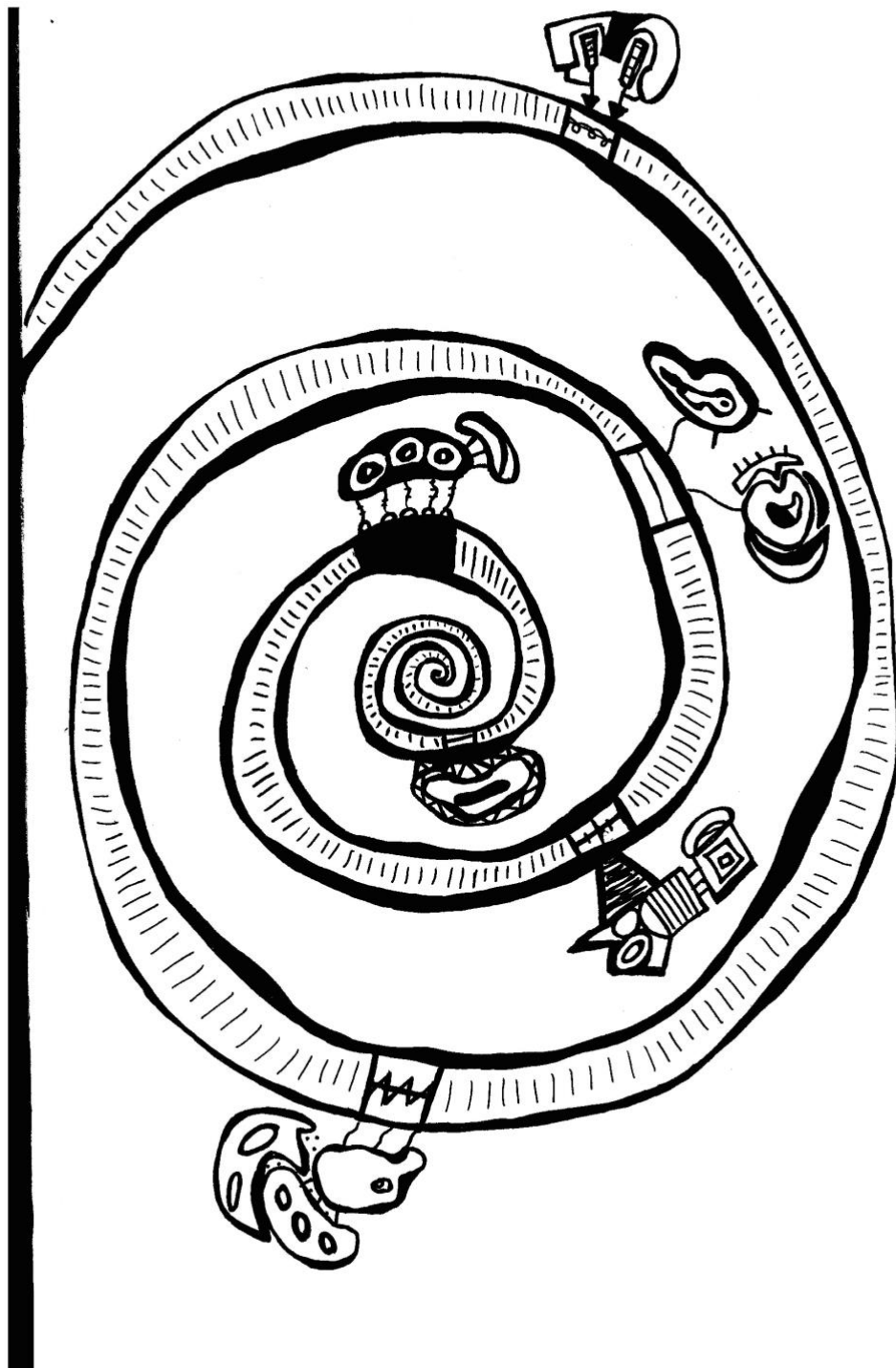
CIS-REGULATORY ELEMENTS CONTROLLING THE IDENTITY OF PANCREATIC PROGENITORS IN VERTEBRATES

Tese de Candidatura ao grau de Doutor em Biologia Molecular e Celular; Programa Doutoral da Universidade do Porto (Instituto de Ciências Biomédicas Abel Salazar e Faculdade de Ciências).

Orientador – Doutor José Carlos Ribeiro Bessa

Categoria– Investigador Principal

Afiliação – Instituto de Biologia Molecular e Celular (IBMC), Instituto de Investigação e Inovação em Saúde (i3S) da Universidade do Porto.



Life as A Neverending Regulatory Cycle – Original hand drawing

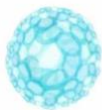
Financial support

This work was financially supported by:

- PhD fellowship SFRH/BD/145110/2019, Fundação para a Ciência e Tecnologia (FCT).
- Programa Graduado em Biologia Molecular e Celular, Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto (ICBAS-UP), Faculdade de Ciências da Universidade do Porto (FCUP).
- European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-2015-StG-680156-ZPR).



ICBAS | INSTITUTO DE CIÊNCIAS
BIOMÉDICAS ABEL SALAZAR
SCHOOL OF MEDICINE AND
BIOMEDICAL SCIENCES



DOCTORAL PROGRAM:
Molecular and Cell Biology
Programa Doutoral em Biologia Molecular e Celular



FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO



IBMC

INSTITUTO DE BIOLOGIA MOLECULAR E CELULAR
INSTITUTE FOR MOLECULAR AND CELL BIOLOGY



INSTITUTO DE INVESTIGAÇÃO
E INOVAÇÃO EM SAÚDE
UNIVERSIDADE DO PORTO

Acknowledgements/Agradecimentos

Agora que se aproxima o fim deste percurso pelos confins do Conhecimento, da Biologia e da própria Vida, chega a altura de agradecer a todos os meus companheiros de viagem e pessoas sem as quais esta não seria possível.

Ao meu orientador, José Bessa, por todos os ensinamentos partilhados e a oportunidade de realizar este trabalho no grupo de investigação VDR. A sua incessante energia e encorajamento para aprender mais e, sobretudo, melhor, resultaram numa aquisição de novos conhecimentos e apetências que considero muito importantes para o meu crescimento pessoal e profissional.

Aos meus colegas e amigos do grupo VDR (os que ainda permanecem e outros que por lá passaram), por estarem sempre disponíveis para partilhar conhecimento, discutir ideias e ajudar na resolução de problemas, sem esquecer os momentos de convívio. Considero ter partilhado o espaço de trabalho com pessoas gentis, de bom coração e espírito de entreajuda. São estes a Ana Gali, Joana Teixeira, Joana Marques, Ana Eufrásio, Marta Duque, Fábio Ferreira, Mafalda Galhardo, Chiara Perrod, Pablo Cabezas-Sainz, Isabel Guedes, Francisco Cadete, Renata Carriço, Leonor Carvalho, Hugo Marcelino, Vítor Silva, Francisco Magalhães, João Nogueira, Sofia Borges, Diogo Sousa, Miguel Sousa, Inês Mota e Nuno Oliveira. Existem também outros membros do i3S a quem não posso deixar de agradecer: Filipa Castro, Renata Freitas, Ana Sofia Figueiredo, Joana Loureiro, Ana Mafalda Rocha, Rob Mensik, Catarina Meireles, Emília Cardoso, Paula Magalhães, Paula Sampaio, Maria Azevedo e María Lázaro.

À minha família e amigos, que me ajudaram a elevar o espírito nas partes acidentadas do caminho. Um agradecimento especial ao João Coelho, Filipe Nunes Vicente e João Manuel Matos, amigos de longa data, também doutorandos ou recém-doutorados, pelas conversas filosóficas e perspectiva fora da caixa.

À Mariana, pelo apoio e companhia durante este percurso.

À minha mãe, por sempre me ter apoiado nas minhas escolhas e por ajudar a fomentar um espírito crítico e dinâmico.

Ao meu pai que, tendo partido cedo, deixou ensinamentos e memórias sem os quais não me teria aventurado tão longe.

Publication list

Original research article, published

Amorim, J.P.*, Macedo-Gali, A.*, Marcelino, H., Bordeira-Carriço, R., Naranjo, S., Rivero-Gil, S., Teixeira, J., et al. (2020) A Conserved Notochord Enhancer Controls Pancreas Development in Vertebrates. *Cell Rep.* 32:107862. doi: 10.1016/j.celrep.2020.107862

*Equal contribution as first author

Methodology article, published

Amorim, J.P., Bordeira-Carriço, R., Gali-Macedo, A., Perrod, C., and Bessa, J. (2020). CRISPR-Cas9-Mediated Genomic Deletions Protocol in Zebrafish. *STAR Protocols* 1, 100208. 10.1016/j.xpro.2020.100208.

Review article, published

Duque, M.*, **Amorim, J.P.***, and Bessa, J.* (2021). Ptf1a function and transcriptional cis-regulation, a cornerstone in vertebrate pancreas development. *The FEBS Journal* 289, 5121–5136. 10.1111/febs.16075.

*Equal contribution as first author

Original research article, in preparation

Duque, M*., **Amorim, J.P.***, Galhardo, M. and Bessa, J. A distal enhancer regulates Ptf1a to ensure proper lineage acquisition and maintenance.

Resumo

O desenvolvimento de identidades celulares em organismos multicelulares é possível, em parte, porque células diferentes utilizam informação regulatória distinta. Esta informação está presente em elementos cis-regulatórios (ECRs). A desregulação destes elementos por mutação ou variação nucleotídica pode ter impacto na doença humana. O pâncreas é um órgão central para doenças como diabetes ou cancro pancreático e, embora os mecanismos envolvidos no desenvolvimento do pâncreas e da doença pancreática sejam conhecidos, o papel dos ECRs permanece indefinido. No Capítulo II desta tese, estudamos ECRs que controlam o desenvolvimento pancreático do peixe-zebra, mais especificamente a regulação do gene *nog2*, que codifica um conhecido inibidor da via de sinalização BMP. Os nossos resultados mostram que Nog é expresso na notocorda, controlado transversalmente em humano, ratinho e peixe-zebra por, pelo menos, um ECR específico de notocorda, e que o ganho e perda de função de *nog2* se correlacionam com um aumento e diminuição da diferenciação pancreática, respectivamente. Estes resultados indicam que Nog é uma molécula de sinalização pró-pancreática que difunde da notocorda para o domínio progenitor pancreático. No Capítulo III, exploramos os mecanismos de regulação da expressão de *ptf1a* no pâncreas de peixe-zebra, com foco num ECR específico, z3'-DpE e sua dupla função, em células progenitoras multipotentes (CPMs) e células acinares diferenciadas. Observamos alterações no padrão e localização do domínio CPM em embriões portadores de uma deleção de z3'-DpE. Além disso, realizamos RNA-seq em CPMs na presença e ausência de z3'-DpE e observamos enriquecimento de componentes da via de sinalização Notch no grupo de genes desregulados, sugerindo que Ptf1a pode ativar um programa de sinalização Notch em CPMs. Finalmente, realizamos ATAC-seq em células acinares, na presença e ausência de z3'-DpE. Descobrimos que regiões com acessibilidade reduzida em mutantes e motivos de factores de transcrição enriquecidos nessas regiões estão associados a genes envolvidos na identidade acinar. Por outro lado, regiões com maior acessibilidade e motivos enriquecidos nestas regiões estão associados a genes relacionados com inflamação e cancro pancreático. Estes resultados destacam o papel crucial de Ptf1a na diferenciação de CPMs e na manutenção da identidade das células acinares, que por sua vez pode prevenir perfis de expressão genética aberrante que podem levar a transformação oncogénica.

As descobertas relatadas nesta tese realçam o papel crucial das ECRs no desenvolvimento do pâncreas e na manutenção da identidade celular diferenciada. Além disso, os nossos resultados suportam a hipótese de que a disrupção de ECRs pancreáticas pode contribuir para o desenvolvimento de cancro pancreático.

Abstract

The development of cellular identities in multicellular organisms is possible, in part, because different cells use distinct cis-regulatory information. This information is present in non-coding cis-regulatory elements (CREs) such as transcriptional enhancers. Dysregulation of regulatory dynamics by mutation or nucleotide variation at the level of CREs might impact human disease. The pancreas is a central organ for diseases such as diabetes or pancreatic cancer and, although the general mechanisms of pancreatic development and disease are fairly known, the role of enhancers in these processes remains elusive. In Chapter II of this thesis, we focused on regulatory elements controlling early pancreatic development in zebrafish, more specifically in CREs controlling the expression of the *nog2* gene, which encodes for a known inhibitor of the BMP signaling pathway. We show that *Nog* is expressed in the notochord, controlled transversely in humans, mice, and zebrafish by at least one notochord enhancer, and that *nog2* gain and loss of function correlates with an increase and a decrease of pancreatic differentiation, respectively. Our results indicate that *Nog* is a crucial pro-pancreatic signal diffusing from the notochord to the dorsal bud-derived pancreatic progenitor domain. In Chapter III, we have explored the mechanisms of regulation of *ptf1a* expression in zebrafish, focusing on a distal enhancer, z3'-DpE and its dual function, early in ventral bud-derived multipotent progenitor cells (MPCs) and late in differentiated acinar cells. We observed alterations in the pattern and localization of the MPC domain in embryos harboring a deletion of z3'-DpE. Furthermore, we performed RNA-seq of z3'-DpE^{+/+} and ^{-/-} MPCs and found downregulated genes to be enriched in Notch signaling pathway components, suggesting that *Ptf1a* might activate a Notch signaling program in MPCs. Finally, we performed ATAC-seq of z3'-DpE^{+/+} and ^{-/-} acinar cells. We found regions with decreased accessibility in z3'-DpE^{-/-} and enriched transcription factor (TF) motifs within these to be associated with genes involved in acinar cell identity. On the other hand, regions with increased accessibility and enriched TF motifs within these were associated with genes related to inflammation and pancreatic cancer. These results emphasize the crucial role of *ptf1a* in proper differentiation of MPCs and in the maintenance of acinar cell identity, which in turn might prevent aberrant gene expression leading to oncogenic transformation.

Altogether, the discoveries reported in this doctoral thesis emphasize the crucial role of enhancers in pancreas development and maintenance of differentiated cell identity. Moreover, our findings support the notion that disruption of enhancers might contribute to the development of pancreatic cancer, highlighting the importance of cis-regulation in the process of pancreatic dysfunction and transformation.

Table of Contents

Financial support	v
Acknowledgements/Agradecimientos	vii
Resumo.....	xi
Abstract.....	xiii
Table of Contents	xv
Abbreviations.....	xvii
Chapter I – General Introduction: Cis-regulation and pancreatic development in vertebrates.....	1
1.1 Transcriptional regulation in developmental biology	4
1.1.1 Promoters	4
1.1.2 Silencers.....	5
1.1.3 Insulators	5
1.1.3. Enhancers	5
1.2 Vertebrate pancreas development	6
1.2.1 Pancreas patterning and specification	6
1.2.2 Intercellular signals involved in pancreas development	8
1.2.3 Transcription Factors involved in pancreas specification	10
1.2.4 Endocrine pancreas determination in mammals	11
1.3. Zebrafish Pancreas Development.....	12
1.3.1 Dorsal bud	13
1.3.2 Ventral Bud.....	14
1.4. The role of Ptf1a in pancreas development	16
1.5. Mechanisms of regulation of Ptf1a expression	20
1.6. Research Aims.....	25
1.7. References	26
Chapter II: A conserved notochord enhancer controls pancreas development in vertebrates.....	35
2.1. Introduction.....	38
2.2. Results.....	40
2.2.1 <i>nog2</i> controls non-autonomously the embryonic endocrine pancreas size .	40
2.2.2 <i>nog2</i> genomic regulatory landscape contains a notochord-specific enhancer	43
2.2.3 The <i>nog2</i> notochord enhancer is required for proper pancreas development	47
2.2.4 The expression of <i>Nog</i> in the notochord is conserved in vertebrates and is mediated by equivalent notochord enhancers	49
2.3. Discussion	51
2.4. Methods.....	55
2.4.1. Materials Availability.....	55
2.4.2. Data and Code Availability	55
2.4.3. Experimental Model and Subject Details	55
2.4.4. Method Details.....	56

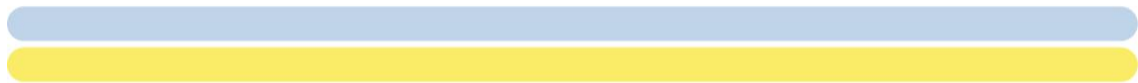
2.5. Supplementary Information	61
2.6 Supplementary datasets	68
2.7. References	69
Chapter III: A distal enhancer regulates <i>ptf1a</i> to ensure proper lineage acquisition and maintenance in pancreatic acinar cells	75
3.1. Introduction	78
3.2. Results and Discussion	82
3.2.1. z3'-DpE enhancer is active in early MPCs and persists in differentiated exocrine cells	82
3.2.2. Deletion of the z3'-DpE enhancer results in an earlier separation of the z3'-DpE-active and <i>nkx6.1</i> -expressing domains and a posterior-oriented elongation of the <i>nkx6.1</i> expression domain	85
3.2.3. Analysis of the transcriptome of MPCs upon deletion of z3'-DpE reveals a <i>ptf1a</i> -dependent developmental gene network	86
3.2.4. Differentially accessible regions in acinar cells of z3'-DpE ^{-/-} larvae indicate compromised acinar identity, concomitant with an upregulation of inflammation and pancreatic cancer associated genes	89
3.3. Methods	97
3.3.1. Experimental Model and Subject Details	97
3.3.2. Method Details	97
3.3.3. Processing and bioinformatics analysis	99
3.4. Supplementary Information	102
3.5. Supplementary datasets	110
3.6. References	148
Chapter IV	153
4.1. Concluding remarks	155
4.2. References	159

Abbreviations

ARX – Aristaless-related homeobox
AP-1 – Activator protein 1
ASCL1 – Achaete-scute family BHLH transcription factor 1
ATAC-seq – Assay for transposase-accessible chromatin using sequencing
BAC – Bacterial artificial chromosome
BHLH – Basic helix-loop-helix
BHLHA15 – Basic helix-loop-helix family member A15
BMP – Bone morphogenetic protein
BP – Base pair
CAC – Centroacinar cell
CAS9 – CRISPR-associated protein 9
CDX4 – Caudal type homeobox 4
ChIP-seq – Chromatin immunoprecipitation followed by sequencing
CPA1 – Carboxypeptidase A1
CRE – Cis-regulatory element
CRISPR – Clustered regularly interspaced short palindromic repeats
CS – Carnegie stage
CTCF – CCCTC-binding factor
DCAS9 – Endonuclease dead Cas9
DNA – Deoxyribonucleic acid
E – Embryonic
EPC – Endocrine progenitor cell
FC – Fold change
FGF – Fibroblast growth receptor
FOXA2 – Forkhead box A2
GATA4 – GATA-binding factor 4
GFP – Green fluorescent protein
GO – Gene ontology
GWAS – Genome-wide association study
HESC – Human embryonic stem cells
HNF1B - Hepatocyte nuclear factor 1 homeobox B
H3K4me1 – Monomethylation of lysine 4 on histone H3
H3K27ac – Acetylation of lysine 27 on histone H3
H3'-DPE – Human ptf1a distal pancreatic enhancer
IDR – Irreproducible discovery rate
KLF10 – Kruppel like factor 10
KRAB – Kruppel-associated box domain
HPF – Hours post fertilization

KRAS – Kirsten ras homolog
 MAF – Musculoaponeurotic fibrossarcoma
 MPC – Multipotent progenitor cell
 MO – Morpholino
 M5'-AR – Mouse ptf1a auto-regulatory enhancer
 NEUROD1 – Neurogenic differentiation 1
 NEUROG3 – Neurogenin 3
 NKX2.2 – Homeobox protein 2.2
 NKX6.1 – Homeobox protein 6.1
 NOG – Noggin
 NR5A2 – Nuclear Receptor Subfamily 5 Group A Member 2
 PANIN – Pancreatic intraepithelial neoplasia
 PAX4 – Paired box 4
 PDAC – Pancreatic ductal adenocarcinoma
 PDX1 – Pancreatic and duodenal homeobox 1
 PNC – Pancreatic notch-responsive cell
 POU3F4 – POU class 3 homeobox 4
 PTF1A – Pancreas-specific transcription factor 1A
 RA – Retinoic acid
 RNA – ribonucleic acid
 RNA-seq – RNA sequencing
 RT – Room temperature
 SGRNA – Single-guide RNA
 SHH – Sonic hedgehog
 SMAD - Mothers against decapentaplegic
 SOX9 – Sry-box9
 SST - Somatostatin
 TAD – Topologically associated domain
 TCF3/4/12 – Transcription factor 3/4/12
 TF – Transcription factor
 TFBS – Transcription factor binding site
 TRE – Trans-regulatory element
 TSS – Transcription starting site
 WNT – Wingless-related integration site
 WT – Wild-type
 Z3'-DPE – Zebrafish ptf1a distal pancreatic enhancer
 Z3'-EA – Zebrafish ptf1a early-acting enhancer
 Z5'-AR – Zebrafish ptf1a auto-regulatory enhancer
 4C-seq - Circular chromosome conformation capture coupled with next-generation sequencing

Chapter I – General Introduction



Cis-regulation and pancreatic development in vertebrates

Part of this chapter is adapted from a review article, published as:

Duque, M. †, **Amorim, J.P.** † and Bessa, J. †* (2021). Ptf1a function and transcriptional cis-regulation, a cornerstone in vertebrate pancreas development. The FEBS Journal 289, 5121–5136. 10.1111/febs.16075.

†These authors contributed equally to this work

*Corresponding author.

Author contributions:

MD, JA and JB writing the original draft, review, and editing. JB supervision and funding acquisition.

1.1 Transcriptional regulation in developmental biology

Cells from the same organism have identical genomes but different transcriptional profiles, implying the existence of transcriptional regulatory information acquired in different cell types and states. This regulatory information plays a major role in driving the faithful execution of major biological processes such as development, proliferation, aging or differentiation, which require precise spatial and temporal expression of genes (Isshiki et al, 2001; Maduro & Rothman, 2002; Pourquié, 2003). Particularly, in multicellular organisms, development of different cellular identities is dependent on this regulatory information, which is mostly present in non-coding cis-regulatory elements (CREs) of the genome.

The eukaryotic transcription machinery, driven mostly by RNA-polymerase II (Pol II), is based on two complementary regulatory components: cis and trans-regulatory elements. Cis-regulatory elements (CREs) are DNA sequences spread over large distances, in coding or non-coding regions of the genome. Trans-regulatory elements (TREs) refers mainly to transcription factors (TFs), which are DNA-binding proteins. These proteins recognize and bind specific sequences within CREs to initiate, enhance or suppress transcription (Maston G.A. et al, 2006; Ong & Corces, 2011). This interaction, modulated by many epigenetic process, is thought to be very dynamic and complex. CREs are controlled by the competitive or cooperative binding of different transcription factors. TREs and CREs can regulate transcription synergistically with the help of transcriptional co-factors, RNA-binding proteins, non-coding RNAs epigenetic and chromatin modifications (Levine, M., & Tjian, R. 2003; Maurano, M.T. et al, 2012). CREs can be classified in two distinct families: promoters, that include a core and proximal promoter regions that reside within 1 kb of the transcription starting site (TSS) and distal regulatory elements, containing silencers, insulators and enhancers (Maston, G.A. et al, 2006; Epstein, D.J. et al 2009; Yang & Corces, 2011).

1.1.1 Promoters

The core promoter locates at the start of a gene and serves as the docking site for the transcriptional machinery. Also, it defines the position of the transcription starting site (TSS) and direction of transcription (Smale S.T. et al 2003). The first described core promoter element was the TATA box, the binding site for the tata-binding protein (TBP) subunit of the eukaryotic general transcription factor (TFIID). Besides the TATA box, metazoan core promoters can be composed of numerous other elements such as: Initiator element (Inr), Downstream Promoter Element (DPE), Downstream Core Element (DCE), TFIIB- Recognition Element (BRE), and Motif Ten Element (MTE) (Maston G.A. et al 2006). The proximal promoter is defined as the region immediately

upstream from the core promoter and typically contains multiple binding sites for activator proteins.

1.1.2 Silencers

Silencers have a negative effect on the transcription of a target gene. These typically function independently of orientation and distance from the promoter. They can be part of a proximal promoter, a distal enhancer or function as an independent distal regulatory module. These sequences may cooperate in binding to DNA, and they can act synergistically. TFs called repressors bind to specific sequences contained in silencers, thus being the functional effectors on the process of silencing gene expression (Maston G.A. et al 2006). Several models have been proposed to explain repressor function. Repressors might block the binding of a nearby activator (Harris, M.B. et al 2005), or directly compete for the same site (Li, L. et al 2004). Alternatively, a repressor may prevent activators from accessing promoters by favoring a repressive chromatin state through recruitment of histone-modifying proteins or chromatin-stabilizing factors (Srinivasan, L. et al 2004).

1.1.3 Insulators

The function of insulators is to limit the action of other CREs to defined domains, dividing the genome into regulatory partitions. They mostly function by blocking genes from being affected by regulatory elements of neighboring genes. Insulators can block enhancer-promoter interaction or prevent the spreading of repressive chromatin (heterochromatin). Although orientation is not necessary for insulator function, their position is crucial to establish domains of regulation. In vertebrates, the only known protein to mediate insulator activity is CTCF (CCCTC-binding factor) (Maston G.A. et al 2006). Currently, two models of insulator function exist: the first explanation is simply that an insulator-bound activator cannot interact with its target promoter. The second model associates insulators with the structural organization of chromatin, proposing that insulators interact with each other and can tether multiple elements to the same foci, resulting in formation of isolated chromatin loops (Capelson, M. et al 2004).

1.1.3. Enhancers

Enhancers typically regulate transcription in a spatial- and/or temporal-specific manner, independently of both the distance from and orientation relative to the promoter. These elements are usually modular, meaning that gene expression driven by a single promoter can be regulated by distinct enhancer elements at different times or in different tissues (Atchinson, M.L. et al 1988). Enhancers are usually composed of a cluster of transcription factor binding sites (TFBS) and the TFs that bind to these sequences are described to work cooperatively to boost transcription (Maston G.A. et al 2006).

Enhancers can be very distal to their target gene promoters. Their ability to still exert their function in these cases can be explained by a DNA-looping model, where the enhancer and promoter are brought into proximity by DNA loops. Enhancers thus play a major role in controlling the levels of gene expression in different cells in a time-dependent manner (Tijian, R. & Maniatis, T., 1994).

Cell-specific expression of TFs, through different steps of development, partially defines cell identity and, ultimately, drives differentiation and tissue compartmentalization. Transitions between several cellular states occur in the developing embryo as multipotent progenitor cells differentiate to generate tissue-specific adult cells (Duque et al, 2022; Zhou et al, 2007). This phenomenon must be tightly controlled, both spatially and temporally, most likely by regulation of gene expression through the activity of non-coding CREs, signaling pathways based on diffusible molecules and cell-cell interactions. Understanding these dynamics is crucial to further our comprehension on the field of developmental biology but also to explain developmental defects underlying human disease.

1.2 Vertebrate pancreas development

The pancreas is a central organ for human diseases such as pancreatic cancer and diabetes, both having a dramatic societal burden. Functionally, it can be separated into two major compartments. The exocrine pancreas includes digestive enzyme-secreting acinar cells arranged into clusters at the end of ducts. Mature duct cells actively secrete bicarbonate and mucins and drain acinar digestive enzymes towards the duodenum (Slack, J.M.W, 1995). The endocrine pancreas secretes hormones into the circulation and is responsible for regulating glucose metabolism. It is arranged in islets containing hormone-producing α cells (glucagon), β cells (insulin), δ cells (somatostatin) and e cells (ghrelin) (Argenton et al, 1999; Baetens et al, 1979; Brissova et al, 2005). The islets of Langerhans are intermingled with blood vessels, neurons, and a mesodermally-derived stromal component. The interaction between endocrine and vascular cells is responsible for the precise regulation of hormone release, which is crucial for the maintenance of glucose homeostasis (Gittes, 2009).

1.2.1 Pancreas patterning and specification

Pancreas development in vertebrates is an intricate process through which various cell lineages develop from multipotent progenitor cells (MPCs) that differentiate into different cell types of the endocrine and exocrine pancreas, organizing themselves to form a single organ (Zhou et al, 2007; Kinkel & Prince, 2009). In mouse, the flat sheet

of endoderm folds to form the primitive gut tube, which is divided into foregut, midgut and hindgut (Jørgensen et al, 2007; Pan & Wright, 2011). The dorsal foregut endoderm evaginates into the surrounding mesenchyme at E9.5 (Embryonic day 9.5) in the mouse (Pan & Wright, 2011), while in humans this phenomenon occurs around Carnegie Stage 13 (CS13) (Jennings et al, 2015; Sambathkumar et al, 2018) (Figure 1.1). This is followed by the emergence of the anlage of the ventral pancreas plus common bile duct, from the ventral foregut endoderm. Mouse pancreas formation is divided in two overlapping waves of development. The first wave is referred to as the primary transition and occurs between E9.5–E12.5 (Pan & Wright, 2011). The first differentiated endocrine cells begin to appear in the dorsal bud at this phase of development (Pictet et al, 1972; Herrera, 2000). Contrasting with this, an early endocrine differentiation corresponding to the mouse primary transition has not been found in human embryos (Villasenor et al, 2010; Jennings et al, 2013, 2015).

Around E11.5 in the mouse and CS15 in humans, the gut tube undergoes a rotation that brings the dorsal and ventral buds into proximity (Pan & Wright, 2011; Sambathkumar et al, 2018). This contact between dorsal and ventral buds results in their fusion to form a single organ (Pan & Wright, 2011; Gittes, 2009) (Figure 1.1). After this, the densely-packed epithelium starts to undergo active plexus remodeling, continued epithelial expansion and production of more plexus, sending finger-like protrusions into the mesenchyme, and over the entire organ (Pan & Wright, 2011).

Starting at E13.5 in the mouse, the epithelium undergoes a morphogenetic transformation in which cells differentiate massively and lineage allocation into the main pancreatic lineages occurs (Pan & Wright, 2011). This is called the secondary transition and is a common phenomenon between human and mouse. In this phase of development, compartmentalization of the pancreatic epithelium occurs, resulting in two distinct cellular domains, referred to as tip and trunk. While the tip domain contains progenitor cells that initially express multipotency markers but later become committed to an acinar fate, the adjacent trunk epithelial region consists of an endocrine-duct bipotential progenitor pool (Jennings et al, 2015; Zhou et al, 2007). These domains are distinguished by the expression of specific transcription factors (Zhou et al, 2007). After the secondary transition, at E16.5, the epithelium expands further and this is thought to be largely driven by acinar proliferation (Pan & Wright, 2011).

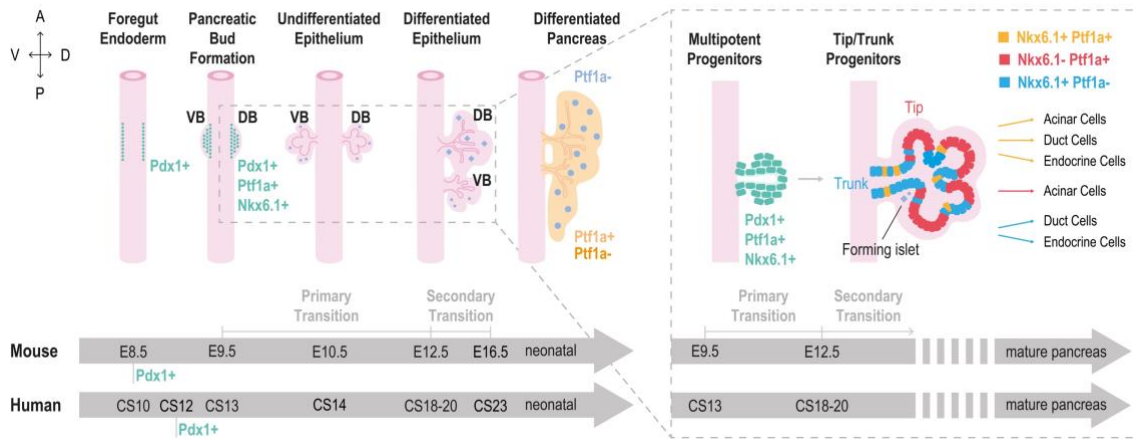
Placental Mammal (human and mouse)

Figure 1.1 - Representation of the major morphogenetic events of pancreatic development in human (Jennings et al, 2015; Sambathkumar et al, 2018) and mouse (Schaffer et al, 2010; Jørgensen et al, 2007; Pan & Wright, 2011), with corresponding developmental times (gray arrows). Primary and secondary transitions (Jørgensen et al, 2007; Pan & Wright, 2011; Gittes, 2009) are annotated. Expression of key MPC TF- encoding genes, as *Ptf1a*, *Pdx1* and *Nkx6.1*, is annotated. In the top right corner, orange cells represent the progressively restricted domain of *Nkx6.1*- and *Ptf1a*-positive cells that will give rise to mutually exclusive domains. Dashed lines highlight an overview of the main developmental processes during primary and secondary transitions in mouse and human. A, anterior; CACs, centroacinar cells; D, dorsal; DB, dorsal pancreatic bud; E, embryonic day; P, posterior; V, ventral; VB, ventral pancreatic duct.

1.2.2 Intercellular signals involved in pancreas development

During vertebrate pancreas development, several gradients of signaling molecules are established, many emanating from adjacent tissues, providing positional cues for the developing pancreas. By binding to cell surface receptors, these ligands restrict the field of MPCs, culminating in the expression of specific TF-encoding genes such as *Pdx1*, *Nkx6.1* or *Ptf1a* (Zhou et al, 2007; Hald et al, 2008; Schaffer et al, 2010). Some of the signaling pathways that are known to be crucial for pancreatic development in vertebrates are Hedgehog signaling, RA (retinoic acid) signaling, Wnt signaling, BMP signaling or the Notch signaling pathway (Kim & Hebrok, 2001).

Mesenchymal cells are non-autonomous factors playing an important role in pancreatic development. Among the signals produced by mesenchymal cells are BMPs emanating from the lateral plate mesoderm, Fgf10, retinoic acid (RA), and Wnt (Bhushan, 2001; Maden, 2001; Chung et al, 2008; Rodriguez-Seguel et al, 2013). Fgf10 is not required for initial bud formation but is important for the proliferation of *Pdx1*⁺ MPCs (Bhushan, 2001; Norgaard et al, 2003). In *Xenopus*, exposure to soluble Wnt5a induced MPC gene expression, specifically *Pdx1* and *Ptf1a* (Rodriguez-Seguel et al, 2013). Supporting the important role of Wnt signaling in the identity of MPCs, it was

observed that cis-regulatory modules active in human MPCs are associated with Wnt signaling target genes (Cebola et al, 2015), in agreement with the pro-pancreatic role observed in mouse (Rodriguez-Seguel et al, 2013). The RA signaling pathway is required for proper pancreatic specification in zebrafish (Stafford & Prince, 2002), being repressed by *Cdx4* in the posterior endoderm (Kinkel et al, 2008) (Figure 1.2).

In most vertebrates, the foregut endoderm is located adjacently to the notochord, an organ that plays a crucial role in pancreatic development (Kim et al, 1997; Jennings et al, 2013; Amorim et al, 2020). In mice, Activin (bB) and FGF2 emanating from the notochord have been shown to repress *Shh* expression in the endoderm (Hebrok et al, 1998). However, ectopic pancreatic tissue is not observed in *Shh* mutant embryos (Hebrok et al, 2000), suggesting that Shh inhibition is permissive, rather than instructive, for pancreas fate (McCracken & Wells, 2012). Experiments of notochord ablation in chicken embryos showed that pancreatic development is impaired (Kim et al, 1997). In Chapter II of the present doctoral thesis, focused on signaling pathways involved in zebrafish pancreatic development, more specifically on *Nog2*, a diffusible molecule known to inhibit Bmp signaling (Fürthauer et al, 1999). We show that *Nog* is expressed in the notochord, controlled transversely in humans, mice, and zebrafish by at least one notochord enhancer, and that *nog2* gain and loss of function in zebrafish correlates with an increase and a decrease of pancreatic differentiation, respectively (Amorim et al, 2020). These results indicate that *Nog* is one of the crucial pro-pancreatic signaling molecules emanating from the notochord (Figure 1.2). In agreement with this, *Nog* is a well-known Bmp signaling inhibitor (Fürthauer et al, 1999) and this signaling pathway is known to antagonize pancreas differentiation (Chung et al, 2008, 2010) while promoting liver differentiation. We postulate that, by counteracting Bmp signaling, *Nog2* diffusion from the notochord to the pancreatic progenitor domain allows proper development of the pancreas (Amorim et al, 2020).

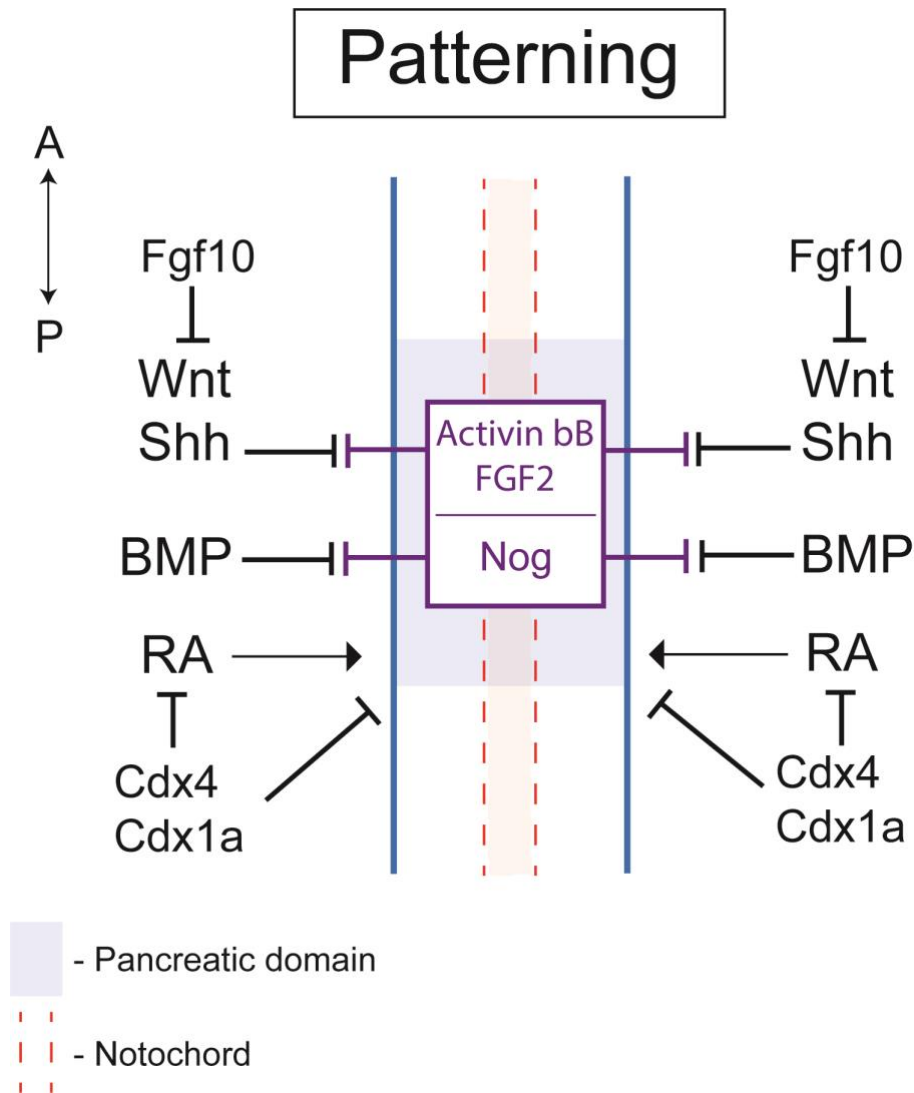


Figure 1.2 - Signaling pathways involved in vertebrate pancreas patterning and specification. Fgf10 is important for the proliferation of pancreatic progenitor cells (Bhushan, 2001)] and interacts with both Wnt and Shh signaling, establishing a cross-regulation by feedback loops (Watson & Francavilla, 2018). Wnt signaling has been observed to promote pancreatic fate, by expression of progenitor markers Pdx1 and Ptf1a (Rodríguez-Seguel et al, 2013). Activin (bB) and FGF2 are released from the notochord and regulate pancreas development, by repression of endodermal Shh (Hebrok et al, 2000). Another signal from the notochord, Nog, inhibits BMPs emanating from the lateral plate mesoderm, controlling the pancreas size and location (Amorim et al, 2020). RA signaling is known to be crucial for pancreatic development (Stafford & Prince, 2002) and the interaction between RA and Cdx factors is thought to be important for correct AP patterning (Kinkel et al, 2008).

1.2.3 Transcription Factors involved in pancreas specification

Pancreatic progenitor cells are commonly defined by the expression of specific TFs across vertebrates. Human multipotent pancreatic progenitor cells (MPCs) show expression of *PTF1A* (Cebola et al, 2015) and its function is crucial for pancreas development (Weedon, Michael N. et al, 2014). MPCs in human ventral and dorsal

pancreatic buds are also marked by expression of *PDX1*, *NKX6.1*, *SOX9*, and *GATA4* encoding TFs (Jennings et al, 2015), essential for pancreatic specification, defining the MPC cell population. These cells are responsible for the formation of the dorsal and ventral pancreatic buds, during initial steps of pancreatic morphogenesis. After this, as differentiation starts, sub-groups of progenitor cells appear at different locations. Central duct-like structures (composed of trunk cells) are involved in the formation of duct and endocrine cells, whereas more peripheral clustered cells (tip cells) differentiate to form acinar cells of the exocrine pancreas (Jennings et al, 2015). Trunk cells are known to express *NKX6.1* and *SOX9* and to have less *GATA4* expression, while tip cells remain *NKX6.1*⁺/*SOX9*⁺/*GATA4*⁺ (Jennings et al, 2015) (Figure 1.1). Shortly after, tip cells lose *NKX6.1* expression (Jennings et al, 2013), suggesting an early segregation of the *PTF1A*⁺ acinar compartment (Schaffer et al, 2010; Esni, 2004; Solar et al, 2009), which initiates prior to the major wave of endocrine differentiation (Jennings et al, 2015). The segregation of *PTF1A* and *NKX6.1* expression into mutually exclusive domains suggests that a mutual repression mechanism is triggered just previous to endocrine and exocrine specification. Indeed, it has been suggested that through co-repression, *Nkx6.1* and *Ptf1a* function as antagonistic lineage determinants in MPCs, in an equilibrium that governs endocrine versus exocrine fate decision (Schaffer et al, 2010) (Figure 1.1).

In mouse, MPCs have been characterized by the expression of several TF-encoding genes: *Ptf1a*, *Pdx1*, *Nkx6.1*, *Sox9*, *Nkx2.2*, *Hnf1b*, and *Cpa1* (Zhou et al, 2007; Solar et al, 2009). During secondary transition, similarly to human, tip cells of the branching epithelium adopt an acinar fate, being marked by the expression of *Ptf1a* and *Cpa1* (Zhou et al, 2007; Solar et al, 2009), while cells in the trunk become restricted to a ductal or endocrine fate and are characterized by the expression of *Nkx6.1* and *Hnf1b* (Zhou et al, 2007; Solar et al, 2009).

1.2.4 Endocrine pancreas determination in mammals

As MPCs segregate into endocrine and exocrine cellular compartments, they acquire expression of specific TF-encoding genes. The dynamics of this process is controlled by Notch signaling (Apelqvist et al, 1999). MPCs are under high levels of Notch signaling; however, when this pathway is inactivated, they become committed to an endocrine fate, becoming endocrine progenitor cells (EPCs) and activating genes encoding TFs required for endocrine differentiation (Apelqvist et al, 1999; Zhu et al, 2006). In mouse, progenitor cells within the central duct-like structures acquire

expression of *Ngn3* (Gradwohl et al, 2000; Schwitzgebel et al, 2000) and these constitute the EPC population. Importantly, in human development, *NEUROG3* expression coincides with the appearance of the first fetal beta cells (Jennings et al, 2015). *SOX9* is one other important player in pancreas development. Its expression in MPCs of the human dorsal and ventral pancreatic buds has been observed to start around Carnegie stage 12 (CS12) (Jennings et al, 2013). It is absent from EPCs expressing *NEUROG3* and from differentiated endocrine cells. However, its expression persists in pancreatic duct cells (Jennings et al, 2013).

Finally, EPCs differentiate into several hormone-producing cells, forming the pancreatic islets: alpha-cells—expressing glucagon, beta-cells—expressing insulin, delta-cells—expressing somatostatin, and epsilon-cells—expressing ghrelin. Preceding differentiation, intermediate states exist where specific TFs control specific endocrine cell fates. Examples of such TFs are *Pax4*, *Arx*, *Mafa*, *Mafb*, *Foxa2*, or *Pou3f4* (Dassaye et al, 2016). Loss of either *Pax4* or *Arx* in mouse has been observed to not affect the total number of endocrine cells but instead changing the relative distribution of endocrine subtypes (Sosa-Pineda et al, 1997; Collombat, 2005).

1.3. Zebrafish Pancreas Development

While many studies on the genetics of pancreatic function and development have been done using the mouse model system, the accessibility and optical clarity of zebrafish embryos allows cell manipulation and sophisticated *in vivo* imaging approaches, which makes the zebrafish a powerful tool to study the pancreas, especially in a developmental setting. Furthermore, the zebrafish pancreas shares a basic structure and cellular makeup with the mammalian pancreas, implying that the data acquired in this model is broadly applicable to the study of the molecular and patterning processes required for development of this organ (Kinkel & Prince, 2009). In pancreatic MPCs, specific TFs are expressed, helping to define how and when these cells will differentiate. Although they are not identical, much of the pancreatic developmental program seems to be conserved between mice and zebrafish, since many key TFs have been found to be expressed in the pancreatic primordium of both species, performing similar functions (Kinkel & Prince, 2009).

The zebrafish pancreas develops from two distinct anlagen arising from foregut endoderm and containing multipotent progenitor cells, called dorsal and ventral buds. The dorsal bud generates the first wave of endocrine cells, clustering at 24 hours post fertilization (hpf) to form the principal islet (Field et al, 2003; Ward et al, 2007) (Figure

1.3). The ventral bud emerges anteriorly to the dorsal bud around 32 hpf and gives rise to acinar, ductal, and a second wave of endocrine cells (Wang et al, 2015). Differentiation of these later-appearing endocrine cells has been proposed to be equivalent to the secondary transition in mammals (Hesselson et al, 2009) and occurs between 48 and 120 hours post fertilization (Pisharath et al, 2007). The secondary islets appear along the pancreatic tail, which forms by growing in a posterior direction, after the envelopment of the principal islet occurs (Figure 1.3).

Similar to what is seen in mouse, TF-encoding genes such as *pdx1*, *nkx6.1* and *ptf1a* are expressed in zebrafish MPCs (Burlison et al, 2008; Ghaye et al, 2015), although the genetic identity of this cellular population is not yet fully understood (Figure 1.3, Figure 1.4). Part of these cells have been reported to be under high levels of Notch signaling. However, when this pathway is inactivated, they transit from a MPC to an EPC state, activating genes encoding TFs required for endocrine differentiation, such as *ascl1b* and *neurod1* (Apelqvist et al, 1999; Zhu et al, 2006; Flasse et al, 2013).

1.3.1 Dorsal bud

The liver and pancreas anlagen have been postulated to derive from a common hepatopancreatic domain within the foregut endoderm of the zebrafish (Chung et al, 2008). In this study, the authors found that, within the monolayer of foregut endoderm, the most medial cells contribute to endocrine pancreas, those lying just 2 cell-widths lateral contribute to liver, and the intermediate zone contributes to exocrine pancreas. This pattern of cell fates correlates with a gradient of *pdx1* expression, where its highest expression is seen in cells fated to become endocrine. These results imply that cells expressing this TF have the capacity to give rise to all pancreatic cell types, while a gradient of Pdx1 seems to dictate the specific fate of these cells. Importantly, *pdx1* starts to be expressed at 14hpf in zebrafish (Biemar et al, 2001) and, in the mouse, it is known to be also expressed in multipotent pancreatic progenitors (Jørgensen et al, 2007).

In mouse pancreatic development, the early-forming dorsal bud contains a mix of endocrine and exocrine progenitors, being able to give origin to both domains of the pancreas. However, this does not happen in the zebrafish, where the dorsal bud exclusively gives rise to endocrine cell types, generating a single principal islet in the first 48 hours of development (Kinkel et al, 2008; Kinkel & Prince, 2009). Also, whereas in the mouse both pancreatic buds require Ptf1a for differentiation (Kawaguchi et al, 2002), the zebrafish protein seems to be only required for the later developing ventral bud (Lin et al, 2004) (Figure 1.3, Figure 1.4). Furthermore, while *Ngn3* is expressed in

endocrine progenitors of the mouse, this is not the case in the zebrafish, where its role has been proposed to be replaced by two others: *ascl1b* and *neurod1* (Flasse et al, 2013). The striking fact that, in the zebrafish, expression of *ascl1b* starts at 10 hpf (Flasse et al, 2013), being the first known pancreatic TF-encoding gene to be expressed, before *pdx1*, suggests that the first cells in the pancreatic domain directly acquire an endocrine identity. This could also mean that *ascl1b*⁺ cells have the capacity to generate multiple pancreatic lineages, although this has not been formally demonstrated.

Another key factor for pancreatic development is the homeobox TF Nkx6.1. In the mouse, it is expressed in multipotent progenitors (Zhou et al, 2007) and, in the zebrafish, *nkx6.1* is expressed early in the pancreatic primordium of the dorsal bud starting at 11.5 hpf (Binot, A-C. et al, 2010). Together, cells expressing these TFs likely represent the initial population of multipotent progenitors, responsible for the formation of the principal islet, deriving from the dorsal bud (Figure 1.3, Figure 1.4).

1.3.2 Ventral Bud

The principal islet is located in the head of the pancreas in the larval zebrafish at 5 days post fertilization (dpf) and it has been suggested that the intrapancreatic duct is composed entirely of Notch-responsive cells – PNCs (Beer et al, 2016). The majority of the endocrine lineage arises from these cells, during an event called the secondary transition, when secondary islets appear along the duct in the tail of the pancreas (Parsons et al, 2009). During this transition, blocking of Notch activity leads to a fate switch of these cells into endocrine progenitors – EPCs (Parsons et al, 2009; Dalgin & Prince, 2015). The Notch signaling pathway controls the differentiation of pancreatic cells both in zebrafish and mice. In mouse, it controls endocrine cell differentiation through the repression of Ngn3. In zebrafish, the control of endocrine cell fate is fulfilled by *Ascl1b* and *NeuroD* TFs, both being repressed by Notch signaling (Flasse et al, 2013; Ghaye et al, 2015). These transcription factors have been reported to define committed endocrine precursor cells for both the dorsal and ventral bud. Following the envelopment of the principal islet, the exocrine part of the pancreas, which is composed of acinar cells, grows in a posterior direction to form the pancreatic tail, which have been shown to derive from *ptf1a*⁺ cells that do not respond to Notch signaling (Wang et al, 2015) (Figure 1.3, Figure 1.4).

It has been suggested that the zebrafish ventral bud contains two distinct progenitor cell populations: a population of Pancreatic Notch-responsive cells (PNCs) and a *ptf1a*-expressing population (Wang et al, 2011) (Figure 1.3, Figure 1.4). Ghaye

and colleagues (Ghaye et al, 2015) hypothesized that Notch signaling responsiveness could be the key factor in the segregation of these cells into endocrine/ductal ($nkx6.1^+/ptf1a^+$ /Notch on) or acinar cells ($nkx6.1^+/ptf1a^+$ /Notch off). In this study, it was hypothesized that $nkx6.1^+$ cells progressively become PNCs, since initiation of $nkx6.1$ expression is independent of Notch signaling (Ghaye et al, 2015). Moreover, after transient co-expression of $nkx6.1$ and $asc11b$ in the pancreatic anlagen, these cells segregate in two different domains, one expressing both genes and the other only $nkx6.1$. Through loss and gain-of-function experiments, the same authors have observed that Notch signaling works to maintain $nkx6.1$ expression in PNCs, while repressing $asc11b$. Therefore, only when Notch is inactivated, Notch- responsive cells transit into an endocrine progenitor state. Altogether, these results allow two different interpretations for how ventral MPCs are defined (Figure 1.4). The first is that there is an early segregation of $ptf1a^+$ cells and PNCs that differentiate to form separate domains (exocrine versus endocrine/ductal, respectively), supported by the fact that a unique progenitor domain has not been detected in the zebrafish (Wang et al, 2011). Another view is that MPCs expressing both $nkx6.1$ and $ptf1a$ might progressively become responsive to Notch signaling, consequently repressing $ptf1a$ expression, leading cells to a specific fate—endocrine/ductal (Ghaye et al, 2015). On the other hand, cells retaining high Ptf1a levels (and Notch-unresponsive) differentiate over time to generate the acinar compartment.

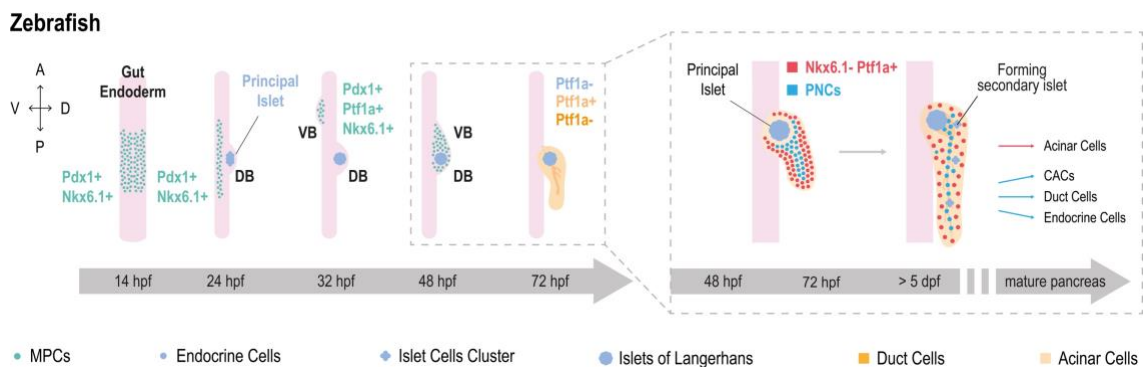


Figure 1.3 - Representation of the major morphogenetic events of pancreatic development in zebrafish (Wang et al, 2011; Ghaye et al, 2015), with corresponding developmental times (gray arrows). The second wave of islet differentiation in zebrafish is annotated ('Forming secondary islets'; (Parsons et al, 2009). Expression of $ptf1a$, as well as other selected key MPCs TF-encoding genes, $pdx1$ and $nkx6.1$, is annotated. A, anterior; CACs, centroacinar cells; D, dorsal; DB, dorsal pancreatic bud; E, embryonic day; P, posterior; V, ventral; VB, ventral pancreatic duct.

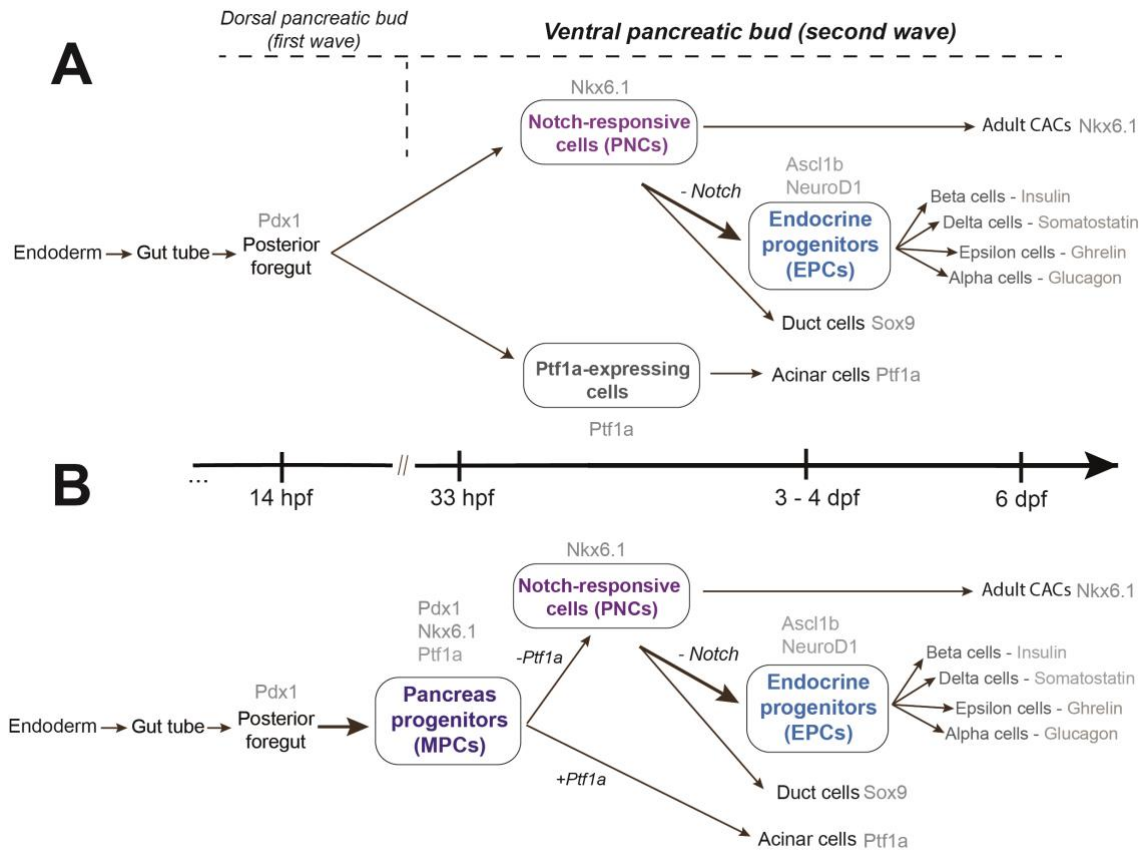


Figure 1.4 – Pancreatic development of the zebrafish, showing two suggested models. Key TFs involved in this process are shown. (A) In (Wang et al, 2011), the authors suggest that there is an early segregation of *ptf1a*-expressing cells and PNCs that differentiate to form separated domains of the pancreas (exocrine versus endocrine/ductal, respectively). (B) Another view, supported by the data in (Ghaye et al, 2015), suggests that multipotent progenitors expressing both *nkx6.1* and *ptf1a* might progressively become responsive to Notch signaling, then starting to repress *ptf1a* expression, leading cell differentiation to a specific fate—endocrine/ductal. In both, the progenitor cells that are known to be maintained in the adult zebrafish pancreas are also represented. These are called CACs—centroacinar cells (Delaspre et al, 2015).

1.4. The role of *Ptf1a* in pancreas development

Ptf1a cDNA was first isolated from a rat exocrine pancreatic cell line and was considered as a transcriptional regulator of digestive enzymes such as amylase and elastase, in adult acinar cells (Krapp et al, 1996). In mouse, *Ptf1a* is expressed in early pancreatic MPCs during the primary transition (Kawaguchi et al, 2002) and later at the secondary transition, where it is restricted to MPC pro-acinar progenitors in the tip epithelium (Zhou et al, 2007). It is also expressed in differentiated acinar cells but not in ductal or endocrine tissue (Krapp et al, 1996). *Ptf1a* is a bHLH TF that forms the trimeric PTF1 complex with two other proteins: An E protein and Rbpj, or its pancreas-restricted paralogue Rbpjl (Beres et al, 2006) (Figure 1.5). In mouse early development, *Ptf1a* is expressed in the MPC domain, along with other TF-encoding genes, as described

above, such as *Pdx1* and *Nkx6.1* (Jonsson, J. et al, 1994; Kawaguchi et al, 2002; Zhou et al, 2007; Schaffer et al, 2010). Rbpj/Ptf1a functions at these earlier stages, activating Rbpjl expression that, later on during acinar cell development, replaces Rbpj function favoring the Rbpjl/Ptf1a complex (Masui et al, 2007, 2008) (Figure 1.5). Unsurprisingly, mice encoding the Ptf1a^{w298a} mutant protein that does not bind to Rbpj (Beres et al, 2006), but does to Rbpjl, phenocopy null mutants for Ptf1a, including absence of pancreatic ventral bud and delay of dorsal bud growth (Masui et al, 2007). Additionally, in Rbpj-deficient pancreata, amylase-expressing acini and islets are formed during late embryonic and postnatal development, suggesting an essential role of Rbpj in early but not late development.

Rbpj can function by a Notch-dependent and Notch-independent way, in the latter case co-binding to Ptf1a (Masui et al, 2007). These functions are non-overlapping, being the Notch dependent related with the maintenance of a progenitor state, inhibiting ectopic endocrine progression (Apelqvist et al, 1999; Fujikura et al, 2006), while the Ptf1a-bound Notch-independent function is required for the growth and morphogenesis of pancreatic epithelia (Masui et al, 2007). Moreover, the inactivation of Notch1 and Notch2 do not inhibit pancreatic development, suggesting that these are not essential for pancreatogenesis, contrary to Rbpj (Nakhai et al, 2008). These results suggest that the early function of Rbpj is more linked with the activity of Ptf1a than to Notch signaling.

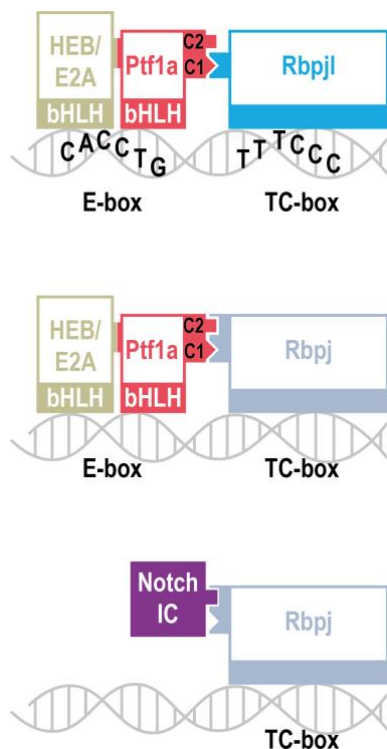


Figure 1.5 - The PTF1 complex. Ptf1a interacts in a trimer complex with an E protein (HEB/E2A) and Rbpjl (top panel) or Rbpj (middle panel), binding to a DNA E-box and a TC-box. Rbpj can operate in a Notch-independent way, by interacting with Ptf1a (middle panel), or in a Notch-dependent way, by interacting with the Notch intracellular domain (Notch IC; bottom panel).

Ventral and dorsal buds have different requirements for early development. In mouse, loss-of-function of *Ptf1a*, *Pdx1*, or double loss-of-function, ablates ventral bud development. However, the dorsal bud is yet able to develop alpha and beta cells (Burlison et al, 2008). In mutants for *Ptf1a*, the dorsal bud generates an extremely reduced endocrine pancreas with a reduced beta cell mass (Kawaguchi et al, 2002; Burlison et al, 2008; Fukuda et al, 2008) but acinar cells are never observed (Krapp et al, 1998; Kawaguchi et al, 2002). Regarding the ventral pancreas, lineage tracing in *Ptf1a* mutant background shows labeling in gut cells, where normally *Ptf1a* is not expressed, suggesting that *Ptf1a* is required for the ventral pancreatic bud identity (Kawaguchi et al, 2002). Additionally, independent double lineage tracing for *Ptf1a*- and *Pdx1*-expressing cells showed that the vast majority of pancreatic cells are derived from double *Ptf1a*⁺ and *Pdx1*⁺ cells. Although some *Pdx1*⁺ *Ptf1a*- derived beta cells were found, these had lower levels of insulin, suggestive of a lesser mature state, compared to the *Pdx1*⁺/*Ptf1a*⁺ derived beta cells, highlighting the importance of *Ptf1a* in the development of mature beta cells (Chen et al, 2019). This role in MPC identity is also corroborated by experiments of ectopic transient expression of *Ptf1a* in mouse embryonic stem cells, moderately inducing all lineages of pancreas development, including mature beta cells (Nair et al, 2014). Additionally, Ptf1a likely contributes to the precise specification of pancreas, since *Ptf1a*-derived cells are detected only in the pancreas, contrasting with *Pdx1*-derived cells, found in the mouse pancreas, gastric antrum, and duodenum (Chen et al, 2019).

After this important early contribution to pancreas development, *Ptf1a* gets restricted to the tip of pancreatic branches formed during the secondary transition (Figure 1.1). Lineage tracing experiments of *Ptf1a*-expressing cells show labeling in all acinar and most of duct and islets cells (Kawaguchi et al, 2002). Conditional lineage tracing shows that at the primary transition, *Ptf1a*⁺ cells give rise to even parts of endocrine, duct, and acinar pancreas. At E13, already in the secondary transition, tip MPCs give rise mainly to acinar cells, with increasing proportions at E14 and E15, becoming almost fully committed to acinar progenitors (Pan et al, 2013). This change in the multipotent potential of MPCs coincide with the mutually exclusive domains of expression of *Ptf1a* and *Nkx6.1*. Although at the primary transition (E10.5), there is co-

expression of *Nkx6.1* with *Ptf1a* (Schaffer et al, 2010), and at E12.5, *Nkx6.1* is more restricted to the trunk and *Ptf1a* to the tip (Schaffer et al, 2010) (Figure 1.1). By E14.5, when tip cells have fully committed to an acinar fate (Zhou et al, 2007), *Nkx6.1* is almost completely excluded from the tip forming a sharp boundary with *Ptf1a* (Schaffer et al, 2010). This is possible by a mutual repression of *Nkx6.1* and *Ptf1a* that becomes effective in the secondary transition (Schaffer et al, 2010). NR5A2 is a member of the nuclear hormone receptor family (Becker-André, M. et al, 1993) and its expression is controlled by the Ptf1 complex, being important for the formation and maintenance of the MPCs of the secondary transition, converting MPCs to an acinar lineage, and for expansion and differentiation of pre-acinar cells (Hale et al, 2014). After differentiation of acinar cells, *Ptf1a* is required to maintain acinar cell identity in adult mice, since its conditional loss-of-function in acinar cells results in loss of expression of digestive enzymes (Rose et al, 2001; Holmstrom et al, 2011; Hoang et al, 2016). Interestingly, also upon *Ptf1a* loss of function in mouse acinar cells, acinar-to-ductal metaplasia and the induction of a KRAS-permissive, PDAC-like (pancreatic ductal adenocarcinoma) gene expression profile was observed (Krah et al, 2015). More recently, the same authors have shown that sustained *Ptf1a* expression in the mouse pancreas is enough to prevent KRAS-driven tumorigenesis and that re-introduction of *Ptf1a* in pancreatic intraepithelial neoplasia (PanINs) reverts them to quiescent acinar cells (Krah et al, 2020). These results suggest that *Ptf1a* could have a tumor suppressor function in the adult exocrine pancreas and this function might be closely linked with its canonic function of acinar cell identity maintenance.

Observations in human genetic alterations have highlighted that *PTF1A* might have similar functions as described in mouse. A mutation in the *PTF1A* gene, resulting in a protein truncation, leads to cerebellar and pancreatic agenesis, that causes neonatal diabetes (Sellick et al, 2004). Additionally, alterations in pancreas-specific cis-regulatory regions of *PTF1A* induce pancreatic agenesis and neonatal diabetes, without clear cerebellar phenotypes (Weedon, Michael N. et al, 2014; Gabbay et al, 2017; Evliyaoğlu et al, 2018; Demirbilek et al, 2020). These observations suggest that *PTF1A* might have an early role in endocrine and exocrine pancreas development, followed by a strong later requirement for exocrine proper differentiation. This is reinforced by the fact that single-cell RNA-seq in human pancreatic cells shows that *PTF1A*⁺/*SOX9*⁺ of the tip have a MPC-like profile (Villani et al, 2019).

In zebrafish, as in mammals, *ptf1a* plays an important role in acinar cell fate (Lin et al, 2004; Zecchin et al, 2004). The dorsal bud never shows expression of *ptf1a*, giving rise only to the endocrine cells composing the principal islet (Argenton et al, 1999). This

first wave of endocrine differentiation has some similarities to the mice primary transition. Interestingly, mouse embryos mutant for *Ptf1a* show a similar potential for early endocrine differentiation arising from the dorsal bud, in a *Ptf1a*-independent manner (Burlison et al, 2008). In zebrafish, *ptf1a* expression starts at 32hpf in the ventral bud, after endocrine differentiation of the principal islet from the dorsal bud (Lin et al, 2004), in cells that are *pdx1*⁺ and *nkx6.1*⁺ (Dong et al, 2008) (Figure 1.3). Later in development, ventral bud cells migrate in a dorsal and posterior direction enveloping the principal islet (Figure 1.3). Similar to the mammal secondary transition, ventral bud cells give rise to endocrine cells that contribute to the principal islet between 48 and 120 hpf (Pisharath et al, 2007). After 5–6 days of development, more endocrine cells appear along the pancreatic duct, forming secondary islets. In *ptf1a* loss-of-function, acinar cells and secondary islets are not detected, being only present the principal islet, while *ptf1a* lineage-labeled cells were converted into gall bladder and other non-pancreatic cell types (Wang et al, 2015). Additionally, reduced levels of *ptf1a* present in a hypomorphic condition show a delay in ventral pancreas specification, accompanied by an exocrine to endocrine fate switch, suggesting that lower levels of Ptf1a can function, in a cell autonomous manner, to promote endocrine fate, whereas high levels repress it (Dong et al, 2008). Accordingly, reduced Ptf1a dosage has been observed to promote a greater contribution toward non-acinar lineages (Wang et al, 2015).

The vital role of *Ptf1a* in vertebrate pancreas organogenesis is indisputable. *Ptf1a* has an early role in endocrine and exocrine pancreas development, followed by a later requirement for exocrine proper differentiation and maintenance of acinar fate. Importantly, these roles are dependent on the levels of expression of *Ptf1a* in specific cell types of the developing pancreas.

1.5. Mechanisms of regulation of Ptf1a expression

The tissue-specific functions of *Ptf1a* require precise spatiotemporal regulation of its expression levels by the activity of multiple non-coding CREs. In mice, *Ptf1a* levels are maintained by an auto-regulatory enhancer (m5'-AR) located 13.4kb upstream of the *Ptf1a* transcriptional start site (TSS) and conserved among mammals (Masui et al, 2008; Meredith et al, 2009) (Figure 1.6). m5'-AR contains two consensus binding sites for PTF1 validated in vitro (Masui et al, 2008; Meredith et al, 2009) and Ptf1a protein binds to this sequence in mouse embryonic neural tube and adult pancreas (Masui et al, 2008; Meredith et al, 2009, 2013). In mouse transgenic reporter assays, m5'-AR drives reporter expression exclusively in *Ptf1a*⁺ cells (Masui et al, 2008; Meredith et al,

2009; Mona et al, 2017). These results strongly suggest that m5'-AR maintains *Ptf1a* levels through a positive auto-regulatory loop. Additionally, m5'-AR fails to be activated in *Ptf1a* loss-of-function mice (Meredith et al, 2009), showing that its activity requires pre-existing Ptf1a protein. Indeed, in luciferase assays, the enhancer is inactive in *Ptf1a*-cells but is activated when cells are co-transfected with plasmids expressing *Ptf1a* (Masui et al, 2008). Likewise, in chick electroporation assays, reporter expression is restricted to *Ptf1a*-expressing domains but is induced in other domains when *Ptf1a* is ectopically expressed (Meredith et al, 2009; Mona et al, 2017). Thus, pre-existing Ptf1a is required and sufficient to activate m5'-AR. A similar CRE is present in zebrafish (z5'-AR; Figure 1.6A). The z5'-AR is located 3 to 5kb upstream of *ptf1a* TSS, containing three consensus PTF1 binding sites, where the two most proximal are necessary and sufficient to drive reporter expression, suggesting that its activity is PTF1-dependent (Pashos et al, 2013). A bacterial artificial chromosome (BAC) spanning the *ptf1a* locus (Park et al, 2008) and expressing morpholino (MO)-resistant *ptf1a* rescues pancreas development in zebrafish *ptf1a* morphants, but rescue fails when the z5'-AR is deleted (Pashos et al, 2013), further illustrating how the z5'-AR is necessary to maintain the *ptf1a* levels required for pancreas development.

While the elements responsible for the trigger of *Ptf1a* expression in mice are unknown, an early-acting enhancer was identified in zebrafish (z3'-EA). z3'-EA is located 1–6 kb downstream of the TSS and drives reporter expression in the cerebellum and pancreas during early development (Pashos et al, 2013). In the pancreas, reporter expression is first detectable at 34 hpf (Pashos et al, 2013), coinciding with the onset of *ptf1a* expression in pancreatic progenitor cells (Park et al, 2008; Pashos et al, 2013). In contrast, z5'-AR- driven reporter expression is only detectable at 42 hpf (Pashos et al, 2013) (Figure 1.6B,C). Moreover, z3'-EA activity decreases as acinar cell differentiation begins (Pashos et al, 2013), suggesting that z3'-EA is the early trigger of *ptf1a* expression, which is subsequently maintained by the z5'-AR- mediated positive auto-regulatory loop. Deletion of z3'-EA from a BAC encoding MO-resistant *ptf1a* fails to rescue normal pancreas development in *ptf1a* morphants. However, the fish still form a hypoplastic pancreas (Pashos et al, 2013). Therefore, while the z3'-EA is required for proper pancreas development, there may be other elements capable of triggering *ptf1a* expression.

Like in mouse, multiple sequences downstream of the zebrafish *ptf1a* gene display enhancer activity in the developing nervous system (Pashos et al, 2013) (Figure 1.6) but, more recently, a zebrafish enhancer crucial for pancreas development was uncovered, 39kb downstream of the *ptf1a* TSS (Bordeira-Carriço et al, 2022). This *ptf1a*

distal pancreatic enhancer (z3'-DpE; Figure 1.6) is active in MPCs and differentiated acinar and duct cells. Similar to the 5'-AR, z3'-DpE-driven reporter expression is weaker in progenitors and stronger in differentiated exocrine cells and the sequence interacts with the *ptf1a* promoter, suggesting that the z3'-DpE also regulates *ptf1a* expression directly (Bordeira-Carriço et al, 2022). However, unlike the z5'-AR, the z3'-DpE does not contain PTF1 consensus binding sequences (Duque et al, 2022).

A similar element is present in the human genome (Figure 1.6A). Recessive mutations within a 400 bp non-coding sequence, 25 kb downstream of *PTF1A*, were identified in families with pancreatic agenesis. Moreover, chromatin conformation capture experiments in hESC (human embryonic stem cells)-derived human MPCs revealed that the human *PTF1A* distal enhancer (h3'-DpE) also interacts directly with the *PTF1A* promoter (Weedon, Michael N. et al, 2014). In contrast with *PTF1A* coding mutations, which result in pancreatic and cerebellar agenesis, neurological features are absent in reported cases of h3'-DpE mutations (Weedon, Michael N. et al, 2014). This illustrates how cis-regulatory mutations can have far greater tissue specificity than their disease-associated coding mutation counterparts.

Despite showing no sequence conservation with the human enhancer (Figure 1.6), z3'-DpE is likely functionally equivalent to h3'-DpE, with deletion of z3'-DpE resulting in pancreatic agenesis (Bordeira-Carriço et al, 2022). Interestingly, h3'-DpE is active in human MPCs but inactive in adult exocrine pancreatic cell lines (Weedon, Michael N. et al, 2014), while z3'-DpE shows activity in zebrafish MPCs but remains active in the adult zebrafish exocrine pancreas (Bordeira-Carriço et al, 2022). Therefore, although the two enhancers have equivalent functions during early pancreas development, the zebrafish sequence also contains cis-regulatory information relevant for adult pancreas function. Collectively, the activity of the zebrafish *ptf1a* downstream enhancers recapitulates the reporter expression pattern observed for the mice 3'-12.4 kb region but also exhibits activity in cerebellum and pancreas, which was absent for the mouse sequence (Masui et al, 2008; Meredith et al, 2009). Interestingly, we have reported the existence of a fragment with high degree of conservation with the h3'-DpE downstream of the mouse *Ptf1a* that may correspond to its functional equivalent (Duque et al, 2022) (Figure 1.6). More recently, the work of (Miguel-Escalada et al, 2022) has shown that the mouse conserved distal *Ptf1a* enhancer is active exclusively in MPCs of the developing mouse pancreas and deleting this CRE resulted in pancreatic hypoplasia and diabetes. Altogether, this suggests a species-specific late function for the zebrafish sequence.

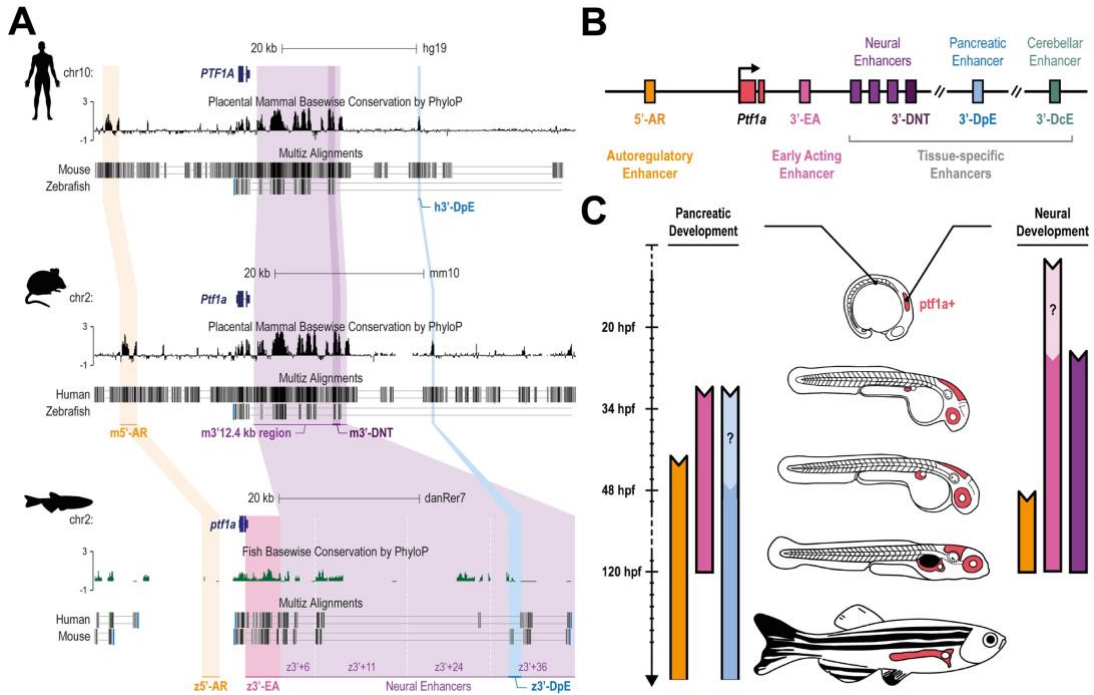


Figure 1.6 - (A) Regulatory landscape of *Ptf1a*. (Top) Human *PTF1A* distal pancreatic enhancer (h3'-DpE, blue). (Middle) Mouse autoregulatory enhancer (m5'-AR, orange), neural enhancers cluster region (m3'-12.4 kb region, light purple), and DNT-specific enhancer (m3'-DNT, dark purple). (Bottom) Zebrafish autoregulatory enhancer (z5'-AR), early-acting enhancer (z3'-EA), neural enhancers (z3' + 6, z3' + 11, z3' + 24, and z3' + 36), and distal pancreatic enhancer (z3'-DpE, blue). Sequence conservation across mammals (black) or fish (green) is shown for each panel and putative ortholog sequences between species are indicated by the colored boxes. z3'-DpE shows no sequence conservation with the human h3'-DpE enhancer; however, they likely represent functional equivalents, as assessed by functional assays. In mouse, there is a fragment with high degree of conservation with the h3'-DpE downstream of *Ptf1a* that could correspond to its functional equivalent, although this hypothesis has not been tested by functional assays. B) Summary of reported *Ptf1a* enhancers. *Ptf1a* expression during vertebrate development is regulated by multiple CREs, including an upstream autoregulatory enhancer (5'-AR), a proximal downstream early- acting enhancer (3'-EA) and multiple distal cell type-specific enhancers spread across a large genomic region. 5'-DpE, distal pancreatic enhancer; 5'-DcE, distal cerebellar enhancer (Adapted from (Mona et al, 2016)). (C) Reported activity of the z5'-AR (orange), z3'-EA (pink), z3'-DpE (blue), and neural enhancers (purple) throughout zebrafish embryonic development. Endogenous *ptf1a* expression is presented in red at 16hpf (14- somites), 35hpf (prim-22), 42hpf (high pec), 120hpf (larva), and adult zebrafish.

A comparison of the findings from humans, mice, and zebrafish reveals an intricate network of CREs that regulate *Ptf1a* transcription during development (Figure 1.6), as well as direct links between noncoding mutations and adverse phenotypic effects resulting from disruption of progenitor cell expansion and cell fate decision. Yet, these findings are only a glimpse into the complex enhancer interactions underlying the spatiotemporal control of *Ptf1a* levels. In Chapter III of this thesis, we have explored the mechanisms of regulation of *ptf1a* expression in zebrafish, focusing on the distal

pancreatic enhancer z3'-DpE and its dual function, early in MPCs of the developing pancreas and late in differentiated acinar cells. By performing enhancer assays and immunohistochemistry in z3'-DpE^{+/+} and ^{-/-} backgrounds, we were able to observe alterations in the pattern and localization of the MPC domain, including a posteriorly-oriented elongation of the MPC domain and the earlier appearance of what most likely corresponds to differentiated acinar cells, suggesting that differentiation into the various pancreatic lineages might be perturbed. Then, we performed RNA-sequencing of z3'-DpE^{+/+} and ^{-/-} MPCs and found downregulated genes to be enriched in Notch signaling pathway components, indicating that Ptf1a might activate an endogenous Notch signaling program in MPCs. To profile chromatin accessibility, we performed ATAC-seq of z3'-DpE^{+/+} and ^{-/-} differentiated acinar cells. We found regions with decreased accessibility in z3'-DpE^{-/-} and enriched TF motifs within these regions to be associated with genes related to acinar cell identity, endocrine function and developmental processes. On the other hand, regions with increased accessibility were found to be associated with genes related to developmental processes, inflammation, proliferation or epithelial-to-mesenchymal (EMT) transition. Moreover, TF motifs enriched within these regions were related to inflammatory processes and the respective TF-encoding genes were enriched within DisGeNET lists of genes associated to pancreatic cancer. Overall, these results highlight Ptf1a as a crucial factor for the proper development of the zebrafish pancreas, enforcing a differentiation program that allows proper differentiation of MPCs into differentiated acinar cells. After this, it functions to maintain the identity of differentiated acinar cells, together with other known acinar-specific TFs, which might prevent aberrant gene expression leading to pancreatic dysfunction and oncogenic transformation, highlighting *ptf1a* as a candidate for further study in the context of pancreatic cancer and possible preventive and therapeutic approaches to this severe disease.

1.6. Research Aims

The general aim of this thesis is to identify essential genes and regulatory modules defining the cellular state of pancreatic progenitors and differentiated acinar cells, focusing on:

- Regulatory elements controlling a known inhibitor of the Bmp signaling pathway, which helps determine the correct location of pancreatic progenitors and dimension of the pancreatic field, early in development;
- Gene networks controlling the expression of *ptf1a* in multipotent progenitor cells (MPCs) of the developing pancreas and in differentiated acinar cells.

1.7. References

- Amorim, J.P., Gali-Macedo, A., Marcelino, H., Bordeira-Carriço, R., Naranjo, S., Rivero-Gil, S., Teixeira, J., Galhardo, M., Marques, J. & Bessa, J. (2020) A Conserved Notochord Enhancer Controls Pancreas Development in Vertebrates. *Cell Reports* 32, 107862
- Apelqvist, Å., Li, H., Sommer, L., Beatus, P., Anderson, D.J., Honjo, T., de Angelis, M.H., Lendahl, U. & Edlund, H. (1999) Notch signalling controls pancreatic cell differentiation. *Nature* 400, 877–881
- Argenton, F., Zecchin, E. & Bortolussi, M. (1999) Early appearance of pancreatic hormone-expressing cells in the zebrafish embryo. *Mech Dev* 87, 217–221
- Atchison, M.L. (1988) Enhancers: Mechanisms of Action and Cell Specificity. *Annu. Rev. Cell. Biol.* 4, 127–153
- Baetens, D., Malaisse-Lagae, F., Perrelet, A. & Orci, L. (1979) Endocrine Pancreas: Three-Dimensional Reconstruction Shows Two Types of Islets of Langerhans. *Science, New Series* 206, 1323–1325
- Becker-André, M., André, E. & DeLamarter, J. (1993) Identification of Nuclear Receptor mRNAs by RT-PCR Amplification of Conserved Zinc-Finger Motif Sequences. *Biochemical and Biophysical Research Communications* 194, 1371–1379
- Beer, R.L., Parsons, M.J. & Rovira, M. (2016) Centroacinar cells: At the center of pancreas regeneration. *Developmental Biology* 413, 8–15
- Beres, T.M., Masui, T., Swift, G.H., Shi, L., Henke, R.M. & MacDonald, R.J. (2006) PTF1 Is an Organ-Specific and Notch-Independent Basic Helix-Loop-Helix Complex Containing the Mammalian Suppressor of Hairless (RBP-J) or Its Parologue, RBP-L. *MCB* 26, 117–130
- Bhushan, A.I., N.; Kato, S.; Thiery, J.; Czernichow, P.; Bellusci, S. and Scharfmann, R. (2001) Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. *Development*, 5109–5117
- Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B. & Driever, W. (2001) Pancreas Development in Zebrafish: Early Dispersed Appearance of Endocrine Hormone Expressing Cells and Their Convergence to Form the Definitive Islet. *Developmental Biology* 230, 189–203
- Binot, A-C., Manfroid, I., Flasse, L., Winandy, M., Motte, P., Martial, J.A., Peers, B. & Voz, M.L. (2010) Nkx6.1 and nkx6.2 regulate α - and β -cell formation in zebrafish by acting on pancreatic endocrine progenitor cells. *Developmental Biology* 340, 397–407
- Bordeira-Carriço, R., Teixeira, J., Duque, M., Galhardo, M., Ribeiro, D., D. Acemel, R., N. Firbas, P., J. Tena, J., Eufrásio, A., Marques, J., et al (2022) Multidimensional chromatin profiling of zebrafish pancreas to uncover and investigate disease- relevant enhancers. *Nature Communications* 13

- Brissova, M., Fowler, M.J., Nicholson, W.E., Chu, A., Hirshberg, B., Harlan, D.M. & Powers, A.C. (2005) Assessment of Human Pancreatic Islet Architecture and Composition by Laser Scanning Confocal Microscopy. *J Histochem Cytochem* 53, 1087–1097
- Burlison, J.S., Long, Q., Fujitani, Y., Wright, C.V.E. & Magnuson, M.A. (2008) Pdx-1 and Ptf1a concurrently determine fate specification of pancreatic multipotent progenitor cells. *Developmental Biology* 316, 74–86
- Capelson, M., and Corces, V.G. (2004) Boundary elements and nuclear organization. *Biology of the Cell* 96, 617–629
- Cebola, I., Rodríguez-Seguí, S.A., Cho, C.H.-H., Bessa, J., Rovira, M., Luengo, M., Chhatriwala, M., Berry, A., Ponsa-Cobas, J., Maestro, M.A., et al (2015) TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nat Cell Biol* 17, 615–626
- Chen, C., Shiota, C., Agostinelli, G., Ridley, D., Jiang, Y., Ma, J., Prasad, K., Xiao, X. & Gittes, G.K. (2019) Evidence of a developmental origin for β -cell heterogeneity using a dual lineage-tracing technology. *Development* 146, dev164913
- Chung, W.-S., Andersson, O., Row, R., Kimelman, D. & Stainier, D.Y.R. (2010) Suppression of Alk8-mediated Bmp signaling cell-autonomously induces pancreatic β -cells in zebrafish. *Proc Natl Acad Sci USA* 107, 1142–1147
- Chung, W.-S., Shin, C.H. & Stainier, D.Y.R. (2008) Bmp2 Signaling Regulates the Hepatic versus Pancreatic Fate Decision. *Developmental Cell* 15, 738–748
- Collombat, P. (2005) The simultaneous loss of Arx and Pax4 genes promotes a somatostatin-producing cell fate specification at the expense of the α - and β -cell lineages in the mouse endocrine pancreas. *Development* 132, 2969–2980
- Dalgin, G. & Prince, V.E. (2015) Differential levels of NeuroD establish zebrafish endocrine pancreas cell fates. *Developmental Biology* 402, 81–97
- Dassaye, R., Naidoo, S. & Cerf, M.E. (2016) Transcription factor regulation of pancreatic organogenesis, differentiation and maturation. *Islets* 8, 13–34
- Delaspre, F., Beer, R.L., Rovira, M., Huang, W., Wang, G., Gee, S., Vitery, M. del C., Wheelan, S.J. & Parsons, M.J. (2015) Centroacinar Cells Are Progenitors That Contribute to Endocrine Pancreas Regeneration. *Diabetes* 64, 3499–3509
- Demirbilek, H., Cayir, A., Flanagan, S.E., Yildirim, R., Kor, Y., Gurbuz, F., Haliloğlu, B., Yıldız, M., Baran, R.T., Akbas, E.D., et al (2020) Clinical Characteristics and Long-term Follow-up of Patients with Diabetes Due To *PTF1A* Enhancer Mutations. *The Journal of Clinical Endocrinology & Metabolism* 105, e4351–e4359
- Dong, P.D.S., Provost, E., Leach, S.D. & Stainier, D.Y.R. (2008) Graded levels of Ptf1a differentially regulate endocrine and exocrine fates in the developing pancreas. *Genes & Development* 22, 1445–1450
- Duque, M., Amorim, J.P. & Bessa, J. (2022) Ptf1a function and transcriptional cis-

regulation, a cornerstone in vertebrate pancreas development. *The FEBS Journal* 289, 5121–5136

Epstein, D.J. (2009) Cis-regulatory mutations in human disease. *Briefings in Functional Genomics and Proteomics* 8, 310–316

Esni, F. (2004) Notch inhibits Ptf1 function and acinar cell differentiation in developing mouse and zebrafish pancreas. *Development* 131, 4213–4224

Evliyaoğlu, O., Ercan, O., Ataloğlu, E., Zübarioğlu, Ü., Özcabı, B., Dağdeviren, A., Erdoğan, H., Franco, E.D. & Ellard, S. (2018) Neonatal Diabetes: Two Cases with Isolated Pancreas Agenesis due to Homozygous PTF1A Enhancer Mutations and One with Developmental Delay, Epilepsy, and Neonatal Diabetes Syndrome due to KCNJ11 Mutation. *Jcrpe* 10, 168–174

Field, H.A., Dong, P.D.S., Beis, D. & Stainier, D.Y.R. (2003) Formation of the digestive system in zebrafish. ii. pancreas morphogenesis☆. *Developmental Biology* 261, 197–208

Flasse, L.C., Pirson, J.L., Stern, D.G., Von Berg, V., Manfroid, I., Peers, B. & Voz, M.L. (2013) *Ascl1b* and *Neurod1*, instead of *Neurog3*, control pancreatic endocrine cell fate in zebrafish. *BMC Biol* 11, 78

Fujikura, J., Hosoda, K., Iwakura, H., Tomita, T., Noguchi, M., Masuzaki, H., Tanigaki, K., Yabe, D., Honjo, T. & Nakao, K. (2006) Notch/Rbp-j signaling prevents premature endocrine and ductal cell differentiation in the pancreas. *Cell Metabolism* 3, 59–65

Fukuda, A., Kawaguchi, Y., Furuyama, K., Kodama, S., Horiguchi, M., Kuhara, T., Kawaguchi, M., Terao, M., Doi, R., Wright, C.V.E., et al (2008) Reduction of *Ptf1a* Gene Dosage Causes Pancreatic Hypoplasia and Diabetes in Mice. *Diabetes* 57, 2421–2431

Fürthauer, M., Thisse, B. & Thisse, C. (1999) Three Different *noggin* Genes Antagonize the Activity of Bone Morphogenetic Proteins in the Zebrafish Embryo. *Developmental Biology* 214, 181–196

Gabbay, M., Ellard, S., Franco, E.D. & Moisés, R.S. (2017) Pancreatic Agenesis due to Compound Heterozygosity for a Novel Enhancer and Truncating Mutation in the PTF1A Gene. *Jcrpe* 9, 274–277

Ghaye, A.P., Bergemann, D., Tarifeño-Saldivia, E., Flasse, L.C., Von Berg, V., Peers, B., Voz, M.L. & Manfroid, I. (2015) Progenitor potential of *nkx6.1*-expressing cells throughout zebrafish life and during beta cell regeneration. *BMC Biol* 13, 70

Gittes, G.K. (2009) Developmental biology of the pancreas: A comprehensive review. *Developmental Biology* 326, 4–35

Gradwohl, G., Dierich, A., LeMeur, M. & Guillemot, F. (2000) *neurogenin3* is required for the development of the four endocrine cell lineages of the pancreas. *Proceedings of the National Academy of Sciences* 97, 1607–1611

Hald, J., Sprinkel, A.E., Ray, M., Serup, P., Wright, C. & Madsen, O.D. (2008) Generation and Characterization of *Ptf1a* Antiserum and Localization of *Ptf1a* in Relation to *Nkx6.1*

and Pdx1 During the Earliest Stages of Mouse Pancreas Development. *J Histochem Cytochem* 56, 587–595

Hale, M.A., Swift, G.H., Hoang, C.Q., Deering, T.G., Masui, T., Lee, Y.-K., Xue, J. & MacDonald, R.J. (2014) The nuclear hormone receptor family member NR5A2 controls aspects of multipotent progenitor cell formation and acinar differentiation during pancreatic organogenesis. *Development* 141, 3123–3133

Harris, M.B., Mostecky, J., and Rothman, P.B. (2005) Repression of an Interleukin-4-responsive Promoter Requires Cooperative BCL-6 Function. *Journal of Biological Chemistry* 280, 13114–13121

Hebrok, M., Kim, S., St-Jacques, B., McMahon, A. & Melton, D. (2000) Regulation of pancreas development by hedgehog signaling. *Development*

Hebrok, M., Kim, S.K. & Melton, D.A. (1998) Notochord repression of endodermal Sonic hedgehog permits pancreas development. *Genes & Development* 12, 1705–1713

Herrera, P.L. (2000) Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. *Development* 127, 2317–2322

Hesselson, D., Anderson, R.M., Beinat, M. & Stainier, D.Y.R. (2009) Distinct populations of quiescent and proliferative pancreatic -cells identified by HOTcre mediated labeling. *Proceedings of the National Academy of Sciences* 106, 14896–14901

Hoang, C.Q., Hale, M.A., Azevedo-Pouly, A.C., Elsässer, H.P., Deering, T.G., Willet, S.G., Pan, F.C., Magnuson, M.A., Wright, C.V.E., Swift, G.H., et al (2016) Transcriptional Maintenance of Pancreatic Acinar Identity, Differentiation, and Homeostasis by PTF1A. *Mol Cell Biol* 36, 3033–3047

Holmstrom, S.R., Deering, T., Swift, G.H., Poelwijk, F.J., Mangelsdorf, D.J., Kliewer, S.A. & MacDonald, R.J. (2011) LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. *Genes & Development* 25, 1674–1679

Isshiki, T., Pearson, B., Holbrook, S. & Doe, C.Q. (2001) *Drosophila* Neuroblasts Sequentially Express Transcription Factors which Specify the Temporal Identity of Their Neuronal Progeny. *Cell* 106, 511–521

Jennings, R.E., Berry, A.A., Kirkwood-Wilson, R., Roberts, N.A., Hearn, T., Salisbury, R.J., Blaylock, J., Piper Hanley, K. & Hanley, N.A. (2013) Development of the Human Pancreas From Foregut to Endocrine Commitment. *Diabetes* 62, 3514–3522

Jennings, R.E., Berry, A.A., Strutt, J.P., Gerrard, D.T. & Hanley, N.A. (2015) Human pancreas development. *Development* 142, 3126–3137

Jonsson, J., Carlsson, L., Edlund, T. & Edlund, H (1994) Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature Letters*

Jørgensen, M.C., Ahnfelt-Rønne, J., Hald, J., Madsen, O.D., Serup, P. & Hecksher-Sørensen, J. (2007) An Illustrated Review of Early Pancreas Development in the Mouse. *Endocrine Reviews* 28, 685–705

- Kawaguchi, Y., Cooper, B., Gannon, M., Ray, M., MacDonald, R.J. & Wright, C.V.E. (2002) The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet* 32, 128–134
- Kim, S.K. & Hebrok, M. (2001) Intercellular signals regulating pancreas development and function. *Genes Dev* 15, 111–127
- Kim, S.K., Hebrok, M. & Melton, D.A. (1997) Notochord to endoderm signaling is required for pancreas development. *Development*, 4243–4252
- Kinkel, M.D., Eames, S.C., Alonzo, M.R. & Prince, V.E. (2008) Cdx4 is required in the endoderm to localize the pancreas and limit -cell number. *Development* 135, 919–929
- Kinkel, M.D. & Prince, V.E. (2009) On the diabetic menu: Zebrafish as a model for pancreas development and function. *BioEssays* 31, 139–152
- Krah, N.M., Narayanan, S.M., Yugawa, D.E., Straley, J.A., Wright, C.V.E., MacDonald, R.J. & Murtaugh, L.C. (2020) Prevention and reversion of pancreatic tumorigenesis through a differentiation-based mechanism. *Developmental Cell* 50
- Krah, N.M., Willet, S.G., Pan, F.C., Cash, G.M., Bronner, M.P., Wright, C.V., MacDonald, R.J. & Murtaugh, L.C. (2015) The acinar differentiation determinant PTF1A inhibits initiation of pancreatic ductal adenocarcinoma. *eLife*
- Krapp, A., Knöfler, M., Frutiger, S., Hughes, G.J., Hagenbüchle, O. & Wellauer, P.K. (1996) The p48 DNA-binding subunit of transcription factor PTF1 is a new exocrine pancreas-specific basic helix-loop-helix protein. *The EMBO Journal* 15, 4317–4329
- Krapp, A., Knöfler, M., Ledermann, B., Burki, K., Berney, C., Zoerkler, N., Hagenbuchle, O. & Wellauer, P.K. (1998) The bHLH protein PTF1-p48 is essential for the formation of the exocrine and the correct spatial organization of the endocrine pancreas. *Genes & Development* 12, 3752–3763
- Levine, M., and Tjian, R. (2003) Transcription regulation and animal diversity. *Nature* 424, 147–151
- Li, L., He, S., Sun, J.-M., and Davie, J.R. (2004) Gene regulation by Sp1 and Sp3. *Biochem. Cell Biol.* 82, 460–471
- Lin, J.W., Biankin, A.V., Horb, M.E., Ghosh, B., Prasad, N.B., Yee, N.S., Pack, M.A. & Leach, S.D. (2004) Differential requirement for ptf1a in endocrine and exocrine lineages of developing zebrafish pancreas. *Developmental Biology* 270, 474–486
- Maden, M. (2001) Role and distribution of retinoic acid during CNS development. *International Review of Cytology* pp 1–77. Elsevier
- Maduro, M.F. & Rothman, J.H. (2002) Making Worm Guts: The Gene Regulatory Network of the *Caenorhabditis elegans* Endoderm. *Developmental Biology* 246, 68–85
- Maston, G.A., Evans, S.K. & Green, M.R. (2006) Transcriptional Regulatory Elements in the Human Genome. *Annu Rev Genom Hum Genet* 7, 29–59
- Masui, T., Long, Q., Beres, T.M., Magnuson, M.A. & MacDonald, R.J. (2007) Early

pancreatic development requires the vertebrate Suppressor of Hairless (RBPJ) in the PTF1 bHLH complex. *Genes & Development* 21, 2629–2643

Masui, T., Swift, G.H., Hale, M.A., Meredith, D.M., Johnson, J.E. & MacDonald, R.J. (2008) Transcriptional Autoregulation Controls Pancreatic Ptf1a Expression during Development and Adulthood. *MCB* 28, 5458–5468

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012) Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science* 337, 1190–1195

McCracken, K.W. & Wells, J.M. (2012) Molecular pathways controlling pancreas induction. *Seminars in Cell & Developmental Biology* 23, 656–662

Meredith, D.M., Borromeo, M.D., Deering, T.G., Casey, B.H., Savage, T.K., Mayer, P.R., Hoang, C., Tung, K.-C., Kumar, M., Shen, C., et al (2013) Program Specificity for Ptf1a in Pancreas versus Neural Tube Development Correlates with Distinct Collaborating Cofactors and Chromatin Accessibility. *Mol Cell Biol* 33, 3166–3179

Meredith, D.M., Masui, T., Swift, G.H., MacDonald, R.J. & Johnson, J.E. (2009) Multiple Transcriptional Mechanisms Control Ptf1a Levels during Neural Development Including Autoregulation by the PTF1-J Complex. *Journal of Neuroscience* 29, 11139–11148

Miguel-Escalada, I., Maestro, M.A., Balboa, D., García-Hurtado, J., Sebé-Pedrós, A. & Ferrer, J. (2022) Pancreas agenesis mutations disrupt a lead enhancer controlling a developmental enhancer cluster. *Developmental Cell* 57

Mona, B., Avila, J.M., Meredith, D.M., Kollipara, R.K. & Johnson, J.E. (2016) Regulating the dorsal neural tube expression of Ptf1a through a distal 3' enhancer. *Developmental Biology* 418, 216–225

Mona, B., Uruena, A., Kollipara, R.K., Ma, Z., Borromeo, M.D., Chang, J.C. & Johnson, J.E. (2017) Repression by PRDM13 is critical for generating precision in neuronal identity. *eLife* 6, e25787

Nair, G.G., Vincent, R.K. & Odorico, J.S. (2014) Ectopic Ptf1a Expression in Murine ESCs Potentiates Endocrine Differentiation and Models Pancreas Development In Vitro: Ptf1a Promotes Endocrine Differentiation. *Stem Cells* 32, 1195–1207

Nakhai, H., Siveke, J.T., Klein, B., Mendoza-Torres, L., Mazur, P.K., Algul, H., Radtke, F., Strobl, L., Zimmer-Strobl, U. & Schmid, R.M. (2008) Conditional ablation of Notch signaling in pancreatic development. *Development* 135, 2757–2765

Norgaard, G.A., Jensen, J.N. & Jensen, J. (2003) FGF10 signaling maintains the pancreatic progenitor cell state revealing a novel role of Notch in organ development. *Developmental Biology* 264, 323–338

Ong, C.-T. & Corces, V.G. (2011) Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet* 12, 283–293

Pan, F.C., Bankaitis, E.D., Boyer, D., Xu, X., Van de Casteele, M., Magnuson, M.A., Heimberg, H. & Wright, C.V.E. (2013) Spatiotemporal patterns of multipotentiality in

- Ptf1a-expressing cells during pancreas organogenesis and injury-induced facultative restoration. *Development* 140, 751–764
- Pan, F.C. & Wright, C. (2011) Pancreas organogenesis: From bud to plexus to gland. *Dev Dyn* 240, 530–565
- Park, S.W., Davison, J.M., Rhee, J., Hruban, R.H., Maitra, A. & Leach, S.D. (2008) Oncogenic KRAS Induces Progenitor Cell Expansion and Malignant Transformation in Zebrafish Exocrine Pancreas. *Gastroenterology* 134, 2080–2090
- Parsons, M.J., Pisharath, H., Yusuff, S., Moore, J.C., Siekmann, A.F., Lawson, N. & Leach, S.D. (2009) Notch-responsive cells initiate the secondary transition in larval zebrafish pancreas. *Mechanisms of Development* 126, 898–912
- Pashos, E., Park, J.T., Leach, S. & Fisher, S. (2013) Distinct enhancers of ptf1a mediate specification and expansion of ventral pancreas in zebrafish. *Developmental Biology* 381, 471–481
- Pictet, R.L., Clark, W.R., Williams, R.H. & Rutter, W.J. (1972) An ultrastructural analysis of the developing embryonic pancreas. *Developmental Biology* 29, 436–467
- Pisharath, H., Rhee, J.M., Swanson, M.A., Leach, S.D. & Parsons, M.J. (2007) Targeted ablation of beta cells in the embryonic zebrafish pancreas using *E. coli* nitroreductase. *Mechanisms of Development* 124, 218–229
- Pourquié, O. (2003) The Segmentation Clock: Converting Embryonic Time into Spatial Pattern. *Science* 301, 328–330
- Rodriguez-Seguel, E., Mah, N., Naumann, H., Pongrac, I.M., Cerda-Esteban, N., Fontaine, J.-F., Wang, Y., Chen, W., Andrade-Navarro, M.A. & Spagnoli, F.M. (2013) Mutually exclusive signaling signatures define the hepatic and pancreatic progenitor cell lineage divergence. *Genes & Development* 27, 1932–1946
- Rose, S.D., Swift, G.H., Peyton, M.J., Hammer, R.E. & MacDonald, R.J. (2001) The Role of PTF1-P48 in Pancreatic Acinar Gene Expression. *J Biol Chem* 276, 44018–44026
- Sambathkumar, R., Migliorini, A. & Nostro, M.C. (2018) Pluripotent Stem Cell-Derived Pancreatic Progenitors and β -Like Cells for Type 1 Diabetes Treatment. *Physiology* 33, 394–402
- Schaffer, A.E., Freude, K.K., Nelson, S.B. & Sander, M. (2010) Nkx6 Transcription Factors and Ptf1a Function as Antagonistic Lineage Determinants in Multipotent Pancreatic Progenitors. *Developmental Cell* 18, 1022–1029
- Schwitzgebel, V.M., Scheel, D., Connors, J., Kalamaras, J., Lee, J., Anderson, D., Sussel, L. & Johnson, J. (2000) Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development*, 3533–3542
- Sellick, G.S., Barker, K.T., Stolte-Dijkstra, I., Fleischmann, C., J Coleman, R., Garrett, C., Gloyn, A.L., Edghill, E.L., Hattersley, A.T., Wellauer, P.K., et al (2004) Mutations in PTF1A cause pancreatic and cerebellar agenesis. *Nat Genet* 36, 1301–1305

- Slack, J.M.W (1995) Developmental biology of the pancreas. *Development* 121, 1569–1580
- Smale, S.T., & Kadonaga, J.T. (2003) The RNA Polymerase II Core Promoter. *Annu. Rev. Biochem.* 72, 449–479
- Solar, M., Cardalda, C., Houbracken, I., Martín, M., Maestro, M.A., De Medts, N., Xu, X., Grau, V., Heimberg, H., Bouwens, L., et al (2009) Pancreatic Exocrine Duct Cells Give Rise to Insulin-Producing β Cells during Embryogenesis but Not after Birth. *Developmental Cell* 17, 849–860
- Sosa-Pineda, B., Chowdhury, K., Torres', M., Oliver, G. & Gruss, P. (1997) The Pax4 gene is essential for differentiation of insulin-producing cells in the mammalian pancreas. *Nature*, 399–402
- Srinivasan, L., and Atchison, M.L. (2004) YY1 DNA binding and PcG recruitment requires CtBP. *Genes Dev.* 18, 2596–2601
- Stafford, D. & Prince, V.E. (2002) Retinoic Acid Signaling Is Required for a Critical Early Step in Zebrafish Pancreatic Development. *Current Biology* 12, 1215–1220
- Tijian, Robert & Maniatis, Tom (1994) Transcriptional Activation: A Complex Puzzle with Few Easy Pieces. *Cell* 77, 5–8
- Villani, V., Thornton, M.E., Zook, H.N., Crook, C.J., Grubbs, B.H., Orlando, G., De Filippo, R., Ku, H.T. & Perin, L. (2019) SOX9+/PTF1A+ Cells Define the Tip Progenitor Cells of the Human Fetal Pancreas of the Second Trimester. *STEM CELLS Translational Medicine* 8, 1249–1264
- Villasenor, A., Chong, D.C., Henkemeyer, M. & Cleaver, O. (2010) Epithelial dynamics of pancreatic branching morphogenesis. *Development* 137, 4295–4305
- Wang, Y., Rovira, M., Yusuff, S. & Parsons, M.J. (2011) Genetic inducible fate mapping in larval zebrafish reveals origins of adult insulin-producing -cells. *Development* 138, 609–617
- Wang, Y.J., Park, J.T., Parsons, M.J. & Leach, S.D. (2015) Fate mapping of *ptf1a* -expressing cells during pancreatic organogenesis and regeneration in zebrafish. *Dev Dyn* 244, 724–735
- Ward, A.B., Warga, R.M. & Prince, V.E. (2007) Origin of the zebrafish endocrine and exocrine pancreas. *Dev Dyn* 236, 1558–1569
- Watson, J. & Francavilla, C. (2018) Regulation of FGF10 Signaling in Development and Disease. *Front Genet* 9, 500
- Weedon, Michael N., Cebola, I., Patch, A.-M., Flanagan, S.E., De Franco, E., Caswell, R., Rodríguez-Seguí, S.A., Shaw-Smith, C., Cho, C.H.-H., Allen, H.L., et al (2014) Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. *Nat Genet* 46, 61–64
- Yang, J. & Corces, V.G. (2011) Chromatin Insulators: A Role in Nuclear Organization

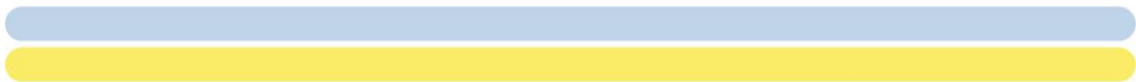
and Gene Expression. In *Advances in Cancer Research* pp 43–76. Elsevier

Zecchin, E., Mavropoulos, A., Devos, N., Filippi, A., Tiso, N., Meyer, D., Peers, B., Bortolussi, M. & Argenton, F. (2004) Evolutionary conserved role of ptf1a in the specification of exocrine pancreatic fates. *Developmental Biology* 268, 174–184

Zhou, Q., Law, A.C., Rajagopal, J., Anderson, W.J., Gray, P.A. & Melton, D.A. (2007) A Multipotent Progenitor Domain Guides Pancreatic Organogenesis. *Developmental Cell* 13, 103–114

Zhu, X., Zhang, J., Tollkuhn, J., Ohsawa, R., Bresnick, E.H., Guillemot, F., Kageyama, R. & Rosenfeld, M.G. (2006) Sustained Notch signaling in progenitors is required for sequential emergence of distinct cell lineages during organogenesis. *Genes & Development* 20, 2739–2753

Chapter II



A conserved notochord enhancer controls pancreas development in vertebrates

This chapter is adapted from an original research article, published as:

Amorim, J.P. †, Macedo-Gali, A. †, Marcelino, H, Bordeira-Carriço, R, Naranjo, S, Rivero-Gil, S, Teixeira, J, Galhardo, M, Marques, J. and José Bessa*. (2020) A Conserved Notochord Enhancer Controls Pancreas Development in Vertebrates. *Cell Reports*. 32:107862. doi: 10.1016/j.celrep.2020.107862

†These authors contributed equally to this work

*Corresponding author.

Author contributions:

J.B. conceived and supervised the project. J.P.A, A.G.-M, H.M, and J.B. conceived, designed, and analyzed the data. R.B.C. performed 4C-seq experiments. J.T. and M.G. performed computational analyses. S.N, S.R.-G, and J.M. contributed with reagents and materials. J.P.A, A.G.-M, and J.B. wrote the paper.

2.1. Introduction

The pancreas is an endoderm-derived organ that has several crucial roles, among them, controlling levels of glucose in the blood by producing insulin from beta cells and aiding digestion by secreting digestive enzymes to the digestive system. The pancreas develops from a group of progenitor cells, originating from the endoderm developing gut, under the control of signaling pathways that determine its correct position and size (Prince et al, 2017; Sakhneny et al, 2019). One of these determinant signaling molecules is Bmp, which has various roles in pancreas development. These roles include the regulation of pancreatic size by controlling liver and pancreas fates, partly by restricting the expression of *pdx1*, an important top hierarchical gene present in pancreatic progenitor cells (Prince et al, 2017; Sakhneny et al, 2019). In zebrafish, and according to this model, *bmp2b* expressed and secreted from the lateral plate mesoderm generates a gradient at which high levels of Bmp2b restrict pancreatic size, favoring liver fate and establishing the development of the pancreas in a region aligned with the midline of the embryo (Chung et al, 2008, 2010). In this context, Bmp signaling works as negative regulator of pancreas development. Interestingly, it has been long postulated that the notochord, located at the midline of the embryo, dorsal to the pancreas, has a pro-pancreatic function (Prince et al, 2017; Sakhneny et al, 2019). Part of these observations derive from studies in which the notochord was surgically removed from chicken embryos, showing that in its absence, pancreatic markers such as *insulin* were defectively expressed, as well as early progenitor markers such as *Pdx1* (Kim et al, 1997). These results suggest that a non-autonomous signal must emanate from the notochord to induce pancreatic fate. One possibility is that activin bB and fibroblast growth factor 2 (*FGF2*), expressed in the notochord, might have a role in pancreatic induction via repression of *SHH* expression in the endoderm (Hebrok et al, 1998). Nevertheless, ectopic pancreatic tissue is not observed in the endoderm of Shh mutant embryos (Hebrok et al, 2000), suggesting that inhibition of Shh signaling is permissive, rather than instructive, for pancreas fate (McCracken and Wells, 2012). A complementary hypothesis to be addressed is the presence of a Bmp antagonist expressed in the notochord that might be required for proper pancreatic induction; however, little is known about how Bmp signaling is modulated in the endoderm to effectively determine pancreatic size and position.

In zebrafish, one of the most notochord-specific Bmp antagonists is Noggin2 (Nog2) (Furthauer et al, 1999). *nog2* starts to be expressed by the end of gastrulation (10 h post-fertilization [hpf]) in the zebrafish axial mesoderm, which shortly gives rise to the notochord (Furthauer et al, 1999). This expression is maintained up to the 18-somite

stage (16 hpf), when the transcript slowly starts to disappear, and persists until 30 hpf in the posterior tip of the notochord. In this work, we disrupted the function of *nog2*, observed consistent impairment in the number of pancreatic progenitor cells and insulin-expressing cells, and showed *nog2* to be an important Bmp antagonist required for pancreatic development. The introduction of an enhancer blocking insulator (Bessa et al, 2014) downstream of *nog2* also impaired beta cell differentiation and reduced the number of pancreatic progenitor cells, suggesting the existence of an enhancer required for the *nog2* pancreatic function. By analyzing histone modifications associated with enhancer activity (histone H3 lysine 27 acetylation [H3K27ac] and histone H3 lysine 4 monomethylation [H3K4me1]) (Bogdanovic et al, 2012; Marlétaz et al, 2018; Rada-Iglesias et al, 2011) and performing circular chromosome conformation capture coupled with next-generation sequencing (4C-seq) (Zhao et al, 2006), we identified a *nog2* notochord enhancer, among others. To further demonstrate that the notochord enhancer is crucial for the *nog2* pancreatic function, we targeted it using CRISPR-Cas9, inducing genomic deletions (Letelier et al, 2018); complementing these assays, we used Cas9 endonuclease dead (dCas9) fused to a KRAB repressor domain (dCas9KRAB) (Thakore et al, 2015) and observed beta cell differentiation impairment. Moreover, in heteroallelic mutant embryos for partial deletion of the notochord enhancer, we observed a reduced number of pancreatic progenitor cells and beta cells when performing immunostaining for Nkx6.1 and insulin, respectively. In addition, we traced Nog2 *in vivo* and observed its diffusion from the notochord to the pancreatic progenitor domain, supporting its non-autonomous function in pancreas development. To understand whether this mechanism could also be present in mammals, we explored the human and mouse *Nog* regulatory landscapes, identifying notochord enhancers. Overall, these results suggest that *Nog* has acquired a notochord enhancer, most likely in the basis of the vertebrate phylogeny, that contributes to the crucial function of the notochord as an important signaling center for the development of complex organs such as the pancreas.

2.2. Results

2.2.1 *nog2* controls non-autonomously the embryonic endocrine pancreas size

nog2 encodes a zebrafish Bmp antagonist that starts to be expressed at 10 hpf in the primordium of the notochord (Furthauer et al, 1999). To identify possible non-coding genetic modifications affecting pancreas development, we mobilized a transposable element containing a potent enhancer blocking insulator (Bessa et al, 2014). We found a novel integration of this construct (ED301) located 2 kb downstream of *nog2* (*nog2*^{ED301}) (Figure 2.1A; Figure S2.1). Homozygous embryos for this integration showed a decrease of *nog2* expression at 10 hpf that was detected by qPCR (Figure 2.1B). To determine whether the downregulation of *nog2* could affect pancreas development, we performed *in situ* hybridization for *insulin* and determined that its area of expression was reduced in *nog2*^{ED301} homozygous embryos (Figures 2.1C and 2.1D) compared with control embryos. Next, to determine whether the phenotype observed in pancreas development was caused by the action of the insulator, we injected Cre-recombinase in homozygous embryos for the *nog2*^{ED301} integration and performed *in situ* hybridization for *insulin*. Cre recombinase targeted two *loxP* cassettes flanking the insulator in the ED301 construct (Figure 2.1A). In this experimental condition, we observed a rescue of the decreased area of *insulin* expression (Figure 2.1D). These results suggest that the ectopic insulator might be blocking a *cis*-regulatory element of *nog2* required for proper pancreas development.

To validate *nog2* as an important gene controlling pancreas development, we generated a *nog2* mutant line (*nog2*^{mut}) using the CRISPR-Cas9 targeting system. This mutant line shows a frameshift mutation in the *nog2* gene (Figure S2.2A). Although *nog2* is not expressed at the shield developmental stage, when the dorsal/ventral axis of the zebrafish embryo is established by a ventral gradient of Bmp, it is known that ectopic expression of *nog2* interferes with Bmp signaling, causing dorsalization phenotypes (Furthauer et al, 1999). Therefore, to test whether the *nog2*^{mut} frameshift mutation is a loss of function of the *nog2* gene, we injected mRNA of the frameshift version of *nog2* in zebrafish embryos. We did not observe a dorsalization phenotype, contrasting the wild-type (WT) version of *nog2* mRNA (Figures S2.2C and S2.2D). These results show that *nog2*^{mut} is most likely a null mutant for *nog2*. We incrossed *nog2*^{+/-mut}, generating a mixture of WT, heterozygous, and homozygous embryos, and outcrossed *nog2*^{+/-mut} with *nog2*^{ED301/ED301}. In both cases, the resulting embryos had a significant decrease

in the area of expression of *insulin*, demonstrating that *nog2* is required for the differentiation of a correctly sized pancreas (Figure 2.1D). The decrease of the area of expression of *insulin* could be caused by a decrease in the number of beta cells or by a decrease in their size. To discriminate between these two possible scenarios, we injected a morpholino against *nog2* in a transgenic reporter line for *insulin* expression (dilorio et al, 2002) and observed a reduced number of beta cells in the *nog2* morpholino-injected animals (Figures S2.2E and S2.2F). In addition, we performed immunostaining for Nkx6.1, to label pancreatic progenitor cells, and for insulin, to label differentiated beta cells, in homozygous embryos for the *nog2*^{ED301} integration. We observed that the numbers of both Nkx6.1 and insulin cells decreased compared with controls. A similar decrease in the average number of Nkx6.1- and insulin-expressing cells was also detected in a 50% mixed population of embryos carrying the heteroallelic combination of *nog2*^{ED301/mut} and *nog2*^{+/ED301} (Figures 2.1E to 2.1G). These results suggest that the reduced number of insulin-expressing cells might be the consequence of a smaller pancreatic progenitor domain, indicating that *nog2* is required for the development of a properly sized pancreas.

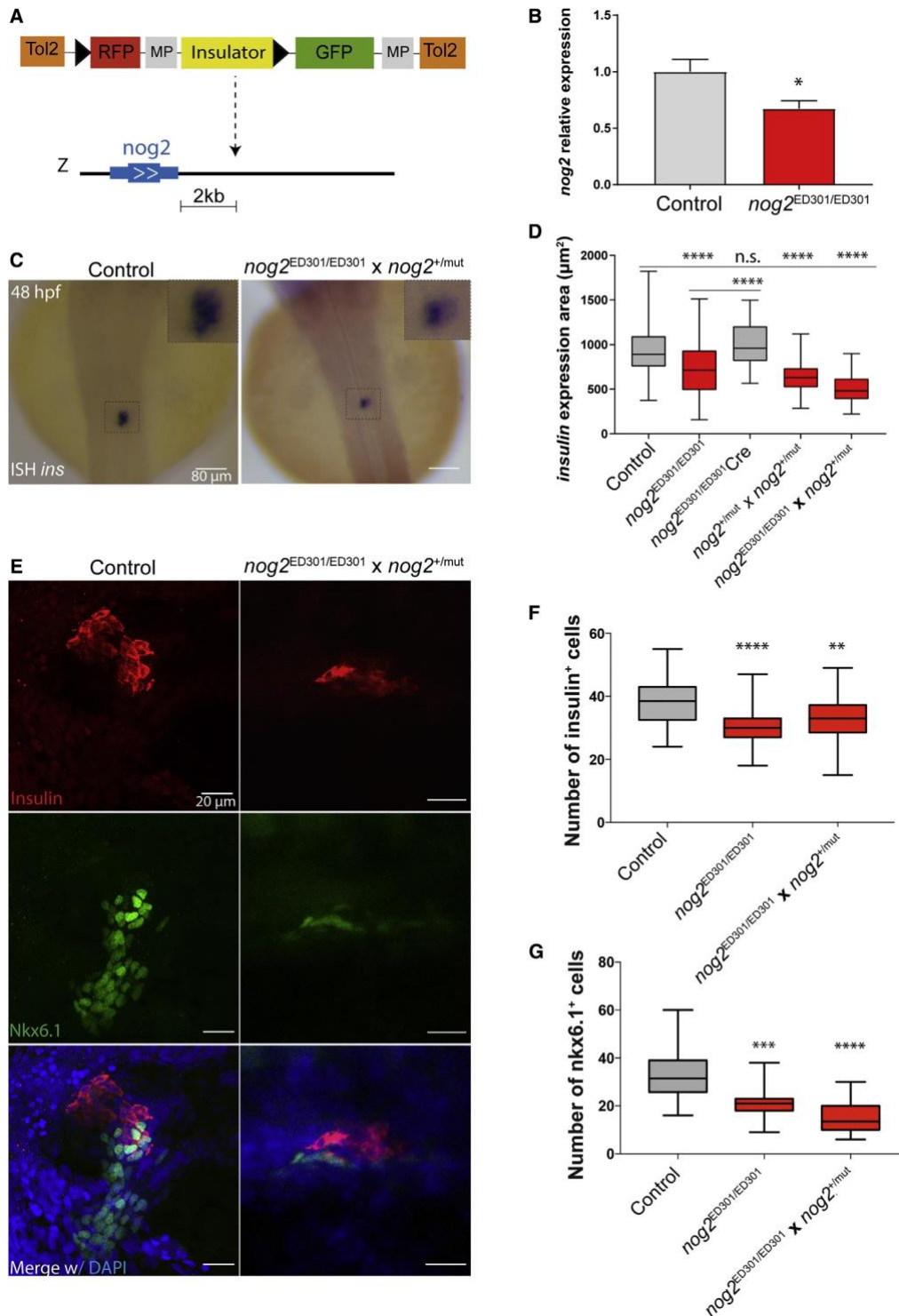


Figure 2.1 - (A) ED301 zebrafish line corresponds to an ED transposon integration containing a potent enhancer blocking insulator (yellow) and mapped 2 kb downstream of *nog2*. The *loxP* sequences are depicted as black triangles (see also Figure S1). Scale bar represents 2 kilobases. (B) In the ED301 zebrafish line, *nog2* expression is downregulated, as detected by qPCR performed in three biological replicates of 30 embryos each (error bars represent SD; *p < 0.05). (C) *In situ* hybridization of *insulin* in 48 hpf representative embryos from an outcross of *nog2*^{ED301/ED301} and *nog2*^{+/-mut} lines compared with control embryos. Scale bar represents 80 μ m. (D) Quantification of the insulin expression area detected by *in situ* hybridization in control (n = 194), *nog2*^{ED301/ED301} (n = 93), *nog2*^{ED301/ED301} injected with Cre

recombinase (n = 32), *nog2*^{+/mut} incross (n = 115), and *nog2*^{ED301/ED301}, *nog2*^{+/mut} outcross embryos (n = 36) at 48 hpf. Error bars represent SD; ****p < 0.0001. (E) Representative confocal images of 48 hpf zebrafish embryos counterstained with a DAPI nuclear marker (blue), an anti-insulin antibody marking beta cells (red), and an anti-Nkx6.1 antibody marking pancreatic progenitor cells (green). Images represent the maximum-intensity z projection of several focal planes obtained in a Leica Sp5 confocal microscope using a 403 objective. Scale bars represent 20 μ m. (F) Quantification of the number of insulin-expressing cells in *nog2*^{ED301/ED301} and *nog2*^{ED301/ED301}, *nog2*^{+/mut} outcross embryos compared with controls (n R 30). Error bars represent SD; ****p < 0.0001, **p < 0.01. (G) Quantification of the number of nkx6.1-expressing cells in *nog2*^{ED301/ED301} and *nog2*^{ED301/ED301}, *nog2*^{+/mut} outcross embryos compared with controls (n R 18). Error bars represent SD; ****p < 0.0001, ***p < 0.001.

2.2.2 *nog2* genomic regulatory landscape contains a notochord-specific enhancer

Nog proteins are well-known diffusible molecules that act as Bmp antagonists (Inomata et al, 2013). *nog2* function in pancreas development might be explained by its expression in the notochord and subsequent diffusion to the endoderm, which has a non-autonomous, pro-pancreatic role. The *nog2*^{ED301} zebrafish line contains an insulator integration in the genomic landscape of *nog2* that causes its transcriptional downregulation. This leads us to postulate that an essential enhancer for the *nog2* pancreatic function should be present in the *nog2* regulatory landscape, downstream of the insulator integration. To identify the possible *cis*-regulatory element that might be triggering the expression of *nog2* in the notochord, we analyzed the landscape of *nog2*, looking for regions of open chromatin, as determined by Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), and enriched for H3K27ac and H3K4me1, two histone marks associated with enhancer activity (Figures 2.2A and 2.2B). Because these histone modifications probably occur before the onset of *nog2* expression, starting in the notochord at 10 hpf, we looked at histone marks at the developmental stages of dome, 80% epiboly, and 24 hpf (Bogdanovic et al, 2012; Marletaz et al, 2018). In addition, we performed 4C-seq using the *nog2* promoter as the point of view to identify the subset of putative *cis*-regulatory sequences that interact with the promoter of *nog2* and therefore belong specifically to the regulatory landscape of this gene (Figure S2.3A). Using these strategies, we defined 3 putative *nog2* enhancers (Figure 2.2B, E1 to E3). We performed enhancer transgenesis assays for these sequences and verified that all 3 sequences are enhancers, together recapitulating the early expression pattern of *nog2* (Figure 2.2C; Figure S2.1E). Of the 3 sequences, E3 was shown to be a strong notochord enhancer; E2 was shown to be is a forebrain, hindbrain, and somite enhancer; and E1 was shown to be a strong ventral forebrain

enhancer (Figure 2.2C).

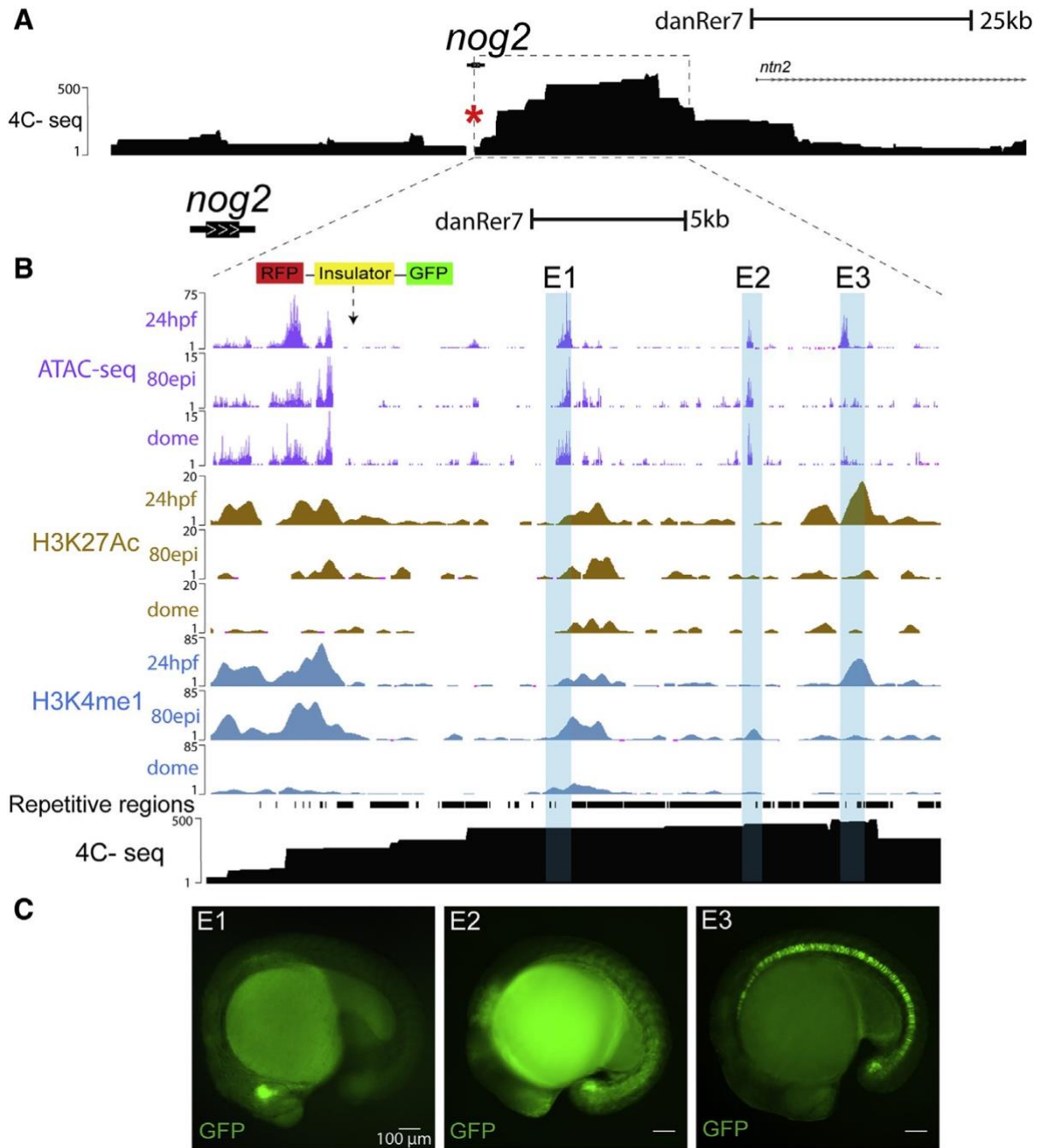


Figure 2.2 - (A) Genomic landscape of *nog2*, represented in the University of California Santa Cruz Genomics Institute (UCSC) genome browser (Kent et al, 2002), comprising more than 100 kb. The black track represents chromatin interaction points detected by 4C-seq using the *nog2* promoter as the viewpoint (red asterisk). Scale bar represents 25 kilobases. (B) Zoom-in in the locus of *nog2* comprising approximately 25 kb. Representation of the ED transposon integration 2 kb downstream of *nog2*. Purple (top), brown (middle), and blue (bottom) tracks represent the ATAC-seq and chromatin immunoprecipitation sequencing (ChIP-seq) signal for H3K27ac and H3K4me1, respectively, at three developmental times: 24 hpf, 80%

epiboly (80 epi), and dome. The black track represents the 4C-seq signal at 24 hpf using the *nog2* promoter as the viewpoint. Pale blue boxes highlight selected sequences for enhancer activity assays (E1 to E3). Two replicates of 4C data are shown in Figure S3. Scale bar represents 5 kilo bases. (C) Representative images of GFP reporter lines for enhancers E1 to E3 (see also Figure S1). Scale bars represent 100 μ m.

Next, we wanted to address whether *nog2* can diffuse from the notochord to the pancreatic progenitor domain. To test this, first we built a construct containing a Nog2-encoded protein fused to GFP. To test whether the GFPNog2 fusion protein maintained its functionality, we injected GFP*nog2* mRNA in 1-cell-stage embryos and observed the presence of dorsalization phenotypes, as in the case of WT *nog2* mRNA (Figures S2.2C and S2.2D). Then, we built a transposable element containing the *nog2*E3 notochord enhancer and a minimal promoter upstream of GFP*nog2* (*nog2*E3:GFP*nog2*) to be able to track Nog2 diffusion (Figure 2.3A; Figure S2.2B). Upon injection of zebrafish embryos with this construct and Tol2 mRNA, we performed staining using an anti-Nkx6.1 antibody that labels pancreatic progenitor cells (Binot et al, 2010; Hesselson et al, 2011). We observed clear co-localization of GFP aggregates in the pancreatic progenitor domain (Figure 2.3A) in 16 of 20 embryos (80%), as analyzed by confocal microscopy, showing that Nog2 can diffuse from the notochord to the pancreatic field in the endoderm. These results contrast those of controls expressing only GFP (Figure 2.3B), in which GFP aggregates were not seen outside of the notochord in 18 analyzed embryos. These results help to explain the non-autonomous effect observed in the pancreas in the *nog2* loss-of-function assays.

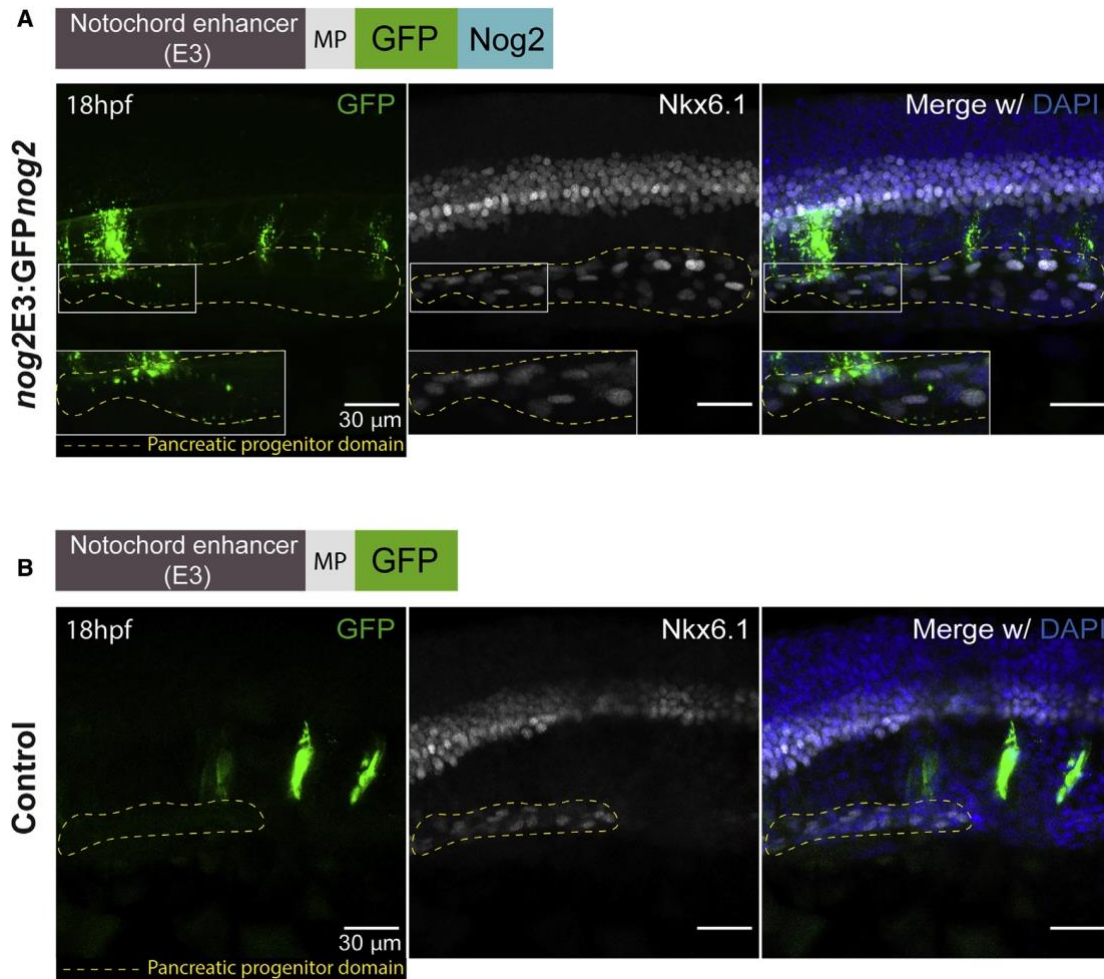


Figure 2.3 - (A) Representative confocal image of a 18 hpf zebrafish embryo injected with the *nog2E3:GFPnog2* construct. Of 20 embryos, 16 (80%) showed co-localization of GFPNog2 aggregates (green) with pancreatic progenitor cells, labeled with anti-Nkx6.1 antibody (white). Details of the colocalization can be observed in the inset (white box). (B) Representative confocal image of a 18 hpf zebrafish embryo injected with *nog2E3:GFP* (control), in which the pattern of GFP expression (green) is restricted to the notochord. None of 18 analyzed embryos displayed co-localization of GFP aggregates with Nkx6.1-labeled cells. Embryos were counterstained with the nuclear marker DAPI (blue), and the endocrine progenitor domain is delimited by a yellow dashed line (see also Figure S2). Images represent the maximum-intensity z projection of several focal planes obtained in a Leica Sp5 confocal microscope using a 403 objective. Scale bars represent 30 μ m.

2.2.3 The *nog2* notochord enhancer is required for proper pancreas development

The *nog2*E3 notochord enhancer most likely controls the expression of *nog2* in the notochord and is required for pancreatic development. To address this hypothesis, we synthesized a pair of single-guide RNAs (sgRNAs) 237 bp apart (*nog2*E3 sg1 and sg2) to target the *nog2*E3 notochord enhancer and generate genomic deletions for this *cis*-regulatory sequence. Embryos injected with Cas9 and the combination of *nog2*E3 sg1 and sg2 showed somatic deletions (Figure 2.4A; Figures S2.3E–S2.3G). *In situ* hybridization for *insulin* was performed in these injected embryos, together with controls injected with Cas9 alone, and the area of *insulin* expression was measured. We observed a decreased area of *insulin* expression for *nog2*E3 sg1 and sg2 (Figures 4A–4C). To validate these results, we injected a combination of *nog2*E3 sg1 and sg2 with dCas9-KRAB (Thakore et al, 2015). When compared with control embryos injected with dCas9-KRAB alone, we observed a reduction in the area of insulin expression (Figures S3B–S3D), suggesting that the *nog2*E3 notochord enhancer is modulating the expression of *nog2* in this tissue and is required for proper pancreas development. Indeed, partial deletion of the *nog2*E3 notochord enhancer shows a tendency for decreased *nog2* expression (Figure S2.4). Furthermore, we observed a decrease in the average number of Nkx6.1- and insulin-expressing cells in homozygous embryos for this deletion (Figures 2.4G–2.4I), recapitulating the phenotype seen in the *nog2*^{ED301} integration mutants. Accordingly, ectopic expression of *nog2* in the notochord should result in increased pancreas size. To test this, we injected the *nog2*E3:GFP*nog2* construct and measured the size of the area of expression of *insulin*, noticing an incremental change in the pancreatic field area (Figures 2.4D–2.4F). These results demonstrate that the *nog2*E3 notochord enhancer is required for the proper function of *nog2* in pancreatic development and to establish a correctly sized pancreas.

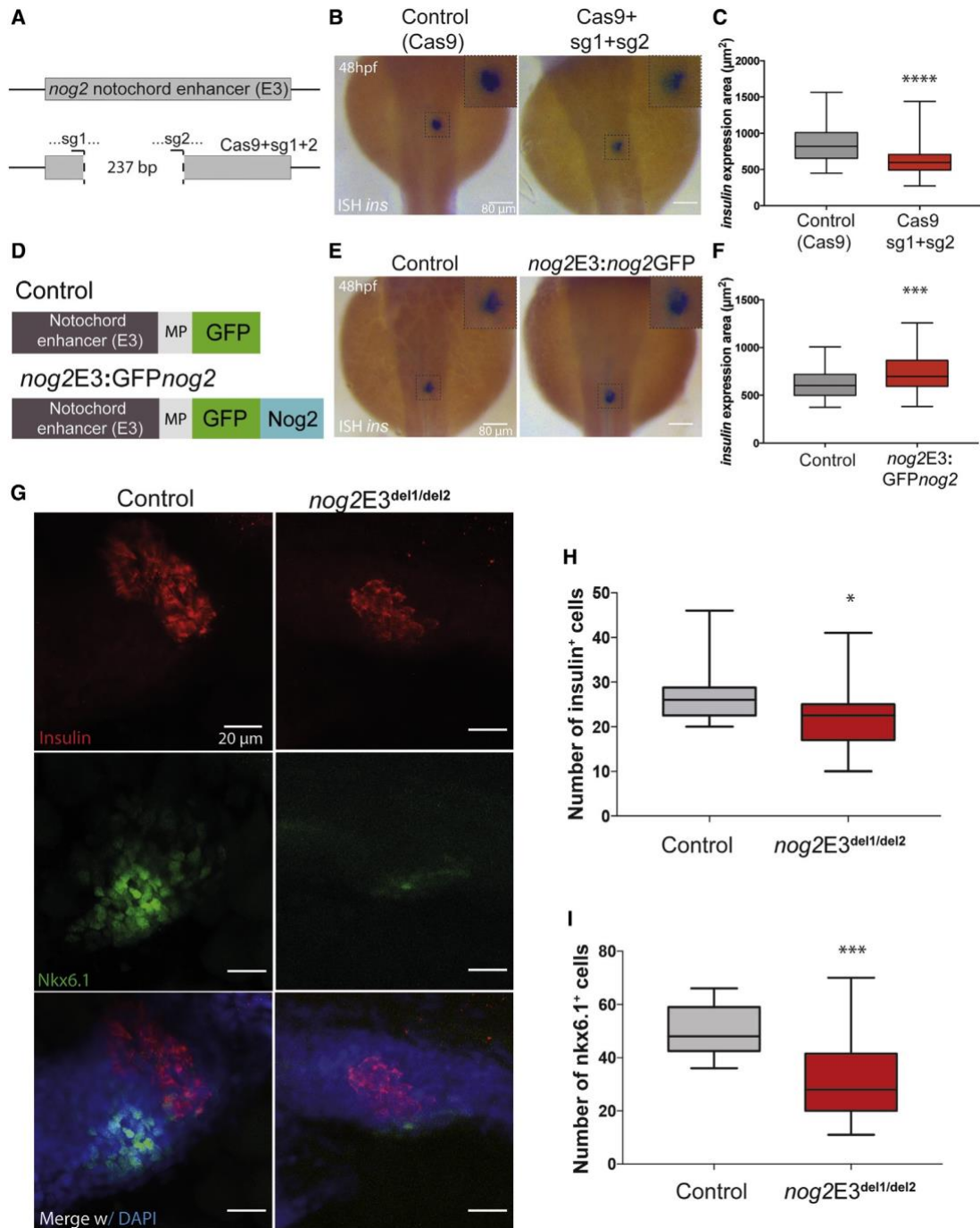


Figure 2.4 - (A) Representation of the WT *nog2E3* enhancer (above) and somatic deletions (below) generated by the injection of the Cas9 protein, together with sg1 and sg2 targeting two regions of the sequence 237 bp apart (see also Figure S3). (B) *In situ* hybridization of *insulin* in 48 hpf representative embryos, injected with the Cas9 protein, and two sgRNAs compared with control embryos, injected with only Cas9. Scale bars represent 80 μ m. (C) Quantification of the *insulin* expression area detected by *in situ* hybridization in injected embryos at 48 hpf compared with controls. Error bars represent SD; ****p < 0.0001. In all cases, n = 97. (D) Diagram of the *nog2E3:GFPnog2* construct used to achieve notochord-specific overexpression of GFPnog2 and the respective control. (E) *In situ* hybridization of *insulin* in 48 hpf representative embryos injected with *nog2E3:GFPnog2* compared with control embryos. Scale bars represent 80 μ m. (F) Quantification of the *insulin* expression area detected by *in situ* hybridization in injected

embryos at 48 hpf compared with controls. Error bars represent SD; *** $p < 0.001$. In all cases, $n \geq 38$. (G) Representative confocal images of 48 hpf zebrafish embryos counterstained with a DAPI nuclear marker (blue), an anti-insulin antibody marking β cells (red), and an anti-Nkx6.1 antibody marking pancreatic progenitor cells (green). Images represent the maximum-intensity z projection of several focal planes obtained in a Leica Sp5 confocal microscope using a 403 objective. Scale bars represent 20 μ m. (H) Quantification of the number of insulin-expressing cells in $\text{nog2E3}^{\text{del1/del2}}$ embryos compared with controls ($n = 19$). Error bars represent SD; * $p < 0.05$. (I) Quantification of the number of nkx6.1-expressing cells in $\text{nog2E3}^{\text{del1/del2}}$ embryos compared with controls ($n \geq 13$). Error bars represent SD; *** $p < 0.001$.

2.2.4 The expression of *Nog* in the notochord is conserved in vertebrates and is mediated by equivalent notochord enhancers

Next, we wanted to understand whether *NOG* expression in humans could be controlled by a notochord enhancer. To clarify this, we screened the human *NOG* genomic landscape for putative enhancers, based on sequence conservation and SNPs associated with type 2 diabetes and hyperglycemic traits (Figure 2.5A) (Morris et al, 2012). We selected 9 candidate human sequences (D12 to D10) (Figure 2.5A) to perform enhancer reporter assays in zebrafish and determined whether any of these sequences is a notochord enhancer. Preliminary results obtained by the generation of mosaic zebrafish transgenic embryos showed that sequence D10 was able to drive expression of GFP in the notochord (Figure S2.5D). We confirmed this result by generating stable zebrafish transgenic reporter lines for the selected sequences (Figure 2.5A). Apart from the D10 notochord enhancer, we found that D5 is a somite, forebrain, and hindbrain enhancer; D6 is a forebrain and somite enhancer; and D9 has broad enhancer activity and does not show high tissue specificity (Figure 2.5B; Figures S2.5A–S2.5D). Enhancers D5, D6, D9, and D10 were confirmed and obtained concordant expression patterns in at least two independent genomic integrations of the reporter constructs (Figures S2.5A–S2.5D). Sequences D7, D11, and D12 showed unreproducible patterns, and D12 and D8 showed no expression of the reporter gene. The expression patterns of the different human enhancers correlate with the known expression of *Nog* in mice, chicken, and *Xenopus* (Figure S2.5E), including the notochord and nervous system, suggesting that the *Nog* vertebrate genomic landscape has functionally similar *cis*-regulatory elements.

We were able to align the human D10 enhancer with the genome of other mammal species, suggesting that this *cis*-regulatory element is conserved. The most distantly related species with which D10 can be aligned are two representatives of marsupials, opossum (*Monodelphis domestica*) and Tasmanian devil (*Sarcophilus*

harrisii), that diverged from humans around 180 million years ago (Figure S2.5F). Interestingly, in four analyzed cases—mouse, dog, armadillo, and opossum—the D10 ortholog sequences are always found in synteny with the respective *Nog* ortholog genes (Figures S2.6A and S2.6B), suggesting that these conserved sequences belong to the regulatory landscape of *Nog*. Furthermore, looking at HiC data from human and mouse tissues, it is possible to appreciate that NOG and D10 enhancer, as well as the corresponding mouse ortholog sequences, are within the same topologically associating domain (TAD) (Dixon et al, 2012) in a genomic region enriched for chromatin interactions (Figures S2.6C–S2.6F). Next, we wanted to understand whether the conserved D10 orthologs correspond to notochord enhancers. In mice, the D10 ortholog sequence is located 157 kb downstream of *Nog* (Figure 2.5C). We isolated this sequence, tested it for enhancer activity using reporter assays, and demonstrated that this sequence is also a notochord enhancer (Figure 2.5D). Overall, these results suggest that the regulatory mechanism for the expression of *Nog* in the notochord might be equivalent in many vertebrate lineages.

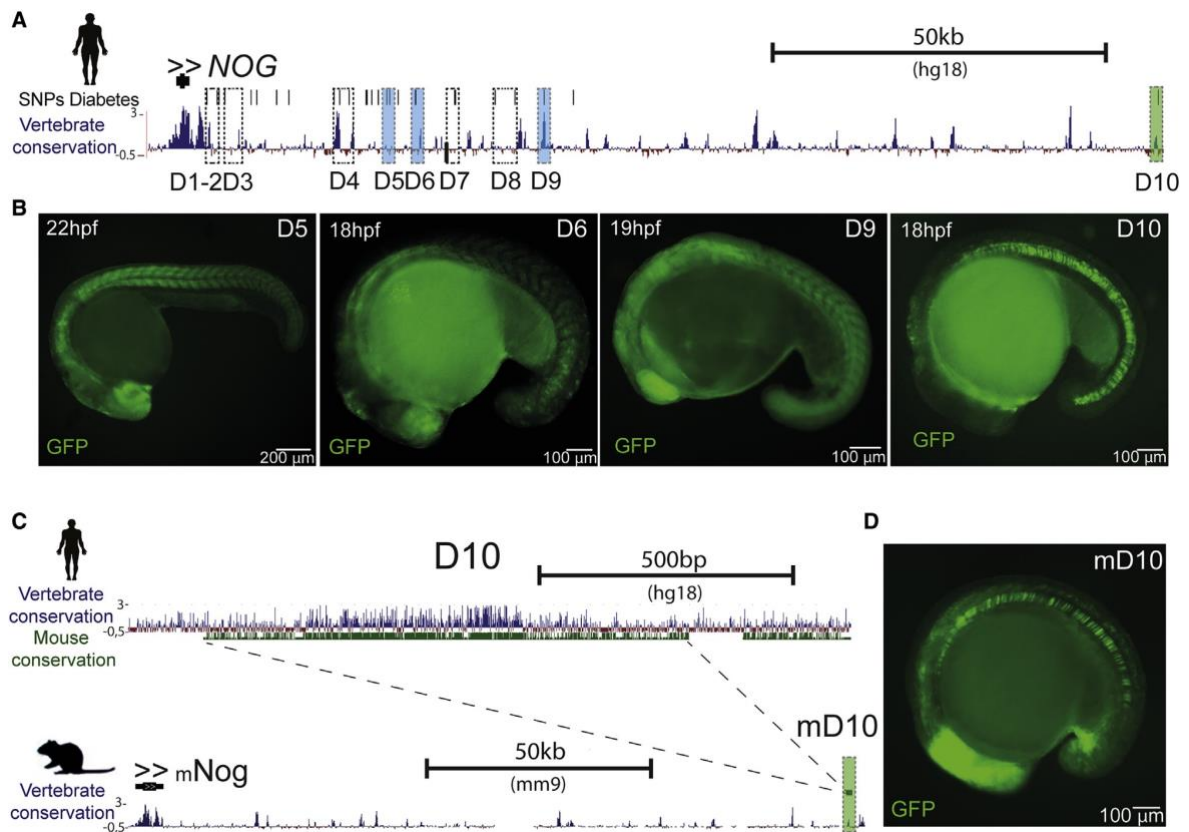


Figure 2.5 - (A) Genomic landscape of human *NOG*, comprising approximately 150 kb. Black vertical lines represent SNPs associated with diabetes from the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium ($p < 0.02$) (Morris et al, 2012; Pasquali et al, 2014). The dark blue track represents vertebrate conservation and shows multiple alignments of 100 vertebrate species and measurements of

evolutionary conservation (phyloP) (Pollard et al, 2010). Dashed line boxes are putative enhancer sequences selected for enhancer assays. Sequences in pale blue shaded boxes showed reproducible expression patterns. D10 (green shaded box) showed enhancer activity in the notochord. Scale bar represents 50 kilobases. (B) Representative images of GFP reporter lines for enhancers D5, D6, D9, and D10 (see also Figure S5). Scale bars represent 100 or 200 μ m. (C) Genomic landscape of the human D10 sequence, showing vertebrate (blue track) and mouse conservation (green track) (above). Genomic landscape of mouse *Nog*, including the conserved mouse D10 (mD10) sequence, located approximately 150 kb downstream of the promoter (see also Figures S5 and S6). Scale bars represent 500 basepairs or 50 kilobases. (D) Representative image of the GFP reporter line for the mouse enhancer mD10. Scale bar represents 100 μ m.

2.3. Discussion

Non-autonomous cell signaling is crucial for proper organogenesis. It has been demonstrated that several signaling pathways regulate pancreatic development in vertebrates. Among them, Hedgehog signaling represses the expression of the pancreatic multipotent progenitor marker *Pdx1* during late gastrulation (Tehrani and Lin, 2011). In mice, activin bB and *Fgf2* expressed in the notochord might repress *Shh* expression in the endoderm (Hebrok et al, 1998). However, ectopic pancreatic tissue is not observed in *Shh* mutant embryos (Hebrok et al, 2000), suggesting that Shh inhibition is permissive, rather than instructive, for pancreas fate (McCracken and Wells, 2012). Experiments of notochord ablation in chicken embryos showed that pancreatic development is impaired (Kim et al, 1997). These results highlight the notochord as an important signaling center for pancreatic development. In subsequent experiments using co-culture of mice tissues, it has been shown that the notochord is sufficient to trigger expression of *Pdx1* in the endoderm, but not *Insulin*, whereas cells from the dorsal aorta are able to trigger the expression of *Pdx1* and *Insulin*. These results further suggest that the notochord and the dorsal aorta have a role in the control of pancreas progenitor cells, whereas aorta endothelial cells have an extra role in the differentiation of beta cells (Lammert et al, 2001).

Here, we show that *Nog* is expressed in the notochord, controlled transversely in humans, mice, and zebrafish by at least one notochord enhancer, and that *nog2* gain and loss of function in zebrafish correlate with an increase and a decrease of pancreatic differentiation, respectively. In addition, we show that Nog2 diffuses from the notochord to the pancreatic domain, as detected indirectly by an ectopic expression system (Figure 2.3). These results indicate that Nog is one of the crucial pro-pancreatic signaling molecules emanating from the notochord. In agreement with this, Nog is a well-known

Bmp signaling inhibitor (Furthauer et al, 1999), and this signaling pathway is known to antagonize pancreas differentiation (Chung et al, 2008, 2010) while promoting liver differentiation. In this model, the Bmp2b gradient emanating from the lateral plate mesoderm is not enough for the establishment of a proper pancreatic progenitor domain by control of the expression of *pdx1*. Nog2 from the notochord, located in the midline, counteracts the lateral signaling of Bmp2b, allowing the proper placement and proper size of the pancreatic progenitor domain in the midline (Figure 2.6). In absence of Nog2 from the notochord, Bmp2b signaling is most likely expanded, resulting in a reduced pancreatic progenitor domain that will cause partial impairment of beta cell development (Figure 2.6). These phenotypes are recapitulated by impairing a *nog2* notochord enhancer, either by the action of an insulator or by the deletion of the enhancer. Nevertheless, in these genetic backgrounds, the downregulation of *nog2* is more robustly detected in the insulator background, suggesting the existence of several notochord enhancers or that partial deletion of the enhancer results in partial loss of expression of *nog2*. In both cases, knockdown of *nog2* is sufficient to induce a pancreatic phenotype.

Although *nog2* has a clear pro-pancreatic role during zebrafish development, null mutants for *nog2* do not show complete loss of pancreas differentiation. Chicken notochord ablation has been shown to induce complete loss of pancreatic differentiation (Kim et al, 1997). Therefore, it is possible that apart from *nog2*, there might be synergistic action with other pro-pancreatic signaling molecules diffusing from the notochord.

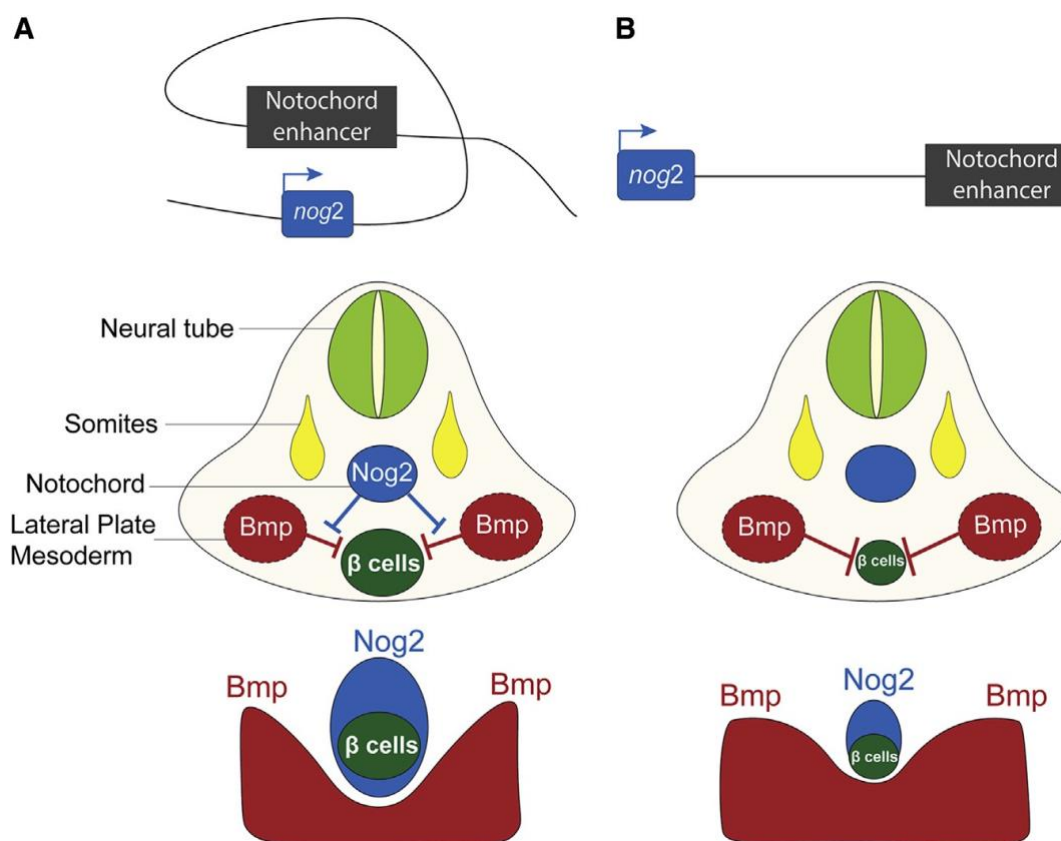


Figure 2.6 - Diagram of the *nog2* notochord enhancer driving expression of *nog2*. *Nog2* diffusion from the notochord, located in the midline of the embryo, counteracts the signaling of Bmp2b from the lateral plate mesoderm, allowing the proper placement and proper size of the endocrine pancreas. (B) When the activity of the *nog2* notochord enhancer is disrupted, *nog2* expression in the notochord is reduced. In this case, Bmp2b signaling is expanded, impairing beta cell development and reducing pancreas size.

In this work, we postulate that the function of *NOG* is conserved in humans. Supporting this, the genomic landscape of *NOG* contains a notochord enhancer, clearly showing the potential of this gene to be expressed in these cells. A study of the human notochordal transcriptome during embryogenesis was able to detect *NOG* expression during development (Rodrigues-Pinto et al, 2018), specifically at a late developmental time in human pancreatic formation. This interpretation is strengthened by advances in *in vitro* differentiation of human beta cells from human embryonic stem cells. Optimization protocols have empirically determined that the addition of *NOG* was able to improve the efficiency of such protocols, clarifying the importance of BMP inhibition in the proper development of the human pancreas (Kroon et al, 2008; Kumar et al, 2014; Mfopou et al, 2010; Zhang et al, 2009). Cai and colleagues explored the role of *NOG* in human beta cell differentiation, showing that the inclusion of *NOG* in the media resulted

in inhibition of phosphorylation of Smad1/5/8, a BMP signaling transducer (Cai et al, 2010). From an evolutionary point of view, it is known that the vertebrate phylum contains at least three distinct *Nog* genes: *Nog1*, *Nog2*, and *Nog3* (Eroshkin et al, 2006). In mammals, only *Nog1* was kept, most likely retaining all functions and expression patterns of the three ancestral genes from the vertebrate phylum. An interesting open question relates to the emergence of the *cis*-regulatory module of the notochord that in humans and mice is present in the landscape of *Nog*, the ortholog of the zebrafish *nog1*, whereas in zebrafish, the notochord enhancer is present in the landscape of *nog2*. This could be explained by the independent recruitment of recently evolved notochord enhancers, as observed in the rapid evolution of *cis*-regulatory sequences of the mammal liver (Villar et al, 2015). An alternative explanation could be the existence of an ancestral regulatory module driving the expression of *Nog* in the notochord that upon gene duplication degenerated in particular *Nog* genes. In the tetrapod lineage, the notochord *cis*-regulatory elements were most likely kept in the landscape of *Nog1*, a scenario supported by (1) the high conservation of the enhancer from humans to marsupials (Figure S2.6), (2) the expression of *Nog1* in the notochord from humans to *Xenopus* (Figure 2.5; Figure S2.5), and (3) the functional characterization of the mice and human *Nog* notochord enhancers. In contrast, the existence of the notochord enhancer in the landscape of zebrafish *nog2* suggests that this enhancer was kept in this locus in the ray-finned fish lineage. Most likely, the acquisition of the *Nog* notochord enhancer occurred after the divergence of urochordates, because expression of *Nog* is not detected in the sea squirt *Ciona intestinalis* or in a cephalochordate, the amphioxus *Branchiostoma lanceolatum*. Interestingly, insulin-producing cells in basal cephalochordates are not organized in a distinct organ and are instead disseminated in the gut (Arntfield and van der Kooy, 2011). Finally, and in light of the axochord hypothesis (Brunet et al, 2015), the *Nog* notochord enhancer might have an even more ancient origin. Supporting this hypothesis, the annelids *Terebratalia transversa* and *Platyeris dumerilii* show expression of *Nog* in the axochord, a structure reminiscent of the notochord that has a mixture of muscle and notochord double properties (Lauri et al, 2014; Passamaneck et al, 2015).

In summary, in this work, we demonstrate the transversal existence of a notochord enhancer in humans, mice, and zebrafish in the landscape of *Nog*, a known Bmp antagonist. We show that this enhancer is required for proper pancreas differentiation in the zebrafish, contributing to understanding of the notochord as an important signaling center required for the complex development of the vertebrate pancreas.

2.4. Methods

2.4.1. Materials Availability

All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

2.4.2. Data and Code Availability

Raw reads from 4C experiments (fastq files of the 2 replicates) are accessible in the European Nucleotide Archive (ENA, <https://www.ebi.ac.uk/ena>, project PRJEB36984, study ERP120260, runs ERR3964955 and ERR3964956). The processed signal is available for visualization in the following UCSC genome browser session http://genome-euro.ucsc.edu/s/vdr/danRer7_4Cseq_nog2. Perl script for processing 4C seq data is available upon request.

2.4.3. Experimental Model and Subject Details

Zebrafish

Zebrafish (*Danio rerio*) fish were handled according to European animal welfare regulations and standard protocols. The Animal Facility of the i3S is licensed by the Portuguese official veterinary department (DGAV) and accredited by the AAALAC (June, 2018), complies with Portuguese law (Portaria 1005/02 and Portaria 1131/97), the 2010/63/EU Directive, and follows the FELASA (Federation of European Laboratory Animal Science Associations) guidelines and recommendations concerning laboratory animal welfare. The current project is licensed by DGAV (reference: 0421/000/000/2016). Wild-Type (WT) zebrafish animals from AB or TU strain were used in this study. The following genotypes were used, as described in the text and Key Resources Table: $\text{nog2}^{\text{ED301}}$, nog2^{mut} , $\text{nog2E3}^{\text{del1/del2}}$, Tg (Ins:GFP), nog2E1:GFP , nog2E2:GFP , nog2E3:GFP , NOGD10:GFP , NOGD5:GFP , NOGD9:GFP and mD10:GFP . Adult animals are kept at a controlled temperature (25°C) and under a 14h light-8h dark photoperiod. Recirculated water is continuously monitored regarding pH (7), salinity (900 μs) and temperature (28°C) values. In order to obtain zebrafish embryos, adult zebrafish were bred in ratios of 3 females to 2 males or 1 male to 1 female, depending on the purpose of the experiment. Embryos ranging from 16hpf to 48hpf were used.

2.4.4. Method Details

Zebrafish transgenesis

ED301 line was generated as described by Bessa et al. (2014). Enhancer reporter assays were performed by cloning the selected genomic fragments into the ZED vector (Bessa et al, 2009). For the mouse sequence mD10 only, the pminiTol2- Z48-Gw-CARGFP vector was used (Bessa et al, 2009). Briefly, genomic regions of interest were PCR amplified using iMAXII polymerase and specific primers (Table S1). TA cloning was used to introduce PCR products in the pCR8/GW/TOPO vector, as described in the pCR8/GW/TOPO TA Cloning Kit (ThermoFisher Scientific). To clone these sequences in the destination vectors, the Gateway LR Clonase II Enzyme mix (ThermoFisher Scientific) and respective protocol was used. Vectors were purified using phenol:chloroform and injected in one-cell-stage zebrafish embryos (AB strain), together with mRNA encoding Tol2 protein, in a 1:1 ratio and a final concentration of 25 ng/ul each. Injected embryos showing reporter expression were grown to adulthood and germinal line transmission was determined by outcrossing injected fish with wild-type animals and isolating F1 embryos with GFP expression. Embryos were kept in E3 medium with 0.003% 1-phenyl-2-thiourea (SigmaAldrich) to prevent pigmentation and were staged according to Kimmel et al. (1995). Embryos showing GFP expression were anesthetized by adding tricaine (ethyl -3-aminobenzoate; SigmaAldrich) to the E3 medium, and expression patterns were documented at 18-24 hpf, using the Leica M205 stereomicroscope. Post-imaging analysis was done in Fiji-ImageJ (Schneider et al, 2012). Stable transgenic lines were selected to grow and maintained by outcrossing with wild-type animals.

Genomic DNA extraction

Batches of embryos were incubated with Chelex 100 sodium form resin (Sigma #C7901 25 gr) 5% in TE buffer and with proteinase K 1 mg/mL (ThermoFisher Scientific), at 56°C for 3 hours. After that, proteinase K activity was inactivated by incubating the samples at 100°C for 10 minutes.

Cas9 targeting, sgRNA synthesis, mutant generation

The sgRNAs targeting the coding regions of *nog2* and the regions flanking *nog2E3* (Table S1) were designed using the Crisprscan software (Moreno-Mateos et al, 2015). These oligonucleotides were annealed *in vitro* by denaturation at 95°C for 5 min followed by slow cooling at RT. Annealed oligonucleotides (1:10) were further inserted into 100ng

of pDR274 vector (Hwang et al, 2013; Addgene plasmid #42250) previously cut with BsaI (Anza, ThermoScientific). pDR274 vectors carrying sgRNA sequences were linearized with HindIII, purified with phenol:chloroform and transcribed with T7 RNA polymerase (ThermoFisher Scientific). sgRNAs were purified following the mRNA purification protocol previously described (Bessa et al, 2009). To achieve *nog2* knockdown, one-cell-stage zebrafish embryos were injected with Cas9 protein (PNA Bio Inc) (150 250 ng/ul) and 1 sgRNA (80125 ng/ul), targeting the coding region of *nog2* (sg1*nog2*, Table S1). To perform the *nog2*E3 enhancer knockdown, embryos were co-injected with a mixture containing a combination of either Cas9 protein or dCas9-KRAB mRNA and 2 sgRNAs (80125 ng/ul each) (sg1E3, sg2E3, Table S1). In both cases, zebrafish mutants for *nog2*E3 were generated using the combination of sgRNA1 and sgRNA2. Enhancer deletion in zebrafish was detected by PCR with HOT FIREPol® DNA Polymerase (Solis BioDyne) using primers flanking the region of interest (see Figure S3 and Table S1). PCR product was analyzed by electrophoresis in 2% agarose gel, and further confirmed by Sanger sequencing. Injected embryos were grown and adult fish fins were cut to perform genomic DNA extraction and genotyping as described above (see Figure S4 and Table S1). One heterozygous male and one female having similar deletions were selected and crossed. Embryos were genotyped by fin clipping and genomic DNA extraction, followed by PCR amplification using the primers depicted in Table S1 (see also Figure S4).

Whole-mount *in situ* hybridization

Whole-mount *in situ* hybridization was performed on zebrafish embryos staged according to Kimmel et al. (1995), using antisense riboprobes targeting *nog2* or *insulin*. *nog2* probe was amplified from zebrafish cDNA, using the primers depicted in Table S1, and cloned into pCR2.1-TOPO TA vector (ThermoFisher Scientific). A vector for generating an *insulin*-targeting probe was kindly sent to us by Francesco Argenton (Argenton et al, 1999). Vectors were linearized using HindIII restriction enzyme (ThermoFisher Scientific). Antisense RNA probes were synthesized, using T7 RNA polymerase (ThermoFisher Scientific) and digoxigenin-labeled nucleotides (Roche) and kept in ISH hybridization buffer (HYB+; Thisse et al, 1993). Zebrafish embryos were collected at 18 or 48hpf and fixed overnight in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). Fixed embryos were washed in PBS-Tween 0.1% (PBT) and dehydrated in methanol (MeOH), in a series of MeOH/PBT solutions. Dehydrated embryos were kept at -20°C in 100% MeOH. After this, embryos were rehydrated through a reverse MeOH/PBT series and then treated with proteinase K 1 mg/mL

(ThermoFisher Scientific), for 1–10 min. Embryos were washed in PBT, post-fixed in 4% PFA for 20 min and washed again in PBT. Pre-hybridization was performed by incubating embryos in HYB+ buffer for 3 h, at 70°C. After that, hybridization buffer was replaced with solution containing riboprobes at the concentration of 2 ng/uL and an overnight incubation was performed, at 70°C. In the next day, embryos were washed with a series of solutions containing washing buffer (HYB-) and sodium saline citrate buffer (SSC), as described by Thisse et al. (1993). Embryos were incubated for 1 h in blocking buffer, containing 2% normal goat serum (ThermoFisher Scientific) and 2 mg/mL bovine serum albumin (BSA; Sigma-Aldrich). Following this, embryos were incubated for 2 h with anti-digoxigenin-AP antibody (Sigma-Aldrich) 1:5000, dissolved in blocking buffer. Several rinses in PBT were performed. In the third day, embryos were washed with alkaline-phosphatase (AP) buffer (Thisse et al, 1993) and AP activity was revealed through incubation in NBT/BCIP substrate (Sigma-Aldrich), in the dark, for at least 30 min. Embryos were rinsed in PBT to stop the reaction and kept in 100% glycerol at 4°C. Coloration associated to the presence of mRNA was observed and photographed, using the Leica M205 stereomicroscope. Post-imaging analysis was done in Fiji-ImageJ.

Immunohistochemistry

Whole-mount immunohistochemistry was performed using mouse anti-rat *nkx6.1* (Developmental Studies Hybridoma Bank) and rabbit anti-insulin (abcam) antisera. Goat antimouse IgG AlexaFluor 488 (Invitrogen), goat anti-rabbit IgG AlexaFluor 568 (Invitrogen) and goat anti-mouse IgG AlexaFluor 647 (Invitrogen) were used as secondary antisera. Briefly, embryos were fixed overnight in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). Fixed embryos were sequentially washed with PBS Triton 0.1%, then treated for 30 min to 2 h with PBS-Triton 1%, washed again with PBS-Triton 0.1% and incubated for 1h with PBS-T+BSA 5% before the incubation with the primary antisera, diluted 1:50 in PBS-T+BSA, overnight at 4°C. After extensive wash, embryos were incubated with the secondary antisera, overnight at 4°C, washed extensively and mounted in 50% glycerol in PBS. Embryos were stained with DAPI and 1,5 µm Z-series stacks were acquired on a Leica SP5 confocal microscope, using a 40x objective. Post-imaging analysis and quantifications were done in Fiji-ImageJ.

Nog2 overexpression vector

The sequence encoding the signaling domain of the *Nog2* protein was PCR amplified from zebrafish genomic DNA, using primers carrying adaptor ends for BamHI and NcoI

restriction enzymes. The mature encoding *nog2* domain was amplified using 5'EcoRI and 3' XhoI restriction sites. The PCR products were sequentially cloned into pCS2GFP (Bessa et al, 2008) by restriction enzyme cloning and the vector was sequenced. The *nog2 signaling_GFP_mature nog2* sequence was introduced into the #237 pMEMCS vector of the tol2 kit (Kwan et al, 2007), by BamHI/ XhoI cloning. The minimal promoter–E3 enhancer cassette was obtained by BamHI/ SmaI digestion from the ZEDnog2E3 vector and cloned into the #228 p5EMCS vector of the tol2 kit. The generated #237 and #228 vectors were recombined with the #302 into the #394 destination vector using recombination driven by clonase II.

Morpholino injections

For the *nog2* knockdown, 25 nL of 1 mM solution of morpholino targeting *nog2* 5'UTR or *Danio rerio* standard control morpholino (Table S1) (Gene tools, LLC) were injected in one-cell-stage Tg(ins:GFP) zebrafish embryos. Prior to the injection, morpholinos were incubated at 65°C for 10 min, vortexed, and diluted to 1mM in HyPure H2O. Injected embryos were grown to 48 hpf and fixed overnight at 4°C in 4% paraformaldehyde in PBS. Fixed embryos were stained with DAPI and 1,5 µm Z-series stacks were acquired on a Leica SP5 confocal microscope, using a 40x objective. GFP + cells were counted using Fiji-ImageJ. Embryos from 3 independent injections were analyzed for each condition.

Nog2 mRNA injections

25 nL of an 8 ng/µL solution of Nog2 WT or Nog2^{mut} mRNAs were injected in one-cell-stage zebrafish embryos. Dorsalization phenotypes were observed and photographed around 24hpf, using the Leica M205 stereomicroscope. Post-imaging analysis was done in Fiji-ImageJ.

4C-seq

4C-seq was performed according with Fernández-Miñán et al. (2016), with minor alterations. For each replicate, at least 200 24hpf zebrafish embryos were collected. Embryos were fixed in 2% formaldehyde for 10 min, and stored at 80°C. Cell lysis was performed on ice, with a 15 mL Tenbroeck Homogenizer, not exceeding 10 min. Ligation was performed by using 60U T4 DNA Ligase (#EL0012, ThermoFisher Scientific). Restriction enzymes used were DpnII (#R0543M, NEB) and Csp6I (#ER0211, ThermoScientific) for the first and second cuts, respectively. Chromatin was purified by

Amicon Ultra15 Centrifugal Filter Device (Milipore). 4C libraries were prepared for Illumina sequencing by using Expand Long Template Polymerase (Roche) with primers targeting the TSSs of *nog2* (Table S1) and including Illumina adapters. Final PCR product was purified by High Pure PCR Product Purification Kit (Roche) and AMPure XP PCR purification kit (Agencourt AMPure XP). 4C-seq libraries were first inspected for quality control using FASTQC v.0.11.5 (Data S1) and demultiplexed using the script “demultiplex.py” from the FourCSeq package (Klein et al, 2015), allowing for 1 mismatch in the primer sequence. A custom perl script was used for subsequent processing, as previously described (Noordermeer et al, 2011, Splinter et al, 2012). Shortly, reads were aligned to the zebrafish genome (Zv9/danRer7) using Bowtie (Langmead et al, 2009), keeping only uniquely mapping reads (v1.1.2, m 1). Reads within fragments flanked by restriction sites of the same enzyme, checked with BedTools (Quinlan and Hall, 2010), or fragments smaller than 40 bp were filtered out. Mapped reads were then converted to readsperfirstenzymefragmentend units, and smoothed using a 30 fragment mean running window algorithm.

Quantitative PCR

Total mRNA was extracted from batches of 30 embryos or single embryos at 12 hpf with TRIzol (Invitrogen) and subsequent phenol:chloroform purification. Total mRNA was used for reverse transcription with random primers (SuperScript II, Invitrogen). Quantitative PCR was performed on a BioRad CFX96 realtime system using Sybr (iSybr Green Supermix, BioRad), using the primers depicted in Table S1. Each biological replicate was quantified in triplicates. *eef1a* and beta-actin or beta-actin alone were used as internal controls, and expression levels were compared using the DDCT method.

Quantification and Statistical Analysis

Assuming that the data are normally distributed, an unpaired t-test with Welch's correction was performed to test the significance of differences among sample averages. For the comparison of *nog2* relative expression between single embryos, a Mann-Whitney test to compare ranks was used. Asterisks in each figure represent significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, n.s. = non-significant. Statistics and plots were done using GraphPad Prism 6.

2.5. Supplementary Information

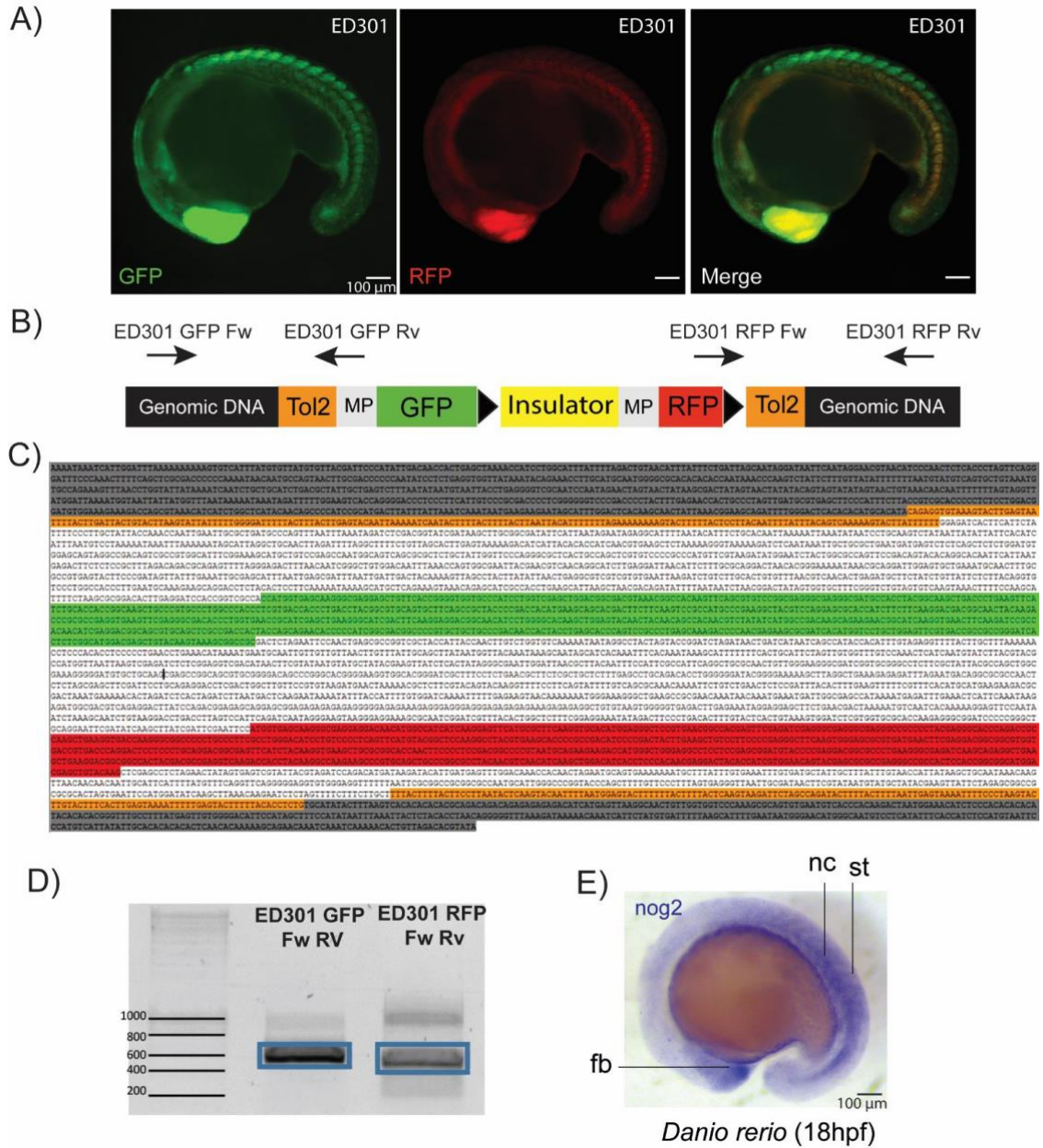


Figure S2.1. Related to Figures 2.1 and 2.2 - A) Representative images of the ED301 line expression pattern at 18hpf. GFP (green) expression is seen in the forebrain, hindbrain, somites and notochord and RFP (red) expression is seen in forebrain and notochord. B) Diagram of the ED transposon located in the *nog2* locus in the ED301 transgenic line. Arrows represent primers used to confirm the ED transposon integration site. C) Sequence of the ED vector and adjacent genomic regions (dark grey) in the ED301 integration. D) Agarose gel showing PCR products obtained using the depicted primers (see Table S1), confirming the ED301 integration. E) In situ hybridization of *nog2*, in a 18hpf representative embryo. fb – forebrain, nc – notochord, st – somites. Scale bar represents 100 μ m.

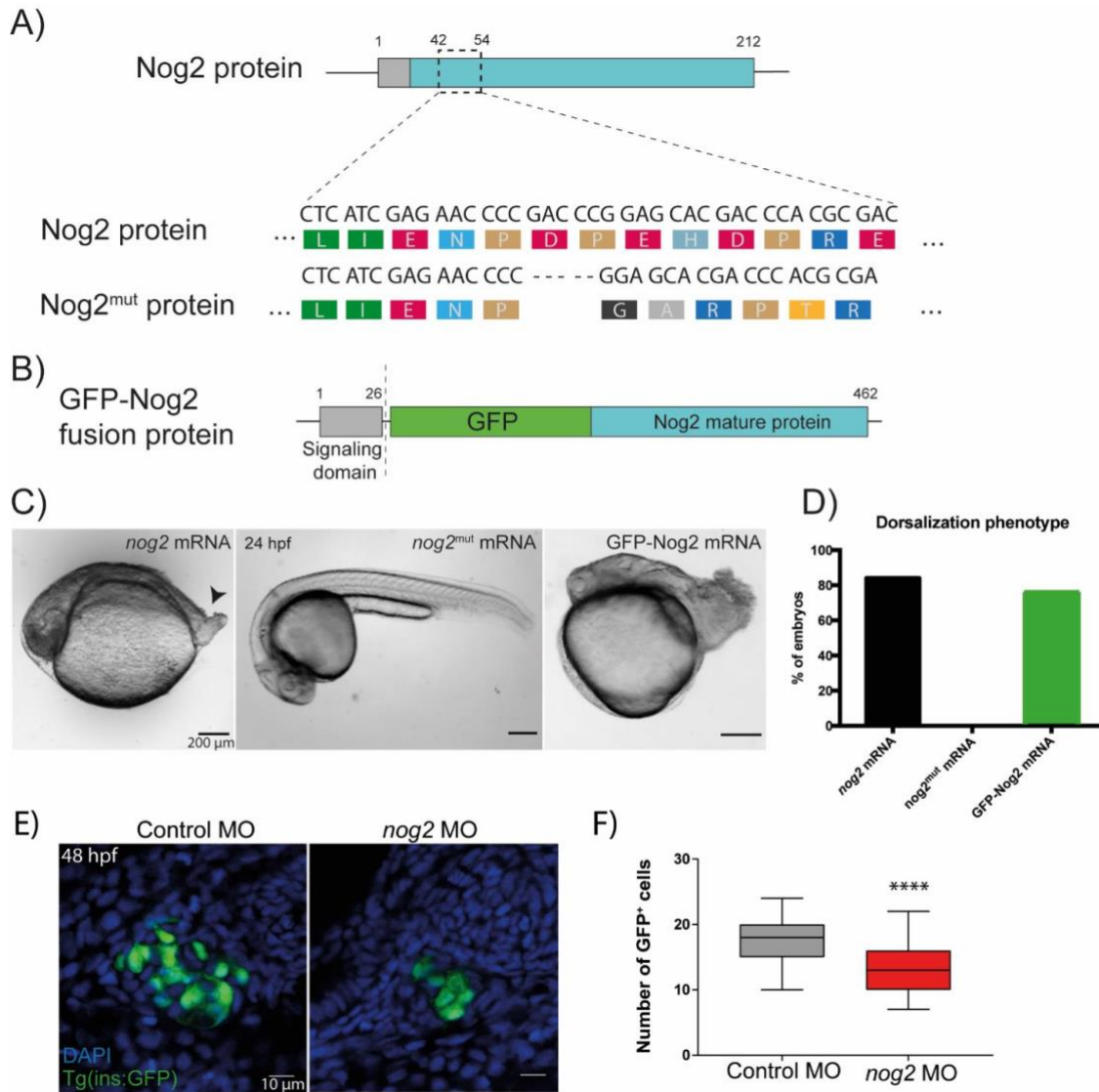


Figure S2.2. Related to Figures 2.1, 2.3 and 2.4 - A) Representation of the nucleotide and corresponding translated amino acid sequences of Nog2 and Nog2^{mut} proteins, focusing on the region of the mutation. B) Diagram of the GFPNog2 fusion protein, depicting the signaling domain of Nog2, which is removed during the maturation of the protein. C) Representative images of 24hpf embryos injected with Nog2 mRNA, Nog2^{mut} mRNA and GFPNog2 mRNA. Only the mutant version of the protein fails to cause hyperdorsalization phenotypes (embryos with ovoid shape and loss of ventral structures such as tail and trunk – arrowhead) Scale bars represent 200 μm. D) Quantification of the number of embryos showing dorsalization phenotypes. In all cases n≥61. E) Representative confocal images of an insulin GFP reporter line (Tg (*ins*:GFP); green) for 48hpf embryos injected with a control morpholino (Control MO) and a *nog2* targeting morpholino (*nog2*MO). Embryos were counterstained with DAPI nuclear marker (blue) Scale bars represent 10 μm. F) Quantification of the number of GFP positive cells from 48hpf Tg (*ins*:GFP) embryos injected with a control morpholino (Control) and a *nog2* targeting morpholino (*nog2*MO). Error bars represent SD and **** denotes p-value < 0.0001. In both cases n=52.

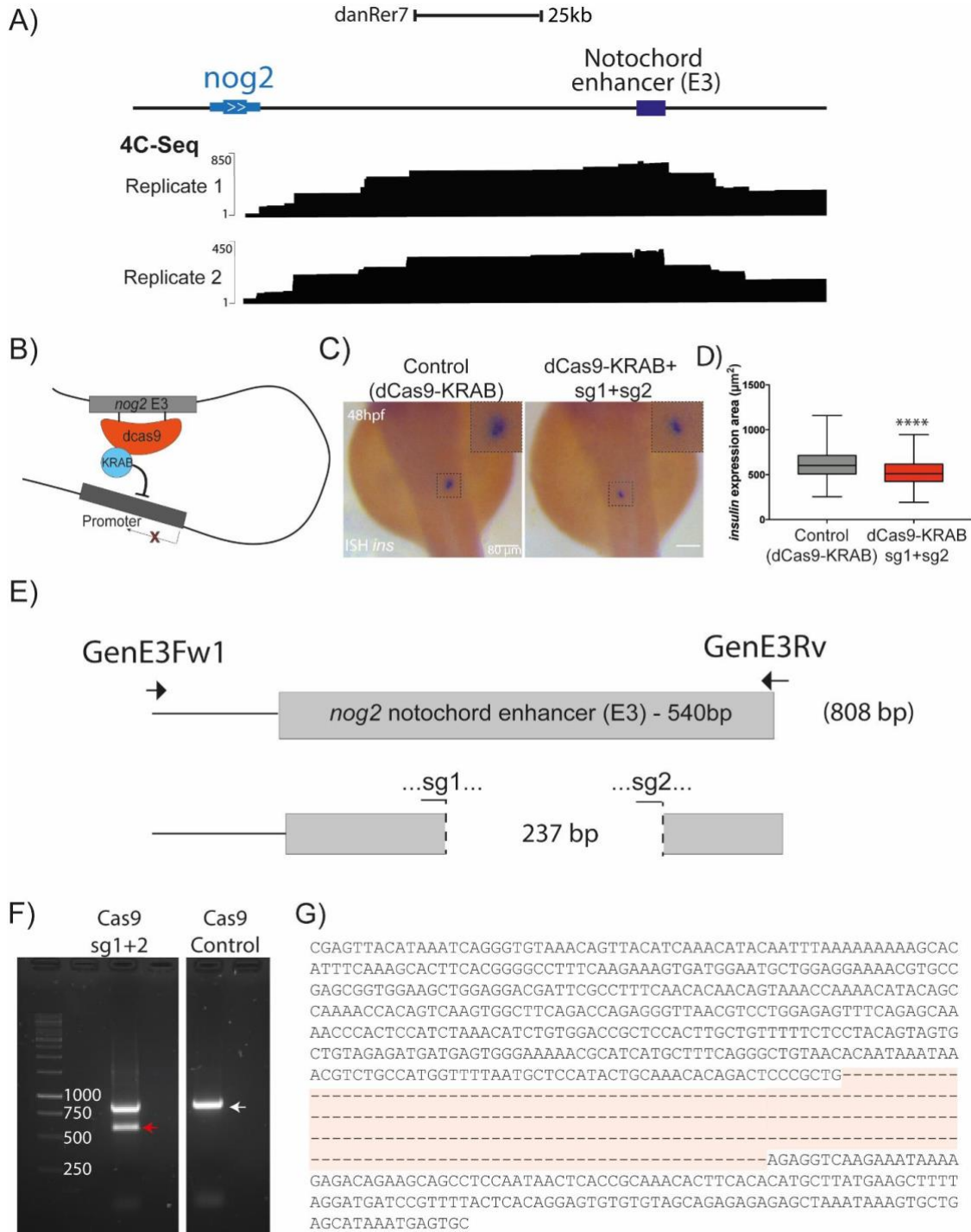


Figure S2.3. Related to Figures 2.2 and 2.4 - A) 4C-seq replicates (black tracks) showing a physical interaction between the *nog2* promoter and a region of approximately 100 kb downstream of the gene, containing the E3 notochord enhancer. Scale bar represents 25 kilobases. B) Representation of the repression of *nog2* expression, mediated by the KRAB domain, fused to a dCas9 protein, which targets the *nog2*E3 enhancer via sg1 and sg2. C) *In situ* hybridization of *insulin* in 48hpf representative embryos injected with dCas9-KRAB mRNA and two sgRNAs, comparing to control embryos, injected with only dCas9-KRAB. Scale bars represent 80 μ m. D) Quantification of the *insulin* expression area, detected by *in situ* hybridization in injected embryos at 48hpf, comparing to controls. Error bars represent SD and **** denotes p-value < 0.0001. In all cases $n \geq 99$. E) Representation of the sg1 and sg2 location within the E3 enhancer and

predicted genomic deletions with an expected size of 237bp. The full size of the E3 sequence is 540 bps and was selected based on ATAC-Seq and Chip-Seq data (see Figure 2 – pale blue box). Primers used to genotype the somatic deletions are also shown (Arrows; GenE3Fw1 and GenE3Rv; see Table S1). F) Upon injection of Cas9 protein, sg1 and sg2 RNAs (Cas9 sg1+2), a batch of 8 embryos was genotyped, using the GenE3Fw1 and GenE3Rv primers, amplifying a shorter band compatible with a 237bp genomic deletion (red arrow). This band was not present in control animals injected with Cas9 alone (Cas9 Control), where only the wild type band was observed (white arrow). Cas9 sg1+2 and Cas9 control are not consecutive lanes from the same gel. G) The short band observed in B) (red arrow) was sequenced, confirming a somatic deletion in the E3 enhancer.

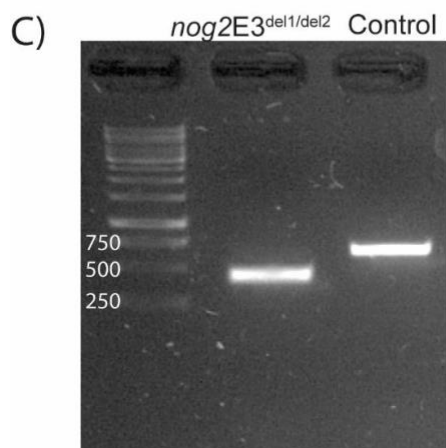
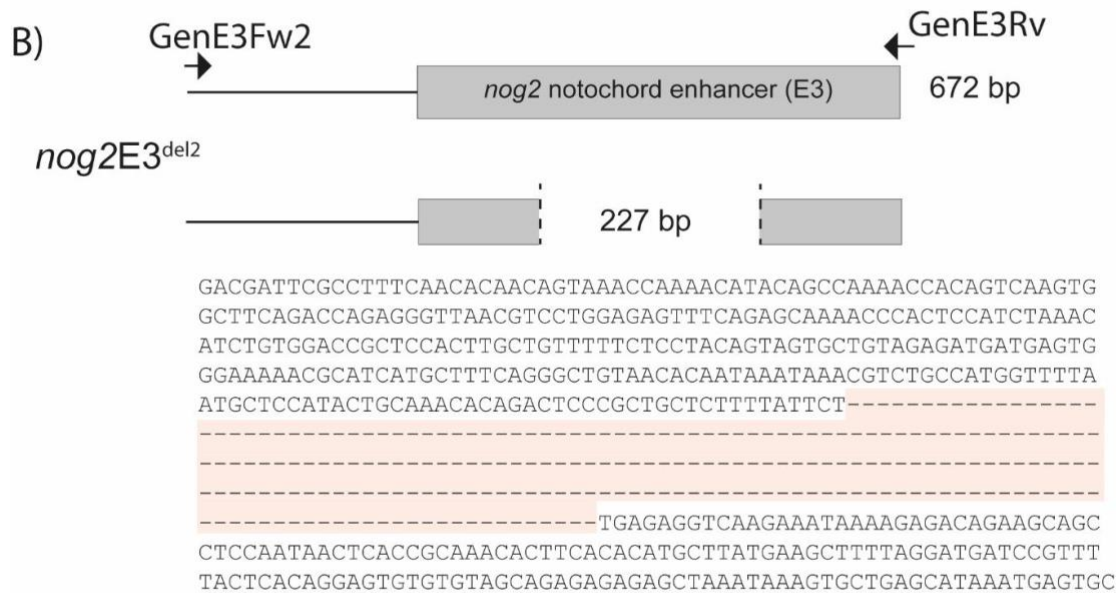
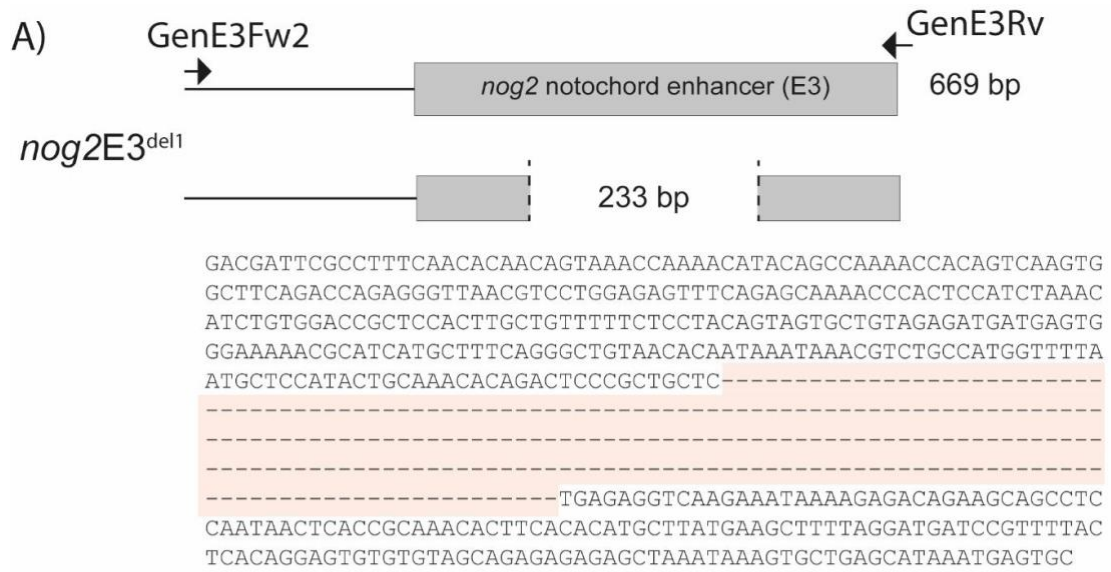


Figure S2.4. Related to Figure 2.4 - A and B) Representation of two generated genomic deletions of the E3 enhancer (del1 and del2), with sizes of 233 and 227bp, respectively. Adult heterozygous fish were genotyped by fin clipping followed by genomic DNA extraction and PCR amplification, using the GenE3Fw2 and GenE3Rv primers (see Table S1). The bands corresponding to alleles containing the deletion were extracted and sequenced. C) One heterozygous male and one female were crossed and their progeny was genotyped to select wt (control) and homozygous embryos for the deletion (*nog2E3*^{del1/del2}). Correspondent amplified bands are shown, in a 1.5% agarose gel.

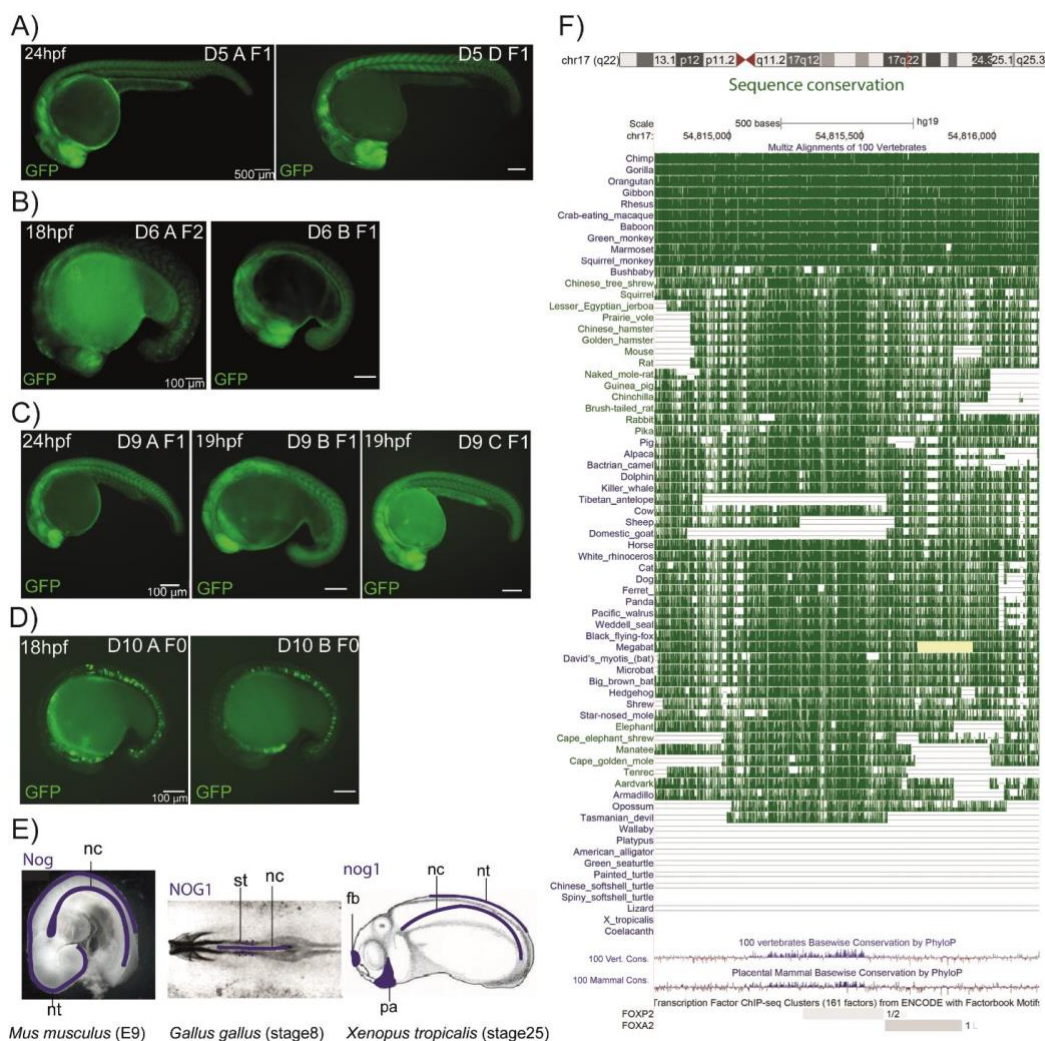
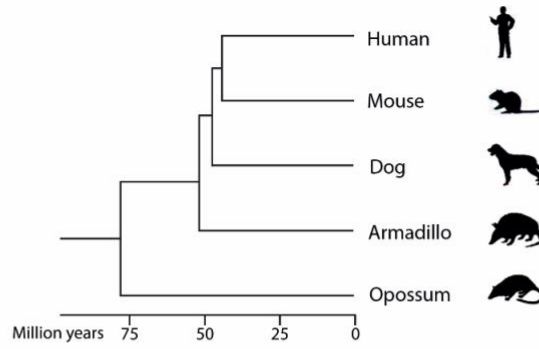
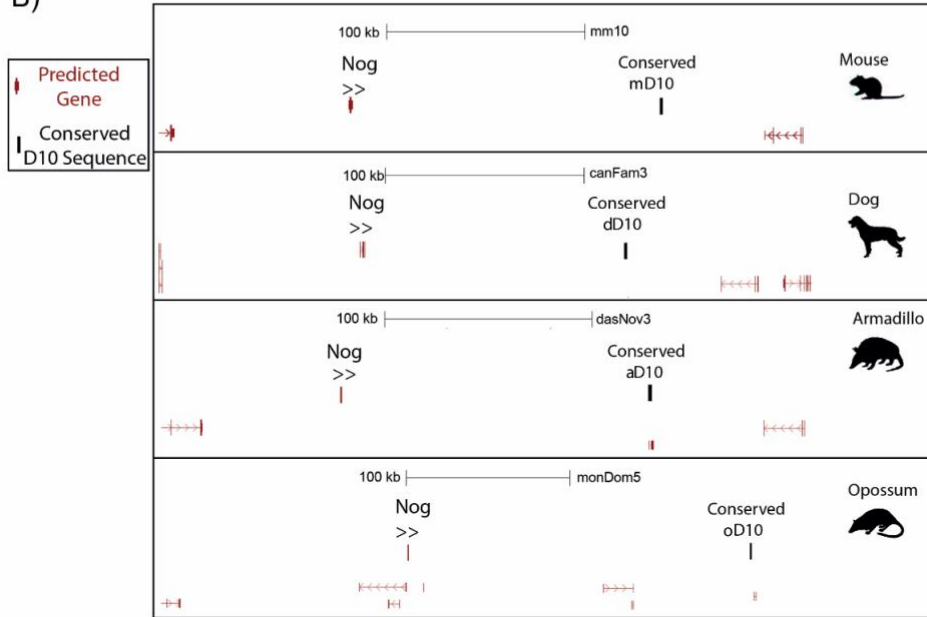


Figure S2.5. Related to Figure 2.5 - Representative images of two or more independent transgenic lines showing a reproducible GFP (green) expression pattern for the A) D5 (somites, forebrain and hindbrain), B) D6 (somites and forebrain) and C) D9 (not restricted) enhancers. D) Representative images of two injected embryos showing reproducible expression pattern of GFP in the notochord, driven by the D10 enhancer. E) Representation of the expression patterns of *Nog* genes (dark blue) of *Mus musculus*, *Gallus gallus* and *Xenopus tropicalis* at early developmental stages. fb – forebrain, nc – notochord, nt – neural tube, pa – pharyngeal arches, st – somites. F) Tracks from USCS genome browser Multiz Alignments of 100 Vertebrates (Blanchette et al, 2004) showing conservation of the human D10 enhancer among mammal species. The most distantly related species that D10 is able to be aligned with, are two representatives of marsupials, Opossum (*Monodelphis domestica*) and Tasmanian Devil (*Sarcophilus harrisii*).

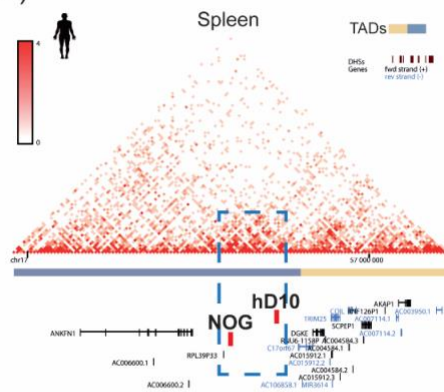
A)



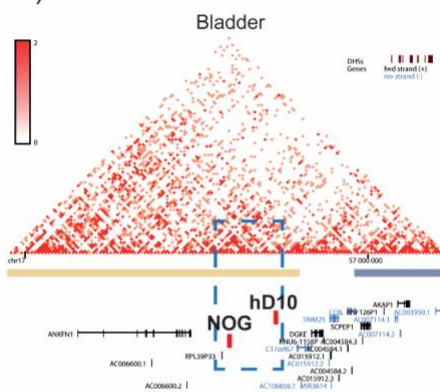
B)



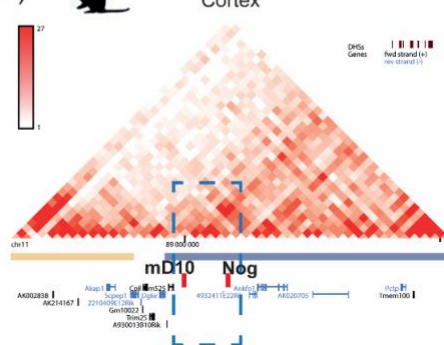
C)



D)



E)



F)

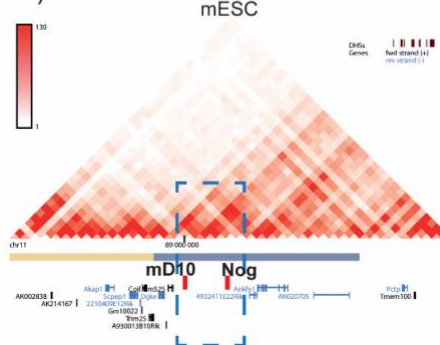


Figure S2.6. Related to Figure 2.5 - A) Phylogenetic tree representing the estimated time of divergence (million years) of selected mammals (humans, mouse, dog, armadillo and opossum) that share sequence conservation detected by alignment with the human D10 enhancer. B) Representation of the genomic landscape of Nog genes of selected mammals, showing synteny blocks containing Nog and ortholog sequences aligned with the human D10 enhancer. C and D) Representation of Hi-C data depicting the interaction between the human NOG gene and the hD10 enhancer (both in red) in spleen and bladder tissues (Schmitt et al, 2016), respectively. E and F) Representation of Hi-C data depicting the interaction between the mouse Nog gene and the mD10 enhancer (both in red) in cortex and mESCs (Dixon et al, 2012), respectively. Images were generated using the 3D Genome Browser (Wang et al, 2018).

2.6 Supplementary datasets

The supplementary dataset described in this chapter are available online, as supplementary material of the publication:

<https://www.sciencedirect.com/science/article/pii/S2211124720308433>

2.7. References

- Argenton, F., Zecchin, E., and Bortolussi, M. (1999). Early appearance of pancreatic hormone-expressing cells in the zebrafish embryo. *Mech. Dev.* 87, 217–221.
- Arntfield, M.E., and van der Kooy, D. (2011). b-Cell evolution: How the pancreas borrowed from the brain: The shared toolbox of genes expressed by neural and pancreatic endocrine cells may reflect their evolutionary relationship. *BioEssays* 33, 582–587.
- Bessa, J., Tavares, M.J., Santos, J., Kikuta, H., Laplante, M., Becker, T.S., Gomez-Skarmeta, J.L., and Casares, F. (2008). *meis1* regulates cyclin D1 and c-myc expression, and controls the proliferation of the multipotent cells in the early developing zebrafish eye. *Development* 135, 799–803.
- Bessa, J., Tena, J.J., de la Calle-Mustienes, E., Fernandez-Miñán, A., Naranjo, S., Fernandez, A., Montoliu, L., Akalin, A., Lenhard, B., Casares, F., and Gómez-Skarmeta, J.L. (2009). Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of *cis*-regulatory regions in zebrafish. *Dev. Dyn.* 238, 2409–2417.
- Bessa, J., Luengo, M., Rivero-Gil, S., Ariza-Cosano, A., Maia, A.H, Ruiz- Ruano, F.J., Caballero, P., Naranjo, S., Carvajal, J.J., and Gómez-Skarmeta, J.L. (2014). A mobile insulator system to detect and disrupt *cis*-regulatory landscapes in vertebrates. *Genome Res.* 24, 487–495.
- Binot, A.C., Manfroid, I., Flasse, L., Winandy, M., Motte, P., Martial, J.A., Peers, B., and Voz, M.L. (2010). Nkx6.1 and nkx6.2 regulate a- and b-cell formation in zebrafish by acting on pancreatic endocrine progenitor cells. *Dev. Biol.* 340, 397–407.
- Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F.A., Roskin, K.M., Baertsch, R, Rosenbloom, K., Clawson, H., Green, E.D., et al. (2004). Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 14, 708–715.
- Bogdanovic, O., Fernandez-Miñán, A., Tena, J.J., de la Calle-Mustienes, E., Hidalgo, C., van Kruysbergen, I., van Heeringen, S.J., Veenstra, G.J., and Gómez-Skarmeta, J.L. (2012). Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. *Genome Res.* 22, 2043–2053.
- Brunet, T., Lauri, A., and Arendt, D. (2015). Did the notochord evolve from an ancient axial muscle? The axochord hypothesis. *BioEssays* 37, 836–850.
- Cai, J., Yu, C., Liu, Y., Chen, S., Guo, Y., Yong, J., Lu, W., Ding, M., and Deng, H. (2010).

Generation of homogeneous PDX1(+) pancreatic progenitors from human ES cell-derived endoderm cells. *J. Mol. Cell Biol.* 2, 50–60.

Chung, W.S., Shin, C.H., and Stainier, D.Y.R. (2008). Bmp2 signaling regulates the hepatic versus pancreatic fate decision. *Dev. Cell* 15, 738–748.

Chung, W.S., Andersson, O., Row, R., Kimelman, D., and Stainier, D.Y.R. (2010). Suppression of Alk8-mediated Bmp signaling cell-autonomously induces pancreatic beta-cells in zebrafish. *Proc. Natl. Acad. Sci. USA* 107, 1142–1147.

dilorio, P.J., Moss, J.B., Sbrogna, J.L., Karlstrom, R.O., and Moss, L.G. (2002). Sonic hedgehog is required early in pancreatic islet development. *Dev. Biol.* 244, 75–84.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.

Eroshkin, F.M., Ermakova, G.V., Bayramov, A.V., and Zarsky, A.G. (2006). Multiple noggins in vertebrate genome: cloning and expression of noggin2 and noggin4 in *Xenopus laevis*. *Gene Expr. Patterns* 6, 180–186.

Fernandez-Miñán, A., Bessa, J., Tena, J.J., and Gómez-Skarmeta, J.L. (2016). Assay for transposase-accessible chromatin and circularized chromosome conformation capture, two methods to explore the regulatory landscapes of genes in zebrafish. *Methods Cell Biol.* 135, 413–430.

Furthauer, M., Thisse, B., and Thisse, C. (1999). Three different noggin genes antagonize the activity of bone morphogenetic proteins in the zebrafish embryo. *Dev. Biol.* 214, 181–196.

Hebrok, M., Kim, S.K., and Melton, D.A. (1998). Notochord repression of endodermal Sonic hedgehog permits pancreas development. *Genes Dev.* 12, 1705–1713.

Hebrok, M., Kim, S.K., St Jacques, B., McMahon, A.P., and Melton, D.A. (2000). Regulation of pancreas development by hedgehog signaling. *Development* 127, 4905–4913.

Hesselton, D., Anderson, R.M., and Stainier, D.Y.R. (2011). Suppression of Ptf1a activity induces acinar-to-endocrine conversion. *Curr. Biol.* 21, 712–717.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-

Cas system. *Nat. Biotechnol.* **31**, 227–229.

Inomata, H., Shibata, T., Haraguchi, T., and Sasai, Y. (2013). Scaling of dorsal-ventral patterning by embryo size-dependent degradation of Spemann's organizer signals. *Cell* **153**, 1296–1311.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res.* **12**, 996–1006.

Kim, S.K., Hebrok, M., and Melton, D.A. (1997). Notochord to endoderm signaling is required for pancreas development. *Development* **124**, 4243–4252.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310.

Klein, F.A., Pakozdi, T., Anders, S., Ghavi-Helm, Y., Furlong, E.E.M., and Huber, W. (2015). FourCSeq: analysis of 4C sequencing data. *Bioinformatics* **31**, 3085–3091.

Kroon, E., Martinson, L.A., Kadoya, K., Bang, A.G., Kelly, O.G., Eliazar, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nat. Biotechnol.* **26**, 443–452.

Kumar, S.S., Alarfaj, A.A., Munusamy, M.A., Singh, A.J., Peng, I.C., Priya, S.P., Hamat, R.A., and Higuchi, A. (2014). Recent developments in b-cell differentiation of pluripotent stem cells induced by small and large molecules. *Int. J. Mol. Sci.* **15**, 23418–23447.

Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant, J.M., Yost, H.J., Kanki, J.P., and Chien, C.B. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–3099.

Lammert, E., Cleaver, O., and Melton, D. (2001). Induction of pancreatic differentiation by signals from blood vessels. *Science* **294**, 564–567.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.

Lauri, A., Brunet, T., Handberg-Thorsager, M., Fischer, A.H., Simakov, O., Steinmetz, P.R., Tomer, R., Keller, P.J., and Arendt, D. (2014). Development of the annelid

axochord: insights into notochord evolution. *Science* **345**, 1365–1368.

Letelier, J., de la Calle-Mustienes, E., Pieretti, J., Naranjo, S., Maeso, I., Nakamura, T., Pascual-Anaya, J., Shubin, N.H., Schneider, I., Martinez-Morales, J.R., and Gómez-Skarmeta, J.L. (2018). A conserved Shh *cis*-regulatory module highlights a common developmental origin of unpaired and paired fins. *Nat. Genet.* **50**, 504–509.

Marlétaz, F., Firbas, P.N., Maeso, I., Tena, J.J., Bogdanovic, O., Perry, M., Wyatt, C.D.R., de la Calle-Mustienes, E., Bertrand, S., Burguera, D., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* **564**, 64–70.

McCracken, K.W., and Wells, J.M. (2012). Molecular pathways controlling pancreas induction. *Semin. Cell Dev. Biol.* **23**, 656–662.

Mfopou, J.K., Chen, B., Mateizel, I., Sermon, K., and Bouwens, L. (2010). Noggin, retinoids, and fibroblast growth factor regulate hepatic or pancreatic fate of human embryonic stem cells. *Gastroenterology* **138**, 2233–2245.

Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting *in vivo*. *Nat. Methods* **12**, 982–988.

Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segre, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network-Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., Laat, W.D., and Duboule, D. (2011). The Dynamic Architecture of Hox Gene Clusters. *Science* **334** (6053), 222–225.

Pasquali, L., Gaulton, K.J., Rodríguez-Seguíl, S.A., Mularoni, L., Miguel-Escalada, I., Akerman, I., Tena, J.J., Morán, I., Gómez-Marín, C., van de Bunt, M., et al. (2014). Pancreatic islet enhancer clusters enriched in type 2 diabetes risk- associated variants. *Nat. Genet.* **46**, 136–143.

Passamaneck, Y.J., Hejnal, A., and Martindale, M.Q. (2015). Mesodermal gene

expression during the embryonic and larval development of the articulate brachiopod *Terebratalia transversa*. *Evodevo* 6, 10.

Pollard, K.S., Hubisz, M.J, Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* 20, 110–121.

Prince, V.E., Anderson, R.M., and Dalgin, G. (2017). Zebrafish Pancreas Development and Regeneration: Fishing for Diabetes Therapies. *Curr. Top. Dev. Biol.* 124, 235–276.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.

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

Rodrigues-Pinto, R., Ward, L., Humphreys, M., Zeef, L.A.H., Berry, A., Hanley, K.P., Hanley, N., Richardson, S.M., and Hoyland, J.A. (2018). Human noto- chordal cell transcriptome unveils potential regulators of cell function in the developing intervertebral disc. *Sci. Rep.* 8, 12866.

Sakhneny, L., Khalifa-Malka, L., and Landsman, L. (2019). Pancreas organogenesis: Approaches to elucidate the role of epithelial-mesenchymal interactions. *Semin. Cell Dev. Biol.* 92, 89–96.

Schmitt, A.D, Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C.L., Li, Y., Lin, S., Lin, Y., Barr, C.L., and Ren, B. (2016). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep.* 17, 2042– 2059.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675.

Splinter, E., Wit, E.D., Werken, H.J., Klous, P, and Laat, W.D. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. *Methods* 58 (3), 221–230.

Tehrani, Z., and Lin, S. (2011). Antagonistic interactions of hedgehog, Bmp and retinoic acid signals control zebrafish endocrine pancreas development. *Development* 138, 631–640.

Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome

editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149.

Thisse, C., Thisse, B., Schilling, T.F., and Postlethwait, J.H. (1993). Structure of the zebrafish *snail1* gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development* **119**, 1203–1215.

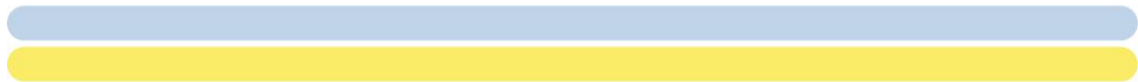
Villar, D., Berthelot, C., Aldridge, S., Rayner, T.F., Lukk, M., Pignatelli, M., Park, T.J., Deaville, R., Erichsen, J.T., Jasinska, A.J., et al. (2015). Enhancer evolution across 20 mammalian species. *Cell* **160**, 554–566.

Wang, Y., Song, F., Zhang, B., Zhang, L., Xu, J., Kuang, D., Li, D., Choudhary, M.N.K., Li, Y., Hu, M., et al. (2018). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol.* **19**, 151.

Zhang, D., Jiang, W., Liu, M., Sui, X., Yin, X., Chen, S., Shi, Y., and Deng, H. (2009). Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells. *Cell Res.* **19**, 429–438.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., et al. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* **38**, 1341–1347.

Chapter III



A distal enhancer regulates *ptf1a* expression to ensure proper lineage acquisition and maintenance of pancreatic acinar cells

This chapter is adapted from an original research article, in preparation:
Duque, M. †, **Amorim, J.P.** †, Galhardo, M. and José Bessa*. (2024) A Distal Enhancer
Regulates *ptf1a* Expression to Ensure Proper Lineage Acquisition and Maintenance of
Pancreatic Acinar Cells

†These authors contributed equally to this work

*Corresponding author.

Author contributions:

J.B. conceived and supervised the project. J.P.A., M.D. and J.B. conceived, designed,
and analyzed the data. J.P.A and M.G. performed computational analyses.

3.1. Introduction

The vertebrate pancreas is a central organ for human diseases such as pancreatic cancer and diabetes. Functionally, it can be divided into the endocrine and exocrine compartments. The endocrine pancreas secretes hormones into the circulation and is responsible for the regulation of glucose metabolism (Baetens et al, 1979; Argenton et al, 1999; Brissova et al, 2005). The exocrine pancreas is composed of acinar cells, which produce and secrete enzymes necessary for digestion and are arranged into clusters at the end of ducts, which in turn deliver those enzymes to the intestine (Slack, J.M.W, 1995). A group of multipotent progenitor cells (MPCs) drives pancreatic development, originating from the endoderm developing gut, under the control of signaling pathways that determine its correct position and size (Sakhneny et al, 2018; Amorim et al, 2020; Duque et al, 2022). MPCs express specific transcription factors (TFs) such as the well-studied *Ptf1a*, *Nkx6.1* and *Pdx1* and these are known to play a crucial role in the proper development of the pancreas in vertebrates (Burlison et al, 2008; Ghaye et al, 2015).

Ptf1a is a bHLH TF that forms the trimeric PTF1 complex with two other proteins: An E protein and *Rbpj*, or its pancreas-restricted paralogue *Rbpjl* (Beres et al, 2006). *Ptf1a* plays a crucial role in the development of the nervous system and pancreas (Kawaguchi et al, 2002; Sellick et al, 2004; Glasgow, 2005; Hoshino et al, 2005; Meredith et al, 2009), being expressed in progenitor cells of both developing tissues (Obata et al, 2001; Kawaguchi et al, 2002). *Ptf1a* is expressed in early pancreatic MPCs of the mouse developing pancreas, during the primary transition (Kawaguchi et al, 2002) and later at the secondary transition, where it is restricted to pro-acinar progenitors in the tip epithelium (Zhou et al, 2007). It is also expressed in differentiated acinar cells but not in ductal or endocrine tissue (Krapp et al, 1996).

In zebrafish, *ptf1a* plays an important role in MPCs of the developing pancreas and, later on, in the establishment of acinar cell fate (Lin et al, 2004; Zecchin et al, 2004). *ptf1a* expression starts at 32hpf in the ventral bud, after endocrine differentiation of the principal islet from the dorsal bud (Lin et al, 2004), in cells that are *pdx1*⁺ and *nkx6.1*⁺ (Dong et al, 2008). As pancreas development unfolds, ventral bud progenitor cells migrate in a dorsal and posterior direction enveloping the principal islet. Around 72hpf, the first differentiated acinar cells start to appear along the forming pancreatic tail (Biemar et al, 2001) and *ptf1a* expression becomes progressively restricted to these cells (Zecchin et al, 2004), suggesting that this TF-encoding gene is involved in the acquisition and maintenance of the acinar cell fate. Upon *ptf1a* loss-of-function in zebrafish, acinar cells and secondary islets were not detected, being only present the principal islet, while

ptf1a lineage-labeled cells were converted into gall bladder and other non-pancreatic cell types (Wang et al, 2015). Additionally, reduced levels of *ptf1a* present in a hypomorphic condition show a delay in ventral pancreas specification, accompanied by an exocrine to endocrine fate switch, suggesting that lower levels of Ptf1a can function, in a cell autonomous manner, to promote endocrine fate, whereas high levels repress it (Dong et al, 2008). The function of Ptf1a as one of the key effectors of the transcriptional program of acinar cells, working for the maintenance of the differentiated acinar cell identity in vertebrates has been reported by several authors (Holmstrom et al, 2011; Hoang et al, 2016; Jiang et al, 2016; Azevedo-Pouly et al, 2023). Interestingly, upon *Ptf1a* loss of function in mouse acinar cells, acinar-to-ductal metaplasia (ADM) and the induction of a KRAS-permissive, PDAC-like (pancreatic ductal adenocarcinoma) gene expression profile was observed (Krah et al, 2015). In acinar-to-ductal metaplasia (ADM), acinar cells are reprogrammed into duct-like cells. Together with oncogenic alterations, ADM can lead to pancreatic intraepithelial neoplasia (PanIN), which in turn can progress to PDAC. During ADM, acinar cells lose their characteristic shape and function and adopt a morphology that is more characteristic of duct cells (Marstrand-Daucé et al, 2023). This process is known to involve changes in the expression of genes controlling cell differentiation, proliferation, survival or inflammation. This shift from acinar cells to duct-like cells is also dependent on a downregulation of acinar gene expression (Jiang et al, 2023; Marstrand-Daucé et al, 2023). Within these are the acinar-specific TF-encoding genes *Ptf1a*, *Bhlha15* and *Nr5a2* and genes expressing digestive enzymes such as *Cpa1* (Shi et al, 2013; vonFigura et al, 2014; Krah et al, 2015; Jiang et al, 2016, 2023; Marstrand-Daucé et al, 2023). Interestingly, loss of *Nr5a2* has been shown to accelerate ADM (vonFigura et al, 2014) and this TF-encoding gene was observed to cooperate with *Ptf1a* in the maintenance of acinar cell identity (Holmstrom et al, 2011). Moreover, recent results showed that induction of *Ptf1a* expression in established PanINs or PDAC can revert pancreatic tumorigenesis (Jakubison et al, 2018; Krah et al, 2020). Taken together, these results suggest that Ptf1a might have a tumor suppressor function in the adult exocrine pancreas and this might be achieved through its canonic function of acinar cell identity maintenance, together with other acinar-specific TFs such as *Bhlha15* or *Nr5a2*, thus inhibiting reprogramming of these cells into aberrant identities that could lead to pancreatic cancer.

The tissue-specific functions of *Ptf1a* require precise spatiotemporal regulation of its expression levels, which is achieved by multiple noncoding CREs. In zebrafish, three regulatory sequences in the landscape of the *ptf1a* gene have been shown to exhibit enhancer activity in pancreatic cells, while others were found to be active in the

developing nervous system (Pashos et al, 2013; Bordeira-Carriço et al, 2022; Duque et al, 2022). One of these pancreatic elements was recently uncovered by members of our lab, located approximately 39kb downstream of the *ptf1a* TSS and exhibiting marks for open chromatin and enrichment for H3K27Ac (Bordeira-Carriço et al, 2022), a histone modification associated to enhancer activity (Rada-Iglesias et al, 2011; Marlétaz et al, 2012). This distal pancreatic enhancer (z3'-DpE) was observed to interact with the *ptf1a* promoter, by 4C-seq and was shown to be active in the zebrafish pancreas, in MPCs and differentiated acinar and duct cells (Bordeira-Carriço et al, 2022). Furthermore, CRISPR/Cas9-mediated deletion of a part of this sequence resulted in a significant downregulation of *ptf1a* expression in MPCs, suggesting that z3'-DpE regulates *ptf1a* expression directly (Bordeira-Carriço et al, 2022). A similar element is present in the human genome. Recessive mutations within a 400 bp non-coding sequence, 25 kb downstream of *PTF1A*, were identified in families with pancreatic agenesis. Moreover, chromatin conformation capture experiments in hESC-derived human MPCs revealed that the human *PTF1A* distal enhancer (h3'-DpE) also interacts directly with the *PTF1A* promoter (Weedon, Michael N. et al, 2014). In contrast with *PTF1A* coding mutations, which result in pancreatic and cerebellar agenesis, neurological features are absent in reported cases of h3'-DpE mutation (Weedon, Michael N. et al, 2014). Despite showing no sequence conservation with the human enhancer, z3'-DpE is likely functionally equivalent to h3'-DpE, with deletion of z3'-DpE resulting in pancreatic hypoplasia (Bordeira-Carriço et al, 2022). Importantly, h3'-DpE is active in human MPCs but inactive in adult exocrine pancreatic cell lines (Weedon, Michael N. et al, 2014; Duque et al, 2022). More recently, the conserved mouse distal *Ptf1a* enhancer was shown to have a similar early function in MPCs as a deletion within this sequence resulted in reduced levels of *Ptf1a* in MPCs but no function was observed in differentiated acinar cells (Miguel-Escalada et al, 2022). Interestingly, z3'-DpE shows an early function in zebrafish MPCs, equivalent to what happens in human and mouse, and a late function in the adult zebrafish exocrine pancreas (Bordeira-Carriço et al, 2022), suggesting a species-specific late function for the zebrafish sequence.

In this work, we have explored the pattern of the z3'-DpE active domain and the impact of deleting this CRE in the chromatin accessibility and transcriptomic profiles of zebrafish MPCs and differentiated acinar cells. By performing enhancer assays and confocal microscopy, we have extensively characterized the GFP expression of a reporter line under the control of z3'-DpE, throughout early pancreatic development, in wild-type (z3'-DpE^{+/+}) and mutant (z3'-DpE^{-/-}) zebrafish embryos. In both conditions, z3'-DpE reporter activity starts at 35hpf and is initially seen co-localizing with cells

expressing *nkx6.1*, which at this stage is known to label the MPC domain (Ghaye et al, 2015). Later, cells expressing GFP but not *nkx6.1* appear, most likely representing the first differentiated acinar cells. We found that, in embryos harboring a deletion of *z3'-DpE*, the MPC domain is shifted posteriorly at 48h and 60hpf, in relation to the islet center. Furthermore, we observed the appearance of cells expressing only GFP 12 hours earlier than in wild-types. These results could be suggestive of an earlier differentiation of MPCs into the various cellular lineages, due to a downregulation of *ptf1a*. We then performed RNA-sequencing of *z3'-DpE*^{+/+} and *-/-* MPCs and found downregulated genes to be enriched in Notch signaling pathway components, suggesting that *Ptf1a* might be responsible for activating an endogenous Notch signaling program that might be important for the differentiation of MPCs. To profile chromatin accessibility, we performed ATAC-sequencing of *z3'-DpE*^{+/+} and *-/-* differentiated acinar cells. We found regions with decreased chromatin accessibility in *z3'-DpE*^{-/-} and enriched TF motifs within these regions to be associated with genes and TFs related to acinar cell identity, endocrine function and developmental processes. On the other hand, regions with increased accessibility were found to be associated with genes related to developmental processes, inflammation, proliferation and epithelial-to-mesenchymal (EMT) transition. Furthermore, TF motifs enriched within regions with increased accessibility were associated to inflammatory processes (such as several members of the AP-1 complex) and the respective TF-encoding genes were enriched within DisGeNET lists of genes associated to pancreatic cancer (Piñero et al, 2015). Overall, these results highlight the importance of *ptf1a* for the correct lineage acquisition and proper maintenance of the acinar cell identity. Moreover, our results suggest a possible role of *ptf1a* as a tumor suppressor gene in acinar cells, possibly preventing aberrant gene expression and alterations in cell identity, highlighting *ptf1a* as a candidate for further study in the context of pancreatic cancer and possible preventive and therapeutic approaches to this severe disease.

3.2. Results and Discussion

3.2.1. z3'-DpE enhancer is active in early MPCs and persists in differentiated exocrine cells

Previous results from our laboratory have shown that a distal sequence in the regulatory landscape of *ptf1a* is able to drive expression of GFP in zebrafish MPCs and differentiated acinar cells (Bordeira-Carriço et al, 2022). Furthermore, CRISPR/Cas9-mediated deletion of this enhancer resulted in pancreatic hypoplasia in adult fish (Bordeira-Carriço et al, 2022). In the current work, we sought to further explore the early and late functions of the distal pancreatic enhancer controlling *ptf1a* expression. First, we focused on the early function of Ptf1a in zebrafish MPCs. We took advantage of a transgenic reporter line in which GFP expression is under the control of the z3'-DpE enhancer and crossed it with another reporter line, expressing mCherry under the control of the *somatostatin* promotor, allowing to locate the endocrine pancreas domain. As *ptf1a* expression is described to start at approximately 32hpf in the zebrafish ventral bud (Lin et al, 2004; Duque et al, 2022), we collected embryos from several time points between 33 and 96hpf and performed immunohistochemistry with an anti-Nkx6.1 antibody (Figure 3.1; Figure S3.1). By confocal microscopy analysis, we were able to observe GFP expression in MPCs, co-localizing with anti-Nkx6.1 staining, from 35hpf onwards (Figure 3.1A), suggesting that enhancer activity of z3'-DpE starts around the same time as the reported initial expression of *ptf1a*. Interestingly, from 35 to 72hpf, concomitant with an increase of the overall number of MPCs, we observed an increase of the proportion of cells that express both *nkx6.1* and GFP, in relation to the total number of MPCs, suggesting that *ptf1a* expression becomes more prevalent in these cells as pancreas development takes place and MPCs start to differentiate (Figure 3.1B). Between 60 and 96hpf, the proportion of cells showing co-expression of GFP and *nkx6.1* diminishes, coinciding with the appearance of cells expressing GFP only, at 72hpf. These cells most likely represent differentiated acinar cells of the exocrine pancreas, in alignment with previous data reporting the detection of the first differentiated cells expressing carboxipeptidase A1 (*cpa1*), at this developmental timing (Biemar et al, 2001). These data points to a segregation of the *nkx6.1* and *ptf1a* expression domains starting around 60hpf. In agreement with this, an antagonistic interaction between Nkx6.1 and Ptf1a at this stage of pancreas development has been previously reported as crucial for the control of lineage acquisition that will drive the establishment a fully differentiated organ (Schaffer et al, 2010; Hesselson et al, 2012) (Figure S3.1). Later, at 10dpf, GFP expression is observed in differentiated acinar cells of zebrafish larvae, shown by co-localization with mCherry expression, under the control of the promotor of

the acinar known marker *elastase* (Figure 3.1).

These results indicate that z3'-DpE controls *ptf1a* expression in the MPC domain and differentiated acinar compartment, suggesting that this regulatory sequence has a dual function in pancreatic cells, consisting of an early and a species-specific late function. We were able to observe z3'-DpE reporter activity from 35hpf onwards, around the same time as the reported initial expression of *ptf1a*. We observed that, between 60 and 96hpf, *nkx6.1*+ and z3'-DpE-active domains progressively separate, suggesting an antagonistic interaction between Nkx6.1 and Pt1a. This interaction has been previously reported to have an important role in the differentiation of MPCs into the several pancreatic compartments occurring at this stage of development (Biemar et al, 2001; Schaffer et al, 2010; Hesselson et al, 2012).

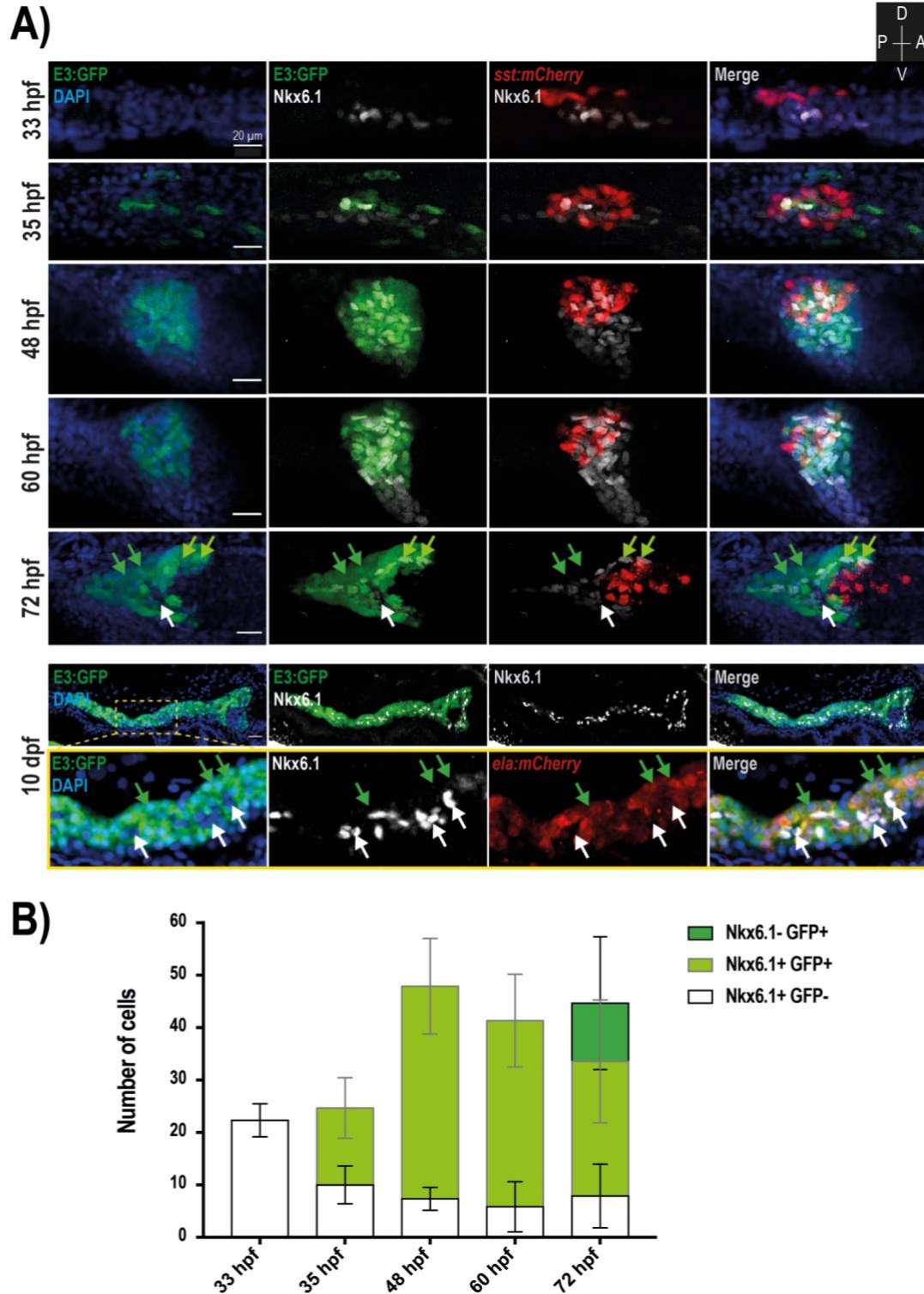


Figure 3.1 – A) GFP expression pattern of the z3'-DpE:GFP transgenic reporter line at 33, 35, 48, 60, 72hpf and 10dpf. From 35hpf onwards, GFP expression co-localizes with anti-nkx6.1 immunostaining in MPCs of the developing pancreas. Between 48and 60hpf, MPCs surround the principal islet, here labeled by the *somatostatin*:mCherry transgenic reporter line. At 72hpf, the first cells expressing GFP but not *nkx6.1* appear. Later, at 10dpf, GFP expression co-localizes with differentiated acinar cells, labeled by the *elastase*:mCherry transgenic reporter. White arrows point to *nkx6.1*+ GFP- cells, light green arrows to *nkx6.1*+GFP+ cells and dark green arrows to *nkx6.1*-GFP+ cells. B) Quantification of the number of cells expressing both *nkx6.1* and GFP or either one in the various developmental timings.

3.2.2. Deletion of the z3'-DpE enhancer results in an earlier separation of the z3'-DpE-active and *nkx6.1*-expressing domains and a posterior-oriented elongation of the *nkx6.1* expression domain

Next, to study the impact of *ptf1a* downregulation in the developing pancreas, we took advantage of the mutant zebrafish line previously established in our laboratory, harboring a partial deletion of the z3'-DpE enhancer (Bordeira-Carriço et al, 2022). We crossed these mutants with the z3'-DpE:GFP transgenic reporter line and were able to follow GFP expression in mutant (z3'-DpE^{-/-}) embryos throughout pancreatic development. Again, we collected embryos from several time points between 33 to 96hpf and performed immunostaining using an anti-Nkx6.1 antibody to label MPCs and anti-insulin to label beta cells of the endocrine pancreas (Figure S3.2). We observed z3'-DpE reporter activity starting at 35hpf, similarly to wild-types. By documenting the number of cells expressing GFP and *nkx6.1*, we observed that, at 48hpf, the total number of MPCs (expressing *nkx6.1*) is significantly reduced in mutants, comparing to wild-type embryos (Figure S3.2B). Moreover, unlike wild-type embryos, mutants showed an increase of cells co-expressing GFP and *nkx6.1* between 60 and 96hpf, in detriment of cells expressing *nkx6.1* only, which were significantly reduced in number (Figure 3.2B; Figure S3.2B). Intriguingly, we observed the appearance of cells expressing GFP only at 60hpf, 12 hours earlier when compared to wild-type embryos, which could indicate an earlier segregation of the *nkx6.1* and *ptf1a* expression domains (Figure 3.2A and B; Figure S3.2). To further explore the hypothesis that the MPC domain is being affected upon the deletion of z3'-DpE, we compared the localization of the *nkx6.1* expression domain between wild-type and mutants at 39, 48 and 60hpf by plotting the *nkx6.1* expression domain along the antero-posterior axis, using the center of the principal islet, labeled by either somatostatin or insulin, as a reference. We were able to observe that, in mutant embryos at 48 and 60hpf, the majority of *nkx6.1*⁺ cells locate in the posterior side, in relation to the islet center, while in wild-types it mostly locates anteriorly (Figure 3.2C). Furthermore, we observed a significant increase in the length of the *nkx6.1* expression domain in mutant embryos, both at 48 and 60hpf (Figure 3.2D), suggestive of an elongated MPC domain, compared to wild-types. Our results suggest that upon deletion of the z3'-DpE enhancer, the MPC domain is affected after 48hpf, showing an altered number of cells expressing *nkx6.1* and a posteriorly-oriented shift of the *nkx6.1* expression domain, in relation to the islet center. These findings, together with the fact the cells expressing only GFP appear earlier in development, could indicate that the process of lineage determination into the various pancreatic cell types occurring at this stage of development is perturbed.

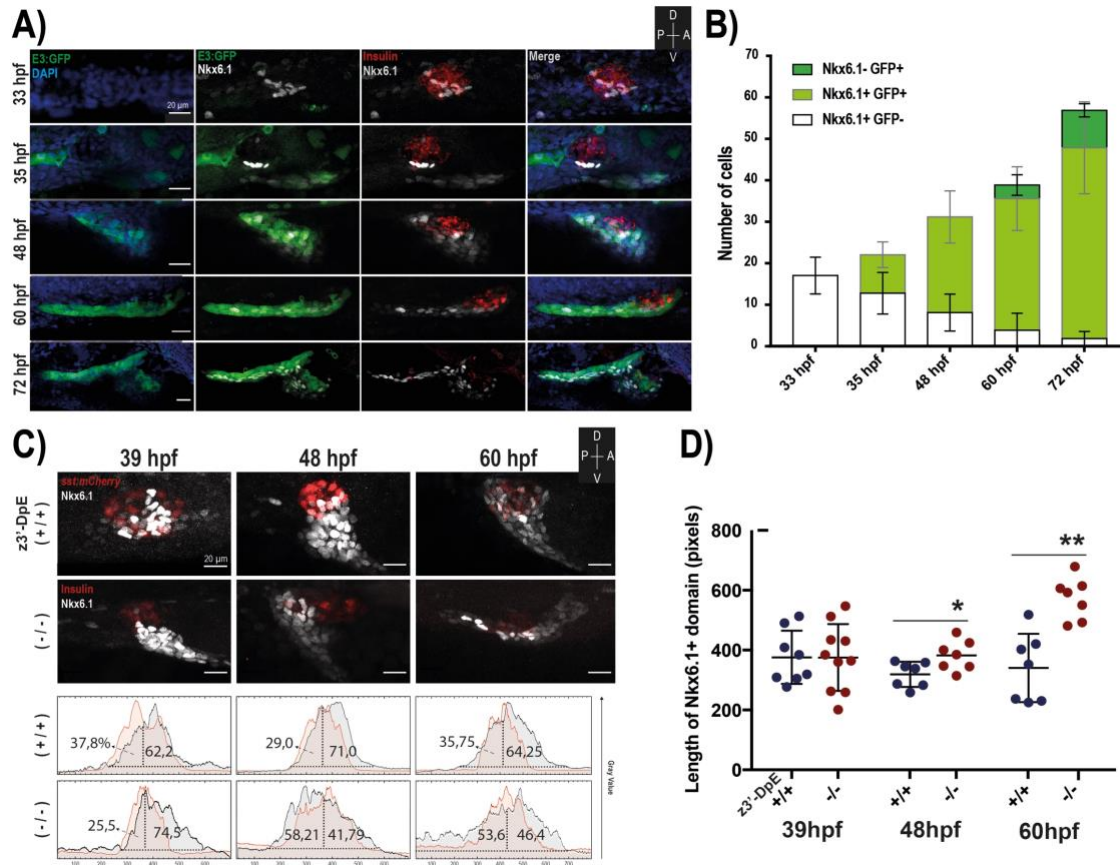


Figure 3.2 - A) GFP expression pattern of the z3'-DpE:GFP transgenic reporter line in embryos harboring a deletion of the z3'-DpE enhancer (z3'-DpE -/-) at 33, 35, 48, 60 and 72hpf. From 35hpf onwards, GFP expression co-localizes with anti-Nkx6.1 immunostaining in MPCs of the developing pancreas. At 48hpf, MPCs surround the principal islet, here labeled by immunostaining with an anti-insulin antibody. At 60hpf, 12 hours earlier than in the wild-type embryos, the first cells expressing GFP but not *nkx6.1* appear. Antero-posterior orientation is shown in the upper right corner. B) Quantification of the number of cells expressing *nkx6.1* and GFP or either one in the various developmental timings. C) Comparison of the localization of the *nkx6.1* expression domain between wild-type and mutants at 39, 48 and 60hpf, in relation to the principal islet, which is located at the center of the image and is represented either by *somatostatin*:mCherry (+/+) or anti-insulin immunostaining (-/-). The majority of anti-Nkx6.1 signal at 48 (58,21%) and 60hpf (53,6%) is located to the posterior side of the islet center, while in wild-types it mostly locates anteriorly (29 and 35,75%, respectively). Antero-posterior orientation is shown in the upper right corner. D) Quantification of the length of the *nkx6.1* expression domain, showing significantly higher pixel length in mutant embryos, at 48 and 60hpf.

3.2.3. Analysis of the transcriptome of MPCs upon deletion of z3'-DpE reveals a *ptf1a*-dependent developmental gene network

In order to access the alterations occurring in the transcriptome of MPCs upon deletion of the z3'-DpE enhancer, we used the z3'-DpE:GFP transgenic reporter line to perform FACS-sorting and collect MPCs of wild-type (z3'-DpE +/+) and mutant (z3'-DpE

-/-) embryos at 48hpf. We used these cells to extract RNA and perform RNA-sequencing of wild-type and mutant MPCs. Using also FACS-sorted wild-type MPCs, we were also able to perform ATAC-sequencing, allowing us to study the accessibility profile of MPCs (Figure 3.3; FigureS3.3; Figure S3.7).

Raw reads from three biological replicates of RNA-seq were first aligned to the zebrafish reference genome (danRer10) using Bowtie2 (Langmead & Salzberg, 2012). Then, differential expression analysis was performed using DEseq2 (Love et al, 2014), revealing 214 downregulated and 90 upregulated genes in mutants, comparing to wild-type MPCs. Lists of differentially expressed genes were tested for enrichment of Gene Ontology (GO) Terms using PANTHER Gene Ontology Enrichment tool (Biological Process) (Thomas et al, 2022) (Figure3.3). While upregulated genes did not show any significant enrichment, the set of downregulated genes in *z3'-DpE -/-* MPCs showed an enrichment of genes involved in biological processes related to proliferation (GO terms “spindle localization” and “cell population proliferation”), “notch signaling pathway”, morphogenesis (“embryonic morphogenesis”, “tissue morphogenesis” and “anatomical structure morphogenesis”) and developmental processes (such as “stem cell differentiation”, “chordate embryonic development” or “embryo/tissue development”). We also observed a strong enrichment of genes associated to the term “positive regulation of the circadian rhythm” (Figure 3.3). These results indicate that the transcriptome of mutant MPCs is significantly altered, with several downregulated genes being involved in developmental processes (Figure3.3; FigureS3.3), suggesting that expression of genes within these GO terms are, directly or indirectly, *Ptf1a*-dependent. While further experiments are needed to demonstrate this, the fact that terms related with patterning and differentiation processes are enriched in our analysis indicates, together with our previous results from enhancer activity assays (Figure 3.2), that patterning/morphogenesis defects might be occurring in mutant MPCs and these could result in altered differentiation into the various pancreatic lineages.

Interestingly, Notch signaling pathway-related genes were observed to be highly enriched in in our differential expression analysis of MPCs transcriptomic data, indicating that the activity of Notch signaling pathway in MPCs is, at least partially, dependent of *Ptf1a*. This is not unexpected, as *Ptf1a* function in development is known to be intrinsically connected with notch signaling (Apelqvist et al, 1999; Fujikura et al, 2006; Masui et al, 2007; Ghaye et al, 2015). Importantly, we also found Notch signaling pathway-related genes within other gene ontology terms such as “stem cell differentiation”, “embryonic morphogenesis” or “embryo/tissue development” (Figure S3.3). These results are in agreement with recent work, where the authors observed a

downregulation of Notch signaling pathway in mouse mutant embryos for the mouse distal *Ptf1a* enhancer (Miguel-Escalada et al, 2022), emphasizing the transversal function of Ptf1a and Notch signaling pathway in vertebrate MPCs and illustrating the equivalent early function of the zebrafish and mouse distal *Ptf1a* enhancers. Active Notch signaling has been shown to function to maintain MPCs in a progenitor state, preventing further differentiation into endocrine progenitor cells (EPCs) (Ghaye et al, 2015). The fact that this pathway is downregulated in MPCs of z3'-DpE *-/-* embryos suggests that Ptf1a function in these cells could include activating a Notch signaling program and favors the hypothesis that lineage acquisition after differentiation of MPCs might be perturbed in the context of *ptf1a* loss of function. This is in line with our observations from enhancer activity assays (Figure 3.2; Figure S3.2), where we saw an earlier appearance of cells exhibiting z3'-DpE activity but not Nkx6.1 expression in mutant background, suggestive of a precocious differentiation of MPCs. Nevertheless, additional experiments would be useful to fully unravel the impact of disturbing the interaction between Ptf1a and Notch signaling in zebrafish MPCs, per example by recurring to known notch inhibitors and observe if the phenotypes arising recapitulate our results from *ptf1a* downregulation.

Taken together, our findings suggest that the transcriptome of MPCs harboring a deletion of the distal *ptf1a* enhancer is significantly altered, highlighting the importance of this regulatory sequence for the proper function of Ptf1a in these cells. Together with our results of enhancer activity assays (Figure 3.2; Figure S3.2), our findings indicate that Ptf1a function in MPCs might include controlling a developmental gene network important for the proper differentiation of progenitors and lineage acquisition of the various differentiated pancreatic lineages. Furthermore, the enrichment of Notch signaling-related genes in the set of downregulated genes in MPCs of z3'-DpE *-/-* embryos suggests that Ptf1a activates an endogenous Notch signaling program in MPCs and stresses the importance of this signaling pathway in pancreas development and differentiation.

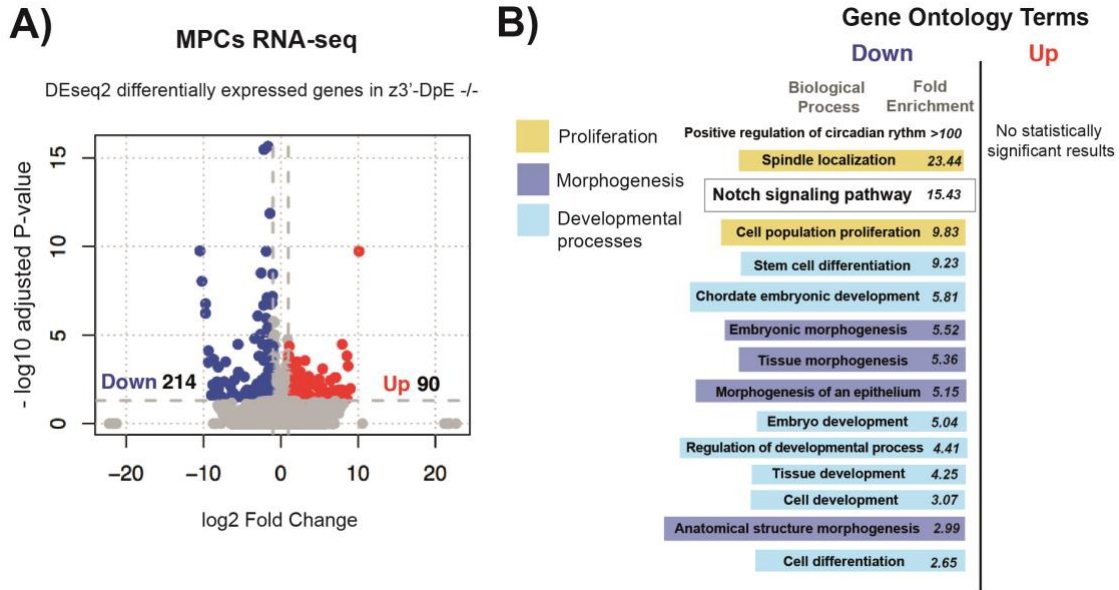


Figure 3.3 – A) Volcano plot of differentially expressed genes, assessed by DESeq2 analysis. Adjusted p-values are shown in the y axis and values of log2 fold change in the x axis. Red dashed lines represent p-value cutoff = 0.05 and log2 fold change cutoff =1. 214 genes are downregulated (blue dots) and 90 genes are upregulated in z3'-DpE -/- MPCs (red dots). B) Gene ontology analysis results using PANTHER Gene Ontology Enrichment tool. In the set of 214 downregulated genes, enrichment was observed in GO terms related to proliferation, notch signaling pathway, morphogenesis and developmental processes. The yellow box represents terms related to proliferation. Dark blue boxes represent terms related to morphogenesis. Light blue boxes represent terms related to developmental processes.

3.2.4. Differentially accessible regions in acinar cells of z3'-DpE -/- larvae indicate compromised acinar identity, concomitant with an upregulation of inflammation and pancreatic cancer associated genes

Next, we focused on the late function of the z3'-DpE enhancer in differentiated acinar cells. To study the impact of the z3'-DpE deletion in the accessibility profiles of acinar cells, we dissected whole pancreases from wild-type (z3'-DpE+/+) and mutant (z3'-DpE-/-) adult zebrafish harboring a transgene that labels in vivo acinar cells Tg(ela:mCherry). We dissociated the respective pancreata and performed FACS-sorting of acinar cells, followed by ATAC-sequencing. For each condition two replicates of ATAC libraries were sequenced. Resulting raw reads were trimmed using Skewer (Jiang et al, 2014) and aligned to the zebrafish reference genome (danRer10), using Bowtie2 (Langmead & Salzberg, 2012). After peak calling for each replicate, done with MACS2 (Zhang et al, 2008), the Irreproducible Discovery Rate (Li et al, 2011) was applied, in order to obtain a reproducible set of peaks, between replicates. Further analysis was performed for peaks with an IDR score reflecting a p-value <=0.1. Taking these parameters in consideration, 35 887 and 30 669 accessible regions were found in z3'-

DpE^{+/+} and z3'-DpE^{-/-} acinar cells, respectively. From these, 20176 were found to be exclusive to ^{+/+}, 14842 exclusive to ^{-/-} and 15711 were common between the two conditions (FigureS3.4A). Interestingly, by intersecting these regions with annotated promoter regions, we found that while 72,2% of common regions are located within promoters, the vast majority of exclusive peaks are found within non-promoter regions, most likely in CREs such as enhancer, repressors or insulators (FigureS3.4A). These results indicate that *ptf1a* is controlling a regulatory network in acinar cells and this is achieved mostly through non-promoter regions, likely CREs.

Peaks presenting an IDR score reflecting a p -value ≤ 0.1 were then analyzed for differential accessibility recurring to EdgeR software (Robinson et al, 2010) with differences considered statistically significant when adjusted p value ≤ 0.05 and $\text{abs}(\log_2\text{FC}) \geq 1$. This analysis resulted in the finding of 5968 regions with decreased accessibility (Down) and 4475 regions with increased accessibility (Up) in mutant acinar cells, compared to wild-type (Figure 3.4A; Figure S3.4B). Using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al, 2010), differentially accessible regions were associated to nearby genes and the resulting lists of genes were fed to PANTHER GO Enrichment Analysis (Thomas et al, 2022) to compare enrichment of genes within gene ontology terms between samples. For the set of genes associated with regions with decreased accessibility in z3'-DpE^{-/-}, we saw an enrichment of associated genes within GO terms related to acinar function (such as “digestive system process”, digestion, “digestive tract morphogenesis” or “digestive tract development”), endocrine function (such as “insulin secretion”, “hormone secretion” or “response to glucose”) and developmental processes (such as “regulation of endodermal cell fate specification”, “determination of pancreatic left/right symmetry” or “pancreas development”) and (Figure 3.4B). These results suggest that, upon downregulation of *ptf1a* in acinar cells, acinar cell identity becomes compromised, with regions exhibiting decreased accessibility enriched nearby genes associated to acinar function (examples shown in Figure 3.4D and Figure S3.5). This is in agreement with previous work showing that Ptf1a is crucial for the establishment and maintenance of the acinar cell identity and that this identity is lost upon downregulation of *ptf1a* (Krah et al, 2015; Hoang et al, 2016; Azevedo-Pouly et al, 2023). However, we did not find a significant difference between the number of regions with decreased and increased accessibility, within gene landscapes that were associated to acinar function in our GO analysis (Figure3.4C), which could mean that a higher number of specific key genes are being affected instead of a general tendency for loss of accessibility in these landscapes. We also observed genes associated to regions with decreased accessibility to be enriched in GO terms

related to developmental processes, which we hypothesize could be related, in part, to differentiation processes that help maintain cell identity by preventing aberrant differentiation. For GO terms within the group of developmental processes, we found that the mean number of regions with decreased accessibility was significantly higher in the landscapes of associated genes, comparing to regions with increased accessibility (Figure 3.4C). Intriguingly, we also found regions with decreased accessibility to be associated with genes involved in endocrine-related processes such as “insulin secretion” or “response to glucose” (Figure 3.4B). While the reason for this is not completely evident, we could hypothesize that *Ptf1a* is controlling a regulatory endocrine program in differentiated acinar cells, possibly representing a promiscuous cellular state with a predisposition to acinar-to-endocrine fate switch.

For the set of regions with increased accessibility, we found associated genes to be enriched within GO terms related to inflammatory processes (such as “regulation of leukocyte differentiation”, “regulation of T cell differentiation” or “macrophage activation involved in immune response”), proliferation (such as “spindle localization” and “endothelial cell proliferation”) or “regulation of epithelial-to-mesenchymal transition (EMT)”. These results suggest that there is an inflammatory response occurring in acinar cells upon deletion of *z3'-DpE*, possibly resulting from a loss of acinar cell identity and acquisition of a pro-inflammatory cellular state, in a *Ptf1a*-dependent manner. Accordingly, previous works showed that loss of *Ptf1a* in acinar cells is sufficient to induce ADM and potentiate inflammation (Krah et al, 2015; Hoang et al, 2016; Azevedo-Pouly et al, 2023). Interestingly, Notch signaling pathway components were found to be similarly enriched in both sets, showing that this pathway is perturbed upon *ptf1a* loss of function.

These results show that a *Ptf1a*-dependent regulatory program is active in acinar cells, responsible for the maintenance of acinar identity and, possibly, allowing a permissive state for acinar-to-endocrine fate switch. Upon downregulation of *ptf1a*, acinar cells present significantly altered accessibility profiles, highlighting the importance of this TF for the homeostasis of the adult exocrine pancreas, which in turn might help prevent the arising of aberrant cellular identities that could lead to pancreatic dysfunction. Importantly, regions with increased chromatin accessibility in mutant acinar cells were associated with genes involved in processes such as inflammation, proliferation and EMT transition, which are known to be hallmarks of oncogenic transformation (Greten & Grivnickov, 2020; Jiang, J et al, 2023; Marstrand-Daucé et al, 2023), suggestive of a potential role of *Ptf1a* as a tumor suppressor in the context of pancreatic cancer.

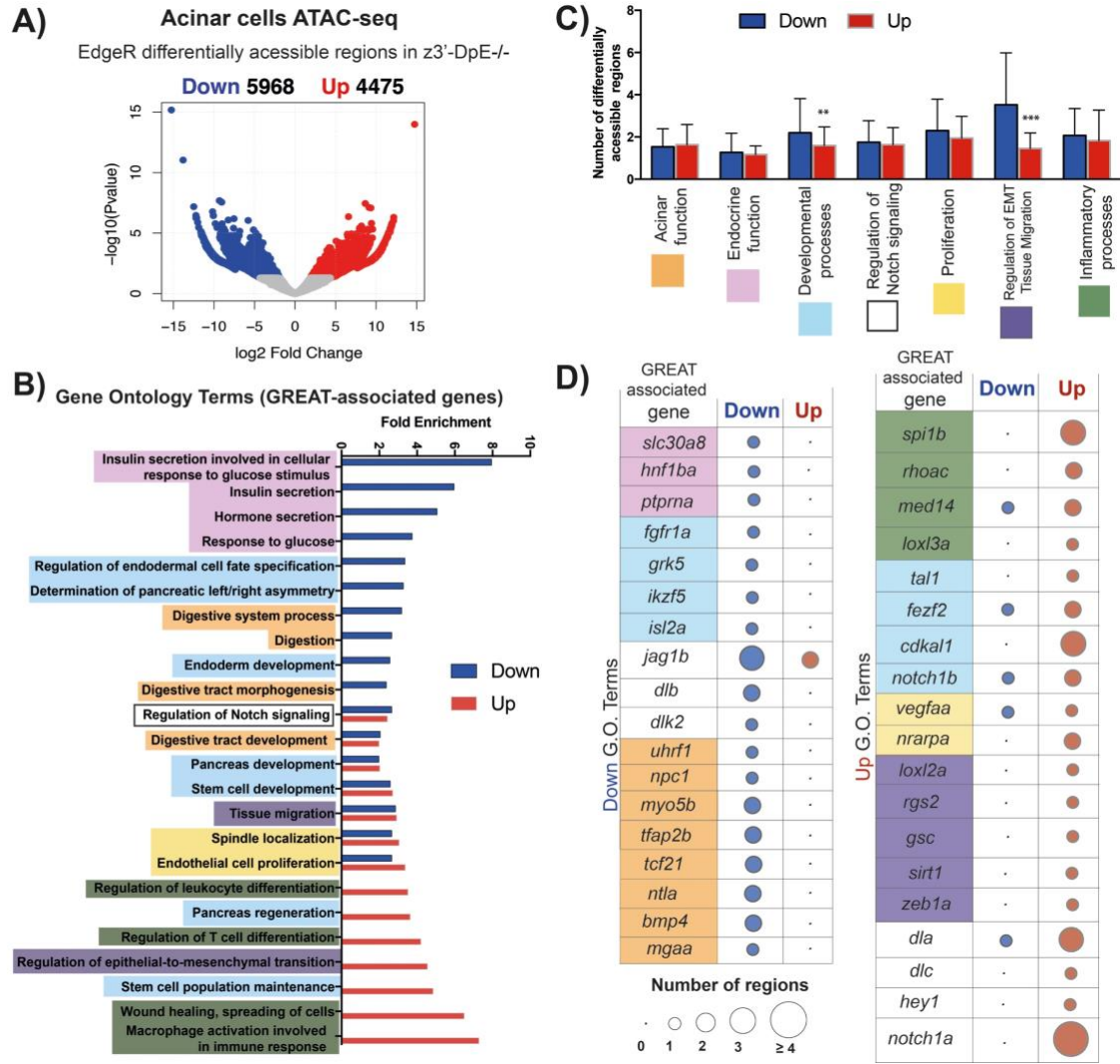


Figure 3.4 – A) VolcanoPlot showing differentially accessible regions in z3'-DpE-/-, comparing to +/+. The y axis represents values of $-\log_{10}(\text{p-value})$ and x axis represents values of $\log_2$ fold change. Downregulated regions are depicted as blue dots and upregulated regions as red dots. B) Selected gene ontology terms significantly enriched in both samples and comparison between them. Fold enrichment is shown in the x axis. C) Mean number of altered regions, as defined by EdgeR analysis, in landscapes of GREAT-associated genes within highlighted GO terms. Error bars represent SD. ***p < 0.001. D) Examples of genes associated to differentially accessible regions within selected GO terms. The number of associated regions is depicted by the size of blue/red circles (see in-figure legend).

Next, we looked at transcription factor motif enrichment in differentially accessible regions in z3-DpE -/-, by performing HOMER analysis (Heinz et al, 2011). To explore and visualize the differences in transcription factor (TF) binding motif predictions between regions with decreased and increased chromatin accessibility, we calculated the difference of percentage of motif enrichment between the two samples (Figure 3.5A). Then, to define which TF motifs are differentially enriched in regions with either decreased or increased accessibility, we selected two sets of TF motifs (Down TFs and

Up TFs) by setting a threshold in delta motif enrichment ($<-0.5\%$ and $>0.5\%$, respectively). Next, we grouped TFs by their described function, according to the literature (Figure 3.5; Table S5) and identified TF-encoding genes that were associated to pancreatic cancer, according to DisGeNET lists of genes associated to pancreatic cancer (Figure 3.5A, B and F; Table S6 and S7). We found acinar-related TFs such as Ptf1a, Bhlha15 and Nr5a2 to be more represented within the group of TF motifs exhibiting higher enrichment in regions with decreased accessibility (Down TFs), comparing to motifs enriched in regions with increased accessibility (Up TFs). We also found Ptf1a co-factors (E2A, Tcf12, TCF4) and endocrine-related factors (Ascl2, Ascl1 and NeuroD1) to be more represented within the Down TFs group. Together with results from GO analysis (Figure 3.4), these results suggest that acinar cell identity is being compromised in z3-DpE^{-/-} exocrine pancreas. The fact that we found enrichment of endocrine-related TF motifs within regions with decreased chromatin accessibility favors the intriguing hypothesis that Ptf1a is controlling an endocrine-related regulatory network in acinar cells, which could mean that acinar cells are predisposed to an acinar-to-endocrine fate switch.

We then looked into the landscapes of differentially enriched TF-encoding genes to understand if CREs within these landscapes showed differential accessibility, as a possible explanation for their enrichment observed in Figure 3.5A. For that, we looked at the number of regions with decreased or increased accessibility that were associated to genes encoding TFs within previously described functional groups of TFs (Figure 3.5C; Table S5). We found no differences between the number of regions with decreased or increased accessibility within landscapes of TF-encoding genes, in any of the groups (Figure 3.5C). Nevertheless, we observed that, for the landscapes of specific genes, regions with increased accessibility correlated with the respective differentially enriched TF motifs (examples shown in Figure 3.5D). The fact that we cannot consistently find regions with decreased accessibility within landscapes of genes encoding TFs from the Down TFs group suggests that these landscapes might not be overall regulated by Ptf1a and therefore, the alterations in the prevalence of TF motifs observed in Figure 3.5A most likely does not derive from alterations in their accessibility profiles. However, we hypothesize that TFs enriched in regions with decreased accessibility (Down TFs) could represent binding partners of Ptf1a and their activity to perform specific functions in differentiated acinar cells might be Ptf1a-dependent. To explore the hypothesis of co-binding, we looked at regions showing decreased accessibility that were predicted to be bound by specific TFs within the Down TFs group and intersected these with regions with decreased accessibility containing Ptf1a motifs. Then, we performed a chi square

to test if predicted co-bound regions were significantly higher than expected for each case (Figure 3.5E). We found that, in the case of acinar-related TFs such as *Bhlha15* and *Nr5a2*, and endocrine-related TFs such as *Ascl1* and *NeuroD1*, a significant number of regions are predicted to be co-bound with *Ptf1a*, although this does not happen for endocrine/progenitor markers such as *Nkx6.1* or *Hnf1b*. These results suggest that, in zebrafish acinar cells, *Ptf1a* regulates exocrine-related genes important for the establishment and maintenance of acinar cell fate and that this function might be achieved through the co-binding with partner TFs such as *Bhlha15* or *Nr5a2*, in agreement with previous work (Holmstrom et al, 2011; Jiang et al, 2016). With the downregulation of *ptf1a*, these co-bound regions become inaccessible, suggesting that *Ptf1a* is required for the maintenance of the accessible state of these chromatin regions. As previously discussed in light of the results from Figure 3.4, an intriguing hypothesis also emerges from our results, where *Ptf1a* might control an endocrine-related regulatory program by co-regulating regions involved in endocrine function, together with TFs such as *Ascl1* and *NeuroD1*. The fact that these regions exhibit decreased accessibility upon *z3-DpE* deletion suggests that *Ptf1a* could work to maintain these regions accessible in the genome of acinar cells and this, in turn, could represent a capacity for trans-differentiation from acinar to endocrine cells, although additional experiments should be performed to confirm this.

Interestingly, we found that, for the set of TF motifs enriched in regions with increased accessibility (Figure 3.5A), the respective TF-encoding genes were significantly enriched within lists of pancreatic cancer-associated genes (DisGeNET lists for Pancreatic neoplasm and Malignant neoplasm of pancreas; Table S6 and S7) (Piñero et al, 2015) (Figure 3.5F). These results emphasize that alterations in chromatin accessibility in the context of *ptf1a* downregulation might favor oncogenic transformation in acinar cells, which in turn highlights the possible role of *ptf1a* as a tumor suppressor gene in these cells. This is in agreement with previous work (Krah et al, 2015), where the authors observed acinar-to-ductal metaplasia (ADM) and the induction of a KRAS-permissive, PDAC-like gene expression profile upon loss of function of *Ptf1a* in mouse acinar cells. Interestingly, we also observed predicted motif enrichment for several AP-1 complex associated TFs within this group. AP-1 complex-related genes have been previously associated with a pro-inflammatory state found in *Nr5a2*-deficient acinar cells of the mouse pancreas (Cobo et al, 2018). In this study, the authors saw that, upon its downregulation, *Nr5a2* undergoes a transcriptional switch from differentiation-specific to inflammation-related genes, promoting AP-1 dependent gene transcription. Another work showed that *Nr5a2* is required for the maintenance of the acinar cell identity and

its loss accelerated ADM in the mouse pancreas (vonFigura et al, 2014). Importantly, *Ptf1a* and *Nr5a2* have been shown to co-regulate exocrine pancreas-specific genes in mice (Holmstrom et al, 2011). In line with this, we show that *Nr5a2* motifs are enriched within regions with decreased accessibility and, on the other hand, motifs for AP-1 complex components are enriched within regions with increased accessibility. These results suggest that *Ptf1a* functions to maintain acinar cell identity and, upon its downregulation, zebrafish acinar cells switch to a pro-inflammatory and pro-oncogenic state, emphasizing the importance of maintaining endogenous cellular identity in differentiated acinar cells to prevent aberrant gene expression that could lead to pancreatic dysfunction and oncogenic transformation. In agreement with this, previous work has shown that *Ptf1a* is crucial for the maintenance of acinar cell identity and that loss of such identity can lead to oncogenic alterations such as ADM (Krah et al, 2015; Hoang et al, 2016; Azevedo-Pouly et al, 2023). Interestingly, previous work also showed that induction of *Ptf1a* expression in established PanINs or PDAC lead to a reversal of pancreatic tumorigenesis (Jakubison et al, 2018; Krah et al, 2020).

As a final approach to our data and taking advantage of our results for ATAC-seq of wild-type MPCs, we decided to intersect regions with an IDR-score reflecting a p-value <0.1 (19 856 regions in MPCs) with the ones from *z3'-DpE+/+* (35 887) and *z3'-DpE/-* (30 669) acinar cells, in order to evaluate the overlap of accessible regions between cell types (Figure S3.7). We observed a similar percentage of common regions between MPCs and wild-type acinar cells (31,64% of 35 887 regions), compared to mutants (32.03% of 30 669 regions). These results show that the number of accessible regions is higher in differentiated cells of the exocrine pancreas, compared to MPCs, suggesting that a number of chromatin regions becomes accessible during differentiation of MPCs into acinar cells and this process seems to be independent of *ptf1a*. Furthermore, while the majority of accessible regions in MPCs are kept in differentiated acinar cells (64,83% of 19 856 regions), the number of regions exclusive to acinar cells is higher in the wild-type condition, suggesting that, within regions shared with MPCs, there are some whose activity is dependent on *ptf1a* expression.

Altogether, our results highlight *Ptf1a* as a crucial TF for the establishment and maintenance of the acinar cell identity in zebrafish, being this function achieved through the regulation of chromatin accessibility within important regulator and effector genes or through co-binding with specific TFs important for either acinar function such as *Bhlha15* and *Nr5a2*. Also, we show that upon deletion of *z3'-DpE*, regions exhibiting increased chromatin accessibility are associated with genes involved in inflammatory processes

and pancreatic cancer development, emphasizing the plausible role of *ptf1a* as a tumor suppressor gene in acinar cells, possibly by preventing aberrant cellular identities.

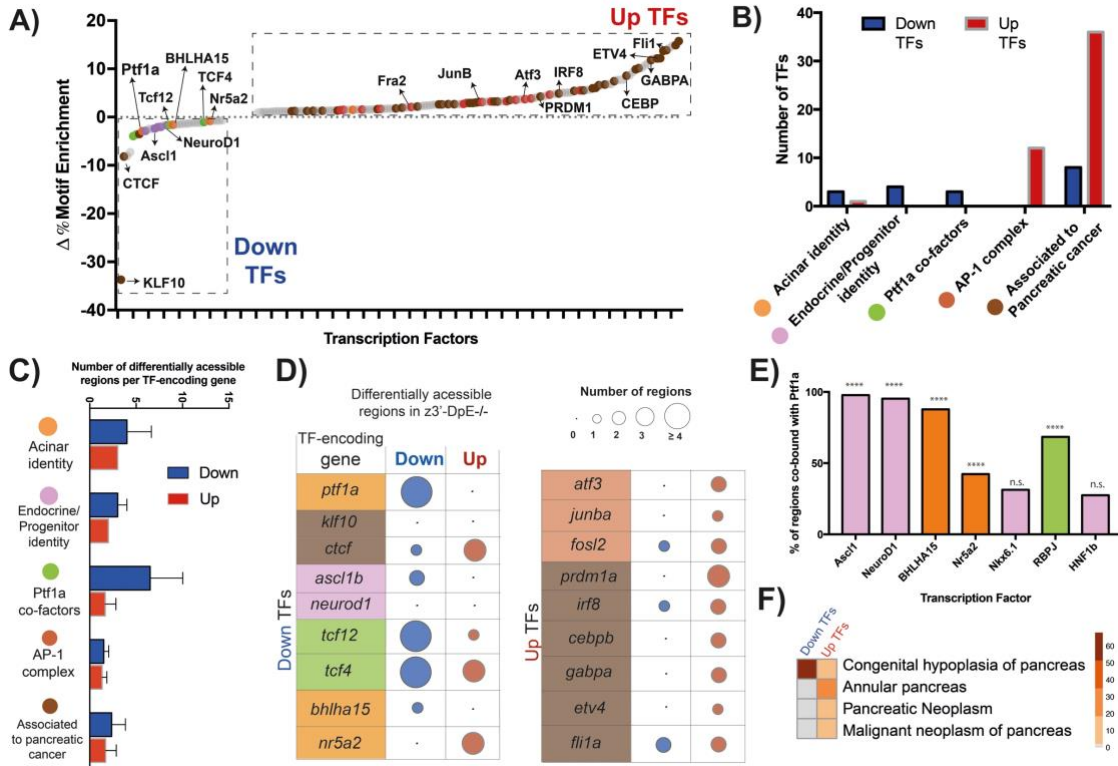


Figure 3.5 – A) Delta motif enrichment (%) of HOMER analysis results between *z3'-DpE+/+* and *-/-* ATAC-seq data. Specific TF motifs are shown. Colored circles represent groups of TFs known to perform similar/related functions, manually curated from the literature (see also Table S5). Orange circles represent acinar-related TFs, green circles represent known *ptf1a* co-factors, pink circles represent endocrine/progenitor-related TFs, red circles represent members of the AP-1 complex and brown circles represent TF-encoding genes that were enriched in DisGeNET lists of genes associated to pancreatic disease (see also F)). B) Number of TF motifs within defined functional groups that were found to be enriched in regions with decreased or increased accessibility. C) Number of differentially accessible regions in landscapes of TF-encoding genes within functional groups of TFs. Error bars represent SD. D) Examples of specific TF-encoding genes, depicting the number of differentially accessible regions within landscapes. E) Percentage of regions predicted by HOMER to be co-bound by specific TFs and Ptf1a. Percentage values are relative to the total number of peaks containing specific TF motifs. F) Heatmap showing fold enrichment of genes encoding differentially enriched TFs, within DisGeNET lists of genes associated to pancreatic diseases.

3.3. Methods

3.3.1. Experimental Model and Subject Details

Zebrafish

Zebrafish (*Danio rerio*) fish were handled according to European regulations and standard protocols. The Animal Facility of the i3S is licensed by the Portuguese official veterinary department (DGAV) and accredited by the AAALAC (June, 2018), complies with Portuguese law (Portaria 1005/02 and Portaria 1131/97), the 2010/63/EU Directive, and follows the FELASA (Federation of European Laboratory Animal Science Associations) guidelines and recommendations concerning laboratory animal welfare. The current project is licensed by DGAV (reference: 0421/000/000/2016). Adult animals are kept at a controlled temperature (25°C) and under a 14h light-8h dark photoperiod. Recirculated water is continuously monitored regarding pH (7), salinity (900 µs) and temperature (28°C) values. In order to obtain zebrafish embryos, adult zebrafish were bred in ratios of 3 females to 2 males or 1 male to 1 female, depending on the purpose of the experiment. Embryos or larvae ranging from 33hpf to 10dpf were used for immunohistochemistry, RNA and ATAC-seq. For ATAC-seq of acinar cells, whole pancreases were extracted from adult fish of 8 to 16 months old.

3.3.2. Method Details

Immunohistochemistry

Whole-mount immunohistochemistry was performed using mouse anti-rat nkx6.1 (Developmental Studies Hybridoma Bank) and rabbit anti-insulin (abcam) antisera. Goat anti-rabbit IgG AlexaFluor 568 (Invitrogen) and goat anti-mouse IgG AlexaFluor 647 (Invitrogen) were used as secondary antisera. Briefly, embryos were fixed overnight in 4% paraformaldehyde (PFA; (Sigma-Aldrich)) in 1x phosphate-buffered saline (PBS) [137 mM NaCl (Fisher Chemical), 2.7 mM KCl (VWR), 10 mM NaHPO₄ (Merk), and 1.8 mM KH₂PO₄ (Merk)]. Fixed embryos were sequentially washed with PBS-Triton 0.1% (PBS-T), then treated for 2h with PBS-T 1%, washed again with PBS-T 0.1% and incubated for 1h with PBS-T+BSA 5% (bovine serum albumin; NZYTech) before the incubation with the primary antisera, diluted 1:50 in PBS-T+BSA, overnight at 4°C. After extensive wash, embryos were incubated with the secondary antisera, overnight at 4°C, washed extensively and mounted in 50% glycerol in PBS. Embryos were stained with DAPI and 1,5 µm Z-series stacks were acquired on a Leica SP5 confocal microscope, using a 40x objective. Post-imaging analysis and quantifications were done in Fiji-ImageJ.

Dissociation and Fluorescence-activated Cell Sorting (FACS)

Whole pancreases were dissected from transgenic adult zebrafish Tg(ela:mCherry) with 8 to 16 months old and fixed using 4% formaldehyde (Sigma-Aldrich) in 1xPBS [137 mM NaCl (Fisher Chemical), 2.7 mM KCl (VWR), 10 mM NaHPO₄ (Merk), and 1.8 mM KH₂PO₄ (Merk)]. Pancreatic cells were extracted from 48 hpf transgenic zebrafish embryos (z3'-DpE:GFP), immediately following euthanasia by rapid chilling, by repeated pipetting up and down in a gentle motion with 300 µL of Ginzburg fish Ringer's solution [55 mM NaCl (Fisher Chemical), 1.8 mM KCl (VWR), 1.25 mM NaHCO₃ (Sigma-Aldrich)]. Embryos were allowed to settle to the bottom and the suspension containing the detached pancreatic progenitor cells and yolk was collected, washed with PBS. For the purpose of ATAC-seq, zebrafish pancreatic progenitors were then fixed for 10min in 4%PFA in 1xPBS. Cells from the whole pancreas and zebrafish pancreatic progenitor cells were then dissociated, on ice, using a 15 mL Dounce homogenizer in 1 mL of ice-cold sort buffer [1% EDTA (VWR), 2 mM HEPES (Sigma-Aldrich) pH 7.0 in 1xPBS), and then passed through a 40-µm cell strainer. Immediately following dissociation, the mCherry and GFP fluorescence was analysed on a BD FACS-ARIATM II cell sorter (BD Biosciences) and, after sorting, cells were collected in sort buffer (for posterior ATAC-seq) or TRK Lysis Buffer (for RNA-seq).

Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq)

ATAC-seq was performed as previously described, with minor changes. For each replicate of ATAC-seq of MPCs, a low number of cells (<5000) was collected. As recommended, in these samples, a scale-down of the tagmentation reaction was performed (10uL of total volume for each tagmentation reaction). After collection from flow cytometry, cells were centrifuged, resuspended and lysed in ice-cold fresh lysis buffer (10mM Tris-HCl, pH7.5; 10mM NaCl; 3mM MgCl₂; 0.1% IGEPAL) and immediately incubated with homemade Tn5 transposase in TAPS-DMF buffer (Picelli et al, 2014) at 37°C for 30 min, followed by proteinase K treatment. Immediately after transposition the samples were purified with MinElute PCR Purification Kit (QIAGEN). We performed qPCR analysis on the samples to determine the appropriate number of PCR cycles to create the library. ATAC libraries were amplified using KAPA HiFi HotStart PCR Kit (Roche) with the primers Ad1 and Ad2 (from 2.1 to 2.13), used for multiplexing. The library was immediately purified using the QIAquick PCR Purification Kit (QIAGEN) and eluted in 20 uL. Successful tagmentation was confirmed with a TAE agarose gel electrophoresis.

RNA extraction

After collection from flow cytometry, RNA was extracted from zebrafish pancreatic progenitor cells, using E.Z.N.A MicroElute Total RNA Kit (OmegaBiotek), following the manufacturer's instructions. Samples were incubated 30 min at 37 °C with 1 µl DNase I (ThermoScientific), 1 µl 10x reaction buffer and 0.5 µl NZY Ribonuclease Inhibitor (NZYTech). Extracted RNA was stored at -80 °C for posterior use.

Quantification and Statistical Analysis

Assuming that the data are normally distributed, an unpaired t-test with Welch's correction was performed to test the significance of differences among sample averages. In the case of the average length of pixels, a Mann-Whitney test to compare ranks was used. A chi-square test was applied in the analysis of transcription factor co-binding predictions. In all analyses, $P < 0.05$ was required for statistical significance and calculated in GraphPad Prism 7 (San Diego, CA, USA) or R.

3.3.3. Processing and bioinformatics analysis

RNA-seq analysis

Total RNA extracted from zebrafish pancreatic progenitors was sequenced on Illumina HiSeq 2500 platform and was inspected for quality control using FASTQC (v.0.11.8) (Andrews, 2010). Sequences were aligned to the reference genome (GRCz10/danRer10) using Bowtie2 (v.2.4.2) (Langmead & Salzberg, 2012) with parameters "--very-sensitive". Read count was performed using featureCounts (Liao et al, 2014), with default parameters. Differential expression analysis was performed with DESeq2 (v.1.32) (Love et al, 2014), with differences considered statistically significant when adjusted p value ≤ 0.05 and $\text{abs}(\log_2\text{FC}) \geq 1$. Gene ontology analysis was performed using PANTHER GO Enrichment Analysis (Thomas et al, 2022). Bam files from each replicate were normalized with bamCoverage, using reads per kilobase per million mapped reads (RPKM) method, from deepTools (v.3.5.1) (Ramirez et al, 2016). For visualization of the expression profiles, peaks from replicates were averaged with WiggleTools (v.1.2) (Zerbino et al, 2014) mean, and the resulting bedGraph files were converted to bigwig files with bedGraphToBigWig (v.377) (Kent et al, 2010). Bigwig tracks were then uploaded to the UCSC browser for visualization.

ATAC-seq analysis

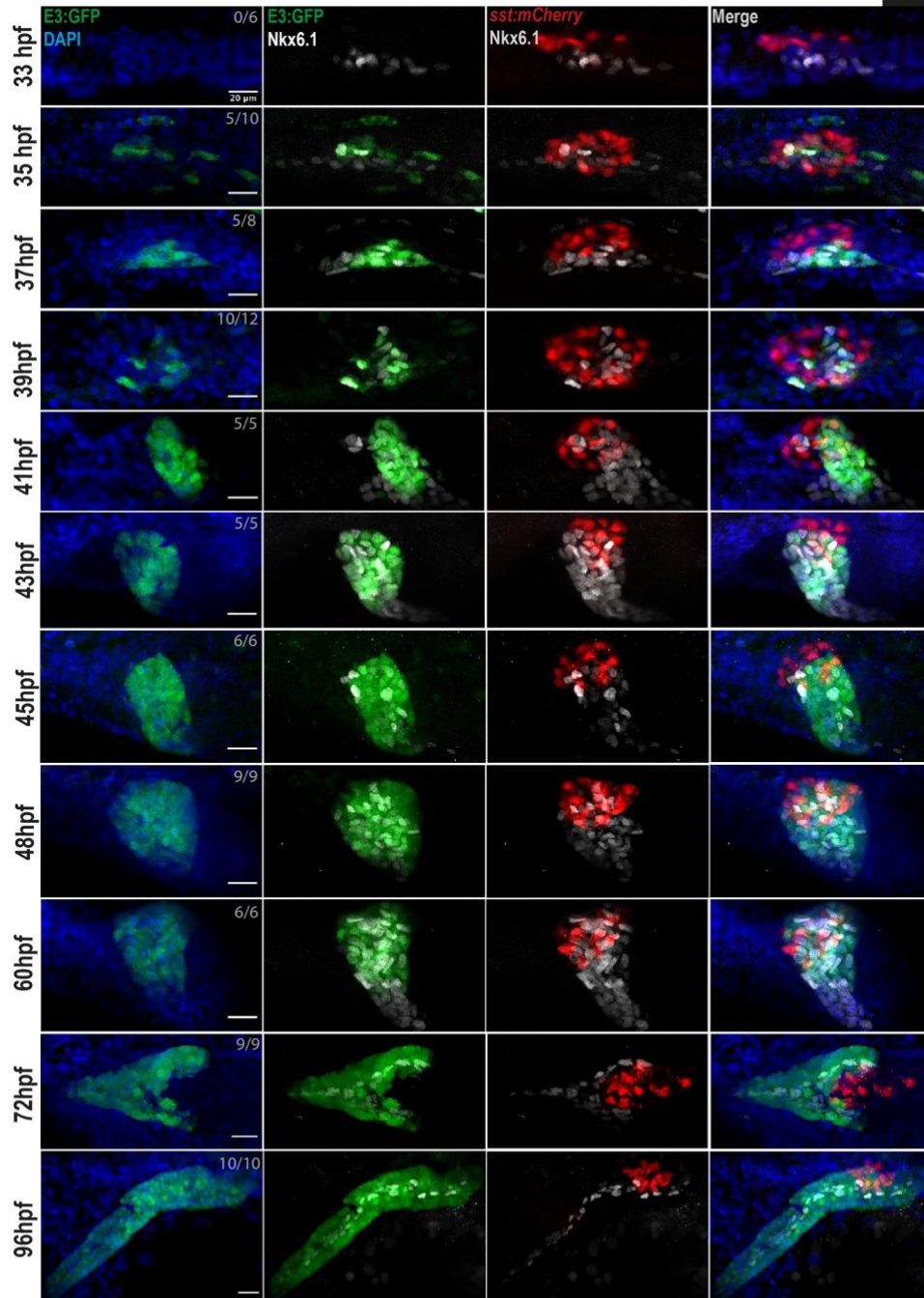
High quality raw reads for replicates of ATAC-seq (FASTQC v.0.11.8) ([Andrews, 2010](#)) were trimmed for adapter sequences using Skewer (v.0.2.1) ([Jiang et al, 2014](#)). All libraries were sequenced on Illumina HiSeq 2500 platform and raw reads were mapped to the reference zebrafish genome (GRCz10/danRer10) using Bowtie2 (v.2.4.2) ([Langmead & Salzberg, 2012](#)). To avoid clonal artefacts, the duplicated mapped reads were removed using Samtools ([Li et al, 2009](#)). Mapped reads were filtered by the fragment size (≤ 120 bp) and mapping quality (≥ 10). To call for enriched regions, MACS2 (v.2.2.7) ([Zhang et al, 2008](#)) was used with the parameters “--nomodel, --keep-dup 1, --llocal 10000, --extsize 74, --shift – 37 and -p 0.07”. Then, the Irreproducible Discovery Rate (IDR, v.2.0.4) was applied, in order to obtain a reproducible set of peaks ([Li et al, 2011](#)). Further analysis was performed for peaks with an IDR score reflecting a p-value ≤ 0.1 . Differential accessibility analysis was performed with EdgeR ([Robinson et al, 2010](#)), with differences considered statistically significant when adjusted p value ≤ 0.05 and $\text{abs}(\log_2\text{FC}) \geq 1$. Bam files from each replicate were normalized with bamCoverage, using reads per kilobase per million mapped reads (RPKM), from deepTools ([Ramirez et al, 2016](#)). For visualization of the accessibility profiles, peaks from replicates were first averaged with WiggleTools (v.1.2) ([Zerbino et al, 2014](#)) mean, and the resulting bedGraph files were converted to bigwig files with bedGraphToBigWig (v.3.77) ([Kent et al, 2010](#)). Bigwig tracks were then uploaded to the UCSC browser for visualization. ATAC-seq heatmaps and profiles were generated with deepTools (v.3.5.1) ([Ramirez et al, 2016](#)), by first computing a matrix with the parameters “--referencePoint center, -b 500, -a 500, --binSize 10, --sortRegions descend” and the bigwig files above mentioned and by using the tools plotHeatmap and plotProfile, with default parameters. Genomic intersections were performed using BEDTools (v.2.29.2) ([Quinlan & Hall, 2010](#)). To associate ATAC-seq peaks to nearby genes, the Genomic Regions Enrichment of Annotations Tool (GREAT) ([McLean et al, 2010](#)) was used. Lists of GREAT-associated genes were fed to PANTHER GO Enrichment Analysis ([Thomas et al, 2022](#)) to compare enrichment of gene ontology terms between samples.

Transcription factor binding motifs enrichment and disease association enrichment of TF-encoding genes

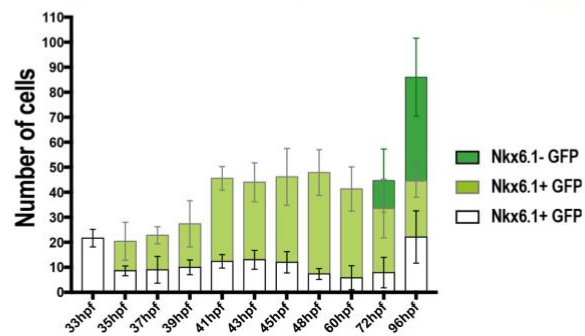
Motif enrichment analysis on differentially accessible peak datasets from ATAC-seq of acinar cells was performed with HOMER (v.4.11) (Heinz et al, 2011), using findMotifsGenome.pl. Enrichment of TF binding motifs in different samples were compared by subtracting the predicted percentage of sequences containing enriched motifs in one sample by the other, by recurring to a custom Python script. Results from the literature were manually curated to define groups of TFs involved in specific functions (see Table S5). To access enrichment of genes associated with pancreatic diseases within our analysis, we used lists of the genes encoding TFs that were predicted to be differentially enriched in ATAC-seq of acinar cells (Down and Up TFs) and took human gene-disease associations from DisGeNET (v.6.0) (Piñero et al, 2015) for the available pancreatic diseases. In detail, pancreatic diseases and their associated genes were selected from the file containing all gene-disease links from DisGeNET (all_gene_disease_associations.tsv, downloaded from the DisGeNET website, v6.0, <http://www.disgenet.org/>, Integrative Biomedical Informatics Group GRIB/IMIM/UPF), filtering for associations with a score > 0.1 to exclude those based only on text-mining. The disease search term used was “pancrea*”, followed by manually filtering for pancreas-related diseases and their human associated genes. Finally, a hypergeometric test was performed to test if the selected TF-encoding genes were enriched in DisGeNET lists of pancreatic diseases-associated genes (see also Tables S6 and S7).

3.4. Supplementary Information

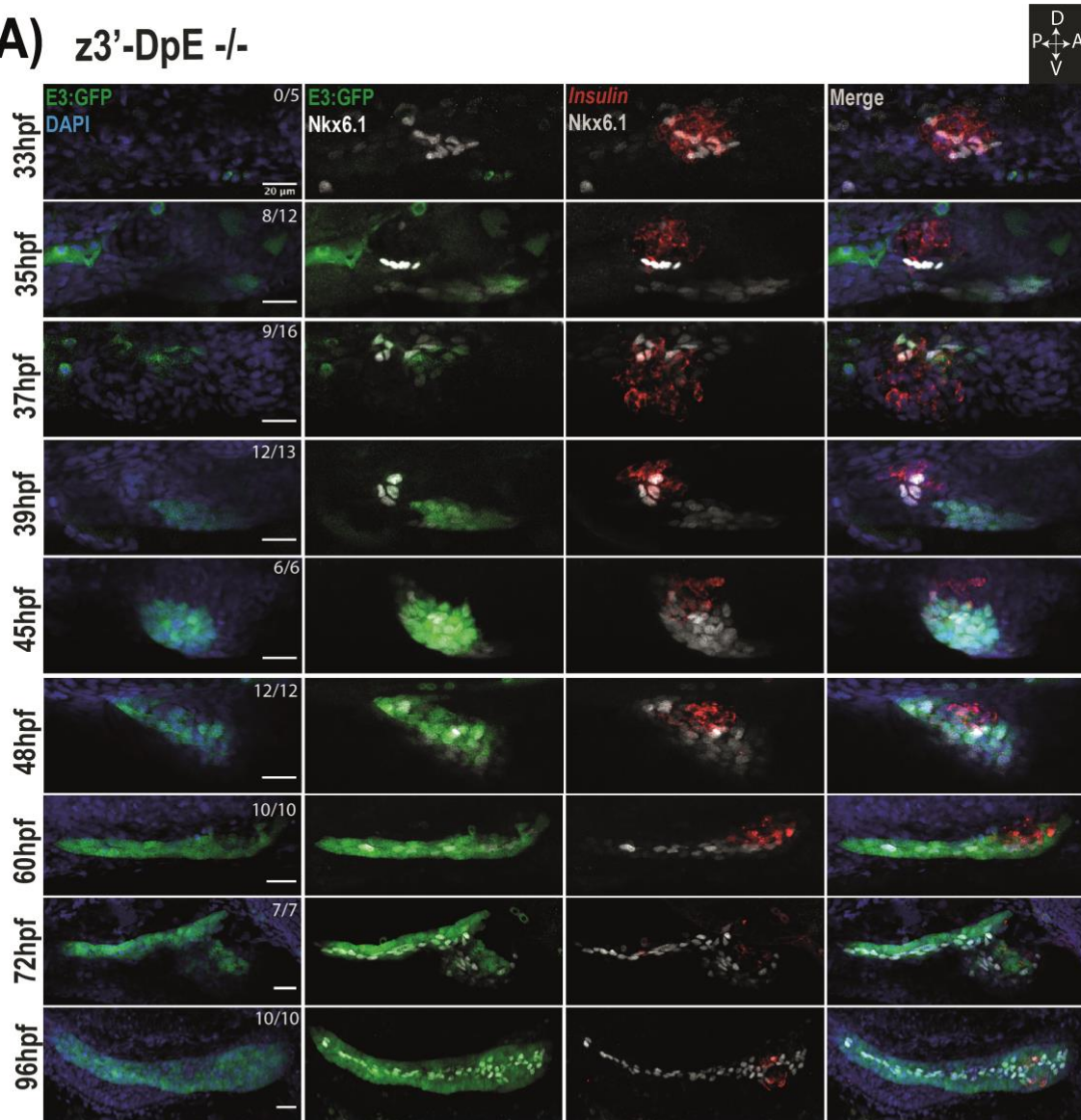
A) z3'-DpE +/+



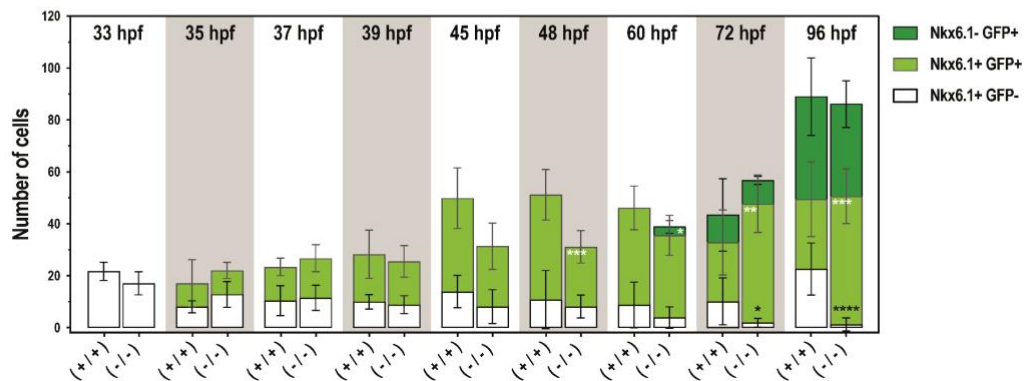
B)



FigureS3.1 (Related to Figure 3.1) – A) GFP expression pattern of the z3'-DpE:GFP transgenic reporter line in z3'-DpE +/+ embryos, from 33 to 96hpf. Nuclei (blue) are stained with DAPI, MPCs (white) are labeled by anti-Nkx6.1 staining and endocrine cells (red) are labeled by the *somatostatin*:mCherry reporter line. Dorsal-ventral and antero-posterior axis are shown. Number of embryos with GFP expression and number of photographed embryos (N) are shown in images of the first column, in the upper right corner. B) Quantification of the number of cells expressing both *nkx6.1* and GFP or either one in the various developmental timings.

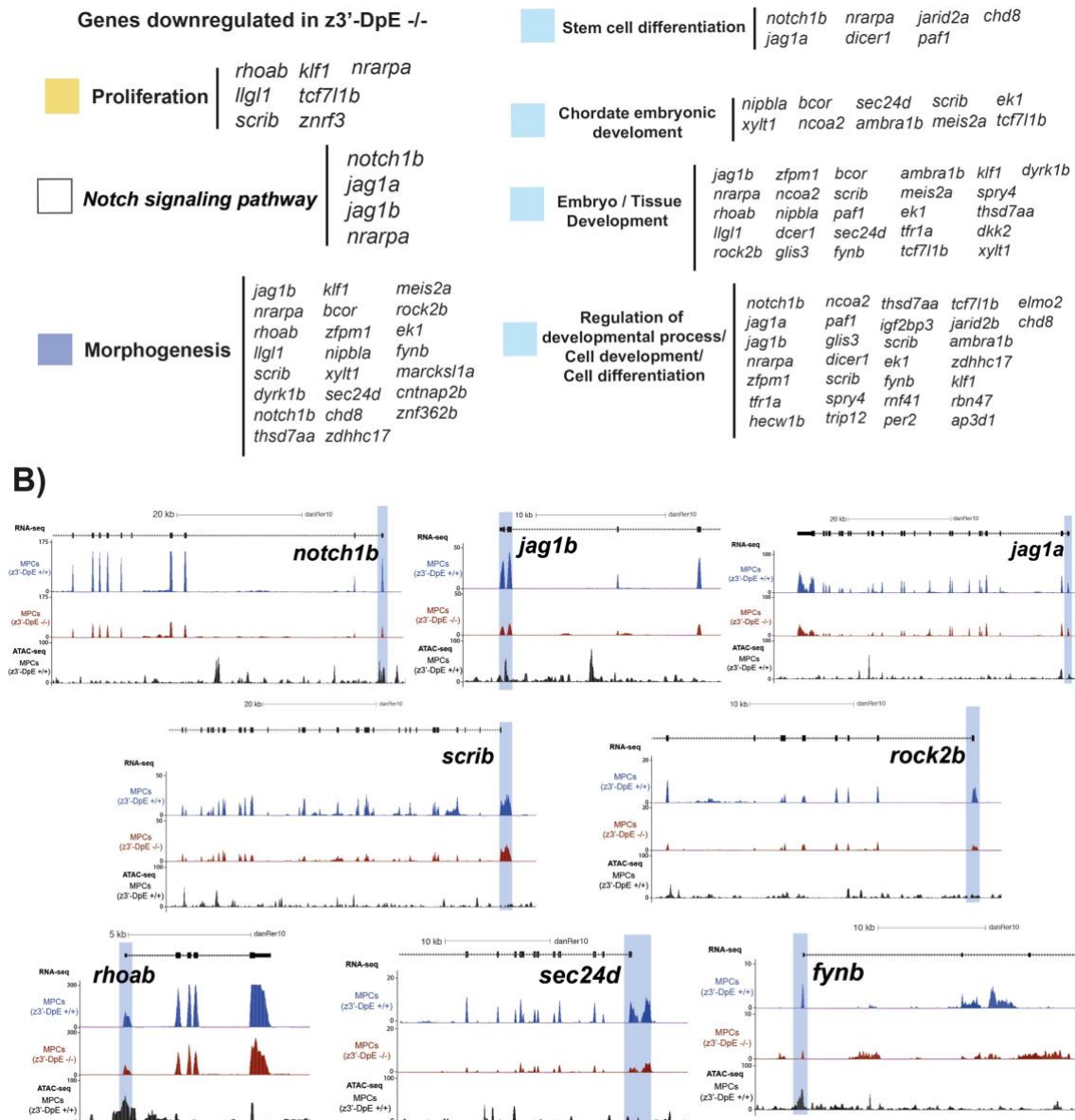
A) *z3'-DpE -/-*

B)

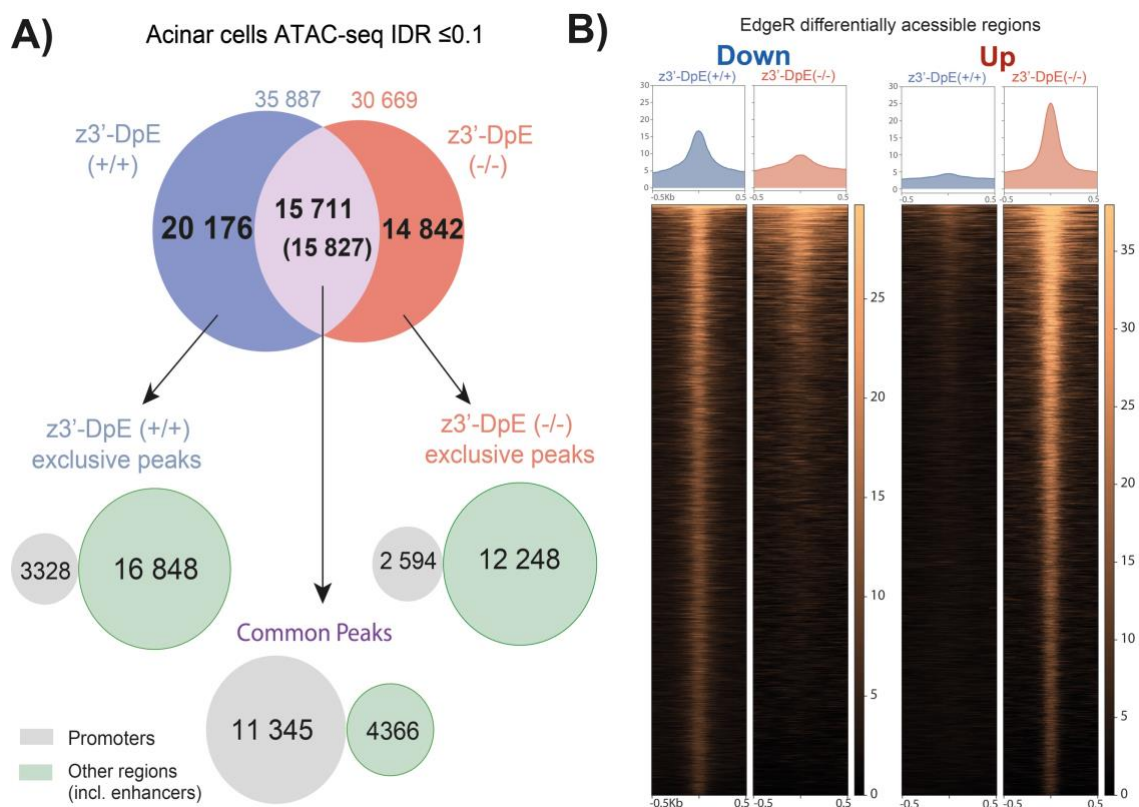


FigureS3.2(Related to Figure 3.2) – A) GFP expression pattern of the *z3'-DpE*:GFP transgenic reporter line in *z3'-DpE -/-* embryos, from 33 to 96hpf. Nuclei (blue) are stained with DAPI, MPCs (white) are labeled by anti-Nkx6.1 staining and endocrine cells (red) are labeled by anti-insulin staining. Dorsal-ventral and antero-posterior axis are shown. Number of embryos with GFP expression and number of photographed embryos (N) are shown in images of the first column, in the upper right corner. B) Quantification of the number of cells expressing both *nrx6.1* and GFP or either one in the various developmental timings,

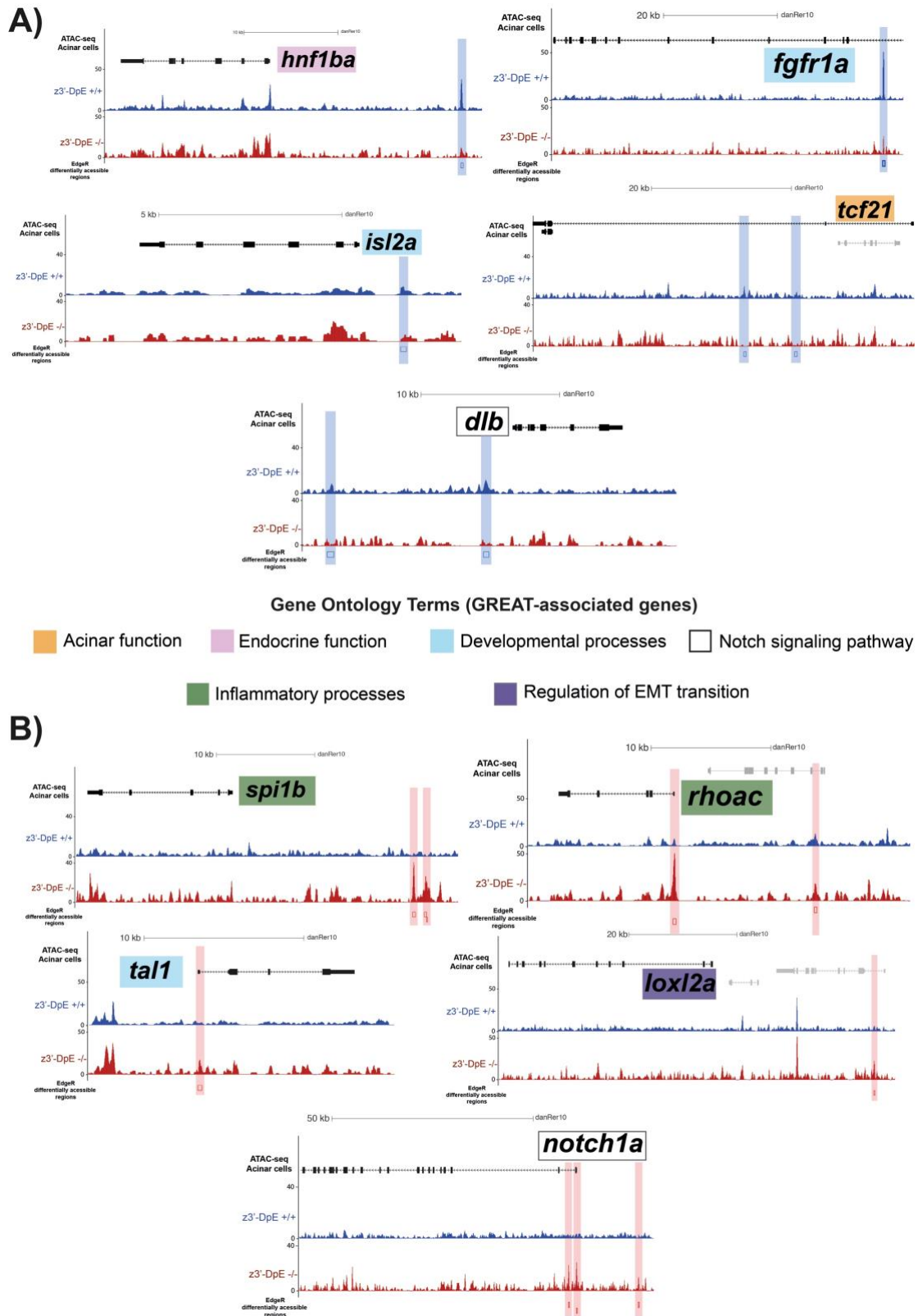
comparing between wild-type and mutants.



FigureS3.3 (Related to Figure 3.3) – A) Downregulated genes among highlighted GO terms in MPCs of *z3'-DpE -/-* embryos, as shown by DEseq2 analysis of RNA-seq results, followed by GO analysis. Notch signaling genes are found within the majority of shown GO terms. **B)** Examples of landscapes of downregulated genes, showing BigWig tracks for RNA-seq data of *z3'-DpE +/+* (blue) and *z3'-DpE -/-* (red) MPCs. Promoters of altered genes according to DEseq2 are highlighted (light blue boxes).



FigureS3.4 (Related to Figure 3.4) – A) Analysis of common and exclusive peaks (with an IDR-score reflecting a p-value ≤ 0.1) within ATAC-seq data for z3'-DpE +/+ (35 887 regions) and -/- (30 669 regions) acinar cells. Common (11 345) and exclusive (20 176 versus 14 842) peaks were intersected with known TSS regions to evaluate the proportion of promoter and regulatory regions within each group, showing that exclusive peaks control mostly non-promoter regions. Gray circles represent promoter regions and green boxes represent non-promoter regions. B) Heatmaps showing differentially accessible peaks resulting from EdgeR analysis and their profiles within ATAC-seq data for z3'-DpE +/+ and -/- acinar cells.



FigureS3.5 (Related to Figure 3.4) – A) Examples of gene landscapes with decreased accessibility, among highlighted GO terms, resulting from EdgeR analysis of ATAC-seq data of z3'-DpE +/- and -/- acinar cells. B) Examples of gene landscapes with increased accessibility among highlighted GO terms. BigWig tracks resulting from ATAC-seq data analysis are shown for +/- (blue) and -/- (red) samples. Differentially accessible peaks are highlighted (blue boxes for regions with decreased accessibility and red boxes for regions with increased accessibility).

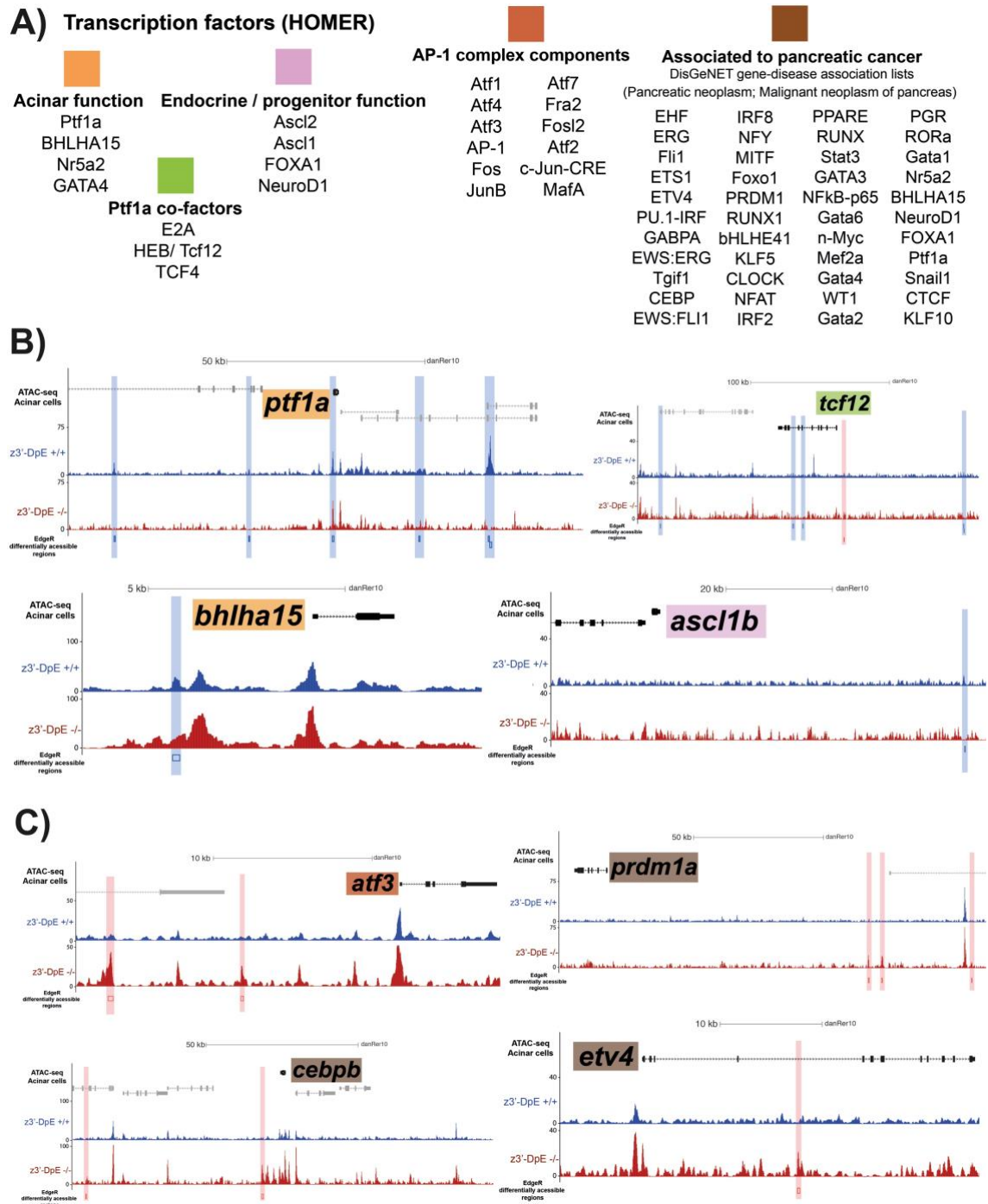
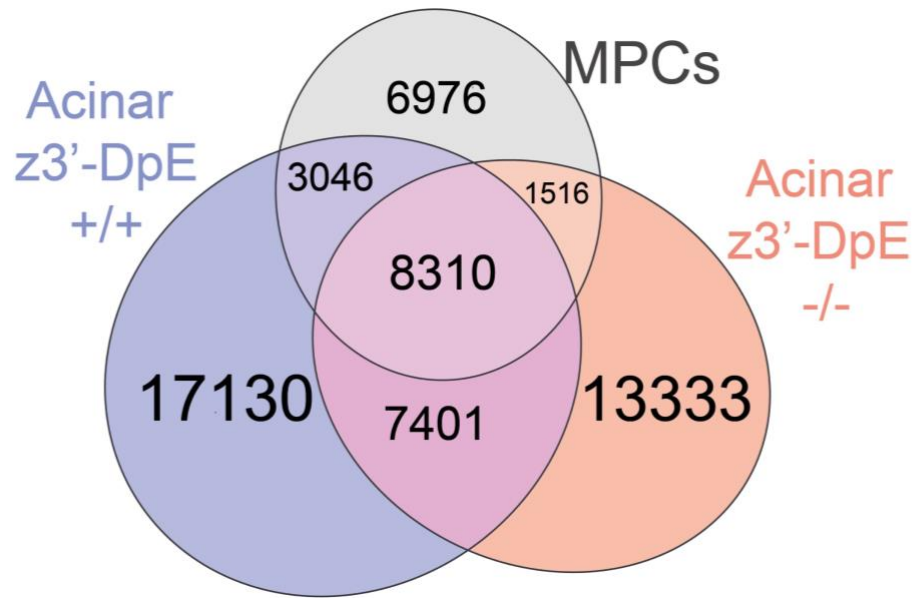


Figure S3.6 (Related to Figure 3.5) – A) Transcription factors found among highlighted functional groups in Figure 3.6, with respective colors being shown. B) Example landscapes of genes encoding TFs that were predicted to be enriched within regions with decreased accessibility in z3'-DpE -/-. B) Example landscapes of genes encoding TFs that were predicted to be enriched within regions with increased accessibility in z3'-DpE -/-. Differentially accessible peaks are highlighted (blue boxes for downregulated regions and red boxes for upregulated regions).



FigureS3.7 (Related to Figure 3.5) – A) Intersection between ATAC-seq peaks of MPCs and wild-type and mutant acinar cells. Peaks with IDR score reflecting a p-value ≤ 0.1 are plotted.

3.5. Supplementary datasets

Primer	Sequence (5'-3')
Ad1	AATGATACGGCGACCACCGAGATCTACACTCGTCGGC AGCGTCAGATGTG
Ad2.1	CAAGCAGAAGACGGCATACGAGATTGCCTTAGTCTC GTGGGCTCGGAGATGT
Ad2.2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTC GTGGGCTCGGAGATGT
Ad2.3	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT
Ad2.4	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTC GTGGGCTCGGAGATGT
Ad2.11	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT
Ad2.13	CAAGCAGAAGACGGCATACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT
Ad2.15	CAAGCAGAAGACGGCATACGAGATCAGATCCAGTCTCGTGGGCTCGGAGATGT

Table S1 – Primers used in this study for ATAC-seq.

Biological Process	Fold Enrichment	FDR
positive regulation of circadian rhythm (GO:0042753)	> 100	1.59E-02
CRD-mediated mRNA stabilization (GO:0070934)	> 100	2.69E-02
thyroid gland development (GO:0030878)	52.08	8.90E-03
midbrain development (GO:0030901)	24.04	3.83E-02
spindle localization (GO:0051653)	23.44	4.00E-02
establishment of spindle localization (GO:0051293)	23.44	3.94E-02
Notch signaling pathway (GO:0007219)	15.43	2.56E-02
regulation of mRNA stability (GO:0043488)	14.20	2.96E-02
regulation of RNA stability (GO:0043487)	13.44	3.53E-02
regulation of mRNA catabolic process (GO:0061013)	12.75	3.98E-02
protein polyubiquitination (GO:0000209)	12.40	1.50E-02
cell population proliferation (GO:0008283)	9.83	2.76E-02
stem cell differentiation (GO:0048863)	9.23	5.28E-03
epithelial tube morphogenesis (GO:0060562)	7.65	2.51E-02
cranial skeletal system development (GO:1904888)	7.44	2.81E-02
protein ubiquitination (GO:0016567)	6.11	1.50E-02
chordate embryonic development (GO:0043009)	5.81	4.77E-03
embryo development ending in birth or egg hatching (GO:0009792)	5.78	4.45E-03
protein modification by small protein conjugation (GO:0032446)	5.54	2.17E-02
embryonic morphogenesis (GO:0048598)	5.52	5.93E-04
actin cytoskeleton organization (GO:0030036)	5.47	1.31E-02

tissue morphogenesis (GO:0048729)	5.36	6.64E-03
actin filament-based process (GO:0030029)	5.20	1.59E-02
tube morphogenesis (GO:0035239)	5.18	8.49E-03
morphogenesis of an epithelium (GO:0002009)	5.15	2.76E-02
embryo development (GO:0009790)	5.04	9.89E-06
embryonic organ development (GO:0048568)	4.83	1.27E-02
head development (GO:0060322)	4.79	2.22E-02
tube development (GO:0035295)	4.70	4.58E-03
protein modification by small protein conjugation or removal (GO:0070647)	4.65	4.39E-02
regulation of developmental process (GO:0050793)	4.41	4.18E-03
epithelium development (GO:0060429)	4.35	2.42E-03
tissue development (GO:0009888)	4.25	8.66E-06
negative regulation of metabolic process (GO:0009892)	4.03	1.27E-02
negative regulation of macromolecule metabolic process (GO:0010605)	3.83	2.61E-02
positive regulation of metabolic process (GO:0009893)	3.67	3.02E-02
circulatory system development (GO:0072359)	3.64	2.24E-02
animal organ development (GO:0048513)	3.51	2.73E-06
cell surface receptor signaling pathway (GO:0007166)	3.39	2.23E-02
regulation of signal transduction (GO:0009966)	3.14	2.49E-02
cell development (GO:0048468)	3.07	6.84E-03
anatomical structure morphogenesis (GO:0009653)	2.99	1.11E-03
multicellular organism development (GO:0007275)	2.78	7.67E-06
protein modification process (GO:0036211)	2.73	4.43E-02
system development (GO:0048731)	2.72	7.55E-05
positive regulation of biological process (GO:0048518)	2.67	1.30E-02
nervous system development (GO:0007399)	2.65	2.83E-02
cell differentiation (GO:0030154)	2.65	7.06E-03
cellular developmental process (GO:0048869)	2.64	6.62E-03
positive regulation of cellular process (GO:0048522)	2.50	4.66E-02
negative regulation of cellular process (GO:0048523)	2.49	3.59E-02
multicellular organismal process (GO:0032501)	2.46	1.14E-05
developmental process (GO:0032502)	2.43	1.03E-05
anatomical structure development (GO:0048856)	2.42	1.61E-05
negative regulation of biological process (GO:0048519)	2.40	4.88E-02
regulation of RNA metabolic process (GO:0051252)	2.31	2.74E-02
regulation of gene expression (GO:0010468)	2.25	1.57E-02
regulation of macromolecule biosynthetic process (GO:0010556)	2.23	1.78E-02
regulation of nucleobase-containing compound metabolic process (GO:0019219)	2.22	3.58E-02
regulation of biosynthetic process (GO:0009889)	2.19	2.12E-02
regulation of metabolic process (GO:0019222)	2.14	1.50E-02
regulation of macromolecule metabolic process (GO:0060255)	2.11	2.55E-02
cellular component organization (GO:0016043)	2.11	2.07E-02

cellular component organization or biogenesis (GO:0071840)	2.02	3.54E-02
regulation of cellular process (GO:0050794)	1.69	3.55E-03
biological regulation (GO:0065007)	1.68	1.16E-03
regulation of biological process (GO:0050789)	1.68	1.88E-03

Table S2 – Gene Ontology Terms associated with significantly downregulated genes in DEseq2 analysis of RNA-seq data. Terms with fold enrichment <1.5 are shown.

Biological Process	Fold Enrichment	FDR
insulin secretion involved in cellular response to glucose stimulus (GO:0035773)	7.94	2.32E-02
vagus nerve development (GO:0021564)	7.94	2.32E-02
late distal convoluted tubule development (GO:0072068)	7.94	2.32E-02
metaphase/anaphase transition of mitotic cell cycle (GO:0007091)	7.94	2.31E-02
distal convoluted tubule development (GO:0072025)	7.94	2.31E-02
negative regulation of synapse assembly (GO:0051964)	7.94	2.31E-02
selective angioblast sprouting (GO:0035474)	7.94	2.31E-02
metaphase/anaphase transition of cell cycle (GO:0044784)	7.94	2.30E-02
regulation of Wnt signaling pathway, calcium modulating pathway (GO:0008591)	7.94	2.30E-02
regulation of nuclear-transcribed mRNA catabolic process, nonsense-mediated decay (GO:2000622)	7.94	2.30E-02
smooth muscle cell proliferation (GO:0048659)	7.94	2.29E-02
DNA ligation involved in DNA repair (GO:0051103)	7.94	2.29E-02
negative regulation of synapse organization (GO:1905809)	7.94	2.29E-02
rhombomere 6 development (GO:0021572)	7.94	3.92E-03
rhombomere 5 development (GO:0021571)	7.94	3.92E-03
motor behavior (GO:0061744)	7.94	3.91E-03
neural plate morphogenesis (GO:0001839)	6.81	5.37E-04
locus ceruleus development (GO:0021703)	6.62	2.83E-03
intracellular sphingolipid homeostasis (GO:0090156)	6.36	1.43E-02
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification (GO:2000053)	6.36	1.43E-02
glial cell fate commitment (GO:0021781)	6.36	1.43E-02
oligodendrocyte cell fate commitment (GO:0021779)	6.36	1.43E-02
renal system pattern specification (GO:0072048)	5.96	1.65E-03
distal tubule development (GO:0072017)	5.96	1.65E-03
negative regulation of endodermal cell fate specification (GO:0042664)	5.96	1.64E-03
motor neuron migration (GO:0097475)	5.96	1.64E-03
proximal/distal pattern formation (GO:0009954)	5.96	1.64E-03
insulin secretion (GO:0030073)	5.96	1.63E-03

pattern specification involved in kidney development (GO:0061004)	5.96	1.63E-03
pattern specification involved in pronephros development (GO:0039017)	5.96	1.63E-03
synaptic transmission, GABAergic (GO:0051932)	5.67	7.50E-03
coronary vasculature development (GO:0060976)	5.67	7.48E-03
pronephric distal tubule development (GO:0035777)	5.30	3.30E-02
anterior/posterior pattern specification involved in kidney development (GO:0072098)	5.30	3.30E-02
peptidyl-histidine modification (GO:0018202)	5.30	3.30E-02
regulation of gene silencing by regulatory ncRNA (GO:0060966)	5.30	3.29E-02
anterior/posterior pattern specification involved in pronephros development (GO:0034672)	5.30	3.29E-02
positive regulation of peptide hormone secretion (GO:0090277)	5.30	3.28E-02
regulation of retinal ganglion cell axon guidance (GO:0090259)	5.30	3.28E-02
regulation of muscle organ development (GO:0048634)	5.30	3.28E-02
pineal gland development (GO:0021982)	5.30	3.27E-02
positive regulation of peptide secretion (GO:0002793)	5.30	3.27E-02
protein histidyl modification to diphthamide (GO:0017183)	5.30	3.26E-02
neuron cell-cell adhesion (GO:0007158)	5.30	3.72E-03
negative regulation of endodermal cell differentiation (GO:1903225)	5.30	3.72E-03
regulation of non-canonical Wnt signaling pathway (GO:2000050)	5.14	2.29E-05
urogenital system development (GO:0001655)	5.11	2.06E-04
hormone secretion (GO:0046879)	5.06	1.87E-03
peptide hormone secretion (GO:0030072)	5.06	1.86E-03
peptide secretion (GO:0002790)	5.06	1.86E-03
gamma-aminobutyric acid signaling pathway (GO:0007214)	4.97	1.03E-04
angioblast cell migration (GO:0035476)	4.97	1.02E-04
DNA topological change (GO:0006265)	4.97	1.58E-02
venous blood vessel morphogenesis (GO:0048845)	4.97	1.58E-02
hypothalamus cell differentiation (GO:0021979)	4.97	1.58E-02
regulation of nodal signaling pathway (GO:1900107)	4.89	9.06E-04
regulation of cell-matrix adhesion (GO:0001952)	4.77	7.50E-03
response to auditory stimulus (GO:0010996)	4.77	7.49E-03
regulation of axon guidance (GO:1902667)	4.77	7.48E-03
hormone transport (GO:0009914)	4.63	3.55E-03
pituitary gland development (GO:0021983)	4.60	9.58E-05
rhombomere development (GO:0021546)	4.54	4.42E-05
epithalamus development (GO:0021538)	4.54	4.41E-05
S-adenosylmethionine metabolic process (GO:0046500)	4.54	1.70E-03
cardioblast differentiation (GO:0010002)	4.41	2.86E-02
angioblast cell migration from lateral mesoderm to midline (GO:0035479)	4.41	2.85E-02
heme transport (GO:0015886)	4.41	2.85E-02
regulation of cell-substrate junction organization (GO:0150116)	4.41	2.85E-02

regulation of focal adhesion assembly (GO:0051893)	4.41	2.84E-02
anterior lateral line development (GO:0048899)	4.41	2.84E-02
regulation of cell-substrate junction assembly (GO:0090109)	4.41	2.84E-02
habenula development (GO:0021986)	4.41	3.68E-04
body morphogenesis (GO:0010171)	4.41	3.67E-04
face morphogenesis (GO:0060325)	4.41	2.83E-02
cell migration in hindbrain (GO:0021535)	4.33	1.35E-02
neural plate development (GO:0001840)	4.33	1.35E-02
vascular endothelial growth factor signaling pathway (GO:0038084)	4.33	1.34E-02
benzene-containing compound metabolic process (GO:0042537)	4.33	1.34E-02
positive regulation of phosphatidylinositol 3-kinase/protein kinase B signal transduction (GO:0051897)	4.33	1.34E-02
inhibitory synapse assembly (GO:1904862)	4.33	1.34E-02
negative regulation of cell fate specification (GO:0009996)	4.33	1.34E-02
cloaca development (GO:0035844)	4.33	1.33E-02
face development (GO:0060324)	4.33	1.33E-02
head morphogenesis (GO:0060323)	4.33	1.33E-02
vascular endothelial growth factor receptor signaling pathway (GO:0048010)	4.24	2.92E-03
cardiac muscle tissue regeneration (GO:0061026)	4.24	2.91E-03
regulation of activin receptor signaling pathway (GO:0032925)	4.21	1.41E-03
muscle cell proliferation (GO:0033002)	4.21	1.40E-03
adenohypophysis development (GO:0021984)	4.21	1.40E-03
rhombomere morphogenesis (GO:0021593)	4.18	6.47E-04
retinoic acid receptor signaling pathway (GO:0048384)	4.15	1.38E-04
anterior/posterior axon guidance (GO:0033564)	3.97	2.20E-02
proximal tubule development (GO:0072014)	3.97	2.19E-02
cellular response to vascular endothelial growth factor stimulus (GO:0035924)	3.97	1.02E-02
response to zinc ion (GO:0010043)	3.97	4.79E-03
pons development (GO:0021548)	3.97	2.27E-03
establishment of tissue polarity (GO:0007164)	3.97	4.56E-02
cardiac muscle tissue growth (GO:0055017)	3.97	4.78E-03
negative regulation of extrinsic apoptotic signaling pathway (GO:2001237)	3.97	4.77E-03
positive regulation of BMP signaling pathway (GO:0030513)	3.97	2.26E-03
Schwann cell development (GO:0014044)	3.97	1.08E-03
venous blood vessel development (GO:0060841)	3.97	2.26E-03
iron coordination entity transport (GO:1901678)	3.97	4.56E-02
embryonic retina morphogenesis in camera-type eye (GO:0060059)	3.84	8.19E-05
hypothalamus development (GO:0021854)	3.83	1.74E-04
cAMP-mediated signaling (GO:0019933)	3.80	7.98E-04
response to glucose (GO:0009749)	3.74	7.55E-03
response to hexose (GO:0009746)	3.74	7.54E-03

response to carbohydrate (GO:0009743)	3.74	7.53E-03
response to monosaccharide (GO:0034284)	3.74	7.51E-03
heart growth (GO:0060419)	3.74	7.50E-03
semaphorin-plexin signaling pathway (GO:0071526)	3.72	3.55E-07
cell volume homeostasis (GO:0006884)	3.71	1.58E-02
steroid hormone mediated signaling pathway (GO:0043401)	3.71	1.58E-02
diencephalon development (GO:0021536)	3.70	7.79E-13
limbic system development (GO:0021761)	3.69	2.66E-04
olfactory placode development (GO:0071698)	3.67	3.28E-02
posterior lateral line neuromast hair cell differentiation (GO:0048923)	3.67	3.27E-02
anterior lateral line system development (GO:0048898)	3.67	3.27E-02
neural nucleus development (GO:0048857)	3.67	5.78E-04
embryonic neurocranium morphogenesis (GO:0048702)	3.67	5.77E-04
negative regulation of signal transduction by p53 class mediator (GO:1901797)	3.67	3.27E-02
mesendoderm development (GO:0048382)	3.67	3.26E-02
hematopoietic or lymphoid organ development (GO:0048534)	3.64	1.25E-03
cranial nerve morphogenesis (GO:0021602)	3.64	1.24E-03
Schwann cell differentiation (GO:0014037)	3.61	2.55E-03
zinc ion transmembrane transport (GO:0071577)	3.58	5.31E-03
negative regulation of intrinsic apoptotic signaling pathway (GO:2001243)	3.58	5.30E-03
midbrain development (GO:0030901)	3.55	3.03E-05
regulation of vascular permeability (GO:0043114)	3.53	1.12E-02
nuclear migration (GO:0007097)	3.53	1.12E-02
nucleus localization (GO:0051647)	3.53	1.12E-02
inner ear receptor cell development (GO:0060119)	3.49	2.12E-05
inner ear receptor cell differentiation (GO:0060113)	3.48	3.07E-06
gland morphogenesis (GO:0022612)	3.48	2.31E-02
peripheral nervous system axon ensheathment (GO:0032292)	3.48	2.31E-02
zinc ion transport (GO:0006829)	3.45	3.73E-03
closure of optic fissure (GO:0061386)	3.45	3.72E-03
cartilage morphogenesis (GO:0060536)	3.44	6.10E-04
regulation of vascular endothelial growth factor receptor signaling pathway (GO:0030947)	3.40	4.76E-02
myelination in peripheral nervous system (GO:0022011)	3.40	4.76E-02
rhombomere formation (GO:0021594)	3.40	4.75E-02
cardiac muscle cell proliferation (GO:0060038)	3.40	4.75E-02
embryonic camera-type eye development (GO:0031076)	3.40	6.42E-07
positive regulation of macroautophagy (GO:0016239)	3.40	4.74E-02
regulation of peptide hormone secretion (GO:0090276)	3.40	4.74E-02
striated muscle cell proliferation (GO:0014855)	3.40	4.73E-02
negative regulation of cell fate commitment (GO:0010454)	3.40	4.73E-02

regulation of peptide transport (GO:0090087)	3.40	4.72E-02
regulation of peptide secretion (GO:0002791)	3.40	4.72E-02
neuromast hair cell development (GO:0035675)	3.40	7.76E-03
regulation of extrinsic apoptotic signaling pathway (GO:2001236)	3.40	7.75E-03
intracellular receptor signaling pathway (GO:0030522)	3.40	3.17E-05
immune system development (GO:0002520)	3.40	1.99E-04
determination of ventral identity (GO:0048264)	3.40	7.74E-03
negative regulation of axon extension (GO:0030517)	3.38	6.76E-05
regulation of endodermal cell differentiation (GO:1903224)	3.35	1.61E-02
central nervous system projection neuron axonogenesis (GO:0021952)	3.35	1.61E-02
ectodermal placode development (GO:0071696)	3.33	4.59E-05
positive regulation of transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0090100)	3.31	5.34E-03
notochord morphogenesis (GO:0048570)	3.29	1.83E-03
thyroid gland development (GO:0030878)	3.27	3.27E-02
regulation of endodermal cell fate specification (GO:0042663)	3.27	3.27E-02
determination of pancreatic left/right asymmetry (GO:0035469)	3.27	3.27E-02
medial fin morphogenesis (GO:0035141)	3.27	3.26E-02
feeding behavior (GO:0007631)	3.27	3.26E-02
positive regulation of axonogenesis (GO:0050772)	3.27	6.09E-04
dorsal/ventral axis specification (GO:0009950)	3.25	1.11E-02
blood vessel endothelial cell migration (GO:0043534)	3.25	1.11E-02
ephrin receptor signaling pathway (GO:0048013)	3.24	3.67E-03
cardiac chamber morphogenesis (GO:0003206)	3.24	3.67E-03
neuron recognition (GO:0008038)	3.23	1.27E-03
midbrain-hindbrain boundary development (GO:0030917)	3.22	1.37E-04
negative regulation of axonogenesis (GO:0050771)	3.22	1.36E-04
embryonic camera-type eye morphogenesis (GO:0048596)	3.21	4.49E-05
embryonic eye morphogenesis (GO:0048048)	3.21	4.45E-06
regulation of actomyosin structure organization (GO:0110020)	3.18	7.49E-03
auditory receptor cell development (GO:0060117)	3.18	2.27E-02
response to alcohol (GO:0097305)	3.18	2.27E-02
digestive system process (GO:0022600)	3.18	7.48E-03
thymus development (GO:0048538)	3.18	2.26E-02
ectodermal placode morphogenesis (GO:0071697)	3.18	8.65E-04
mechanoreceptor differentiation (GO:0042490)	3.18	2.86E-07
negative regulation of axon extension involved in axon guidance (GO:0048843)	3.13	1.74E-03
optic cup morphogenesis involved in camera-type eye development (GO:0002072)	3.12	5.09E-03
otic placode formation (GO:0043049)	3.11	1.53E-02
negative regulation of developmental growth (GO:0048640)	3.11	1.30E-04
cardiac chamber development (GO:0003205)	3.11	1.29E-04

establishment of spindle orientation (GO:0051294)	3.11	1.52E-02
optic nerve development (GO:0021554)	3.09	4.47E-02
activin receptor signaling pathway (GO:0032924)	3.09	4.46E-02
myeloid leukocyte activation (GO:0002274)	3.09	4.46E-02
organ growth (GO:0035265)	3.08	3.48E-03
cyclic-nucleotide-mediated signaling (GO:0019935)	3.08	3.47E-03
ear morphogenesis (GO:0042471)	3.06	5.19E-09
embryonic skeletal system morphogenesis (GO:0048704)	3.06	1.80E-17
thigmotaxis (GO:0001966)	3.06	1.03E-02
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay (GO:0000184)	3.06	1.03E-02
morphogenesis of a branching structure (GO:0001763)	3.06	1.02E-02
establishment of mitotic spindle localization (GO:0040001)	3.06	1.02E-02
establishment of mitotic spindle orientation (GO:0000132)	3.03	3.05E-02
retinal cone cell differentiation (GO:0042670)	3.03	3.04E-02
peptide transport (GO:0015833)	3.03	3.04E-02
semicircular canal morphogenesis (GO:0048752)	3.03	3.04E-02
embryonic viscerocranium morphogenesis (GO:0048703)	3.02	1.01E-09
regulation of phosphatidylinositol 3-kinase/protein kinase B signal transduction (GO:0051896)	3.01	7.01E-03
positive regulation of neuron projection development (GO:0010976)	3.01	7.00E-03
chloride transmembrane transport (GO:1902476)	3.01	2.64E-05
neuromast hair cell differentiation (GO:0048886)	3.01	6.99E-03
semicircular canal development (GO:0060872)	3.01	6.98E-03
inner ear morphogenesis (GO:0042472)	3.01	1.50E-08
posterior lateral line neuromast development (GO:0048919)	3.01	1.63E-03
determination of dorsal/ventral asymmetry (GO:0048262)	3.01	1.63E-03
skeletal system morphogenesis (GO:0048705)	3.00	2.86E-19
morphogenesis of embryonic epithelium (GO:0016331)	2.99	3.68E-06
neural tube patterning (GO:0021532)	2.98	5.43E-05
central nervous system neuron axonogenesis (GO:0021955)	2.98	2.09E-02
rostrocaudal neural tube patterning (GO:0021903)	2.98	2.39E-04
hair cell differentiation (GO:0035315)	2.98	1.15E-05
regulation of fibroblast growth factor receptor signaling pathway (GO:0040036)	2.98	4.69E-03
endothelial cell migration (GO:0043542)	2.98	4.68E-03
embryonic cranial skeleton morphogenesis (GO:0048701)	2.97	8.18E-15
negative regulation of cell projection organization (GO:0031345)	2.94	1.09E-04
inner ear auditory receptor cell differentiation (GO:0042491)	2.94	1.38E-02
negative regulation of chemotaxis (GO:0050922)	2.93	2.13E-03
embryonic skeletal system development (GO:0048706)	2.92	7.31E-17
ear development (GO:0043583)	2.92	1.86E-15
neural precursor cell proliferation (GO:0061351)	2.91	9.27E-03

ectodermal placode formation (GO:0060788)	2.91	9.26E-03
cardiac muscle tissue development (GO:0048738)	2.91	2.57E-08
posterior lateral line development (GO:0048916)	2.91	6.82E-06
positive regulation of neurogenesis (GO:0050769)	2.91	6.79E-06
negative chemotaxis (GO:0050919)	2.91	1.46E-03
retinal ganglion cell axon guidance (GO:0031290)	2.89	9.84E-04
medial fin development (GO:0033338)	2.89	3.97E-02
negative regulation of apoptotic signaling pathway (GO:2001234)	2.87	6.48E-04
neural tube development (GO:0021915)	2.87	1.50E-08
cranial nerve development (GO:0021545)	2.87	6.62E-05
inner ear development (GO:0048839)	2.86	1.49E-14
nephron tubule development (GO:0072080)	2.86	2.71E-02
vesicle cytoskeletal trafficking (GO:0099518)	2.86	2.70E-02
motor neuron axon guidance (GO:0008045)	2.86	4.32E-04
striated muscle tissue development (GO:0014706)	2.85	4.73E-08
tissue migration (GO:0090130)	2.85	2.85E-04
negative regulation of growth (GO:0045926)	2.84	1.92E-04
posterior lateral line system development (GO:0048915)	2.84	1.70E-06
regulation of axon extension involved in axon guidance (GO:0048841)	2.84	1.91E-03
cartilage development (GO:0051216)	2.83	6.17E-12
mechanosensory lateral line system development (GO:0048881)	2.83	1.10E-06
neuron projection guidance (GO:0097485)	2.82	1.13E-19
axon guidance (GO:0007411)	2.82	1.11E-19
brain morphogenesis (GO:0048854)	2.80	8.11E-03
negative regulation of neuron projection development (GO:0010977)	2.80	5.77E-04
negative regulation of cell growth (GO:0030308)	2.80	5.76E-04
protein secretion (GO:0009306)	2.79	5.41E-03
establishment of protein localization to extracellular region (GO:0035592)	2.79	5.40E-03
animal organ formation (GO:0048645)	2.79	5.40E-03
regulation of chemotaxis (GO:0050920)	2.79	1.66E-05
morphogenesis of an epithelial sheet (GO:0002011)	2.78	4.04E-07
carbohydrate homeostasis (GO:0033500)	2.78	3.66E-03
glucose homeostasis (GO:0042593)	2.78	3.66E-03
neuromast development (GO:0048884)	2.77	3.08E-06
epithelial tube formation (GO:0072175)	2.77	7.68E-05
regulation of axonogenesis (GO:0050770)	2.76	8.26E-07
connective tissue development (GO:0061448)	2.76	5.50E-12
enteric nervous system development (GO:0048484)	2.76	1.13E-03
lateral line development (GO:0048882)	2.76	8.81E-09
cellular response to steroid hormone stimulus (GO:0071383)	2.75	3.44E-02
renal tubule development (GO:0061326)	2.75	3.44E-02

swim bladder development (GO:0048794)	2.75	3.43E-02
forebrain development (GO:0030900)	2.74	3.00E-10
embryonic organ morphogenesis (GO:0048562)	2.73	1.51E-30
embryonic epithelial tube formation (GO:0001838)	2.72	1.05E-02
protein localization to extracellular region (GO:0071692)	2.72	7.07E-03
bone morphogenesis (GO:0060349)	2.72	7.06E-03
cellular response to BMP stimulus (GO:0071773)	2.71	3.16E-03
response to BMP (GO:0071772)	2.71	3.15E-03
BMP signaling pathway (GO:0030509)	2.71	3.15E-03
epithelium migration (GO:0090132)	2.70	2.13E-03
autonomic nervous system development (GO:0048483)	2.69	2.82E-04
hindbrain development (GO:0030902)	2.69	5.59E-09
embryonic heart tube development (GO:0035050)	2.68	3.37E-11
chloride transport (GO:0006821)	2.68	4.45E-06
mesenchymal cell migration (GO:0090497)	2.67	3.41E-07
skeletal system development (GO:0001501)	2.67	4.95E-24
heart looping (GO:0001947)	2.65	5.21E-08
neural tube formation (GO:0001841)	2.65	2.04E-02
regulation of neuron projection development (GO:0010975)	2.65	2.08E-10
spindle localization (GO:0051653)	2.65	2.03E-02
regulation of cell junction assembly (GO:1901888)	2.65	2.03E-02
establishment of spindle localization (GO:0051293)	2.65	2.03E-02
sex differentiation (GO:0007548)	2.65	2.03E-02
otic vesicle formation (GO:0030916)	2.65	4.42E-02
hindbrain morphogenesis (GO:0021575)	2.65	8.99E-03
endothelial cell proliferation (GO:0001935)	2.65	4.41E-02
very long-chain fatty acid metabolic process (GO:0000038)	2.65	4.41E-02
neural crest cell migration (GO:0001755)	2.65	6.77E-07
GPI anchor biosynthetic process (GO:0006506)	2.65	4.40E-02
positive regulation of proteolysis involved in protein catabolic process (GO:1903052)	2.65	4.40E-02
regulation of Notch signaling pathway (GO:0008593)	2.65	5.54E-04
protein localization to lysosome (GO:0061462)	2.65	4.39E-02
otolith development (GO:0048840)	2.65	5.52E-04
heart formation (GO:0060914)	2.65	2.98E-02
otolith morphogenesis (GO:0032474)	2.65	4.39E-02
receptor localization to synapse (GO:0097120)	2.65	1.34E-02
receptor internalization (GO:0031623)	2.65	2.98E-02
digestion (GO:0007586)	2.65	2.97E-02
lateral mesoderm development (GO:0048368)	2.65	4.38E-02
synapse assembly (GO:0007416)	2.65	6.03E-03
positive regulation of cell development (GO:0010720)	2.62	7.48E-06

tube formation (GO:0035148)	2.61	1.37E-04
lateral line system development (GO:0048925)	2.61	1.57E-08
negative regulation of nervous system development (GO:0051961)	2.60	7.02E-04
endocrine system development (GO:0035270)	2.60	4.79E-07
epiboly involved in gastrulation with mouth forming second (GO:0055113)	2.59	5.08E-03
sensory perception of sound (GO:0007605)	2.59	5.07E-03
positive regulation of nervous system development (GO:0051962)	2.59	3.44E-05
epithelial cell migration (GO:0010631)	2.59	7.54E-03
epithelial tube morphogenesis (GO:0060562)	2.58	1.18E-14
neural crest cell development (GO:0014032)	2.58	1.12E-08
presynaptic endocytosis (GO:0140238)	2.58	1.69E-02
stem cell development (GO:0048864)	2.58	1.11E-08
synaptic vesicle endocytosis (GO:0048488)	2.58	1.68E-02
inorganic anion transmembrane transport (GO:0098661)	2.58	1.76E-04
embryonic heart tube morphogenesis (GO:0003143)	2.57	2.59E-08
regulation of cell size (GO:0008361)	2.57	3.52E-06
transmission of nerve impulse (GO:0019226)	2.57	2.51E-02
regulation of axon extension (GO:0030516)	2.57	3.89E-04
retina morphogenesis in camera-type eye (GO:0060042)	2.57	1.45E-07
regulation of postsynaptic membrane neurotransmitter receptor levels (GO:0099072)	2.56	3.69E-02
positive regulation of vasculature development (GO:1904018)	2.56	3.69E-02
gastrulation with mouth forming second (GO:0001702)	2.55	1.95E-03
endoderm development (GO:0007492)	2.55	1.95E-03
neural crest cell differentiation (GO:0014033)	2.55	2.21E-09
ossification (GO:0001503)	2.55	9.98E-05
monoatomic anion transmembrane transport (GO:0098656)	2.54	3.26E-04
cranial skeletal system development (GO:1904888)	2.53	1.56E-13
retina homeostasis (GO:0001895)	2.53	9.49E-03
synaptic vesicle recycling (GO:0036465)	2.53	9.48E-03
cell population proliferation (GO:0008283)	2.53	2.32E-08
gland development (GO:0048732)	2.52	8.33E-14
camera-type eye photoreceptor cell differentiation (GO:0060219)	2.52	1.42E-02
cellular ketone metabolic process (GO:0042180)	2.52	1.42E-02
limb development (GO:0060173)	2.51	1.85E-04
regulation of cell morphogenesis (GO:0022604)	2.51	1.84E-04
purine nucleoside bisphosphate biosynthetic process (GO:0034033)	2.51	2.13E-02
ribonucleoside bisphosphate biosynthetic process (GO:0034030)	2.51	2.13E-02
otic vesicle development (GO:0071599)	2.51	2.42E-03
embryonic camera-type eye formation (GO:0060900)	2.51	2.13E-02
actin filament-based movement (GO:0030048)	2.51	2.13E-02
nucleoside bisphosphate biosynthetic process (GO:0033866)	2.51	2.12E-02

hormone-mediated signaling pathway (GO:0009755)	2.50	3.59E-03
pronephros development (GO:0048793)	2.50	1.06E-07
oocyte differentiation (GO:0009994)	2.50	3.11E-02
positive regulation of cell projection organization (GO:0031346)	2.49	1.04E-04
response to mechanical stimulus (GO:0009612)	2.49	5.30E-03
bone mineralization (GO:0030282)	2.48	4.57E-02
metencephalon development (GO:0022037)	2.47	2.02E-03
epidermis development (GO:0008544)	2.47	1.37E-05
tissue regeneration (GO:0042246)	2.47	4.73E-07
response to fibroblast growth factor (GO:0071774)	2.47	1.18E-02
regulation of BMP signaling pathway (GO:0030510)	2.47	8.64E-05
cellular response to fibroblast growth factor stimulus (GO:0044344)	2.47	1.17E-02
central nervous system neuron development (GO:0021954)	2.47	1.17E-02
ameboidal-type cell migration (GO:0001667)	2.47	5.69E-12
fin regeneration (GO:0031101)	2.47	1.28E-04
negative regulation of neurogenesis (GO:0050768)	2.47	2.99E-03
epiboly (GO:0090504)	2.46	7.73E-04
cardiocyte differentiation (GO:0035051)	2.46	7.71E-04
mesenchymal cell differentiation (GO:0048762)	2.46	2.87E-09
epidermal cell differentiation (GO:0009913)	2.46	1.90E-04
liver development (GO:0001889)	2.46	1.40E-07
sensory perception of mechanical stimulus (GO:0050954)	2.46	4.44E-03
sensory organ morphogenesis (GO:0090596)	2.45	2.43E-16
regulation of cellular response to growth factor stimulus (GO:0090287)	2.45	4.50E-08
mitotic cytokinesis (GO:0000281)	2.44	2.60E-02
smoothened signaling pathway (GO:0007224)	2.44	2.59E-02
response to growth factor (GO:0070848)	2.44	1.04E-07
oligodendrocyte differentiation (GO:0048709)	2.44	6.62E-03
hepaticobiliary system development (GO:0061008)	2.44	1.04E-07
nerve development (GO:0021675)	2.44	4.24E-04
regulation of cell shape (GO:0008360)	2.44	1.70E-03
embryo development ending in birth or egg hatching (GO:0009792)	2.44	1.64E-26
embryonic organ development (GO:0048568)	2.44	2.61E-31
heart morphogenesis (GO:0003007)	2.44	6.43E-12
sprouting angiogenesis (GO:0002040)	2.44	9.22E-06
calcium ion transmembrane import into cytosol (GO:0097553)	2.43	9.74E-03
axon extension (GO:0048675)	2.43	9.73E-03
monoatomic anion transport (GO:0006820)	2.43	7.51E-06
regulation of synaptic transmission, glutamatergic (GO:0051966)	2.43	3.77E-02
regulation of ubiquitin-dependent protein catabolic process (GO:2000058)	2.43	3.76E-02
regulation of epithelial cell proliferation (GO:0050678)	2.43	3.76E-02

epithelial cell proliferation (GO:0050673)	2.43	3.75E-02
axonogenesis (GO:0007409)	2.42	5.69E-19
notochord development (GO:0030903)	2.42	3.50E-04
response to steroid hormone (GO:0048545)	2.42	1.44E-02
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0090092)	2.42	5.61E-06
chordate embryonic development (GO:0043009)	2.41	2.70E-25
cellular response to growth factor stimulus (GO:0071363)	2.41	2.11E-07
peripheral nervous system development (GO:0007422)	2.41	2.06E-03
fibroblast growth factor receptor signaling pathway (GO:0008543)	2.40	2.16E-02
regulation of neurogenesis (GO:0050767)	2.40	5.54E-08
regeneration (GO:0031099)	2.40	6.73E-09
transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0007178)	2.39	2.36E-04
embryonic hemopoiesis (GO:0035162)	2.39	2.36E-04
embryonic morphogenesis (GO:0048598)	2.39	8.88E-34
axon development (GO:0061564)	2.39	4.89E-20
kidney development (GO:0001822)	2.38	1.01E-09
pectoral fin development (GO:0033339)	2.38	1.72E-03
cell morphogenesis involved in neuron differentiation (GO:0048667)	2.37	6.01E-19
mesenchyme development (GO:0060485)	2.37	6.83E-11
renal system development (GO:0072001)	2.37	1.87E-09
morphogenesis of an epithelium (GO:0002009)	2.37	1.07E-21
regulation of proteolysis involved in protein catabolic process (GO:1903050)	2.37	1.76E-02
pronephric nephron development (GO:0039019)	2.37	1.76E-02
endosome to lysosome transport (GO:0008333)	2.36	4.58E-02
digestive tract morphogenesis (GO:0048546)	2.36	4.57E-02
regulation of developmental growth (GO:0048638)	2.36	4.86E-05
negative regulation of canonical Wnt signaling pathway (GO:0090090)	2.35	9.76E-03
tissue morphogenesis (GO:0048729)	2.35	2.64E-25
positive regulation of kinase activity (GO:0033674)	2.35	1.99E-04
embryonic pattern specification (GO:0009880)	2.35	3.08E-03
neuron migration (GO:0001764)	2.35	2.59E-02
neural retina development (GO:0003407)	2.35	1.07E-04
determination of heart left/right asymmetry (GO:0061371)	2.34	7.95E-09
positive regulation of Wnt signaling pathway (GO:0030177)	2.34	4.48E-03
positive regulation of transferase activity (GO:0051347)	2.34	1.57E-04
tube development (GO:0035295)	2.33	4.54E-34
positive regulation of phosphorylation (GO:0042327)	2.33	2.11E-07
one-carbon metabolic process (GO:0006730)	2.33	3.75E-02
regulation of extent of cell growth (GO:0061387)	2.33	1.24E-03
camera-type eye morphogenesis (GO:0048593)	2.32	2.65E-08

dorsal/ventral pattern formation (GO:0009953)	2.32	1.96E-06
transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)	2.32	2.44E-10
animal organ morphogenesis (GO:0009887)	2.32	4.50E-36
non-proteinogenic amino acid metabolic process (GO:0170041)	2.32	2.14E-02
inorganic anion transport (GO:0015698)	2.32	3.54E-05
enzyme-linked receptor protein signaling pathway (GO:0007167)	2.32	1.25E-14
plasma membrane bounded cell projection morphogenesis (GO:0120039)	2.31	1.12E-18
negative regulation of developmental process (GO:0051093)	2.31	3.52E-07
cell projection morphogenesis (GO:0048858)	2.31	8.33E-19
mesoderm development (GO:0007498)	2.30	2.64E-04
axis specification (GO:0009798)	2.30	2.58E-03
central nervous system neuron differentiation (GO:0021953)	2.30	4.22E-05
neuron projection morphogenesis (GO:0048812)	2.30	2.16E-18
convergent extension involved in gastrulation (GO:0060027)	2.30	1.41E-03
transforming growth factor beta receptor superfamily signaling pathway (GO:0141091)	2.30	7.44E-04
cardiac muscle cell differentiation (GO:0055007)	2.29	2.62E-02
lymph vessel development (GO:0001945)	2.29	7.12E-03
anterior/posterior pattern specification (GO:0009952)	2.28	6.33E-10
anatomical structure homeostasis (GO:0060249)	2.28	1.10E-03
tube morphogenesis (GO:0035239)	2.28	8.93E-24
angiogenesis (GO:0001525)	2.28	3.57E-11
regulation of cell projection organization (GO:0031344)	2.27	4.08E-09
tissue homeostasis (GO:0001894)	2.27	1.61E-03
retina development in camera-type eye (GO:0060041)	2.27	8.05E-11
columnar/cuboidal epithelial cell differentiation (GO:0002065)	2.27	4.49E-02
neuron fate commitment (GO:0048663)	2.27	3.73E-02
positive regulation of phosphorus metabolic process (GO:0010562)	2.26	2.28E-07
positive regulation of phosphate metabolic process (GO:0045937)	2.26	2.27E-07
Wnt signaling pathway (GO:0016055)	2.26	3.12E-06
negative regulation of cell differentiation (GO:0045596)	2.26	8.80E-05
cell migration involved in gastrulation (GO:0042074)	2.26	2.28E-03
regulation of plasma membrane bounded cell projection organization (GO:0120035)	2.26	7.00E-09
negative regulation of cell development (GO:0010721)	2.25	4.21E-03
growth (GO:0040007)	2.25	2.06E-11
developmental growth (GO:0048589)	2.25	2.04E-11
positive regulation of transcription by RNA polymerase II (GO:0045944)	2.25	8.73E-13
negative regulation of cellular response to growth factor stimulus (GO:0090288)	2.25	1.50E-02
appendage development (GO:0048736)	2.25	1.27E-04

eye morphogenesis (GO:0048592)	2.25	2.87E-09
lysosomal transport (GO:0007041)	2.25	2.85E-02
brain development (GO:0007420)	2.25	2.98E-20
negative regulation of response to external stimulus (GO:0032102)	2.25	1.01E-03
swimming behavior (GO:0036269)	2.24	6.16E-03
regulation of mRNA processing (GO:0050684)	2.24	6.15E-03
regulation of hormone levels (GO:0010817)	2.23	7.97E-04
cell-cell signaling by wnt (GO:0198738)	2.23	4.21E-06
regulation of nervous system development (GO:0051960)	2.23	2.53E-07
stem cell differentiation (GO:0048863)	2.23	1.12E-08
pattern specification process (GO:0007389)	2.22	7.58E-24
transition metal ion transport (GO:0000041)	2.22	8.93E-03
regionalization (GO:0003002)	2.21	5.56E-23
head development (GO:0060322)	2.21	8.86E-20
central nervous system development (GO:0007417)	2.21	7.91E-27
regulation of apoptotic signaling pathway (GO:2001233)	2.20	3.03E-03
epithelium development (GO:0060429)	2.20	2.49E-34
positive regulation of cell migration (GO:0030335)	2.20	4.20E-04
cell fate commitment (GO:0045165)	2.19	3.68E-06
synapse organization (GO:0050808)	2.19	1.35E-04
regulation of cell development (GO:0060284)	2.19	1.31E-09
determination of bilateral symmetry (GO:0009855)	2.18	5.02E-11
specification of symmetry (GO:0009799)	2.18	4.98E-11
RNA catabolic process (GO:0006401)	2.18	2.88E-03
establishment of cell polarity (GO:0030010)	2.18	4.40E-02
fin development (GO:0033333)	2.18	4.67E-04
segmentation (GO:0035282)	2.18	1.94E-04
microtubule cytoskeleton organization involved in mitosis (GO:1902850)	2.18	7.99E-03
cellular response to organic cyclic compound (GO:0071407)	2.18	7.98E-03
determination of left/right symmetry (GO:0007368)	2.18	4.49E-10
embryo development (GO:0009790)	2.18	4.49E-43
cell morphogenesis (GO:0000902)	2.17	6.90E-19
left/right pattern formation (GO:0060972)	2.17	3.52E-10
neuron projection development (GO:0031175)	2.17	7.68E-19
somitogenesis (GO:0001756)	2.16	1.29E-03
glial cell differentiation (GO:0010001)	2.16	1.09E-04
cellular response to external stimulus (GO:0071496)	2.15	2.69E-02
heart development (GO:0007507)	2.15	1.66E-19
convergent extension (GO:0060026)	2.15	3.32E-05
blood vessel morphogenesis (GO:0048514)	2.15	7.62E-12
nephron development (GO:0072006)	2.15	2.35E-03

vascular process in circulatory system (GO:0003018)	2.15	9.24E-03
neuron differentiation (GO:0030182)	2.15	4.26E-30
cell differentiation in spinal cord (GO:0021515)	2.14	2.14E-02
regulation of growth (GO:0040008)	2.14	1.17E-04
regulation of cellular component size (GO:0032535)	2.14	8.02E-07
neuron development (GO:0048666)	2.14	1.22E-23
tissue development (GO:0009888)	2.13	4.30E-51
positive regulation of RNA biosynthetic process (GO:1902680)	2.13	9.60E-16
positive regulation of DNA-templated transcription (GO:0045893)	2.13	9.51E-16
muscle tissue development (GO:0060537)	2.13	5.67E-07
glial cell development (GO:0021782)	2.13	2.58E-03
cellular response to extracellular stimulus (GO:0031668)	2.13	3.79E-02
generation of neurons (GO:0048699)	2.13	2.42E-30
cell fate specification (GO:0001708)	2.12	2.02E-03
cellular anatomical entity morphogenesis (GO:0032989)	2.12	7.90E-18
synaptic vesicle cycle (GO:0099504)	2.12	4.78E-03
vesicle-mediated transport in synapse (GO:0099003)	2.12	4.77E-03
regulation of lipid metabolic process (GO:0019216)	2.12	3.04E-02
signal release (GO:0023061)	2.12	1.55E-02
gliogenesis (GO:0042063)	2.12	7.95E-05
blood vessel development (GO:0001568)	2.12	6.45E-13
mRNA catabolic process (GO:0006402)	2.11	1.17E-02
circulatory system development (GO:0072359)	2.11	7.69E-31
nervous system development (GO:0007399)	2.10	3.93E-54
neuron projection extension (GO:1990138)	2.09	4.32E-02
neurogenesis (GO:0022008)	2.09	1.43E-32
anatomical structure formation involved in morphogenesis (GO:0048646)	2.08	2.61E-24
regulation of mRNA splicing, via spliceosome (GO:0048024)	2.08	3.39E-02
vasculature development (GO:0001944)	2.08	1.47E-14
vacuolar transport (GO:0007034)	2.08	2.56E-03
regulation of anatomical structure size (GO:0090066)	2.08	8.27E-08
kidney epithelium development (GO:0072073)	2.07	3.05E-02
nuclear-transcribed mRNA catabolic process (GO:0000956)	2.07	3.04E-02
ceramide metabolic process (GO:0006672)	2.06	1.78E-02
protein lipidation (GO:0006497)	2.06	1.37E-02
positive regulation of cell differentiation (GO:0045597)	2.05	5.53E-05
positive regulation of cell motility (GO:2000147)	2.05	1.46E-03
regulation of locomotion (GO:0040012)	2.05	1.23E-07
regulation of cell migration (GO:0030334)	2.05	1.57E-06
positive regulation of developmental process (GO:0051094)	2.05	1.78E-07
glomerulus development (GO:0032835)	2.04	3.28E-02

digestive tract development (GO:0048565)	2.04	4.43E-03
receptor-mediated endocytosis (GO:0006898)	2.04	2.51E-02
negative regulation of locomotion (GO:0040013)	2.04	1.97E-02
gastrulation (GO:0007369)	2.04	1.05E-06
regulation of Wnt signaling pathway (GO:0030111)	2.03	1.09E-04
regulation of cell differentiation (GO:0045595)	2.03	1.97E-13
anatomical structure morphogenesis (GO:0009653)	2.03	3.98E-65
cell migration (GO:0016477)	2.03	8.50E-17
negative regulation of intracellular signal transduction (GO:1902532)	2.03	3.74E-04
positive regulation of RNA metabolic process (GO:0051254)	2.03	4.60E-15
positive regulation of protein phosphorylation (GO:0001934)	2.02	3.66E-03
cellular response to endogenous stimulus (GO:0071495)	2.02	3.37E-10
regulation of canonical Wnt signaling pathway (GO:0060828)	2.02	2.25E-03
positive regulation of locomotion (GO:0040017)	2.02	1.42E-03
regulation of anatomical structure morphogenesis (GO:0022603)	2.02	1.11E-08
response to hypoxia (GO:0001666)	2.01	3.46E-02
lipoprotein biosynthetic process (GO:0042158)	2.01	1.68E-02
negative regulation of Wnt signaling pathway (GO:0030178)	2.01	1.68E-02
regulation of cell growth (GO:0001558)	2.01	6.70E-03
response to organic cyclic compound (GO:0014070)	2.00	2.84E-03
positive regulation of nucleobase-containing compound metabolic process (GO:0045935)	2.00	2.04E-15
sensory organ development (GO:0007423)	2.00	2.84E-19
positive regulation of cellular metabolic process (GO:0031325)	1.99	1.20E-23
cell growth (GO:0016049)	1.99	4.84E-02
spinal cord development (GO:0021510)	1.99	2.11E-02
response to decreased oxygen levels (GO:0036293)	1.99	3.77E-02
cell motility (GO:0048870)	1.98	8.78E-17
negative regulation of response to stimulus (GO:0048585)	1.98	2.42E-13
positive regulation of cell population proliferation (GO:0008284)	1.97	4.55E-04
pancreas development (GO:0031016)	1.97	5.95E-04
photoreceptor cell differentiation (GO:0046530)	1.97	6.27E-03
germ cell development (GO:0007281)	1.97	1.01E-02
secretion (GO:0046903)	1.97	1.16E-04
response to endogenous stimulus (GO:0009719)	1.96	1.89E-10
negative regulation of transcription by RNA polymerase II (GO:0000122)	1.96	3.72E-06
response to oxygen levels (GO:0070482)	1.96	3.26E-02
positive regulation of cellular component organization (GO:0051130)	1.96	2.25E-05
regulation of actin cytoskeleton organization (GO:0032956)	1.95	1.63E-04
regulation of multicellular organismal development (GO:2000026)	1.95	1.15E-10
chromosome segregation (GO:0007059)	1.94	1.95E-03
calcium ion transmembrane transport (GO:0070588)	1.94	1.95E-03

negative regulation of cellular component organization (GO:0051129)	1.94	2.25E-04
positive regulation of macromolecule biosynthetic process (GO:0010557)	1.94	4.10E-15
regulation of actin filament organization (GO:0110053)	1.94	2.92E-03
neurotransmitter transport (GO:0006836)	1.94	3.86E-02
regulation of actin filament-based process (GO:0032970)	1.93	1.93E-04
canonical Wnt signaling pathway (GO:0060070)	1.93	4.58E-02
monosaccharide metabolic process (GO:0005996)	1.93	4.57E-02
regulation of transferase activity (GO:0051338)	1.93	2.45E-04
membrane lipid metabolic process (GO:0006643)	1.93	2.05E-03
negative regulation of signal transduction (GO:0009968)	1.93	5.26E-10
positive regulation of cellular biosynthetic process (GO:0031328)	1.93	5.00E-15
negative regulation of cell communication (GO:0010648)	1.93	3.57E-10
negative regulation of signaling (GO:0023057)	1.93	3.54E-10
sodium ion transport (GO:0006814)	1.92	4.03E-03
epithelial cell differentiation (GO:0030855)	1.92	1.05E-05
positive regulation of metabolic process (GO:0009893)	1.92	1.91E-23
positive regulation of biosynthetic process (GO:0009891)	1.92	5.98E-15
somite development (GO:0061053)	1.92	1.74E-03
regulation of developmental process (GO:0050793)	1.91	1.50E-19
positive regulation of nitrogen compound metabolic process (GO:0051173)	1.91	1.58E-18
secretion by cell (GO:0032940)	1.91	6.21E-04
system development (GO:0048731)	1.90	4.80E-80
regulation of secretion (GO:0051046)	1.90	4.99E-03
regulation of membrane potential (GO:0042391)	1.89	8.71E-04
regulation of cell motility (GO:2000145)	1.89	2.45E-05
phosphatidylinositol biosynthetic process (GO:0006661)	1.89	4.42E-02
cell-cell signaling (GO:0007267)	1.89	4.71E-11
establishment of organelle localization (GO:0051656)	1.89	1.03E-03
developmental growth involved in morphogenesis (GO:0060560)	1.89	9.89E-03
establishment or maintenance of cell polarity (GO:0007163)	1.88	1.26E-02
regulation of phosphorylation (GO:0042325)	1.88	1.25E-05
sensory system development (GO:0048880)	1.88	1.47E-13
plasma membrane bounded cell projection organization (GO:0120036)	1.88	4.57E-18
regulation of cell population proliferation (GO:0042127)	1.88	4.10E-06
wound healing (GO:0042060)	1.87	3.72E-02
multicellular organism development (GO:0007275)	1.87	3.11E-91
cellular response to organonitrogen compound (GO:0071417)	1.87	2.63E-02
muscle structure development (GO:0061061)	1.87	2.30E-07
regulation of RNA splicing (GO:0043484)	1.87	2.63E-02
positive regulation of proteolysis (GO:0045862)	1.87	4.72E-02
regulation of actin filament polymerization (GO:0030833)	1.86	3.33E-02

positive regulation of macromolecule metabolic process (GO:0010604)	1.86	1.44E-18
actin filament organization (GO:0007015)	1.86	5.80E-05
cell development (GO:0048468)	1.86	2.27E-37
endocytosis (GO:0006897)	1.86	1.97E-05
cell projection organization (GO:0030030)	1.86	7.95E-18
muscle cell differentiation (GO:0042692)	1.86	3.75E-05
sphingolipid metabolic process (GO:0006665)	1.86	2.22E-02
cellular developmental process (GO:0048869)	1.86	5.80E-53
regulation of supramolecular fiber organization (GO:1902903)	1.86	2.42E-03
cell differentiation (GO:0030154)	1.86	7.47E-53
blood circulation (GO:0008015)	1.85	2.25E-04
regulation of kinase activity (GO:0043549)	1.85	1.57E-03
animal organ development (GO:0048513)	1.85	9.61E-53
developmental process involved in reproduction (GO:0003006)	1.85	1.96E-03
positive regulation of protein modification process (GO:0031401)	1.84	9.23E-03
cell surface receptor signaling pathway (GO:0007166)	1.83	4.72E-20
regulation of actin polymerization or depolymerization (GO:0008064)	1.83	3.75E-02
mitotic cell cycle process (GO:1903047)	1.82	1.59E-04
camera-type eye development (GO:0043010)	1.82	1.39E-08
membrane lipid biosynthetic process (GO:0046467)	1.82	3.44E-02
actin cytoskeleton organization (GO:0030036)	1.82	3.06E-08
multicellular organismal-level homeostasis (GO:0048871)	1.82	1.84E-04
glycerophospholipid biosynthetic process (GO:0046474)	1.82	2.24E-02
actin filament-based process (GO:0030029)	1.82	1.71E-08
regulation of secretion by cell (GO:1903530)	1.81	1.47E-02
cytoskeleton-dependent intracellular transport (GO:0030705)	1.81	4.01E-02
anatomical structure development (GO:0048856)	1.81	9.90E-107
small GTPase-mediated signal transduction (GO:0007264)	1.81	3.45E-03
export from cell (GO:0140352)	1.81	1.61E-03
circulatory system process (GO:0003013)	1.80	4.50E-04
developmental process (GO:0032502)	1.80	1.59E-108
digestive system development (GO:0055123)	1.80	9.07E-03
eye development (GO:0001654)	1.79	3.44E-09
visual system development (GO:0150063)	1.79	3.42E-09
regulation of actin filament length (GO:0030832)	1.79	4.36E-02
phospholipid biosynthetic process (GO:0008654)	1.79	1.14E-02
regulation of protein polymerization (GO:0032271)	1.79	3.75E-02
cellular response to nitrogen compound (GO:1901699)	1.79	2.23E-02
cell division (GO:0051301)	1.78	2.07E-04
regulation of response to external stimulus (GO:0032101)	1.78	7.94E-05
muscle organ development (GO:0007517)	1.77	2.92E-03
calcium ion transport (GO:0006816)	1.77	1.01E-02

regulation of cell communication (GO:0010646)	1.76	6.90E-22
response to wounding (GO:0009611)	1.75	2.10E-02
regulation of signal transduction (GO:0009966)	1.75	5.25E-19
regulation of signaling (GO:0023051)	1.75	2.90E-21
organelle localization (GO:0051640)	1.75	6.49E-04
carbohydrate derivative catabolic process (GO:1901136)	1.75	2.29E-02
regulation of phosphate metabolic process (GO:0019220)	1.75	8.28E-05
regulation of phosphorus metabolic process (GO:0051174)	1.75	8.26E-05
positive regulation of multicellular organismal process (GO:0051240)	1.74	1.45E-04
behavior (GO:0007610)	1.74	1.31E-03
nucleobase-containing compound catabolic process (GO:0034655)	1.74	4.69E-03
multicellular organismal process (GO:0032501)	1.74	1.06E-88
mRNA metabolic process (GO:0016071)	1.74	1.08E-05
regulation of biological quality (GO:0065008)	1.74	2.51E-16
cellular homeostasis (GO:0019725)	1.74	3.23E-04
positive regulation of catabolic process (GO:0009896)	1.74	3.28E-02
regulation of response to stimulus (GO:0048583)	1.74	6.40E-24
aromatic compound catabolic process (GO:0019439)	1.73	1.11E-03
cell cycle process (GO:0022402)	1.72	4.21E-06
homeostatic process (GO:0042592)	1.72	2.55E-09
cellular response to hormone stimulus (GO:0032870)	1.72	8.68E-03
heterocycle catabolic process (GO:0046700)	1.72	2.53E-03
cellular nitrogen compound catabolic process (GO:0044270)	1.72	2.53E-03
modulation of chemical synaptic transmission (GO:0050804)	1.71	2.31E-02
regulation of trans-synaptic signaling (GO:0099177)	1.71	2.31E-02
cell junction organization (GO:0034330)	1.71	2.30E-03
synaptic signaling (GO:0099536)	1.71	3.08E-04
negative regulation of RNA biosynthetic process (GO:1902679)	1.71	3.72E-05
negative regulation of DNA-templated transcription (GO:0045892)	1.71	3.70E-05
cell-cell adhesion (GO:0098609)	1.71	1.11E-04
phosphatidylinositol metabolic process (GO:0046488)	1.71	3.87E-02
regulation of GTPase activity (GO:0043087)	1.71	1.59E-02
RNA splicing (GO:0008380)	1.70	1.76E-03
regulation of cellular component biogenesis (GO:0044087)	1.70	7.78E-04
multicellular organism reproduction (GO:0032504)	1.70	1.21E-02
chemical homeostasis (GO:0048878)	1.70	4.13E-05
import into cell (GO:0098657)	1.69	2.40E-05
apoptotic process (GO:0006915)	1.69	1.92E-03
regulation of multicellular organismal process (GO:0051239)	1.69	7.03E-11
regulation of transcription by RNA polymerase II (GO:0006357)	1.69	2.03E-24
cell cycle (GO:0007049)	1.68	3.83E-08
organic cyclic compound catabolic process (GO:1901361)	1.68	2.02E-03

negative regulation of RNA metabolic process (GO:0051253)	1.68	3.31E-05
organic acid catabolic process (GO:0016054)	1.68	2.63E-02
carboxylic acid catabolic process (GO:0046395)	1.68	2.63E-02
regulation of cellular component organization (GO:0051128)	1.68	6.34E-10
metal ion transport (GO:0030001)	1.67	4.08E-06
positive regulation of cellular process (GO:0048522)	1.67	4.19E-25
cell adhesion (GO:0007155)	1.67	3.82E-07
inorganic ion transmembrane transport (GO:0098660)	1.67	2.07E-06
hemopoiesis (GO:0030097)	1.67	1.13E-05
microtubule-based movement (GO:0007018)	1.67	6.52E-03
chemical synaptic transmission (GO:0007268)	1.67	1.23E-03
anterograde trans-synaptic signaling (GO:0098916)	1.67	1.23E-03
regulation of cytoskeleton organization (GO:0051493)	1.67	2.14E-03
mRNA processing (GO:0006397)	1.67	6.42E-04
cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)	1.67	1.49E-02
regulation of system process (GO:0044057)	1.66	4.01E-02
multicellular organismal reproductive process (GO:0048609)	1.66	2.75E-02
negative regulation of multicellular organismal process (GO:0051241)	1.66	9.49E-03
response to oxygen-containing compound (GO:1901700)	1.66	2.93E-04
striated muscle cell differentiation (GO:0051146)	1.66	1.02E-02
monoatomic ion transport (GO:0006811)	1.65	4.61E-09
mitotic cell cycle (GO:0000278)	1.65	1.22E-03
trans-synaptic signaling (GO:0099537)	1.64	1.83E-03
response to hormone (GO:0009725)	1.64	6.65E-03
regulation of proteolysis (GO:0030162)	1.64	3.56E-02
monoatomic ion transmembrane transport (GO:0034220)	1.63	4.73E-06
regulation of cell cycle process (GO:0010564)	1.63	1.24E-02
DNA-templated transcription (GO:0006351)	1.63	2.27E-02
regulation of mRNA metabolic process (GO:1903311)	1.63	4.86E-02
cell death (GO:0008219)	1.63	2.91E-03
programmed cell death (GO:0012501)	1.63	2.91E-03
regulation of catabolic process (GO:0009894)	1.63	5.25E-03
regulation of cellular metabolic process (GO:0031323)	1.62	6.24E-40
lipid biosynthetic process (GO:0008610)	1.62	4.55E-04
regulation of catalytic activity (GO:0050790)	1.62	2.24E-05
regulation of metabolic process (GO:0019222)	1.62	1.64E-42
regulation of RNA metabolic process (GO:0051252)	1.61	1.35E-27
positive regulation of biological process (GO:0048518)	1.61	8.79E-24
positive regulation of protein metabolic process (GO:0051247)	1.61	4.65E-03
regulation of protein phosphorylation (GO:0001932)	1.61	1.85E-02
regulation of DNA-templated transcription (GO:0006355)	1.61	1.32E-24

organophosphate biosynthetic process (GO:0090407)	1.61	9.22E-04
regulation of RNA biosynthetic process (GO:2001141)	1.61	1.87E-24
dephosphorylation (GO:0016311)	1.60	1.69E-02
regulation of gene expression (GO:0010468)	1.60	1.44E-30
positive regulation of catalytic activity (GO:0043085)	1.60	2.49E-03
regulation of biosynthetic process (GO:0009889)	1.60	1.50E-31
regulation of nucleobase-containing compound metabolic process (GO:0019219)	1.60	2.26E-27
response to lipid (GO:0033993)	1.60	3.68E-02
glycerophospholipid metabolic process (GO:0006650)	1.60	2.80E-02
regulation of macromolecule biosynthetic process (GO:0010556)	1.60	8.68E-31
negative regulation of nucleobase-containing compound metabolic process (GO:0045934)	1.60	1.77E-04
carbohydrate metabolic process (GO:0005975)	1.59	4.90E-03
regulation of cellular biosynthetic process (GO:0031326)	1.59	7.85E-31
intracellular signaling cassette (GO:0141124)	1.59	9.80E-04
monoatomic cation transport (GO:0006812)	1.59	1.80E-05
sulfur compound metabolic process (GO:0006790)	1.59	2.89E-02
regulation of molecular function (GO:0065009)	1.59	1.54E-06
RNA biosynthetic process (GO:0032774)	1.59	3.39E-02
negative regulation of macromolecule biosynthetic process (GO:0010558)	1.59	7.13E-06
regulation of primary metabolic process (GO:0080090)	1.59	8.07E-34
regulation of macromolecule metabolic process (GO:0060255)	1.59	1.49E-35
negative regulation of biosynthetic process (GO:0009890)	1.58	6.30E-06
regulation of response to stress (GO:0080134)	1.58	7.05E-03
monoatomic cation transmembrane transport (GO:0098655)	1.58	2.10E-04
intracellular chemical homeostasis (GO:0055082)	1.58	2.37E-02
muscle cell development (GO:0055001)	1.58	2.88E-02
negative regulation of cellular biosynthetic process (GO:0031327)	1.58	7.81E-06
positive regulation of molecular function (GO:0044093)	1.58	1.29E-03
plasma membrane bounded cell projection assembly (GO:0120031)	1.58	3.24E-03
regulation of nitrogen compound metabolic process (GO:0051171)	1.58	4.29E-32
inorganic cation transmembrane transport (GO:0098662)	1.57	2.67E-04
transmembrane transport (GO:0055085)	1.57	1.46E-09
cytoskeleton organization (GO:0007010)	1.57	5.32E-08
small molecule catabolic process (GO:0044282)	1.57	3.85E-02
negative regulation of cellular process (GO:0048523)	1.57	1.53E-17
inorganic ion homeostasis (GO:0098771)	1.57	3.58E-02
phospholipid metabolic process (GO:0006644)	1.56	1.32E-02
cellular component organization (GO:0016043)	1.56	6.09E-35
cell projection assembly (GO:0030031)	1.56	3.57E-03
endomembrane system organization (GO:0010256)	1.56	2.16E-02

negative regulation of biological process (GO:0048519)	1.55	6.56E-18
cellular response to oxygen-containing compound (GO:1901701)	1.55	2.63E-02
regulation of protein modification process (GO:0031399)	1.55	1.82E-02
cilium assembly (GO:0060271)	1.55	8.67E-03
cellular component organization or biogenesis (GO:0071840)	1.54	6.23E-35
response to nitrogen compound (GO:1901698)	1.53	3.65E-02
reproductive process (GO:0022414)	1.53	1.12E-02
positive regulation of cell communication (GO:0010647)	1.53	7.64E-04
positive regulation of signaling (GO:0023056)	1.53	7.62E-04
transport (GO:0006810)	1.53	1.57E-22
response to abiotic stimulus (GO:0009628)	1.52	5.85E-03
RNA metabolic process (GO:0016070)	1.52	2.54E-07
reproduction (GO:0000003)	1.52	1.52E-02
establishment of localization (GO:0051234)	1.51	1.15E-22
establishment of localization in cell (GO:0051649)	1.51	3.68E-07
regulation of transport (GO:0051049)	1.51	1.09E-03
localization (GO:0051179)	1.51	1.43E-24
regulation of intracellular signal transduction (GO:1902531)	1.51	2.86E-04
regulation of localization (GO:0032879)	1.51	2.90E-04

Table S3 – Gene ontology terms enriched in a list of GREAT-associated genes that were significantly downregulated peaks in EdgeR analysis of acinar cells ATAC-seq. Terms with fold enrichment <1.5 are shown.

Biological process	Fold Enrichment	FDR
negative regulation of reactive oxygen species biosynthetic process (GO:1903427)	9.10	2.25E-02
late distal convoluted tubule development (GO:0072068)	9.10	2.25E-02
distal convoluted tubule development (GO:0072025)	9.10	2.25E-02
octopamine biosynthetic process (GO:0006589)	9.10	2.24E-02
motor behavior (GO:0061744)	9.10	3.48E-03
pore complex assembly (GO:0046931)	9.10	2.24E-02
clathrin-dependent synaptic vesicle endocytosis (GO:0150007)	9.10	2.23E-02
norepinephrine biosynthetic process (GO:0042421)	9.10	2.23E-02
dopamine catabolic process (GO:0042420)	9.10	2.22E-02
positive regulation of mitochondrial calcium ion concentration (GO:0051561)	9.10	2.22E-02
octopamine metabolic process (GO:0046333)	9.10	2.22E-02
nuclear pore complex assembly (GO:0051292)	9.10	2.21E-02
response to selenium ion (GO:0010269)	9.10	2.21E-02
regulation of motor neuron axon guidance (GO:1905812)	9.10	2.20E-02

negative regulation of macrophage differentiation (GO:0045650)	9.10	2.20E-02
regulation of reactive oxygen species biosynthetic process (GO:1903426)	7.28	1.26E-02
macrophage activation involved in immune response (GO:0002281)	7.28	1.26E-02
skeletal muscle satellite cell migration (GO:1902766)	7.28	1.26E-02
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification (GO:2000053)	7.28	1.26E-02
neuron cell-cell adhesion (GO:0007158)	7.08	1.96E-04
negative regulation of myeloid leukocyte differentiation (GO:0002762)	6.83	1.15E-03
epiboly involved in wound healing (GO:0090505)	6.50	6.10E-03
wound healing, spreading of cells (GO:0044319)	6.50	6.09E-03
presynaptic membrane assembly (GO:0097105)	6.50	6.07E-03
postsynaptic membrane assembly (GO:0097104)	6.50	6.06E-03
presynaptic membrane organization (GO:0097090)	6.50	6.04E-03
postsynapse assembly (GO:0099068)	6.50	6.03E-03
phenol-containing compound catabolic process (GO:0019336)	6.50	6.01E-03
L-proline biosynthetic process (GO:0055129)	6.07	2.83E-02
proline biosynthetic process (GO:0006561)	6.07	2.83E-02
sphingolipid mediated signaling pathway (GO:0090520)	6.07	2.82E-02
catecholamine catabolic process (GO:0042424)	6.07	2.82E-02
positive regulation of cell morphogenesis (GO:0010770)	6.07	2.81E-02
swim bladder inflation (GO:0048798)	6.07	2.81E-02
swim bladder maturation (GO:0048796)	6.07	2.80E-02
regulation of retinoic acid biosynthetic process (GO:1900052)	6.07	2.80E-02
regulation of isoprenoid metabolic process (GO:0019747)	6.07	2.79E-02
catechol-containing compound catabolic process (GO:0019614)	6.07	2.79E-02
skin epidermis development (GO:0098773)	6.07	2.78E-02
amine catabolic process (GO:0009310)	5.79	1.19E-03
cellular biogenic amine catabolic process (GO:0042402)	5.79	1.19E-03
heart rudiment development (GO:0003313)	5.79	1.19E-03
negative regulation of leukocyte differentiation (GO:1902106)	5.79	1.18E-03
negative regulation of hemopoiesis (GO:1903707)	5.79	1.18E-03
head morphogenesis (GO:0060323)	5.79	1.18E-03
distal tubule development (GO:0072017)	5.69	1.27E-02
motor neuron migration (GO:0097475)	5.69	1.27E-02
kynurenine metabolic process (GO:0070189)	5.69	1.26E-02
retinoic acid receptor signaling pathway (GO:0048384)	5.54	5.93E-07
establishment of tissue polarity (GO:0007164)	5.46	5.58E-03
response to hydrogen peroxide (GO:0042542)	5.46	5.56E-03
regulation of axon guidance (GO:1902667)	5.46	5.55E-03
iron coordination entity transport (GO:1901678)	5.46	5.54E-03
negative regulation of myeloid cell differentiation (GO:0045638)	5.31	2.31E-03
establishment of planar polarity (GO:0001736)	5.06	2.26E-02

heme transport (GO:0015886)	5.06	2.26E-02
regulation of cell-substrate junction organization (GO:0150116)	5.06	2.25E-02
regulation of focal adhesion assembly (GO:0051893)	5.06	2.25E-02
regulation of cell-substrate junction assembly (GO:0090109)	5.06	2.25E-02
body morphogenesis (GO:0010171)	5.06	1.72E-04
face morphogenesis (GO:0060325)	5.06	2.24E-02
benzene-containing compound metabolic process (GO:0042537)	4.97	1.00E-02
face development (GO:0060324)	4.97	9.98E-03
postsynaptic membrane organization (GO:0001941)	4.85	1.75E-03
stem cell population maintenance (GO:0019827)	4.85	1.74E-03
maintenance of cell number (GO:0098727)	4.82	7.48E-04
anterior/posterior axon guidance (GO:0033564)	4.55	1.62E-02
regulation of cell-matrix adhesion (GO:0001952)	4.55	3.60E-02
proximal tubule development (GO:0072014)	4.55	1.62E-02
urogenital system development (GO:0001655)	4.55	6.98E-03
regulation of epithelial to mesenchymal transition (GO:0010717)	4.55	6.96E-03
presynapse organization (GO:0099172)	4.55	1.61E-02
presynapse assembly (GO:0099054)	4.55	1.61E-02
pronephric nephron tubule development (GO:0039020)	4.28	4.74E-03
dorsal aorta development (GO:0035907)	4.25	1.09E-02
regulation of T cell differentiation (GO:0045580)	4.20	2.41E-02
regulation of nodal signaling pathway (GO:1900107)	4.20	2.41E-02
mesendoderm development (GO:0048382)	4.20	2.40E-02
regulation of glial cell differentiation (GO:0045685)	4.20	2.40E-02
dorsal/ventral axis specification (GO:0009950)	4.14	1.28E-03
activin receptor signaling pathway (GO:0032924)	4.05	7.07E-03
tRNA wobble base modification (GO:0002097)	4.05	7.06E-03
animal organ regeneration (GO:0031100)	4.01	8.45E-04
tRNA wobble uridine modification (GO:0002098)	3.98	1.61E-02
formation of anatomical boundary (GO:0048859)	3.98	1.61E-02
rhombomere formation (GO:0021594)	3.90	3.49E-02
phosphate ion transport (GO:0006817)	3.90	4.58E-03
rhombomere morphogenesis (GO:0021593)	3.83	1.05E-02
regulation of lymphocyte differentiation (GO:0045619)	3.83	1.04E-02
hematopoietic or lymphoid organ development (GO:0048534)	3.79	2.81E-03
postsynapse organization (GO:0099173)	3.79	2.80E-03
regulation of activin receptor signaling pathway (GO:0032925)	3.75	2.24E-02
synaptic vesicle recycling (GO:0036465)	3.72	1.38E-05
presynaptic endocytosis (GO:0140238)	3.69	1.29E-04
synaptic vesicle endocytosis (GO:0048488)	3.69	1.28E-04
intracellular receptor signaling pathway (GO:0030522)	3.68	3.33E-05
nephron tubule development (GO:0072080)	3.64	4.08E-03

immune system development (GO:0002520)	3.64	2.99E-04
cell volume homeostasis (GO:0006884)	3.64	4.92E-02
monoatomic anion homeostasis (GO:0055081)	3.64	1.44E-02
artery morphogenesis (GO:0048844)	3.64	1.44E-02
pancreas regeneration (GO:1990798)	3.64	1.44E-02
thymus development (GO:0048538)	3.64	1.43E-02
positive regulation of axon extension (GO:0045773)	3.64	4.92E-02
midbrain development (GO:0030901)	3.59	1.84E-04
cellular response to oxygen levels (GO:0071453)	3.54	3.06E-02
mitochondrion localization (GO:0051646)	3.54	3.06E-02
aorta development (GO:0035904)	3.54	3.05E-02
regulation of leukocyte differentiation (GO:1902105)	3.52	6.79E-05
renal tubule development (GO:0061326)	3.50	5.65E-03
regulation of myeloid leukocyte differentiation (GO:0002761)	3.50	5.64E-03
rhombomere development (GO:0021546)	3.47	1.99E-02
epithalamus development (GO:0021538)	3.47	1.99E-02
endothelial cell proliferation (GO:0001935)	3.37	7.57E-03
peroxisomal membrane transport (GO:0015919)	3.35	4.12E-02
morphogenesis of a polarized epithelium (GO:0001738)	3.35	4.12E-02
regulation of endothelial cell migration (GO:0010594)	3.35	4.11E-02
peroxisomal transport (GO:0043574)	3.35	4.10E-02
ventral spinal cord interneuron differentiation (GO:0021514)	3.28	1.63E-02
neuron fate specification (GO:0048665)	3.23	6.23E-03
brain morphogenesis (GO:0048854)	3.21	3.85E-03
establishment of spindle orientation (GO:0051294)	3.17	3.32E-02
neuron fate commitment (GO:0048663)	3.16	3.20E-04
establishment of mitotic spindle localization (GO:0040001)	3.15	2.14E-02
positive regulation of neuron projection development (GO:0010976)	3.14	1.32E-02
notochord morphogenesis (GO:0048570)	3.14	1.31E-02
extracellular transport (GO:0006858)	3.12	5.11E-03
epithelial cilium movement involved in extracellular fluid movement (GO:0003351)	3.12	5.09E-03
artery development (GO:0060840)	3.11	3.03E-03
spindle localization (GO:0051653)	3.03	1.06E-02
establishment of spindle localization (GO:0051293)	3.03	1.06E-02
response to yeast (GO:0001878)	3.03	4.27E-02
inner ear receptor cell differentiation (GO:0060113)	3.03	9.09E-04
inner ear auditory receptor cell differentiation (GO:0042491)	3.03	2.68E-02
hypothalamus development (GO:0021854)	3.03	2.67E-02
limbic system development (GO:0021761)	2.93	3.33E-02
tissue migration (GO:0090130)	2.92	9.11E-04
inner ear receptor cell development (GO:0060119)	2.89	6.44E-03

diencephalon development (GO:0021536)	2.86	1.21E-05
semicircular canal development (GO:0060872)	2.83	4.16E-02
regulation of mRNA processing (GO:0050684)	2.82	1.55E-04
regulation of neuron projection development (GO:0010975)	2.81	2.41E-10
regulation of alternative mRNA splicing, via spliceosome (GO:0000381)	2.80	1.29E-02
neural crest cell migration (GO:0001755)	2.78	1.39E-06
epithelial cell proliferation (GO:0050673)	2.78	2.09E-02
regulation of cell morphogenesis (GO:0022604)	2.77	8.91E-05
establishment or maintenance of epithelial cell apical/basal polarity (GO:0045197)	2.76	3.20E-02
negative regulation of axon extension (GO:0030517)	2.73	1.60E-02
mesenchymal cell migration (GO:0090497)	2.73	2.25E-06
regulation of hemopoiesis (GO:1903706)	2.72	6.51E-04
ventral spinal cord development (GO:0021517)	2.71	2.23E-03
positive regulation of cell development (GO:0010720)	2.71	2.32E-05
pronephric nephron development (GO:0039019)	2.71	7.58E-03
neural crest cell development (GO:0014032)	2.71	2.52E-08
stem cell development (GO:0048864)	2.71	2.50E-08
regulation of chemotaxis (GO:0050920)	2.71	3.02E-04
regulation of peptidyl-tyrosine phosphorylation (GO:0050730)	2.71	2.47E-02
inorganic anion transmembrane transport (GO:0098661)	2.71	3.01E-04
neural crest cell differentiation (GO:0014033)	2.70	2.88E-09
hormone-mediated signaling pathway (GO:0009755)	2.70	3.64E-03
positive regulation of neurogenesis (GO:0050769)	2.69	4.93E-04
carbohydrate transport (GO:0008643)	2.69	1.23E-02
regulation of axonogenesis (GO:0050770)	2.68	2.90E-05
regulation of axon extension (GO:0030516)	2.68	8.20E-04
positive regulation of axonogenesis (GO:0050772)	2.68	3.93E-02
chloride transmembrane transport (GO:1902476)	2.67	2.74E-03
kidney epithelium development (GO:0072073)	2.64	9.93E-04
bone morphogenesis (GO:0060349)	2.64	2.99E-02
central nervous system neuron differentiation (GO:0021953)	2.63	2.46E-06
mesenchymal cell differentiation (GO:0048762)	2.62	1.73E-09
neural tube patterning (GO:0021532)	2.60	5.52E-03
regulation of cell size (GO:0008361)	2.60	2.49E-05
cell differentiation in spinal cord (GO:0021515)	2.60	2.55E-03
establishment or maintenance of apical/basal cell polarity (GO:0035088)	2.60	4.74E-02
skeletal system morphogenesis (GO:0048705)	2.60	6.06E-11
establishment or maintenance of bipolar cell polarity (GO:0061245)	2.60	4.73E-02
negative regulation of axonogenesis (GO:0050771)	2.60	2.30E-02
synapse assembly (GO:0007416)	2.60	2.30E-02
negative regulation of cell development (GO:0010721)	2.58	9.09E-04

positive regulation of transcription by RNA polymerase II (GO:0045944)	2.58	1.56E-16
synapse organization (GO:0050808)	2.58	2.78E-06
mitotic cytokinesis (GO:0000281)	2.57	3.60E-02
regulation of myeloid cell differentiation (GO:0045637)	2.57	3.59E-02
ear development (GO:0043583)	2.57	4.12E-09
nephron epithelium development (GO:0072009)	2.56	6.55E-03
regulation of extent of cell growth (GO:0061387)	2.55	5.14E-04
epithelial cell migration (GO:0010631)	2.54	2.78E-02
regulation of mRNA splicing, via spliceosome (GO:0048024)	2.54	5.06E-03
response to decreased oxygen levels (GO:0036293)	2.53	1.75E-03
negative regulation of cell projection organization (GO:0031345)	2.53	1.05E-02
negative regulation of canonical Wnt signaling pathway (GO:0090090)	2.53	1.05E-02
neural tube development (GO:0021915)	2.53	4.67E-05
epithelium migration (GO:0090132)	2.52	2.17E-02
inorganic anion transport (GO:0015698)	2.50	1.89E-05
embryonic cranial skeleton morphogenesis (GO:0048701)	2.50	1.06E-07
negative regulation of neuron projection development (GO:0010977)	2.50	1.62E-02
negative regulation of cell growth (GO:0030308)	2.50	1.62E-02
establishment of cell polarity (GO:0030010)	2.50	1.62E-02
inner ear development (GO:0048839)	2.50	2.60E-08
retinal ganglion cell axon guidance (GO:0031290)	2.48	3.26E-02
response to oxygen levels (GO:0070482)	2.48	1.60E-03
embryonic skeletal system morphogenesis (GO:0048704)	2.48	5.07E-08
ameboidal-type cell migration (GO:0001667)	2.48	3.51E-10
regulation of cell development (GO:0060284)	2.47	5.62E-12
mechanoreceptor differentiation (GO:0042490)	2.47	3.38E-03
response to hypoxia (GO:0001666)	2.44	4.08E-03
positive regulation of nervous system development (GO:0051962)	2.43	1.10E-03
endocrine system development (GO:0035270)	2.43	7.33E-05
stem cell differentiation (GO:0048863)	2.43	1.41E-09
MAPK cascade (GO:0000165)	2.43	3.87E-02
Notch signaling pathway (GO:0007219)	2.43	3.01E-03
regulation of lipid metabolic process (GO:0019216)	2.43	1.12E-02
regulation of Notch signaling pathway (GO:0008593)	2.43	1.12E-02
otolith development (GO:0048840)	2.43	1.12E-02
limb development (GO:0060173)	2.42	2.27E-03
pronephros development (GO:0048793)	2.42	9.92E-06
mesenchyme development (GO:0060485)	2.42	8.89E-10
axon extension (GO:0048675)	2.42	2.94E-02
enteric nervous system development (GO:0048484)	2.42	2.93E-02
embryonic skeletal system development (GO:0048706)	2.41	3.54E-08
positive regulation of cell differentiation (GO:0045597)	2.40	4.40E-07

axis specification (GO:0009798)	2.40	3.61E-03
embryonic eye morphogenesis (GO:0048048)	2.40	1.72E-02
sprouting angiogenesis (GO:0002040)	2.40	1.95E-04
somitogenesis (GO:0001756)	2.39	5.74E-04
embryonic heart tube development (GO:0035050)	2.38	1.14E-06
spinal cord development (GO:0021510)	2.38	2.30E-03
regulation of developmental growth (GO:0048638)	2.38	2.99E-04
chloride transport (GO:0006821)	2.37	1.66E-03
negative regulation of developmental growth (GO:0048640)	2.37	4.52E-02
skeletal system development (GO:0001501)	2.37	3.65E-14
nephron development (GO:0072006)	2.37	1.18E-03
negative regulation of mitotic cell cycle phase transition (GO:1901991)	2.37	3.42E-02
motor neuron axon guidance (GO:0008045)	2.37	3.42E-02
embryonic viscerocranium morphogenesis (GO:0048703)	2.36	4.58E-04
carbohydrate biosynthetic process (GO:0016051)	2.35	2.06E-02
angiogenesis (GO:0001525)	2.35	2.30E-10
segmentation (GO:0035282)	2.35	1.29E-04
embryonic hemopoiesis (GO:0035162)	2.35	1.79E-03
Golgi organization (GO:0007030)	2.34	1.17E-02
hindbrain development (GO:0030902)	2.34	5.03E-05
alcohol biosynthetic process (GO:0046165)	2.34	1.34E-02
forebrain development (GO:0030900)	2.34	1.93E-05
synaptic vesicle cycle (GO:0099504)	2.33	2.71E-03
vesicle-mediated transport in synapse (GO:0099003)	2.33	2.71E-03
regulation of neurogenesis (GO:0050767)	2.32	7.39E-06
definitive hemopoiesis (GO:0060216)	2.31	2.30E-02
primitive hemopoiesis (GO:0060215)	2.31	2.29E-02
regulation of anatomical structure morphogenesis (GO:0022603)	2.31	1.49E-11
swimming behavior (GO:0036269)	2.31	1.43E-02
embryonic heart tube morphogenesis (GO:0003143)	2.31	4.98E-05
convergent extension involved in gastrulation (GO:0060027)	2.30	5.63E-03
fin regeneration (GO:0031101)	2.30	4.15E-03
regulation of cell growth (GO:0001558)	2.30	1.18E-03
neuron projection guidance (GO:0097485)	2.29	6.55E-09
axon guidance (GO:0007411)	2.29	6.50E-09
positive regulation of RNA biosynthetic process (GO:1902680)	2.28	1.68E-16
positive regulation of DNA-templated transcription (GO:0045893)	2.28	1.66E-16
monoatomic anion transmembrane transport (GO:0098656)	2.28	1.55E-02
embryonic organ morphogenesis (GO:0048562)	2.27	1.12E-14
kidney development (GO:0001822)	2.25	1.08E-06
regulation of cell differentiation (GO:0045595)	2.25	4.19E-16
cell migration involved in gastrulation (GO:0042074)	2.25	9.39E-03

regulation of cell projection organization (GO:0031344)	2.25	2.19E-07
chordate embryonic development (GO:0043009)	2.25	6.45E-17
microtubule cytoskeleton organization involved in mitosis (GO:1902850)	2.24	1.67E-02
renal system development (GO:0072001)	2.24	1.14E-06
tRNA modification (GO:0006400)	2.24	2.25E-02
tube development (GO:0035295)	2.24	3.74E-25
regulation of cell shape (GO:0008360)	2.24	2.99E-02
heart looping (GO:0001947)	2.24	5.79E-04
gland development (GO:0048732)	2.24	4.92E-08
embryo development ending in birth or egg hatching (GO:0009792)	2.23	8.37E-17
regulation of RNA splicing (GO:0043484)	2.23	2.32E-03
heart morphogenesis (GO:0003007)	2.23	1.45E-07
regulation of plasma membrane bounded cell projection organization (GO:0120035)	2.23	3.76E-07
tissue regeneration (GO:0042246)	2.23	3.29E-04
positive regulation of cell projection organization (GO:0031346)	2.22	7.60E-03
notochord development (GO:0030903)	2.22	1.05E-02
axonogenesis (GO:0007409)	2.22	4.14E-12
regeneration (GO:0031099)	2.22	1.03E-05
axon development (GO:0061564)	2.21	3.05E-13
embryonic organ development (GO:0048568)	2.20	3.79E-19
bone development (GO:0060348)	2.20	4.39E-02
cilium movement (GO:0003341)	2.20	6.45E-03
regulation of growth (GO:0040008)	2.20	2.96E-04
negative regulation of Wnt signaling pathway (GO:0030178)	2.19	1.16E-02
epithelial tube morphogenesis (GO:0060562)	2.19	2.44E-07
neuron projection morphogenesis (GO:0048812)	2.18	3.92E-13
regulation of nervous system development (GO:0051960)	2.18	9.87E-06
ear morphogenesis (GO:0042471)	2.18	5.57E-03
plasma membrane bounded cell projection morphogenesis (GO:0120039)	2.18	4.35E-13
response to oxidative stress (GO:0006979)	2.17	3.56E-02
positive regulation of RNA metabolic process (GO:0051254)	2.17	3.36E-16
cell morphogenesis involved in neuron differentiation (GO:0048667)	2.17	6.23E-12
retina morphogenesis in camera-type eye (GO:0060042)	2.17	8.35E-04
blood vessel morphogenesis (GO:0048514)	2.16	3.83E-10
cell projection morphogenesis (GO:0048858)	2.16	1.03E-12
liver development (GO:0001889)	2.16	3.34E-04
establishment or maintenance of cell polarity (GO:0007163)	2.16	2.74E-03
epidermis development (GO:0008544)	2.15	5.17E-03
negative regulation of cell differentiation (GO:0045596)	2.15	1.66E-03
cartilage development (GO:0051216)	2.15	2.61E-04
transition metal ion transport (GO:0000041)	2.14	3.95E-02

positive regulation of developmental process (GO:0051094)	2.14	2.80E-07
somite development (GO:0061053)	2.14	4.92E-04
regulation of multicellular organismal development (GO:2000026)	2.14	1.63E-12
cellular response to hormone stimulus (GO:0032870)	2.12	7.31E-05
regulation of BMP signaling pathway (GO:0030510)	2.12	1.69E-02
cranial skeletal system development (GO:1904888)	2.12	2.18E-06
negative regulation of developmental process (GO:0051093)	2.12	1.35E-04
blood vessel development (GO:0001568)	2.12	6.65E-11
morphogenesis of embryonic epithelium (GO:0016331)	2.11	4.37E-02
central nervous system development (GO:0007417)	2.11	1.63E-19
inner ear morphogenesis (GO:0042472)	2.11	1.34E-02
positive regulation of nucleobase-containing compound metabolic process (GO:0045935)	2.11	1.46E-15
regulation of developmental process (GO:0050793)	2.11	1.93E-23
tube morphogenesis (GO:0035239)	2.10	3.07E-15
regulation of cellular component size (GO:0032535)	2.10	2.30E-05
hepaticobiliary system development (GO:0061008)	2.10	4.58E-04
positive regulation of macromolecule biosynthetic process (GO:0010557)	2.10	6.10E-17
transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)	2.10	3.45E-06
organic hydroxy compound biosynthetic process (GO:1901617)	2.09	1.09E-02
connective tissue development (GO:0061448)	2.09	2.27E-04
positive regulation of cellular biosynthetic process (GO:0031328)	2.08	7.01E-17
animal organ morphogenesis (GO:0009887)	2.08	3.93E-21
pectoral fin development (GO:0033339)	2.08	4.86E-02
head development (GO:0060322)	2.08	1.38E-13
sensory organ morphogenesis (GO:0090596)	2.08	5.37E-08
positive regulation of biosynthetic process (GO:0009891)	2.08	8.28E-17
developmental pigmentation (GO:0048066)	2.07	4.26E-02
cell fate specification (GO:0001708)	2.07	1.18E-02
brain development (GO:0007420)	2.07	4.48E-13
epithelium development (GO:0060429)	2.07	2.17E-23
pigmentation (GO:0043473)	2.07	1.65E-03
left/right pattern formation (GO:0060972)	2.06	5.13E-07
determination of left/right symmetry (GO:0007368)	2.06	7.16E-07
neuron projection development (GO:0031175)	2.06	3.07E-13
convergent extension (GO:0060026)	2.06	8.41E-04
camera-type eye morphogenesis (GO:0048593)	2.05	1.07E-04
morphogenesis of an epithelium (GO:0002009)	2.05	2.12E-11
cell population proliferation (GO:0008283)	2.05	1.12E-03
eye morphogenesis (GO:0048592)	2.05	1.58E-05
regulation of actin cytoskeleton organization (GO:0032956)	2.04	1.95E-04

anatomical structure formation involved in morphogenesis (GO:0048646)	2.04	5.58E-19
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0090092)	2.04	5.50E-03
pattern specification process (GO:0007389)	2.04	5.41E-15
retina development in camera-type eye (GO:0060041)	2.03	3.79E-06
monoatomic anion transport (GO:0006820)	2.03	7.19E-03
regulation of small GTPase mediated signal transduction (GO:0051056)	2.03	7.17E-03
heart development (GO:0007507)	2.03	2.15E-13
pancreas development (GO:0031016)	2.03	1.34E-03
anterior/posterior pattern specification (GO:0009952)	2.03	1.38E-05
regionalization (GO:0003002)	2.03	2.07E-14
transforming growth factor beta receptor superfamily signaling pathway (GO:0141091)	2.02	2.92E-02
regulation of actin filament-based process (GO:0032970)	2.02	3.34E-04
determination of bilateral symmetry (GO:0009855)	2.02	3.92E-07
specification of symmetry (GO:0009799)	2.02	3.89E-07
tissue morphogenesis (GO:0048729)	2.01	1.09E-12
determination of heart left/right asymmetry (GO:0061371)	2.01	1.66E-04
microtubule-based transport (GO:0099111)	2.00	1.71E-02
enzyme-linked receptor protein signaling pathway (GO:0007167)	2.00	1.40E-07
response to hormone (GO:0009725)	2.00	5.20E-05
RNA catabolic process (GO:0006401)	2.00	3.17E-02
cell migration (GO:0016477)	2.00	1.73E-13
positive regulation of cellular metabolic process (GO:0031325)	2.00	3.29E-20
neural retina development (GO:0003407)	1.99	2.24E-02
digestive tract development (GO:0048565)	1.99	2.23E-02
cell morphogenesis (GO:0000902)	1.99	7.96E-12
positive regulation of phosphorylation (GO:0042327)	1.99	1.18E-03
regulation of angiogenesis (GO:0045765)	1.99	4.17E-02
response to growth factor (GO:0070848)	1.98	2.22E-03
embryo development (GO:0009790)	1.98	3.96E-26
cardiac muscle tissue development (GO:0048738)	1.98	2.96E-02
circulatory system development (GO:0072359)	1.98	1.31E-20
cellular anatomical entity morphogenesis (GO:0032989)	1.98	9.43E-12
tRNA processing (GO:0008033)	1.98	2.34E-02
appendage development (GO:0048736)	1.97	1.20E-02
regulation of cellular response to growth factor stimulus (GO:0090287)	1.97	1.97E-03
cellular response to endogenous stimulus (GO:0071495)	1.96	1.03E-07
neuron development (GO:0048666)	1.96	1.38E-14
positive regulation of phosphorus metabolic process (GO:0010562)	1.96	8.48E-04
positive regulation of phosphate metabolic process (GO:0045937)	1.96	8.46E-04
positive regulation of cell migration (GO:0030335)	1.95	2.09E-02

neuron differentiation (GO:0030182)	1.95	5.85E-18
striated muscle tissue development (GO:0014706)	1.94	3.26E-02
vasculature development (GO:0001944)	1.94	1.27E-09
embryonic morphogenesis (GO:0048598)	1.94	3.44E-14
nervous system development (GO:0007399)	1.94	2.66E-34
cellular response to growth factor stimulus (GO:0071363)	1.94	5.39E-03
growth (GO:0040007)	1.94	1.23E-05
developmental growth (GO:0048589)	1.94	1.23E-05
response to endogenous stimulus (GO:0009719)	1.94	2.25E-08
negative regulation of cellular component organization (GO:0051129)	1.93	1.84E-03
cell motility (GO:0048870)	1.93	7.79E-13
negative regulation of transcription by RNA polymerase II (GO:0000122)	1.93	9.27E-05
generation of neurons (GO:0048699)	1.92	1.34E-17
RNA splicing (GO:0008380)	1.92	1.70E-04
tissue development (GO:0009888)	1.92	4.61E-29
microtubule-based movement (GO:0007018)	1.91	5.15E-04
myeloid cell homeostasis (GO:0002262)	1.91	2.14E-02
developmental process involved in reproduction (GO:0003006)	1.91	3.00E-03
modulation of chemical synaptic transmission (GO:0050804)	1.90	7.57E-03
regulation of trans-synaptic signaling (GO:0099177)	1.90	7.56E-03
positive regulation of protein phosphorylation (GO:0001934)	1.90	3.06E-02
neurogenesis (GO:0022008)	1.90	1.78E-19
regulation of anatomical structure size (GO:0090066)	1.90	1.36E-04
dorsal/ventral pattern formation (GO:0009953)	1.89	7.97E-03
glial cell differentiation (GO:0010001)	1.89	1.66E-02
cell fate commitment (GO:0045165)	1.88	4.76E-03
behavior (GO:0007610)	1.88	4.43E-04
gliogenesis (GO:0042063)	1.88	1.06E-02
positive regulation of macromolecule metabolic process (GO:0010604)	1.88	3.32E-16
positive regulation of metabolic process (GO:0009893)	1.88	5.32E-18
RNA modification (GO:0009451)	1.88	2.30E-02
positive regulation of cell population proliferation (GO:0008284)	1.87	7.37E-03
positive regulation of nitrogen compound metabolic process (GO:0051173)	1.87	1.93E-14
anatomical structure morphogenesis (GO:0009653)	1.87	9.91E-40
negative regulation of multicellular organismal process (GO:0051241)	1.86	1.88E-03
plasma membrane bounded cell projection organization (GO:0120036)	1.85	1.58E-14
regulation of canonical Wnt signaling pathway (GO:0060828)	1.85	3.92E-02
regulation of locomotion (GO:0040012)	1.85	2.79E-04
regulation of multicellular organismal process (GO:0051239)	1.85	1.55E-13
cell development (GO:0048468)	1.85	8.72E-31
Wnt signaling pathway (GO:0016055)	1.84	1.39E-02

actin filament organization (GO:0007015)	1.84	5.10E-04
inorganic ion homeostasis (GO:0098771)	1.84	2.32E-03
system development (GO:0048731)	1.84	4.57E-59
cell projection organization (GO:0030030)	1.84	2.17E-14
aromatic compound catabolic process (GO:0019439)	1.84	8.49E-04
lateral line development (GO:0048882)	1.84	3.39E-02
sensory organ development (GO:0007423)	1.83	1.75E-11
regulation of cellular component organization (GO:0051128)	1.82	5.10E-12
cell-cell signaling by wnt (GO:0198738)	1.82	1.50E-02
positive regulation of cell motility (GO:2000147)	1.82	4.39E-02
cell-cell adhesion (GO:0098609)	1.82	5.91E-05
positive regulation of multicellular organismal process (GO:0051240)	1.81	1.95E-04
epithelial cell differentiation (GO:0030855)	1.81	8.75E-04
hemopoiesis (GO:0030097)	1.81	8.37E-07
regulation of mRNA metabolic process (GO:1903311)	1.81	1.96E-02
digestive system development (GO:0055123)	1.81	2.96E-02
regulation of cytoskeleton organization (GO:0051493)	1.81	5.76E-04
regulation of actin filament organization (GO:0110053)	1.81	4.02E-02
multicellular organism development (GO:0007275)	1.80	7.03E-66
regulation of supramolecular fiber organization (GO:1902903)	1.80	1.35E-02
regulation of cell migration (GO:0030334)	1.80	2.07E-03
animal organ development (GO:0048513)	1.79	2.02E-39
negative regulation of RNA metabolic process (GO:0051253)	1.79	9.94E-06
mRNA processing (GO:0006397)	1.79	2.06E-04
regulation of GTPase activity (GO:0043087)	1.79	1.73E-02
organic cyclic compound catabolic process (GO:1901361)	1.78	1.60E-03
homeostasis of number of cells (GO:0048872)	1.78	3.15E-02
positive regulation of protein modification process (GO:0031401)	1.78	3.62E-02
cellular developmental process (GO:0048869)	1.77	3.82E-37
cell differentiation (GO:0030154)	1.77	4.97E-37
negative regulation of intracellular signal transduction (GO:1902532)	1.77	3.26E-02
negative regulation of biosynthetic process (GO:0009890)	1.77	3.14E-08
actin cytoskeleton organization (GO:0030036)	1.77	3.44E-06
gamete generation (GO:0007276)	1.77	2.06E-02
cell junction organization (GO:0034330)	1.77	4.36E-03
negative regulation of RNA biosynthetic process (GO:1902679)	1.76	6.18E-05
negative regulation of DNA-templated transcription (GO:0045892)	1.76	6.15E-05
cell surface receptor signaling pathway (GO:0007166)	1.76	2.26E-14
cellular homeostasis (GO:0019725)	1.76	1.11E-03
small GTPase-mediated signal transduction (GO:0007264)	1.76	1.72E-02
positive regulation of cellular component organization (GO:0051130)	1.76	5.11E-03
mitotic cell cycle process (GO:1903047)	1.75	2.94E-03

methylation (GO:0032259)	1.75	1.19E-02
regulation of Wnt signaling pathway (GO:0030111)	1.75	2.20E-02
multicellular organismal reproductive process (GO:0048609)	1.75	2.19E-02
multicellular organism reproduction (GO:0032504)	1.74	2.45E-02
endocytosis (GO:0006897)	1.74	1.70E-03
mRNA metabolic process (GO:0016071)	1.74	1.17E-04
negative regulation of macromolecule biosynthetic process (GO:0010558)	1.73	2.00E-07
actin filament-based process (GO:0030029)	1.72	9.35E-06
heterocycle catabolic process (GO:0046700)	1.72	9.84E-03
cellular nitrogen compound catabolic process (GO:0044270)	1.72	9.82E-03
negative regulation of cellular biosynthetic process (GO:0031327)	1.72	2.98E-07
monoatomic ion homeostasis (GO:0050801)	1.72	3.96E-03
positive regulation of cellular process (GO:0048522)	1.72	8.41E-24
synaptic signaling (GO:0099536)	1.72	1.41E-03
response to oxygen-containing compound (GO:1901700)	1.72	4.11E-04
anatomical structure development (GO:0048856)	1.71	1.64E-71
developmental process (GO:0032502)	1.71	1.92E-73
muscle cell differentiation (GO:0042692)	1.71	4.61E-03
chemical homeostasis (GO:0048878)	1.71	1.96E-04
eye development (GO:0001654)	1.71	2.68E-06
visual system development (GO:0150063)	1.71	2.67E-06
negative regulation of nucleobase-containing compound metabolic process (GO:0045934)	1.71	3.60E-05
regulation of cell population proliferation (GO:0042127)	1.70	1.68E-03
plasma membrane bounded cell projection assembly (GO:0120031)	1.70	1.04E-03
regulation of cell motility (GO:2000145)	1.70	7.20E-03
homeostatic process (GO:0042592)	1.70	2.43E-07
vesicle organization (GO:0016050)	1.68	4.34E-02
mitotic cell cycle (GO:0000278)	1.68	2.57E-03
cell projection assembly (GO:0030031)	1.68	8.76E-04
multicellular organismal process (GO:0032501)	1.67	1.69E-62
regulation of transferase activity (GO:0051338)	1.67	4.06E-02
positive regulation of biological process (GO:0048518)	1.67	2.55E-23
chromatin organization (GO:0006325)	1.67	2.31E-03
regulation of phosphorylation (GO:0042325)	1.66	6.94E-03
cell division (GO:0051301)	1.66	9.39E-03
cell adhesion (GO:0007155)	1.66	9.58E-06
chromatin remodeling (GO:0006338)	1.66	3.17E-02
sensory system development (GO:0048880)	1.66	1.21E-06
regulation of transcription by RNA polymerase II (GO:0006357)	1.65	6.79E-19
cell-cell signaling (GO:0007267)	1.65	2.44E-05
chemical synaptic transmission (GO:0007268)	1.65	7.24E-03

anterograde trans-synaptic signaling (GO:0098916)	1.65	7.22E-03
small molecule biosynthetic process (GO:0044283)	1.65	7.20E-03
camera-type eye development (GO:0043010)	1.65	1.33E-04
mRNA splicing, via spliceosome (GO:0000398)	1.65	4.47E-02
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile (GO:0000377)	1.65	4.46E-02
RNA splicing, via transesterification reactions (GO:0000375)	1.65	4.45E-02
response to abiotic stimulus (GO:0009628)	1.64	1.57E-03
regulation of organelle organization (GO:0033043)	1.64	3.61E-04
negative regulation of cellular metabolic process (GO:0031324)	1.63	3.52E-07
regulation of response to external stimulus (GO:0032101)	1.63	8.68E-03
negative regulation of response to stimulus (GO:0048585)	1.63	5.53E-05
trans-synaptic signaling (GO:0099537)	1.63	1.07E-02
dephosphorylation (GO:0016311)	1.63	3.20E-02
small molecule catabolic process (GO:0044282)	1.62	4.39E-02
reproduction (GO:0000003)	1.62	8.38E-03
negative regulation of signal transduction (GO:0009968)	1.62	5.13E-04
organic hydroxy compound metabolic process (GO:1901615)	1.61	3.42E-02
negative regulation of cell communication (GO:0010648)	1.61	4.84E-04
negative regulation of signaling (GO:0023057)	1.61	4.82E-04
negative regulation of metabolic process (GO:0009892)	1.61	2.79E-07
muscle structure development (GO:0061061)	1.61	2.76E-03
reproductive process (GO:0022414)	1.60	1.08E-02
muscle cell development (GO:0055001)	1.60	4.29E-02
monoatomic cation homeostasis (GO:0055080)	1.60	3.45E-02
regulation of signaling (GO:0023051)	1.60	9.26E-12
cilium assembly (GO:0060271)	1.59	1.15E-02
cellular component organization (GO:0016043)	1.59	7.26E-33
response to organic substance (GO:0010033)	1.59	6.38E-06
regulation of cell communication (GO:0010646)	1.59	1.64E-11
positive regulation of molecular function (GO:0044093)	1.59	4.56E-03
protein-DNA complex organization (GO:0071824)	1.59	6.06E-03
import into cell (GO:0098657)	1.58	2.16E-03
regulation of biological quality (GO:0065008)	1.58	1.04E-08
microtubule-based process (GO:0007017)	1.58	5.15E-04
cellular component organization or biogenesis (GO:0071840)	1.58	9.60E-34
negative regulation of gene expression (GO:0010629)	1.58	3.74E-02
gastrulation (GO:0007369)	1.58	3.74E-02
cytoskeleton organization (GO:0007010)	1.58	7.94E-07
intracellular signaling cassette (GO:0141124)	1.58	4.82E-03
regulation of RNA metabolic process (GO:0051252)	1.58	7.85E-21
regulation of response to stimulus (GO:0048583)	1.57	1.08E-12

apoptotic process (GO:0006915)	1.57	4.11E-02
transmembrane transport (GO:0055085)	1.57	7.07E-08
negative regulation of macromolecule metabolic process (GO:0010605)	1.57	3.61E-06
cell cycle (GO:0007049)	1.57	6.54E-05
negative regulation of cellular process (GO:0048523)	1.57	1.36E-14
regulation of biosynthetic process (GO:0009889)	1.56	8.37E-24
negative regulation of biological process (GO:0048519)	1.56	1.52E-15
regulation of DNA-templated transcription (GO:0006355)	1.56	5.90E-18
negative regulation of nitrogen compound metabolic process (GO:0051172)	1.56	7.47E-05
regulation of RNA biosynthetic process (GO:2001141)	1.56	6.05E-18
RNA processing (GO:0006396)	1.56	9.42E-05
regulation of nucleobase-containing compound metabolic process (GO:0019219)	1.56	3.21E-20
membrane organization (GO:0061024)	1.56	9.17E-03
regulation of cellular metabolic process (GO:0031323)	1.55	5.39E-27
supramolecular fiber organization (GO:0097435)	1.55	2.05E-03
regulation of cellular biosynthetic process (GO:0031326)	1.55	9.65E-23
regulation of macromolecule biosynthetic process (GO:0010556)	1.55	1.60E-22
regulation of gene expression (GO:0010468)	1.55	4.21E-22
phospholipid metabolic process (GO:0006644)	1.55	4.03E-02
sexual reproduction (GO:0019953)	1.55	4.02E-02
carbohydrate metabolic process (GO:0005975)	1.54	3.38E-02
regulation of signal transduction (GO:0009966)	1.54	1.17E-08
regulation of metabolic process (GO:0019222)	1.54	2.98E-28
positive regulation of catalytic activity (GO:0043085)	1.54	2.23E-02
regulation of phosphate metabolic process (GO:0019220)	1.54	2.29E-02
regulation of phosphorus metabolic process (GO:0051174)	1.54	2.29E-02
glycoprotein metabolic process (GO:0009100)	1.53	4.54E-02
regulation of catalytic activity (GO:0050790)	1.53	1.57E-03
regulation of primary metabolic process (GO:0080090)	1.53	2.27E-23
cell cycle process (GO:0022402)	1.52	6.42E-03
cellular response to organic substance (GO:0071310)	1.52	2.56E-03
cilium organization (GO:0044782)	1.52	2.60E-02
regulation of macromolecule metabolic process (GO:0060255)	1.52	1.17E-23
regulation of nitrogen compound metabolic process (GO:0051171)	1.52	4.43E-22
cellular catabolic process (GO:0044248)	1.51	3.55E-04
regulation of molecular function (GO:0065009)	1.51	4.25E-04

Table S4 – Gene ontology terms enriched in a list of GREAT-associated genes that were significantly upregulated in EdgeR analysis of acinar cells ATAC-seq. Terms with fold enrichment <1.5 are shown.

Acinar Identity	Endocrine/Progenitor Identity	Ptf1a co-factors	AP-1 complex components
Ptf1a	Neurog3	Tcf3/E2a	c-Jun
Nr5a2	NeuroD1	Tcf4	JunB
Bhlha15	Ascl1	Tcf12/Heb	JunD
Gata4	Ascl2	Rbpj	Atf1
	Isl1		Atf2
	Pdx1		Atf3
	Nkx6.1		Atf4
	Nkx2.2		Atf5
	FoxA1		Atf6
	FoxA2		Atf7
	Sox9		Batf
	Hnf1a		Batf3
	Hnf1b		c-Fos
	Pax4		FosB
	Pax6		Fra1
			Fra2
			c-Maf
			MafA
			MafB
			MafG

Table S5 – Lists of TFs within groups defined in the context of Figure 3.6, manually curated from the literature.

Table S6 – DisGeNET lists of genes associated to “Pancreatic Neoplasm”, downloaded in March 2024. This table can be found in the following link:

https://drive.google.com/drive/folders/17W_9HQJHQdLEu2hivFMdT7x9rUeZihRw?usp=drive_link

Table S7 – DisGeNET lists of genes associated to “Malignant Neoplasm of Pancreas”, downloaded in March 2024. This table can be found in the following link:

https://drive.google.com/drive/folders/17W_9HQJHQdLEu2hivFMdT7x9rUeZihRw?usp=drive_link

3.6. References

- Amorim, J.P., Gali-Macedo, A., Marcelino, H., Bordeira-Carriço, R., Naranjo, S., Rivero-Gil, S., Teixeira, J., Galhardo, M., Marques, J. & Bessa, J. (2020) A Conserved Notochord Enhancer Controls Pancreas Development in Vertebrates. *Cell Reports* 32, 107862
- Apelqvist, Å., Li, H., Sommer, L., Beatus, P., Anderson, D.J., Honjo, T., de Angelis, M.H., Lendahl, U. & Edlund, H. (1999) Notch signalling controls pancreatic cell differentiation. *Nature* 400, 877–881
- Argenton, F., Zecchin, E. & Bortolussi, M. (1999) Early appearance of pancreatic hormone-expressing cells in the zebrafish embryo. *Mech Dev* 87, 217–221
- Azevedo-Pouly, A., Hale, M.A., Swift, G.H., Hoang, C.Q., Deering, T.G., Xue, J., Wilkie, T.M., Murtaugh, L.C. & MacDonald, R.J. (2023) Key transcriptional effectors of the pancreatic acinar phenotype and oncogenic transformation. *PLOS ONE*
- Baetens, D., Malaisse-Lagae, F., Perrelet, A. & Orci, L. (1979) Endocrine Pancreas: Three-Dimensional Reconstruction Shows Two Types of Islets of Langerhans. *Science, New Series* 206, 1323–1325
- Beres, T.M., Masui, T., Swift, G.H., Shi, L., Henke, R.M. & MacDonald, R.J. (2006) PTF1 Is an Organ-Specific and Notch-Independent Basic Helix-Loop-Helix Complex Containing the Mammalian Suppressor of Hairless (RBP-J) or Its Parologue, RBP-L. *MCB* 26, 117–130
- Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B. & Driever, W. (2001) Pancreas Development in Zebrafish: Early Dispersed Appearance of Endocrine Hormone Expressing Cells and Their Convergence to Form the Definitive Islet. *Developmental Biology* 230, 189–203
- Bordeira-Carriço, R., Teixeira, J., Duque, M., Galhardo, M., Ribeiro, D., D. Acemel, R., N. Firbas, P., J. Tena, J., Eufrásio, A., Marques, J., et al (2022) Multidimensional chromatin profiling of zebrafish pancreas to uncover and investigate disease- relevant enhancers. *Nature Communications* 13
- Brissova, M., Fowler, M.J., Nicholson, W.E., Chu, A., Hirshberg, B., Harlan, D.M. & Powers, A.C. (2005) Assessment of Human Pancreatic Islet Architecture and Composition by Laser Scanning Confocal Microscopy. *J Histochem Cytochem* 53, 1087–1097
- Burlison, J.S., Long, Q., Fujitani, Y., Wright, C.V.E. & Magnuson, M.A. (2008) Pdx-1 and Ptf1a concurrently determine fate specification of pancreatic multipotent progenitor cells. *Developmental Biology* 316, 74–86
- Cobo, I., Martinelli, P., Flández, M., Bakiri, L., Zhang, M., Carrillo-de-Santa-Pau, E., Jia, J., Lobo, V.J.S.-A., Megías, D. & Felipe, I. (2018) Transcriptional regulation by NR5A2 links differentiation and inflammation in the pancreas.
- Dong, P.D.S., Provost, E., Leach, S.D. & Stainier, D.Y.R. (2008) Graded levels of Ptf1a differentially regulate endocrine and exocrine fates in the developing pancreas. *Genes*

& Development 22, 1445–1450

Duque, M., Amorim, J.P. & Bessa, J. (2022) Ptf1a function and transcriptional cis-regulation, a cornerstone in vertebrate pancreas development. *The FEBS Journal* 289, 5121–5136

Fujikura, J., Hosoda, K., Iwakura, H., Tomita, T., Noguchi, M., Masuzaki, H., Tanigaki, K., Yabe, D., Honjo, T. & Nakao, K. (2006) Notch/Rbp-j signaling prevents premature endocrine and ductal cell differentiation in the pancreas. *Cell Metabolism* 3, 59–65

Ghaye, A.P., Bergemann, D., Tarifeño-Saldivia, E., Flasse, L.C., Von Berg, V., Peers, B., Voz, M.L. & Manfroid, I. (2015) Progenitor potential of nkx6.1-expressing cells throughout zebrafish life and during beta cell regeneration. *BMC Biol* 13, 70

Glasgow, S.M. (2005) Ptf1a determines GABAergic over glutamatergic neuronal cell fate in the spinal cord dorsal horn. *Development* 132, 5461–5469

Greten, F.R. & Grivnickov, S.I. (2020) Inflammation and Cancer: Triggers, Mechanisms and Consequences.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C. & Singh, H. (2011) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.

Hesselson, D., Anderson, R.M. & Stainier, D.Y.R. (2012) Suppression of Ptf1a activity induces acinar to endocrine conversion.

Hoang, C.Q., Hale, M.A., Azevedo-Pouly, A.C., Elsässer, H.P., Deering, T.G., Willet, S.G., Pan, F.C., Magnuson, M.A., Wright, C.V.E., Swift, G.H., et al (2016) Transcriptional Maintenance of Pancreatic Acinar Identity, Differentiation, and Homeostasis by PTF1A. *Mol Cell Biol* 36, 3033–3047

Holmstrom, S.R., Deering, T., Swift, G.H., Poelwijk, F.J., Mangelsdorf, D.J., Kliewer, S.A. & MacDonald, R.J. (2011) LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. *Genes & Development* 25, 1674–1679

Hoshino, M., Nakamura, S., Mori, K., Kawauchi, T., Terao, M., Nishimura, Y.V., Fukuda, A., Fuse, T., Matsuo, N., Sone, M., et al (2005) Ptf1a, a bHLH Transcriptional Gene, Defines GABAergic Neuronal Fates in Cerebellum. *Neuron* 47, 201–213

Jakubison, B.L., Schweickert, P.G., Moser, S.E., Yang, Y., Gao, H., Scully, K., Itkin-Ansari, P., Liu, Y. & Konieczny, S.F. (2018) Induced PTF1a expression in pancreatic ductal adenocarcinoma cells activates acinar gene networks, reduces tumorigenic properties, and sensitizes cells to gemcitabine treatment. *Molecular Oncology*

Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. (2014) Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads.

Jiang, J., Hakimjavadi, H., Bray, J.K., Perkins, C. & Gosling, A. (2023) Transcriptional Profile of Human Pancreatic Acinar Ductal Metaplasia.

Jiang, M., Azevedo-Pouly, A.C., Deering, T.G., Hoang, C.Q., DiRenzo, D., Hess, D.A.,

- Konieczny, S.F., Swift, G.H. & MacDonald, R.J. (2016) MIST1 and PTF1 Collaborate in Feed-Forward Regulatory Loops That Maintain the Pancreatic Acinar Phenotype in Adult Mice. *Molecular and Cellular Biology* 36
- Kawaguchi, Y., Cooper, B., Gannon, M., Ray, M., MacDonald, R.J. & Wright, C.V.E. (2002) The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet* 32, 128–134
- Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. & Karolchik, D. (2010) BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26
- Krah, N.M., Narayanan, S.M., Yugawa, D.E., Straley, J.A., Wright, C.V.E., MacDonald, R.J. & Murtaugh, L.C. (2020) Prevention and reversion of pancreatic tumorigenesis through a differentiation-based mechanism. *Developmental Cell* 50
- Krah, N.M., Willet, S.G., Pan, F.C., Cash, G.M., Bronner, M.P., Wright, C.V., MacDonald, R.J. & Murtaugh, L.C. (2015) The acinar differentiation determinant PTF1A inhibits initiation of pancreatic ductal adenocarcinoma. *eLife*
- Krapp, A., Knöfler, M., Frutiger, S., Hughes, G.J., Hagenbüchle, O. & Wellauer, P.K. (1996) The p48 DNA-binding subunit of transcription factor PTF1 is a new exocrine pancreas-specific basic helix-loop-helix protein. *The EMBO Journal* 15, 4317–4329
- Langmead, B. & Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25
- Li, Q., Brown, J.B., Huang, H. & Bickel, P.J. (2011) Measuring reproducibility of high-throughput experiments. *The Annals of Applied Statistics* 5
- Liao, Y., Smyth, G.K. & Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*
- Lin, J.W., Biankin, A.V., Horb, M.E., Ghosh, B., Prasad, N.B., Yee, N.S., Pack, M.A. & Leach, S.D. (2004) Differential requirement for ptf1a in endocrine and exocrine lineages of developing zebrafish pancreas. *Developmental Biology* 270, 474–486
- Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550
- Marstrand-Daucé, L., Lorenzo, D., Chassac, A., Nicole, P., Couvelard, A. & Haumaitre, C. (2023) Acinar-to-Ductal Metaplasia (ADM): On the Road to Pancreatic Intraepithelial Neoplasia (PanIN) and Pancreatic Cancer. *Int J Mol Sci*
- Masui, T., Long, Q., Beres, T.M., Magnuson, M.A. & MacDonald, R.J. (2007) Early pancreatic development requires the vertebrate Suppressor of Hairless (RBPJ) in the PTF1 bHLH complex. *Genes & Development* 21, 2629–2643
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger,

- A.M. & Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology* 28
- Meredith, D.M., Masui, T., Swift, G.H., MacDonald, R.J. & Johnson, J.E. (2009) MultiAutoregulation by the PTF1-J Complex. *Journal of Neuroscience* 29, 11139–11148
- Miguel-Escalada, I., Maestro, M.A., Balboa, D., García-Hurtado, J., Sebé-Pedrós, A. & Ferrer, J. (2022) Pancreas agenesis mutations disrupt a lead enhancer controlling a developmental enhancer cluster. *Developmental Cell* 57
- Obata, J., Yano, M., Mimura, H., Goto, T., Nakayama, R., Mibu, Y., Oka, C. & Kawaichi, M. (2001) p48 subunit of mouse PTF1 binds to RBP-J κ /CBF-1, the intracellular mediator of Notch signalling, and is expressed in the neural tube of early stage embryos: PTF1p48 as an RBP- J κ binding protein. *Genes to Cells* 6, 345–360
- Pashos, E., Park, J.T., Leach, S. & Fisher, S. (2013) Distinct enhancers of ptf1a mediate specification and expansion of ventral pancreas in zebrafish. *Developmental Biology* 381, 471–481
- Piñero, J., Queralt-Rosinach, N., Bravo, À., Deu-Pons, J., Bauer-Mehren, A., Baron, M., Sanz, F. & Furlong, L. (2015) DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database tool*
- Quinlan, A.R. & Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. 26
- Ramírez, F., Ryan, D.P., Richter, A.S. & Heyne, S. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* 44
- Robinson, M.D., McCarthy, D.J. & Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140
- Sakhneny, L., Khalifa-Malka, L. & Landsman, L. (2018) Pancreas organogenesis: Approaches to elucidate the role of epithelial-mesenchymal interactions. *Seminars in Cell & Developmental Biology*
- Schaffer, A.E., Freude, K.K., Nelson, S.B. & Sander, M. (2010) Nkx6 Transcription Factors and Ptf1a Function as Antagonistic Lineage Determinants in Multipotent Pancreatic Progenitors. *Developmental Cell* 18, 1022–1029
- Sellick, G.S., Barker, K.T., Stolte-Dijkstra, I., Fleischmann, C., J Coleman, R., Garrett, C., Gloyn, A.L., Edghill, E.L., Hattersley, A.T., Wellauer, P.K., et al (2004) Mutations in PTF1A cause pancreatic and cerebellar agenesis. *Nat Genet* 36, 1301–1305
- Shi, G., DiRenzo, D., Qu, C., Barney, D. & Miley, D. (2013) Maintenance of Acinar Cell Organization is Critical to Preventing Kras-Induced Acinar-Ductal Metaplasia.
- Slack, J.M.W (1995) Developmental biology of the pancreas. *Development* 121, 1569–1580
- Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.-P. & Huaiyu, M.

(2022) PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Science* 31

vonFigura, G., Morris IV, J.P., Wright, C.V.E. & Hebrok, M. (2014) Nr5a2 maintains acinar cell differentiation and constrains oncogenic Kras-mediated pancreatic neoplastic initiation. *Gut* 63

Wang, Y.J., Park, J.T., Parsons, M.J. & Leach, S.D. (2015) Fate mapping of *ptf1a* -expressing cells during pancreatic organogenesis and regeneration in zebrafish. *Dev Dyn* 244, 724–735

Weedon, Michael N., Cebola, I., Patch, A.-M., Flanagan, S.E., De Franco, E., Caswell, R., Rodríguez-Seguí, S.A., Shaw-Smith, C., Cho, C.H.-H., Allen, H.L., et al (2014) Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. *Nat Genet* 46, 61–64

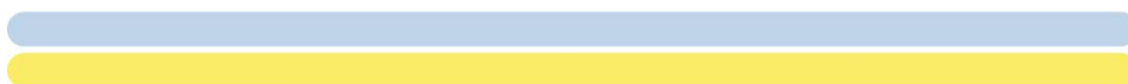
Zecchin, E., Mavropoulos, A., Devos, N., Filippi, A., Tiso, N., Meyer, D., Peers, B., Bortolussi, M. & Argenton, F. (2004) Evolutionary conserved role of *ptf1a* in the specification of exocrine pancreatic fates. *Developmental Biology* 268, 174–184

Zerbino, D.R., Johnson, N., Juettemann, T., Wilder, S.P. & Flicek, P. (2014) WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics* 30

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137

Zhou, Q., Law, A.C., Rajagopal, J., Anderson, W.J., Gray, P.A. & Melton, D.A. (2007) A Multipotent Progenitor Domain Guides Pancreatic Organogenesis. *Developmental Cell* 13, 103–114

Chapter IV



Concluding remarks

4.1. Concluding remarks

In multicellular organisms, all the cells contain identical genomes while expressing distinct genes, at various possible timings. This implies the existence of regulatory information within the genome and that this information can be radically different between cell types. These regulatory networks play a crucial role in driving the faithful execution of major biological processes such as development, proliferation, aging or differentiation, which require precise spatial and temporal expression of genes (Isshiki et al, 2001; Maduro & Rothman, 2002; Pourquié, 2003). Particularly, development of different cellular identities is dependent on this regulatory information, which is mostly present in non-coding cis-regulatory elements (CREs), spread over large genomic distances (Maston et al, 2006; Ong & Corces, 2011).

The pancreas is a central organ for human diseases such as pancreatic cancer and diabetes, both having a dramatic societal burden. Development of the vertebrate pancreas is an intricate process through which various cell lineages develop from multipotent progenitor cells (MPCs) (Zhou et al, 2007; Kinkel & Prince, 2009). During this process, several intercellular signals are known to play a crucial role in the correct spatial and temporal expression of genes. By binding to cell surface receptors, these ligands restrict the field of MPCs, culminating in the expression of specific transcription factors (TFs) such as Pdx1, Nkx6.1 or Ptf1a (Zhou et al, 2007; Hald et al, 2008; Schaffer et al, 2010), which are widely used as markers for MPCs. Some of the signaling pathways that are known to be crucial for pancreatic development in vertebrates are Hedgehog signaling, RA (retinoic acid) signaling, Wnt signaling, BMP signaling or the Notch signaling pathway (Kim & Hebrok, 2001). While many studies on the genetics of pancreatic function and development have been done using the mouse model system, the accessibility and optical clarity of zebrafish embryos allows cell manipulation and sophisticated *in vivo* imaging approaches, which makes the zebrafish a powerful tool to study the pancreas, especially in a developmental setting.

In Chapter II of this thesis, we have used the zebrafish model to study cis-regulatory elements in the context of pancreatic dorsal bud development, by focusing on a diffusible molecule and a known BMP signaling inhibitor, Nog2. We show that *Nog* is expressed in the notochord, controlled transversely in zebrafish, human and mice by at least one notochord enhancer. We postulate that, while there is no conservation at sequence level between zebrafish and human/mouse sequences, these represent functionally equivalent sequences, responsible for driving expression of *Nog* genes in the notochord of vertebrates. Furthermore, we demonstrate that *nog2* gain and loss of function in zebrafish correlates with an increase and a decrease of pancreatic

differentiation, respectively. Specifically, we observed that the deletion of the *nog2* notochord enhancer in zebrafish, achieved through CRISPR/Cas9, is enough to cause a significant decrease in the number of insulin and *nkx6.1*-expressing cells. On the other hand, a notochord-specific overexpression of *nog2* resulted in a significant increase of the area of insulin expression. These results highlight Nog as a crucial pro-pancreatic signaling molecule emanating from the notochord, counteracting BMP signaling to allow proper placement and size of the pancreatic progenitor domain. Our work from Chapter II of this doctoral thesis emphasizes the important role of signaling molecules in development, more specifically, the role of the notochord as a crucial signaling center in pancreatic development. This implies that further studies in the context of notochord-derived intercellular signals and Nog function might be of interest for a more comprehensive knowledge of the phenomena revolving pancreas development. Importantly, these results also highlight the impact of regulatory alterations in major biological processes, which in turn emphasizes the importance of studying regulatory regions of the DNA in the context of development and disease. In the past, most of the studies in the biological areas focused almost exclusively on the expression of genes and alterations in the coding genome. However, more recent work from several laboratories around the world has shown that non-coding CREs are crucial for the regulation of gene expression, both spatially and temporally, bringing into focus the importance of regulatory networks in the fields of genetics and biology.

In Chapter III of the current thesis, we focused on the function and regulation of *Ptf1a*, which encodes for a crucial TF in vertebrate pancreas development and function. After the pancreatic MPC domain is established, with the help of the above mentioned intercellular signals, it is commonly defined by the expression of specific TFs across vertebrates. Human, mouse and zebrafish MPCs all show expression of *Ptf1a* (Cebola et al, 2015; Kawaguchi et al, 2002; Lin et al, 2004; Wang et al, 2015) and its function was shown to be crucial for pancreas development (Weedon, Michael N. et al, 2014; Fukuda et al, 2008; Bordeira-Carriço et al, 2022; Duque et al, 2022). After differentiation of MPCs into acinar cells, *Ptf1a* is required to maintain acinar cell identity, together with other acinar-specific TFs (Rose et al, 2001; Holmstrom et al, 2011; Hoang et al, 2016). Importantly, *Ptf1a* loss of function in mouse acinar cells lead to ADM (acinar-to-ductal metaplasia) and the induction of a KRAS-permissive, PDAC-like (pancreatic ductal adenocarcinoma) gene expression profile (Krah et al, 2015). In this thesis, we have explored the mechanisms of regulation of *ptf1a* expression in zebrafish, focusing on the distal pancreatic enhancer z3'-DpE and its dual function. We found this enhancer to be active early in MPCs of the developing pancreas and later, in differentiated acinar cells,

indicating a species-specific late function for the zebrafish sequence. To study the impact of deleting this CRE in the zebrafish pancreas, we took advantage of a mutant line previously established in our laboratory (Bordeira-Carriço et al, 2022). By performing enhancer assays and immunohistochemistry, comparing between z3'-DpE^{+/+} and -/- backgrounds, we were able to observe alterations in the pattern and localization of the MPC domain, suggesting that differentiation into the various pancreatic lineages could be perturbed. Then, we performed RNA-sequencing of z3'-DpE^{+/+} and -/- MPCs and found downregulated genes to be enriched in Notch signaling pathway components, indicating that Ptf1a activates an endogenous Notch signaling program in MPCs. To profile chromatin accessibility, we performed ATAC-seq of z3'-DpE^{+/+} and -/- differentiated acinar cells. We found regions with decreased accessibility in z3'-DpE^{-/-} and enriched TF motifs within these regions to be associated with genes related to acinar cell identity, endocrine function and developmental processes. On the other hand, regions with increased accessibility were found to be associated with genes related to developmental processes, inflammation, proliferation or epithelial-to-mesenchymal (EMT) transition. Moreover, TF motifs enriched within regions with increased accessibility were related to inflammatory processes and the respective TF-encoding genes were enriched within DisGeNET lists of genes associated to pancreatic cancer (Piñero et al, 2015). These results highlight Ptf1a as a crucial TF for the establishment and maintenance of the acinar cell identity in zebrafish, being this function achieved through the regulation of chromatin accessibility within important regulator and effector genes of exocrine function or through co-binding with specific TFs important for acinar cell identity such as Bhlha15 or Nr5a2. Also, we show that upon deletion of z3'-DpE, regions exhibiting increased chromatin accessibility associated with genes related to inflammation and pancreatic cancer, emphasizing the role of *ptf1a* as a tumor suppressor gene in acinar cells, in agreement with previous results from other labs (Krah et al, 2015; Hoang et al, 2016; Jakubison et al, 2018; Krah et al, 2020).

Overall, in this doctoral thesis, we focused on transcriptional CREs involved in early pancreas development and in the establishment and maintenance of the adult acinar cell identity, more specifically on regulatory elements controlling *nog2* and *ptf1a* genes. We show that, in vertebrates, Nog is a crucial pro-pancreatic signaling molecule emanating from the notochord, counteracting BMP signaling to allow proper placement and size of the pancreatic progenitor domain and that Ptf1a is an important TF for the differentiation of MPCs into acinar cells and for the maintenance of the differentiated acinar cell identity, possibly preventing the arising of aberrant cellular identities that could lead to pancreatic dysfunction and cancer. The discoveries reported in this thesis

emphasize the crucial role of enhancers in pancreas development and maintenance of differentiated cell identity. Moreover, our findings support the notion that disruption of pancreatic enhancers might contribute to the development of pancreatic cancer. While most of the studies focusing on the genetic causes of pancreatic cancer have focused on the coding part of genome, the contribution of non-coding regions has been vastly disregarded. In this context, a more in-depth study of such regulatory regions could lead to the discovery of novel genetic players, which, in turn, might be important for a better understanding of the phenomena involved in pancreatic tumorigenesis and in the development of novel and better therapeutic strategies for this severe disease.

4.2. References

- Bordeira-Carriço, R., Teixeira, J., Duque, M., Galhardo, M., Ribeiro, D., D. Acemel, R., N. Firbas, P., J. Tena, J., Eufrásio, A., Marques, J., et al (2022) Multidimensional chromatin profiling of zebrafish pancreas to uncover and investigate disease-relevant enhancers. *Nature Communications* 13
- Cebola, I., Rodríguez-Seguí, S.A., Cho, C.H.-H., Bessa, J., Rovira, M., Luengo, M., Chhatriwala, M., Berry, A., Ponsa-Cobas, J., Maestro, M.A., et al (2015) TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nat Cell Biol* 17, 615–626
- Duque, M., Amorim, J.P. & Bessa, J. (2022) Ptf1a function and transcriptional cis-regulation, a cornerstone in vertebrate pancreas development. *The FEBS Journal* 289, 5121–5136
- Fukuda, A., Kawaguchi, Y., Furuyama, K., Kodama, S., Horiguchi, M., Kuhara, T., Kawaguchi, M., Terao, M., Doi, R., Wright, C.V.E., et al (2008) Reduction of Ptf1a Gene Dosage Causes Pancreatic Hypoplasia and Diabetes in Mice. *Diabetes* 57, 2421–2431
- Hald, J., Sprinkel, A.E., Ray, M., Serup, P., Wright, C. & Madsen, O.D. (2008) Generation and Characterization of Ptf1a Antiserum and Localization of Ptf1a in Relation to Nkx6.1 and Pdx1 During the Earliest Stages of Mouse Pancreas Development. *J Histochem Cytochem* 56, 587–595
- Hoang, C.Q., Hale, M.A., Azevedo-Pouly, A.C., Elsässer, H.P., Deering, T.G., Willet, S.G., Pan, F.C., Magnuson, M.A., Wright, C.V.E., Swift, G.H., et al (2016) Transcriptional Maintenance of Pancreatic Acinar Identity, Differentiation, and Homeostasis by PTF1A. *Mol Cell Biol* 36, 3033–3047
- Holmstrom, S.R., Deering, T., Swift, G.H., Poelwijk, F.J., Mangelsdorf, D.J., Kliewer, S.A. & MacDonald, R.J. (2011) LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. *Genes & Development* 25, 1674–1679
- Isshiki, T., Pearson, B., Holbrook, S. & Doe, C.Q. (2001) *Drosophila* Neuroblasts Sequentially Express Transcription Factors which Specify the Temporal Identity of Their Neuronal Progeny. *Cell* 106, 511–521
- Jakubison, B.L., Schweickert, P.G., Moser, S.E., Yang, Y., Gao, H., Scully, K., Itkin-Ansari, P., Liu, Y. & Konieczny, S.F. (2018) Induced PTF1a expression in pancreatic ductal adenocarcinoma cells activates acinar gene networks, reduces tumorigenic properties, and sensitizes cells to gemcitabine treatment. *Molecular Oncology*
- Kawaguchi, Y., Cooper, B., Gannon, M., Ray, M., MacDonald, R.J. & Wright, C.V.E. (2002) The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet* 32, 128–134
- Kim, S.K. & Hebrok, M. (2001) Intercellular signals regulating pancreas development and function. *Genes Dev* 15, 111–127
- Kinkel, M.D. & Prince, V.E. (2009) On the diabetic menu: Zebrafish as a model for pancreas development and function. *BioEssays* 31, 139–152

- Krah, N.M., Narayanan, S.M., Yugawa, D.E., Straley, J.A., Wright, C.V.E., MacDonald, R.J. & Murtaugh, L.C. (2020) Prevention and reversion of pancreatic tumorigenesis through a differentiation-based mechanism. *Developmental Cell* 50
- Krah, N.M., Willet, S.G., Pan, F.C., Cash, G.M., Bronner, M.P., Wright, C.V., MacDonald, R.J. & Murtaugh, L.C. (2015) The acinar differentiation determinant PTF1A inhibits initiation of pancreatic ductal adenocarcinoma. *eLife*
- Lin, J.W., Biankin, A.V., Horb, M.E., Ghosh, B., Prasad, N.B., Yee, N.S., Pack, M.A. & Leach, S.D. (2004) Differential requirement for *ptf1a* in endocrine and exocrine lineages of developing zebrafish pancreas. *Developmental Biology* 270, 474–486
- Maduro, M.F. & Rothman, J.H. (2002) Making Worm Guts: The Gene Regulatory Network of the *Caenorhabditis elegans* Endoderm. *Developmental Biology* 246, 68–85
- Maston, G.A., Evans, S.K. & Green, M.R. (2006) Transcriptional Regulatory Elements in the Human Genome. *Annu Rev Genom Hum Genet* 7, 29–59
- Ong, C.-T. & Corces, V.G. (2011) Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet* 12, 283–293
- Piñero, J., Queralt-Rosinach, N., Bravo, À., Deu-Pons, J., Bauer-Mehren, A., Baron, M., Sanz, F. & Furlong, L. (2015) DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database tool*
- Pourquié, O. (2003) The Segmentation Clock: Converting Embryonic Time into Spatial Pattern. *Science* 301, 328–330
- Rose, S.D., Swift, G.H., Peyton, M.J., Hammer, R.E. & MacDonald, R.J. (2001) The Role of PTF1-P48 in Pancreatic Acinar Gene Expression. *J Biol Chem* 276, 44018–44026
- Schaffer, A.E., Freude, K.K., Nelson, S.B. & Sander, M. (2010) Nkx6 Transcription Factors and Ptf1a Function as Antagonistic Lineage Determinants in Multipotent Pancreatic Progenitors. *Developmental Cell* 18, 1022–1029
- Wang, Y.J., Park, J.T., Parsons, M.J. & Leach, S.D. (2015) Fate mapping of *ptf1a* -expressing cells during pancreatic organogenesis and regeneration in zebrafish. *Dev Dyn* 244, 724–735
- Weedon, Michael N., Cebola, I., Patch, A.-M., Flanagan, S.E., De Franco, E., Caswell, R., Rodríguez-Seguí, S.A., Shaw-Smith, C., Cho, C.H.-H., Allen, H.L., et al (2014) Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. *Nat Genet* 46, 61–64
- Zhou, Q., Law, A.C., Rajagopal, J., Anderson, W.J., Gray, P.A. & Melton, D.A. (2007) A Multipotent Progenitor Domain Guides Pancreatic Organogenesis. *Developmental Cell* 13, 103–114