



Exploring zebrafish spinal cord development and tissue regeneration through transcriptomic data analysis

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und
Biowissenschaften des Karlsruher Instituts für
Technologie (KIT) genehmigte

DISSERTATION

von

Fushun Chen, M.Sc.

aus Hebei, China

1. Referent: Dr. Sepand Rastegar

2. Referent: Prof. Dr. Lennart Hilbert

Tag der mündlichen Prüfung: 15.07.2024

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet. Die Satzung des Karlsruher Instituts für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis wurde beachtet. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht. Die Richtigkeit der vorstehenden Erklärungen bestätige ich. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Datum: _____

Unterschrift: _____

Summary

This thesis studies the molecular mechanisms involved in spinal cord neuronal development and tissue regeneration in zebrafish. We focus mainly on the identification and characterisation of transcription factors (TFs), marker genes and signalling pathways throughout the developmental and regenerative process in zebrafish using transcriptional and epigenomic data analysis.

In the first part of this thesis, the neuronal diversity of *sox1a* transgenic line in zebrafish ventral spinal cord was explored using single-cell RNA sequencing. Distinct cell populations expressing *sox1a* were identified, including KA', KA'', V2-pre, and V2s neurons and a new *sox1a*-positive cell type termed ISNs. Our results further show the different cell clusters within the ISN cells, including precursor cell clusters (*fev*⁺, *tph2*⁻) and mature neuron cell clusters (*fev*⁺, *tph2*⁺). Notably, It was discovered that ISNs originate from the lateral floor plate (LFP) based on the pseudotime analysis of our *sox1a:gfp* and published *olig2:gfp* datasets, and our data indicate a negative regulatory role for Notch signaling during the differentiation of LFP cells into ISNs.

In the second part, the aim was to characterize both the shared transcriptional regulation programs and tissue-specific regulatory mechanisms during fin, heart, retina, and spinal cord regeneration processes in zebrafish through Omics data analysis. The results revealed a conserved set of 159 upregulated genes in all four tissues during regeneration through bulk RNA-seq analysis, particularly enriched in pathways associated with cell cycle regulation, DNA replication, and regeneration. Importantly, these genes are associated with cell types exhibiting regenerative potential and immune-related functions in different tissues, as indicated by scRNA-seq analysis. For example, *tubb5* and *cdh2* are critical marker genes in the regeneration process. Following injury, these genes exhibit higher expression levels in specific cell types, including mesenchymal cells in the fin, fibroblasts and epicardial cells in the heart, activated Müller glial cells in the retina,

and ependymo-radial glial cells in the spinal cord. Furthermore, we identified 1085 shared chromatin accessibility regions after fin, heart, and retina injury through ATAC-seq analysis. Within these regions, we observed enrichment of motifs such as Fos, AP-1, Fosl2, Jun-AP, NF-E2, and Nrf2, suggesting potential binding sites for transcription factors that may regulate the regeneration response.

Zusammenfassung

Diese Dissertation untersucht die molekularen Mechanismen, die an der Entwicklung von Rückenmarksneuronen und der Geweberegeneration bei Zebrafischen beteiligt sind. Wir konzentrieren uns hauptsächlich auf die Identifizierung und Charakterisierung von Transkriptionsfaktoren (TFs), Markergenen und Signalwegen während des Entwicklungs- und Regenerationsprozesses bei Zebrafischen unter Verwendung von Transkriptions- und Epigenom-Datenanalysen.

Im ersten Teil dieser Dissertation wurde die neuronale Vielfalt der *sox1a*-transgenen Linie im ventralen Rückenmark von Zebrafischen unter Verwendung der Einzelzell-RNA-Sequenzierung untersucht. Es wurden verschiedene Zellpopulationen identifiziert, die Sox1a exprimieren, darunter KA', KA'', V2-pre und V2s Neuronen sowie ein neuer *sox1a*-positiver Zelltyp namens ISNs. Unsere Ergebnisse zeigen weiterhin die verschiedenen Zellcluster innerhalb der ISN-Zellen, darunter Vorläuferzellcluster (*fev*+, *tph2*-) und reife Neuronenzellcluster (*fev*+, *tph2*+). Insbesondere wurde entdeckt, dass ISNs aus der lateralen Bodenplatte (LFP) stammen, basierend auf der Pseudozeitanalyse unserer *sox1a:gfp* und veröffentlichten *olig2:gfp* Datensätze, und unsere Daten weisen auf eine negative regulatorische Rolle für die Notch-Signalgebung während der Differenzierung von LFP-Zellen in ISNs hin.

Im zweiten Teil war es das Ziel, sowohl die gemeinsamen Transkriptionsregulationsprogramme als auch die gewebespezifischen Regulationsmechanismen während der Regenerationsprozesse von Flosse, Herz, Netzhaut und Rückenmark im Zebrafisch durch die Analyse von Omics-Daten zu charakterisieren. Die Ergebnisse zeigten eine konsistente Gruppe von 159 hochregulierten Genen in allen vier Geweben während der Regeneration durch Massen-RNA-seq-Analyse, die besonders in Pfaden angereichert sind, die mit der Regulierung des Zellzyklus, der DNA-Replikation und der Regeneration verbunden sind. Wichtig ist, dass diese Gene mit Zelltypen verbunden sind, die ein regeneratives Potenzial und

immunbezogene Funktionen in verschiedenen Geweben aufweisen, wie die scRNA-seq-Analyse zeigt. So sind beispielsweise *tubb5* und *cdh2* wichtige Markergene im Regenerationsprozess. Nach einer Verletzung weisen diese Gene höhere Expressionswerte in bestimmten Zelltypen auf, darunter mesenchymale Zellen in der Flosse, Fibroblasten und epikardiale Zellen im Herzen, aktivierte Müller-Gliazellen in der Retina und ependymo-radiale Gliazellen im Rückenmark. Darüber hinaus identifizierten wir durch ATAC-seq-Analyse 1085 gemeinsame Chromatin-Zugänglichkeitsregionen nach Verletzungen der Flossen, des Herzens und der Netzhaut. Innerhalb dieser Regionen beobachteten wir eine Anreicherung von Motiven wie Fos, AP-1, Fosl2, Jun-AP, NF-E2 und Nrf2, was auf potenzielle Bindungsstellen für Transkriptionsfaktoren hinweist, die die Regenerationsreaktion regulieren können.

Table of Contents

Erklärung.....	ii
Summary.....	iv
Zusammenfassung	vi
Table of Contents	viii
List of tables.....	xiii
List of Figures	xiv
List of Abbreviations	xvi
Publications	xix
Acknowledgements.....	xx
Chapter 1: General introduction.....	22
1.1 Gene regulatory's key role in development and regeneration	22
1.2 Zebrafish as a valuable model for studying development and tissue regeneration	23
1.3 High throughput genomic technology in development and regeneration.....	24
1.3.1 Bulk RNA Sequencing	24
1.3.2 Single-Cell RNA Sequencing	26
Chapter 2: Exploring neuronal diversity in the zebrafish ventral spinal cord using single- cell RNA sequencing of the <i>sox1a:gfp</i> lineage	30
2.1 Introduction.....	30
2.1.1 Nervous system.....	30

2.1.2	The development of central nervous system in zebrafish	31
2.1.3	The dorsal-ventral axis in the neural tube	33
2.1.4	The organization of the embryonic spinal cord along the dorsal-ventral axis in zebrafish	34
2.1.5	Sonic Hedgehog signaling	36
2.1.6	Notch signaling pathway	37
2.1.7	The Notch signaling mediates neuronal determination.....	39
2.1.8	The p2 domain gives rise to three distinct V2 neuron subtypes.....	40
2.1.9	The generation of KA'' and V3 from lateral floor plate in zebrafish.....	41
2.1.10	The Intraspinal serotonergic neurons in zebrafish.....	43
2.1.11	The transcription factor Sox1	44
2.2	The aim	45
2.3	The Method and Material	46
2.3.1	Fish Stock Management and Embryo Collection	46
2.3.2	Cell dissociation.....	46
2.3.3	Single-Cell RNA Sequencing and Generation of Count Matrix	47
2.3.4	Preprocessing Single-Cell RNA Sequencing Data	47
2.3.5	Preliminary quality control.....	48
2.3.6	Integration of Datasets	48
2.3.7	Identification of main cell types	49
2.3.8	Iterative clustering refinement of neurons and neural progenitor cells.....	49
2.3.9	Subclustering of ISN	50
2.3.10	Trajectory inference and Pseudotime analysis of ISN cell lineage	50

2.3.11	Reanalysis of LFP, V3 precursors and V3 cells using Tg <i>olig2</i> scRNAseq data	51
2.3.12	Integration of ISN and V3 cell lineage	51
2.3.13	Data and code availability	52
2.4	Result	53
2.4.1	Analysis of scRNA-seq data from the <i>sox1a</i> transgenic lines of zebrafish spinal cord	53
2.4.2	The identification of cell clusters based on transcription markers.....	55
2.4.3	Refined Analysis of Neurons and Neural Progenitor Cells.....	57
2.4.4	The identification of cell subtypes within neural progenitors and neurons	58
2.4.5	The ISN cluster is made of precursor and mature neurons	61
2.4.6	ISN origins from LFP	62
2.4.7	Notch signaling controls LFP cells differentiation into ISNs	65
2.4.8	ISN-pre cells were found in the Tg(<i>olig2:GFP</i>) scRNA-seq dataset.....	67
2.4.9	The comprehensive analysis of integrated scRNA-seq dataset from <i>olig2:eGFP+</i> and <i>sox1a:eGFP+</i> cells confirms the common origin of V3 and ISN .	69
2.5	Discussion	73
2.5.1	LFP progenitor cells generate three distinct ventral neuronal populations with unique molecular profiles and functions.....	73
2.5.2	Diverse genetic programs for the development of KA", V3, and ISN from LFP progenitor cells.....	74
2.5.3	Limitation of the study.....	75
2.6	Supplementary Figure	76

Chapter 3: Integrative analysis of transcriptomes and chromatin accessibility reveals insights into gene regulatory networks across tissues during regeneration	78
3.1 Introduction.....	78
3.1.1 The importance of studying regeneration.....	78
3.1.2 The mechanisms of tissue regeneration in zebrafish.....	79
3.1.3 Fin regeneration in zebrafish.....	80
3.1.4 Heart regeneration in zebrafish.....	80
3.1.5 Retina regeneration in zebrafish.....	81
3.1.6 Spinal cord regeneration in zebrafish	82
3.1.7 Genomics research in zebrafish regeneration.....	82
3.2 The Aim.....	85
3.3 The Method	86
3.3.1 The data collection on regenerating tissues including fin, heart, retina, and spinal cord	86
3.3.2 Bulk RNA-seq analysis on datasets from fin, heart, retina, and spinal cord	87
3.3.3 scRNA-seq analysis for fin, heart, retina, and spinal cord.....	88
3.3.4 ATAC-seq analysis for fin, heart, retina, and spinal cord	90
3.4 Result	92
3.4.1 Bulk RNA-seq profiling of regenerating tissues including fin, heart, retina, and spinal cord	92
3.4.2 Tissue specific patterns in zebrafish regeneration.....	95

3.4.3	scRNA-seq profiling of regenerating tissues including fin, heart, retina, and spinal cord	97
3.4.4	ATAC-seq profiling of regenerating tissues including fin, heart, and retina	113
3.5	Discussion	118
3.5.1	Zebrafish employ varied mechanisms to regenerate distinct tissues	118
3.5.2	Limitation of the study.....	119
	References	120

List of tables

Table 3.1 Sequence Data Sources	87
Table 3.2 The number of differentially expressed genes.....	92
Table 3.3 The GO: BP items related the regeneration.....	95
Table 3.4 The number of differentially expressed regions	114

List of Figures

Figure 1.3.1 The workflow of experimental steps in bulk RNA-seq.	25
Figure 1.3.2 The workflow of scRNA-seq.....	27
Figure 2.1.1 Neural tube formation during neurulation involves two distinct strategies.	32
Figure 2.1.2 Diagrams illustrate spinal cord neuronal development guided by opposing morphogen gradients along the dorsal-ventral axis.	35
Figure 2.1.3 The Sonic Hedgehog signaling pathway.	37
Figure 2.1.4 The Notch signaling pathway in both zebrafish and mouse.	39
Figure 2.4.1 The workflow of scRNA-seq and preliminary data analysis.....	54
Figure 2.4.2 Comprehensive analysis of cell types and developmental dynamics in <i>sox1a:gfp</i> line.....	57
Figure 2.4.3 Quantification, clustering, and developmental distribution of neural progenitors and neurons.	58
Figure 2.4.4 Annotation of cell subtypes within neural progenitors and neurons.	60
Figure 2.4.5 The sub-cluster analysis of ISNs reveals the existence of two distinct groups: ISN-pre and mature ISNs.	62
Figure 2.4.6 Analysis of transcription factor expression, trajectory, and pseudotime dynamics in LFP, ISN-pre, and ISN Cells.	64
Figure 2.4.7 Notch signaling is involved in the differentiation of LFP into ISN.....	66
Figure 2.4.8 Confirmation of ISN origin from LFP using <i>olig2:gfp</i> scRNA-seq data.	68
Figure 2.4.9 Establishing cell lineages of V3 and ISN through pseudo-time analysis using scRNA-seq data from <i>olig2:eGFP+</i> and <i>sox1a:eGFP+</i>	70

Figure 2.4.10 Comparison of pseudotime-dependent gene expression profiles between V3 and ISN cell lineages derived from LFP progenitors.....	72
Figure S2.6.1 Cell type distribution and gene expression across developmental stages.	77
Figure 3.4.1 Comparison of upregulated genes across tissues.	93
Figure 3.4.2 Functional analysis of 159 commonly upregulated genes.	94
Figure 3.4.3 Specific upregulated genes in zebrafish fin regeneration.	96
Figure 3.4.4 Specific upregulated genes in zebrafish heart, spinal cord, and retina regeneration.....	97
Figure 3.4.5 The identification of cell type in zebrafish fin.....	99
Figure 3.4.6 The identification of cell type in zebrafish heart.....	100
Figure 3.4.7 The identification of cell type in zebrafish retina.....	101
Figure 3.4.8 The identification of cell type in zebrafish spinal cord.	102
Figure 3.4.9 AUCCell analysis of commonly upregulated genes across cell types in fin, heart, retina, and spinal cord scRNA-seq datasets.	104
Figure 3.4.10 The expression of <i>lepb</i> , <i>tubb5</i> and <i>cdh2</i> in zebrafish fin.	107
Figure 3.4.11 The expression of <i>lepb</i> , <i>tubb5</i> and <i>cdh2</i> in zebrafish heart.....	109
Figure 3.4.12 The expression of <i>lepb</i> , <i>tubb5</i> and <i>cdh2</i> in zebrafish retina.....	110
Figure 3.4.13 The expression of <i>lepb</i> , <i>tubb5</i> and <i>cdh2</i> in zebrafish spinal cord.	112
Figure 3.4.14 The expression of <i>krt93</i> and <i>krt94</i> in zebrafish fin.	113
Figure 3.4.15 Comparative analysis of upregulated regions in fin, heart, and retina....	116
Figure 3.4.16 The enriched motifs in 1085 commonly upregulated regions.....	117

List of Abbreviations

Ascl	Achaete-scute like
BMP	bone morphogenetic protein
BMP	Bone morphogenetic protein
cDNA	the complimentary DNA
CNS	central nervous system
ddc	dopa decarboxylase
dl	dorsal interneuron
dP	dorsal progenitor
dpf	days post-fertilization
ECM	extra cellular matrix
EGF	Epidermal Growth Factor
EMT	epithelial-mesenchymal transition
ERG	ependymal radial glial cells
FACS	fluorescence-activated cell sorting
FGF	Fibroblast growth factor
gch1	GTP cyclohydrolase 1
GLU IN	glutamatergic inhibitory neurons
GO	Gene Ontology
Hes/Her	hairy/enhancer of the split family of transcription factors
Hh	Hedgehog
hpf	hours post-fertilization
IGF	Insulin-like growth factor

ISN-pre	ISN precursor cells
ISNs	Intraspinal serotonergic neurons
KC	keratinocyte cells
KEGG	Kyoto Encyclopedia of Genes and Genomes
lepb	leptin b
LFP	lateral floor plate
Imx1bb	LIM homeobox transcription factor 1 beta b
MC	mesodermal cells
MFP	medial floor plate
MI	Myocardial infarction
MN	motor neurons
MNs	motor neurons
NICD	Notch intracellular domain
NMDA	N-methyl-D-aspartate
NPC	neural precursor cells
NPC	neural precursor cells
NPro	Neural Progenitors
OPC	oligodendrocyte precursor cells
opr1	opiate receptor-like 1
PCR	polymerase chain reaction
PCs	principal components
PCs	principal components
PDGF	Platelet-Derived Growth Factor
pMN	motoneuronal progenitors

PNS	peripheral nervous system
Pre-V3 IN	V3 precursors interneurons
RA	retinoic acid
scRNA-seq	single-cell RNA sequencing
Shh	sonic hedgehog
Sox	the SRY (sex-determining region Y)-box
syt1a	synaptotagmin Ia
TFs	transcription factors
TGF- β	Transforming Growth Factor beta
tph2	tryptophan hydroxylase 2
UMIs	unique molecule identifiers
V2-pre	V2 precursor cells
V3 IN	V3 interneurons
wpi	weeks post transection injury
WT	Wild-type

Publications

1. **Chen, F.**, Köhler, M., Cucun, G., Takamiya, M., Kizil, C., Cosacak, M. & Rastegar, S. *sox1a:eGFP* transgenic line and single-cell transcriptomics reveal the origin of zebrafish intraspinal serotonergic neurons. *iScience*. 2023 Jul 11;26(8):107342.
2. Zhang, G., Lübke, L., **Chen, F.**, Beil, T., et al. Neuron-radial Glial cell communication via BMP/ID1 signaling is key to long-term maintenance of the regenerative capacity of the adult Zebrafish Telencephalon. *Cells* 10, 5–8 (2021).

Acknowledgements

First, I would like to sincerely thank my mentor, Dr. Sepand Rastegar, for providing me with the opportunity to join your team. Your guidance and support have been invaluable throughout my three-year PhD career.

I would also like to express my gratitude to all our current and former lab members: Tanja, Anne, Jincan, Gökhan, Melina, Sabrina, Christian, Daniel, Masa, Marcel, Agnes, Luisa and Gaoqun for creating a harmonious laboratory environment with their kindness.

I would like to express my sincere gratitude to my Thesis Advisory Committee (TAC) members, Dr. Thomas Dickmeis and Dr. Sanchez Iranzo Hector, for their exceptional guidance and insights shared during TAC meetings and lab discussions. Their support has greatly enriched my research work.

I am deeply grateful to my PhD Examination Committee, Prof. Dr. Martin Bastmeyer, Prof. Dr. Lennart Hilbert, Prof. Dr. Jörg Kämper, Prof. Dr. Natalia Requena, and Priv.-Doz. Dr. Carsten Weiss, for dedicating their valuable time to read my thesis and participate in my examination. Your guidance and support throughout this process have significantly improve my research.

I extend my appreciation to Pro. Dr. Ferenc Mueller, Dr. Joseph Wragg, Dr. Yavor Hadzhiev, Dr. Damir Baranasic, Dr. Ildem Akerman, and Matthias Hörtenhuber, the collaborators of the brain regeneration project. Although it is a pity that this project was not completed in the end, I am grateful for the wealth of knowledge I gained about ATAC-seq and CAGE-seq through their expertise and guidance. Their kindness and helpfulness throughout the entire process was invaluable.

I am very grateful to Dr. Mehmet Ilyas Cosacak and Pro. Dr. Caghan Kizil, collaborators on the *sox1a* project, for their invaluable assistance in sequencing data, preliminary data processing, and I appreciate their insightful data analysis advice and methods.

I would like to thank Master Teachers, Dr. Zhiguang Li, Dr. Yalu Wen, Dr. Xiaodi Deng, and Dr. Dekang Lv for bringing me to the field of bioinformatics and offering invaluable support during my initial studies. I am especially grateful to Dr. Li and Dr. Wen for their letters of recommendation, which were of great help to my doctoral application.

Finally, I would like to express my deepest gratitude to my family for their unwavering selflessness and caring. They always encourage me to chase my dreams. Their unwavering support has provided a solid foundation for me. I would especially like to thank my girlfriend, Lu Zhang. During the most difficult moments of my life, her constant encouragement was the source of my strength. Her support was crucial in motivating me to pursue my PhD and embark on my current academic journey.

Chapter 1: General introduction

1.1 Gene regulatory's key role in development and regeneration

Gene regulation plays a central role in development and regeneration processes. This regulatory process relies on signaling pathways and downstream transcription factors, acting in both temporal and spatial dimensions (Goldman & Poss, 2020; Ho, 2012). Together, these components coordinate cell differentiation and tissue formation and are essential components of mechanisms that control development, repair, and regeneration (Goldman & Poss, 2020).

During development, gene regulatory networks drive cell fate decisions and morphogenesis through the activation of specific genes while repressing other genes that constitute the complex architecture of tissues and organs, including the spinal cord (Sagner & Briscoe, 2019; Mendoza-Parra *et al*, 2016). This complexity is underscored during spinal cord development along the dorsoventral axis, which is controlled by the combinatorial interaction of morphogen gradients and transcription factors (Kutejova *et al*, 2016; Sagner & Briscoe, 2019). Key signaling pathways, such as sonic hedgehog (Shh), bone morphogenetic protein (BMP), and Wnt, play important roles in directing the differentiation of neural progenitor cells into distinct neuronal subtypes and specifying their positional identity along the dorsoventral axis (Sagner & Briscoe, 2019; Manzari-Tavakoli *et al*, 2022). At the same time, transcription factors coordinate the formation of specific neuronal populations, such as motor neurons, interneurons, and sensory neurons, contributing to the functional diversity of the spinal cord (Sagner & Briscoe, 2019; Andrews *et al*, 2019).

Additionally, The gene regulatory networks are essential for coordinating cellular responses and promoting tissue growth and repair processes (Goldman & Poss, 2020). For instance, spinal cord regeneration in zebrafish is governed by dynamic changes in

gene expression, characterized by the reactivation of developmental pathways and induction of regeneration-specific genes like *hb-egfa* (Cigliola *et al*, 2023). Signaling molecules such as Shh, Fgf3 and TNF- α intricately regulate cell proliferation and differentiation during spinal cord regeneration (Tsata & Wehner, 2021; Becker & Becker, 2020).

Therefore, this thesis investigates the complex mechanisms of gene regulation during spinal cord development and tissue regeneration, utilizing high-throughput genomic techniques such as bulk RNA-seq and scRNA-seq. By employing the zebrafish model, our aim is to gain a deeper understanding of the gene regulatory mechanisms underlying these biological processes. This not only provides valuable insights into spinal cord development but also holds great potential for improving the limited regenerative capabilities of higher vertebrates, including humans.

1.2 Zebrafish as a valuable model for studying development and tissue regeneration

The zebrafish is an excellent model for studying tissue development and regeneration, thanks to several key features. Firstly, zebrafish embryos are transparent, making it easy to visualize neuronal development and tissue growth *in vivo* (Shi *et al*, 2015; Schmidt *et al*, 2013). Moreover, zebrafish possess a high degree of genetic similarity to higher vertebrates, including humans, since around 70% of human genes are found in zebrafish (Howe *et al*, 2013). The zebrafish is highly reproductive and has numerous offspring, making it an ideal organism for high-throughput screening (Hajam *et al*, 2021). The efficient and productive nature of zebrafish not only saves time and helps improve the efficiency of experimental design, but also helps improve the statistical robustness of research results, thereby enhancing the reliability and significance of research results (Shi *et al*, 2015). In addition, Zebrafish have remarkable regenerative capabilities, including regeneration of fin, heart, spinal cord, retina, brain, and other tissues (Shi *et al*, 2015). The zebrafish's ability to efficiently repair and replace damaged structures enables us to

observe and understand the cellular and molecular mechanisms that govern tissue regeneration.

1.3 High throughput genomic technology in development and regeneration

1.3.1 Bulk RNA Sequencing

Currently, high-throughput RNA-seq is the most common and important method for studying the complete profile of gene expression at the genomic level due to its quantitative precision (Hegenbarth *et al*, 2022). It has become the widely used tool in investigating the complexities of cancer biology, developmental biology, and clinical research (Peymani *et al*, 2022; Best *et al*, 2015; White *et al*, 2017).

Bulk RNA-seq is used to analyze and quantify the overall transcriptional profile of a sample. The main application of bulk RNA-seq is to study average gene expression patterns and identify genes differentially expressed under different experimental conditions and in different tissues. In general, the RNA-seq workflow comprises several steps: RNA extraction, conversion of RNA to cDNA, adapter ligation, amplification, sequencing and data analysis (Figure 1.3.1) (Head *et al*, 2014).

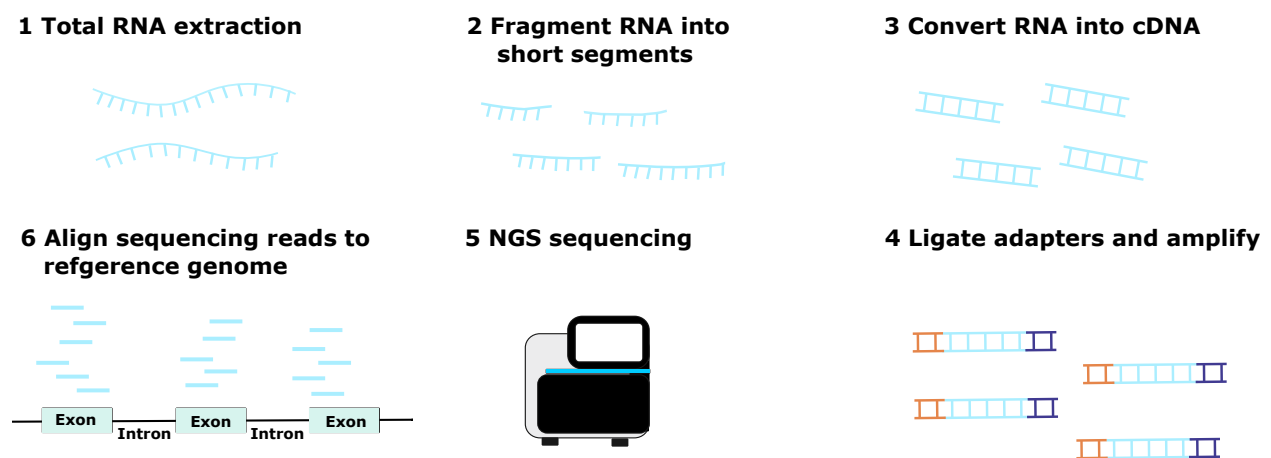


Figure 1.3.1 The workflow of experimental steps in bulk RNA-seq.

RNA is extracted from the sample of interest and fragmented. Subsequently, the fragmented RNA is converted into complementary DNA (cDNA). The cDNA fragments are then ligated to sequencing adapters to create a library. Following library preparation, amplification is performed to increase the amount of DNA for sequencing. Finally, the generated reads are mapped against a reference genome or transcriptome. Modified from (Van den Berge *et al*, 2019)

The main application of bulk RNA-seq is to identify genes that are differentially expressed between different conditions, such as treated versus untreated samples. In developmental biology, studying changes in gene expression at different stages of development can help us understand the molecular mechanisms underlying the developmental process. In this study by White *et al*, a complete time course of mRNA expression during zebrafish development was generated over 18 time points, from 1 cell to 5 days post-fertilization. They characterized the temporal expression profiles of 23,642 genes. They also identified a temporal and functional co-variation of transcripts that associates 5024 unnamed genes with distinct developmental time points (White *et al*, 2017). In addition, bulk RNA-seq has also been widely used in the study of regeneration. It can help identify injury-specific genes during regeneration. For example, researchers have identified 7762 genes differentially expressed during spinal cord regeneration, which are enriched in axon regeneration and the cell cycle (Shen *et al*, 2021).

Bulk RNA-seq uses a population of cells or tissues as starting material to provide an average gene expression profile. It cannot account for the complexity of cellular heterogeneity within a tissue or tumor. It masks individual variations present in specific cell subtypes. The disadvantage of bulk RNA analysis is that it can mask signals from rare cell populations or specific cell types. This can prevent a complete understanding of biological processes. To overcome this drawback, single-cell RNA sequencing (scRNA-

seq) has emerged as a powerful alternative for examining gene expression patterns in greater detail at the level of individual cells (Li & Wang, 2021).

1.3.2 Single-Cell RNA Sequencing

The advent of single-cell RNA sequencing (scRNA-seq) represents a revolutionary breakthrough in genomic research and marks an important milestone in our evolving understanding of the genome. This cutting-edge technology enables researchers to analyze gene expression at the level of individual cells, providing unprecedented insights into cellular diversity, heterogeneity, and dynamics in complex biological systems. It can help us explore cellular behavior and function more precisely and comprehensively. scRNA-seq is primarily employed to assess transcriptional similarities and differences within a population of individual cells. Consequently, the analysis of cellular heterogeneity stands as a fundamental motivation for undertaking scRNA-seq studies (Haque *et al*, 2017). In addition to providing profound insights into cell composition, single-cell studies offer insight into cellular plasticity during dynamic processes such as development and differentiation (Lafzi *et al*, 2018). This innovation greatly enhances our ability to unravel the complexity of cellular processes, contributing to advance in various fields including developmental biology, disease research and personalized medicine (Li & Wang, 2021; Hegenbarth *et al*, 2022).

scRNA-seq enables increasingly larger-scale experiments. Rapid advances in scRNA-seq experimental scalability have been attributed to the integration of automation and complexed microfluidics systems (Zhou *et al*, 2021). This advance not only improves the efficiency of data generation but also opens new possibilities for in-depth exploration of cellular heterogeneity and dynamics at the single-cell level (Lafzi *et al*, 2018). To obtain meaningful scientific results from single-cell datasets, it is important to develop tailored experimental designs. This includes considerations such as sample type, cell number, and preparation methods, sequencing parameters, and computational analysis strategies. The effectiveness of single-cell studies depends on the careful tailoring of these key elements to ensure results are reliable and interpretable (Lafzi *et al*, 2018).

The scRNA-seq workflow usually involved these keys steps: extraction and preparation of tissue, the isolation of single-cell and library preparation, sequencing and primary analysis, and the visualization and interpretation of the obtained data (Fig 1.3.2) (Nguyen *et al*, 2018).

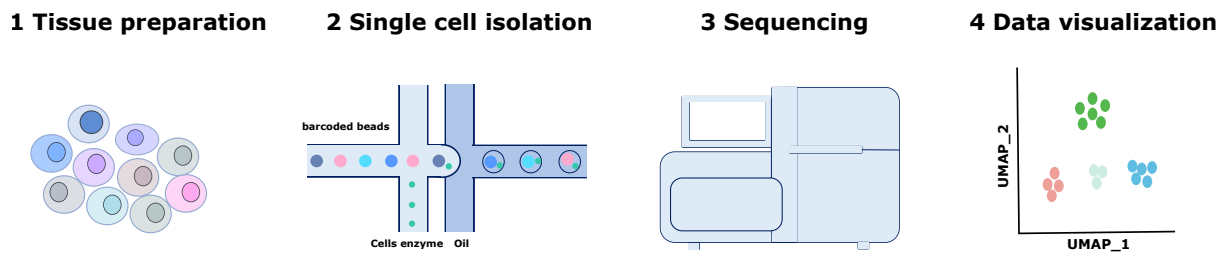


Figure 1.3.2 The workflow of scRNA-seq.

The single cell sequencing workflow includes several steps: beginning with initial tissue preparation, followed by the isolation of individual cells and library preparation. Then sequencing and primary data analysis like alignments to a genomic reference, followed by visualization and interpretation of the results to elucidate cellular heterogeneity and gene expression pattern.

Single-cell RNA sequencing (scRNA-seq) has become a widely adopted and popular technology for dissecting the complexity and heterogeneity of RNA transcripts within individual cells. This cutting-edge technology reveals not only the composition of different cell types, but also their functions in complex tissues, organs, or organisms. Researchers are also increasingly using scRNA-seq to explore how genes work together in individual cells, helping identify groups of genes that are regulated together. This approach could also help understand the connections between different cell types and trace lineages during processes such as embryonic development and cancer. As a powerful and versatile tool, scRNA-seq has applications in various fields of biological and medical research (Jovic *et al*, 2022). The primary application of scRNA-seq is to explore the

cellular heterogeneity, accomplishing this by evaluating transcriptional similarities and differences within a population of individual cells (Haque *et al*, 2017).

In developmental and regenerative research, scRNA-seq has been widely used to reveal the complexity of gene expression and cellular diversity. scRNA-seq has made significant contributions in several key areas of development, including embryonic and neural development, among others. In a study by Qiu *et al*, they used single-cell RNA-seq datasets covering mouse gastrulation and organogenesis. Covering 19 successive stages from E3.5 to E13.5, the researchers created a directed acyclic graph known as TOME (Trajectories of Mammalian Embryogenesis). TOME provides a comprehensive representation of the main cellular trajectories during mammalian development, linking cellular states between different stages. Researchers have used TOME to identify candidate transcription factors (TFs) that regulate cell type specification and "cell type homologs" across vertebrate evolution in a systematic manner (Qiu *et al*, 2022). Another study in zebrafish development, Farnsworth *et al.*, generated an atlas employing single-cell RNA-sequencing methods to investigate gene expression across 44,102 cells from the pharyngula to early larval stages. This dataset provides a valuable resource for exploring the transcriptional profiles of diverse cell types throughout zebrafish developmental stages (Farnsworth *et al*, 2020). Another two studies used single-cell mRNA sequencing to analyze the developing spinal cord in mice and humans, providing valuable insights into neural development. In mice, analysis spanning embryonic days 9.5 to 13.5 identified novel markers that revealed previously unnoticed temporal aspects of generating neuronal diversity in spinal cord (Delile *et al*, 2019). Meanwhile, human spinal cord studies covering 4-7 weeks of gestation have revealed dozens of different cell types, reconstructing subtype-specific differentiation pathways (Rayon *et al*, 2021). Both studies provide a comprehensive catalog of gene expression and cell type identification critical to understanding spinal cord development. These findings highlight the importance of single-cell RNA sequencing in revealing the complexity of development. In regeneration research, scRNA-seq is used to identify and characterize the cell types while checking the dynamic changes in gene expression. For example, Hoang *et al.* employed scRNA-

seq to check changes in gene expression at the cellular level before and after injury in the retina. Their findings revealed selective gene expression in reactive Müller glia which contributes the transition to a proliferative and neurogenic progenitor state in zebrafish (Hoang *et al*, 2020).

We are therefore using scRNA-seq technology to explore zebrafish spinal cord development (Chen *et al*, 2023) and tissue regeneration. This state-of-the-art approach can reveal dynamic changes in gene expression patterns at the cellular level during these biological processes.

Chapter 2: Exploring neuronal diversity in the zebrafish ventral spinal cord using single-cell RNA sequencing of the *sox1a:gfp* lineage

2.1 Introduction

2.1.1 Nervous system

The nervous system consists of a complex network of neurons that transmit signals between different parts of the body. It plays an important role in the control and coordination of various physiological processes in the body. It can enable the body to respond to internal and external stimuli. The nervous system is made up of two parts: the central nervous system (CNS) and the peripheral nervous system (PNS) (Xu *et al*, 2020; Ludwig *et al*, 2023).

The central nervous system (CNS) is the main part of the nervous system and is responsible for processing, interpreting, and coordinating information received from the peripheral nervous system (PNS) (Ludwig *et al*, 2023). It consists of two main parts: the brain and the spinal cord. The brain is the center of our body. It can help us think, understand things, move, and learn. It is a complex biological organ, connected through a complex neural network. It is divided into different regions; each region is responsible for performing different tasks. This complex organization allows the brain to efficiently process information, enabling us to perceive the world, think and act, and express complex behaviors such as language and emotion (Wilson & McNeill, 2007). The spinal cord is an elongated, tubular structure made of nervous tissue that originates in the medulla oblongata of the brainstem and extends into the lumbar region of the spine column, surrounding a central canal that contains cerebrospinal fluid (Ohtaki & Shioda, 2015). The spinal cord serves as a conduit for nerve signals to and from the brain and is critical for reflex movements and basic motor functions (Ludwig *et al*, 2023). Its main

function is to transmit nerve signals, transmitting signals from the motor cortex to the body, and at the same time transmitting signals from sensory neurons to the sensory cortex (Morimoto & Nakajima, 2019).

The peripheral nervous system (PNS) includes all nerves and ganglia outside the central nervous system (Lanigan *et al*, 2021). It helps the central nervous system communicate with the rest of the body. The peripheral nervous system transmits information from sensory receptors (such as touch, vision, and temperature) to the CNS, helping the brain to perceive and explore the surrounding external environment (Lanigan *et al*, 2021). The peripheral nervous system also transmits movement signals from the CNS to muscles and glands, allowing for voluntary and involuntary movement (Catala & Kubis, 2013). For example, it helps us pick up and move objects.

The central nervous system and the peripheral nervous system work together to ensure the precise control and coordination of diverse organs and systems, creating a harmonious unity within the body (Thau *et al*, 2022). Therefore, in-depth studies of nervous system development and regeneration are critical to enhance our understanding and treatment of neurological diseases.

2.1.2 The development of central nervous system in zebrafish

The development of the vertebrate central nervous system (CNS) begins with the formation of the neural plate from the ectoderm (Cearns *et al*, 2016; Blader, 2000). During neurulation, the neural plate undergoes major morphological transformations, culminating in the formation of the neural tube. In zebrafish, the neural plate first transforms into a solid structure known as the neural keel (Figure 2.1.1). Subsequently, a neurocoel or central canal develops through the secondary detachment of cells from the center of the neural tube (Blader, 2000). The neural tube extends along the longitudinal axis of the zebrafish embryo and is the precursor of the entire central nervous system. This tubular

structure extends from the anterior to the posterior end of the embryo, differentiating into different regions, including the brain and spinal cord (Gilbert, 2000).

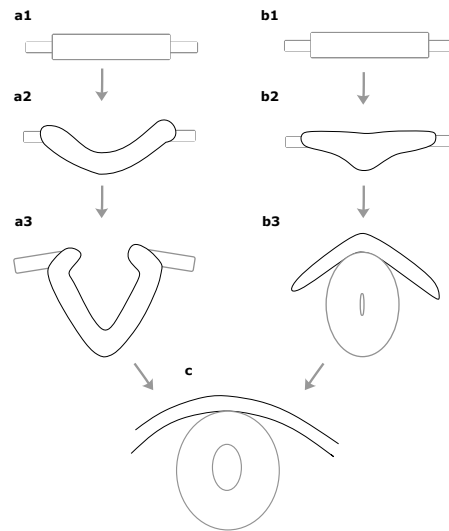


Figure 2.1.1 Neural tube formation during neurulation involves two distinct strategies.

(a1, a2, a3, c) In most vertebrates, the neural plate folds in on itself to create the neural tube. (b1, b2, b3, c) In teleosts such as zebrafish, neural plate cells fuse to form the neural keel, which then reorganizes into the neural stem before cavitation occurs to form the neural tube. Modified from (Cearns *et al*, 2016).

During early development, the anterior part of the neural tube undergoes significant changes. In this anterior region, the neural keel undergoes transformations that give rise to three major vesicles: the forebrain (prosencephalon), midbrain (mesencephalon), and hindbrain (rhombencephalon) (Gilbert, 2000). These primary vesicles serve as fundamental elements, serving as important building blocks that lay the foundation for the subsequent development of specialized brain regions. For example, the forebrain gives rise to structures such as the telencephalon and diencephalon.

Concurrently, the caudal segment of the neural tube elongates to give rise to the spinal cord. As development proceeds, the zebrafish spinal cord continues to elongate, and

different regions along its length become specialized. The spinal cord becomes segmented into distinct units (Kuzmicz-Kowalska & Kicheva, 2021; Gilbert, 2000).

Then the neural stem cells within the neural tube proliferate rapidly. During this process, the proliferating cells differentiate into various cell types, including neurons and glial cells. Subsequently, these differentiated cells migrate to their appropriate locations and contribute to the formation of specific brain regions and the spinal cord (Kuzmicz-Kowalska & Kicheva, 2021; Johansson *et al*, 1999).

Following the development of the brain and spinal cord, these intricate structures undergo further maturation and specialization, contributing to the overall functionality of the central nervous system (CNS). Neurons within the brain and spinal cord extend axons and dendrites, forming intricate connections known as synapses. These synapses create neural circuits, enabling communication between different regions of the CNS. Axons in both the brain and spinal cord undergo myelination, a process where a fatty substance called myelin is wrapped around the axons. Myelination enhances the speed and efficiency of nerve signal transmission (Simons & Nave, 2016). Different regions of the brain become specialized for specific functions, such as the processing of sensory information, motor coordination, and higher cognitive functions. In the spinal cord, different segments correspond to specific body regions, regulating motor and sensory functions (Silbereis *et al*, 2016). As the central nervous system matures, neural circuits become more elaborate. Functional integration allows coordinated responses to stimuli and the development of complex behaviors.

2.1.3 The dorsal-ventral axis in the neural tube

The neural tube exhibits polarity along its dorsal-ventral axis (Zheng *et al*, 2021). This axis is an important developmental framework that significantly influences the organizational structure of the central nervous system (CNS). Tissue differentiates into different cell types along the dorsal-ventral axis, contributing to the intricate functional diversity within the CNS (Kuzmicz-Kowalska & Kicheva, 2021).

Along this axis, specific regions are specialized to fulfill different functions. In the spinal cord, for instance, the dorsal part is associated with sensory functions, processing information related to touch, temperature, and pain. The ventral part is associated with motor functions, transmitting signals to control muscle movements (Gilbert, 2000). Cells that process these sensory and motor functions are organized along the dorsal-ventral axis. The different dorsal root ganglion sensory neurons can detect distinct types of sensory information and send afferent projections to the motor neurons or spinal dorsal horn, where they can synapse to spinal interneurons (Lupo *et al*, 2006). Then these dorsal interneurons process and forward this sensory information to the brain or motor neurons to regulate reflex motor neuron output. Motor neurons in the ventral horn manage motor output. These motor circuits involve motor neurons, whose axons exit the spinal cord through the ventral root to synapse with specific muscles (Andrews *et al*, 2019).

2.1.4 The organization of the embryonic spinal cord along the dorsal-ventral axis in zebrafish

The differentiation of cells along the dorsal-ventral axis is influenced by signals from its immediate environment during embryonic spinal cord development. The ventral spinal cord is influenced by the action of a key growth factor, Sonic hedgehog (Shh) which is produced by the notochord and floor plate, and it can induce ventral neural cell types in concentration-dependent manner. In contrast, the dorsal spinal cord is influenced by two key factors, one is the bone morphogenetic proteins (BMPs) which are produced by the roof plate and are responsible for the induction of dorsal neural cell types in a concentration-dependent manner (Figure 2.1.2A). Another one is Wnt families which also contribute the specification of dorsal neural cell types (Andrews *et al*, 2019).

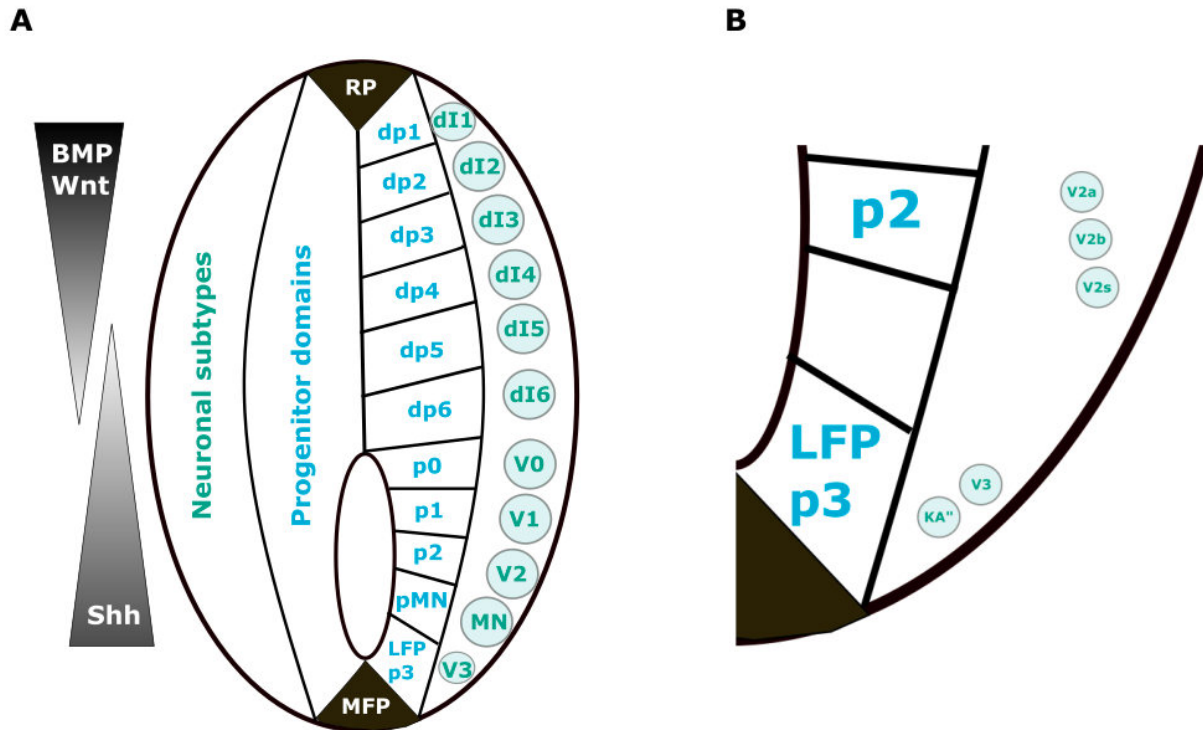


Figure 2.1.2 Diagrams illustrate spinal cord neuronal development guided by opposing morphogen gradients along the dorsal-ventral axis.

(A). The spinal cord is divided into six dorsal and five ventral progenitor domains along the dorsal-ventral axis. Sonic Hedgehog from the floor plate and Bone Morphogenetic Protein/Wnt from the roof plate guide this process. Each of these neural progenitor domains gives rise to distinct neuronal subtypes.

(B). LFP progenitors give rise to neuronal populations termed KA and V3. The p2 progenitor domain gives rise to three V2 interneuron subtypes, termed V2a, V2b, and V2s. Modified from (Cucun *et al*, 2024).

These signals compete and cooperate to specify different domains of progenitors along the dorsal-ventral axis during embryonic spinal cord development. The dorsal spinal cord is specified into six dorsal progenitor (dP) domains, dP1–dP6, which differentiate into corresponding dorsal interneuron (dI) populations, dI1–dI6. The ventral spinal cord is

specified into five progenitor (p) domains, p0-p2, pMN, and p3 or lateral floor plate (LFP), which differentiate into corresponding five distinct cell types: ventral interneurons (V) and motor neurons (MNs), further labeled as V0-V3 and MNs (Figure 2.1.2A). The expression of specific subsets of transcription factors in each neuronal population is influenced by the position of the cells along the dorsal-ventral axis. This spatial regulation of transcription factor activation contributes to the diversification of neuronal identities during development (Andrews *et al*, 2019). The same progenitor cell domain can give rise to different neuronal subtypes at different time points. For example, the p2 domain can generate three different V2 subtypes, known as V2a, V2b and V2s (referred to as V2c in mice). In addition, LFP can give rise to KA⁺ and V3 neurons (Figure 2.1.2A and B) (Cucun *et al*, 2024).

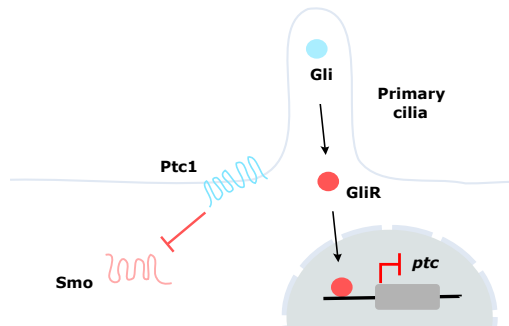
2.1.5 Sonic Hedgehog signaling

Sonic hedgehog (Shh) plays an important role as a morphogen in the spinal cord development which coordinating cell fate decisions along the dorsoventral axis (Zuñiga & Stoeckli, 2017). The spinal cord exhibits positional differences in cell fate that regulated by opposing gradients of morphogenes (Kuzmicz-Kowalska & Kicheva, 2021). Specifically, in the ventral spinal cord, the middle floor plate and notochord serve as the primary source of Shh, which acts as a ventral morphogen (Cucun *et al*, 2024).

The Sonic hedgehog signaling pathway influences cell differentiation through its interaction with the transmembrane receptor Patched1 (Ptc1) (Jacobs & Huang, 2021; Zuñiga & Stoeckli, 2017). In the absence of Shh, The Ptc1 inhibits the Shh signaling by preventing the localisation of Smoothened (Smo) to the primary cilium (Figure 2.1.3A). It can result in the processing of Gli proteins into a repressor form (Jacobs & Huang, 2021). After Shh binding to Ptc1, Smo is allowed to translocate to the primary cilium, inhibiting the proteolytic processing of Gli proteins and leading to the activation of Gli transcription factors (Jacobs & Huang, 2021). This activation triggers the transcription of

downstream target genes (e.g., *ptc*), which regulates the cellular response to Shh signaling (Figure 2.1.3B).

A Shh absent



B Shh present

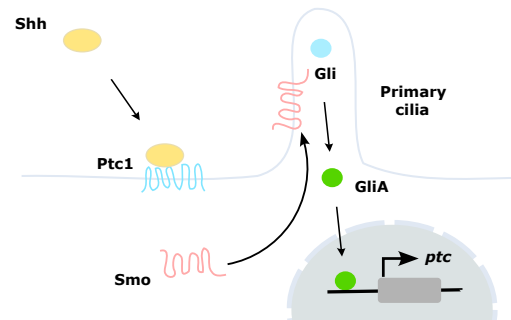


Figure 2.1.3 The Sonic Hedgehog signaling pathway.

A. In the absence of a Sonic hedgehog (Shh) ligand, Patched1 (Ptc1) is localized to the primary cilium and prevents the localization of Smoothened (Smo). This results in the processing of Gli transcription factors into a repressor form (GliR), which inhibits the transcription of downstream targets, such as *ptc*.

B. In the presence of a Shh ligand to Ptc1, Smo can localize to the primary cilium, where it inhibits the proteolytic processing of Gli proteins. This allows the activation of Gli transcription factors as transcriptional activators (GliA), driving transcription of downstream targets. Modified from (Jacobs & Huang, 2021).

2.1.6 Notch signaling pathway

The Notch signaling pathway is involved in regulating development and disease processes, especially in cell fate decisions, affecting differentiation and proliferation. Specifically, in embryonic neurogenesis, the Notch pathway coordinates binary cell fate decisions between adjacent cells, which is important for coordinating the complex process of neural development (Shin *et al*, 2007; Diotel *et al*, 2020; Kageyama *et al*, 2008). The pathway begins with Notch receptors which are transmembrane proteins found on the

surface of cells. Notch receptors are activated by binding to transmembrane ligands present on neighboring cells (Figure 2.1.4). Major ligands include the Delta and Jagged families (Andersson *et al*, 2011). Upon ligand binding, the Notch receptor undergoes a conformational change. The Notch receptor is further cleaved within the transmembrane domain by γ -secretase, resulting in the release of the Notch intracellular domain (NICD) (van Tetering & Vooijs, 2011). This NICD translocates to the nucleus and forms a complex with the Notch signaling mediator Rbpj (Diotel *et al*, 2020). This complex acts as a transcriptional activator to activate downstream genes, such as those in the hairy/enhancer of the split family of transcription factors (Hes/Her) (Yeo *et al*, 2007). Hes/Her proteins act as dimers and regulate neural fate decisions through transcriptional repression of genes. Specifically, they bind to DNA regulatory sequences associated with proneural genes like Achaete-scute like (Ascl), effectively suppressing their expression (Scholpp *et al*, 2009). Additionally, Hes/Her proteins exhibit negative auto-regulation, in which they can bind to regulatory elements within the promoters of their own genes, leading to the repression of their own expression (Figure 2.1.4) (Boareto *et al*, 2017). This mechanism serves to fine-tune intracellular levels of HES proteins, ensuring precise control of Notch signaling responses and promoting proper cell differentiation (Boareto *et al*, 2017; Diotel *et al*, 2020).

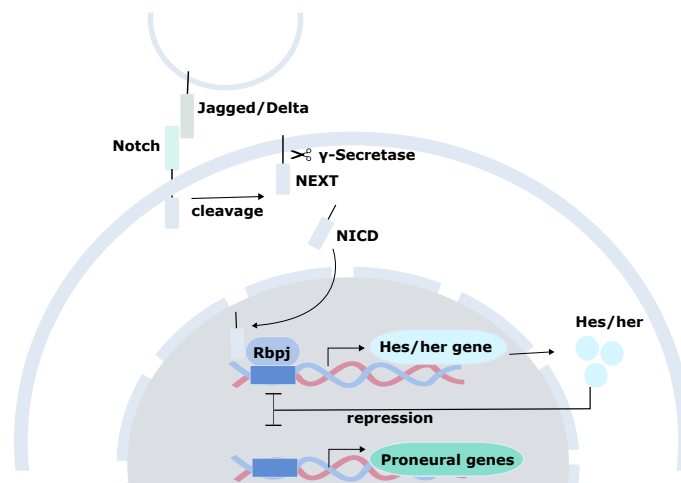


Figure 2.1.4 The Notch signaling pathway in both zebrafish and mouse.

Notch receptor interacts with its ligand Jagged/Delta. Then a sequential series of cleavage steps results in the release of the Notch intracellular domain (NICD), which then migrates into the nucleus. Inside the nucleus, the NICD collaborates with cofactors such as Rbpj to activate the expression of downstream genes, including Hes in mouse and her in zebrafish. The HES/Her proteins form homodimers, which then bind to the promoters of proneural genes to suppress neuronal fate. Furthermore, this dimer has the capacity to bind to its own promoter, leading to negative auto-regulation. Modified from (Diotel *et al*, 2020).

2.1.7 The Notch signaling mediates neuronal determination

The Notch signaling pathway is a fundamental and highly conserved cell signaling mechanism that regulates various aspects of embryonic development, including neurogenesis (Zhou *et al*, 2022). Notch signaling is characterized by its ability to influence binary cell fate decisions, such as promoting the maintenance of undifferentiated progenitor cells (Notch-active) or inducing their differentiation into specific neuronal (Notch-inactive) (Jacobs *et al*, 2022). This dual role of Notch signaling in maintaining progenitor cells and promoting differentiation ensures an appropriate balance between cell proliferation and differentiation during neurogenesis. Importantly, Notch signaling is dynamically regulated spatially and temporally during embryonic development, allowing for precision control over the timing and patterning of neurogenesis in different regions of the nervous system (Tokunaga *et al*, 2004).

During the early stages of spinal cord development in zebrafish embryos, progenitor cells give rise to different neuronal subtypes through a process called neuronal determination. Notch signaling plays a crucial role in this process by influencing the fate decisions of progenitor cells and promoting the generation of specific neuronal subtypes. Notch signaling has important effects on the generation of interneurons in the ventral spinal cord. Active Notch signaling promotes the differentiation of progenitor cells into V2a

interneurons, while sustained Notch activity contributes to the development of V2b interneurons (Cucun *et al*, 2024). Additionally, motor neuron precursors (pMN) exhibiting elevated levels of Notch signaling undergo differentiation into Kolmer-Agduhr' (KA') interneurons, while those with lower Notch activity give rise to primary motor neurons (MN) (Jacobs *et al*, 2022). Moreover, Notch signaling controls the fate of cells originating from the lateral floor plate (LFP) of the spinal cord. High Notch levels sustain LFP cells as progenitors, while V3 cells exhibit intermediate Notch activity, and KA" cells lack Notch response at early stage. V3 interneurons not only receive a greater level of Notch signaling compared to KA" interneurons, but they also accumulate this signaling over an extended duration. Continuous Notch signaling blocks the formation of KA" and V3 interneurons while expanding the LFP domain. Early termination favors KA" cells, whereas late termination enhances V3 interneurons (Jacobs *et al*, 2022; Cucun *et al*, 2024). Overall, these findings demonstrate that Notch signaling duration plays a crucial role in determining the fate of different cell types in the spinal cord, contributing to its cell type diversification during development.

2.1.8 The p2 domain gives rise to three distinct V2 neuron subtypes

The p2 domain can give rise to three different types of V2 neuron: V2a, V2b and V2s (Figure 2.1.2A and B). The p2 domain expresses several key TF genes, including *vsx1* and *foxn4*, which play a key role in controlling the cell differentiation and fate determination. p2 domain can give rise to two distinct sister progenitor groups. The first is comprised of V2a/b intermediate progenitor cells and it can generate both V2a and V2b interneurons. The second group consists of V2b/s intermediate progenitor cells. It contributes the development of both V2b and V2s interneurons. Shh and Notch signaling pathways play important roles in the differentiation V2 neuron subtypes (Batista *et al*, 2008; Cucun *et al*, 2024; Bello-Rojas & Bagnall, 2022).

V2a interneurons are identified as glutamatergic interneurons and play a crucial role in motor coordination and locomotion. These neurons distinguish themselves by distinctive

gene expression patterns, marked by activation of the *vsx1*, *vsx2*, and *slc17a6b* genes, emphasizing their special contribution to the neural circuits that control movement (Kimura *et al*, 2008; Cucun *et al*, 2024). V2b interneurons are characterized as GABAergic neurons. It mainly originates V2a/b intermediate progenitor cells. A small amount is generated from V2b/s intermediate progenitor cells. These V2b interneurons express specific genes, including *gata3*, *gata2a*, *tal1*, and *tal2* (Kimura *et al*, 2008; Cucun *et al*, 2024). V2s interneurons are recognized as inhibitory glycinergic neurons. It originates from V2b/s intermediate progenitor cells. These interneurons clearly express the *sox1a*, *sox1b*, and *nkx1.2lb* genes. And the expression and function of *sox1a* and *sox1b* play an important role in the development and differentiation of V2s interneurons (Gerber *et al*, 2019; Cucun *et al*, 2024; Panayi *et al*, 2010).

2.1.9 The generation of KA" and V3 from lateral floor plate in zebrafish

The floor plate is an important structure in the ventral neural tube of vertebrates. It plays a crucial role in the formation and development of spinal cord neurons. It has two different components: lateral floor plate (LFP) and the medial floor plate (MFP). The MFP is positioned at the center and the LFP is located on both sides (Schäfer *et al*, 2007). The cells in MFP can secrete signals and molecules that influence midline neuronal differentiation (Schäfer *et al*, 2005). The LFP is induced by Sonic Hedgehog (Shh) signaling which is produced by the notochord and MFP and spreads ventrally to form a gradient (Strähle *et al*, 2004; Odenthal *et al*, 2000). The LFP and MFP can be distinguished through the expression patterns of specific marker genes, including sonic hedgehog (*shh*), *foxa*, *nkx2.2a/b*, and *nkx2.9*. These genes expressed in LFP differ from those in MFP, as *nkx2.2a/b* and *nkx2.9* aren't expressed in MFP (Cucun *et al*, 2024; Norton *et al*, 2005). In zebrafish, the LFP is like the p3 progenitor domain of mouse and chick. It is involved in the development of interneurons. It can give rise to the GABAergic Kolmer–Agduhr neurons (KA") interneurons and V3 interneurons (Figure 2.1.2A and B)

(Cucun *et al*, 2024). Our finding shows that LFP is also one of the origins of ISN in zebrafish (Chen *et al*, 2023).

KA cells are categorized into two subtypes: KA' located dorsally and KA'' located ventrally. Both are GABAergic interneurons and shared several transcription factors, including *gata2a*, *gata3*, *tal1/2*, and *gad1/2*. But each has its specific marker genes; for example, *sst1.1* is expressed in KA', while *urp1* and *urp2* are expressed in KA''. The different spatial locations and distinct expression of transcription factors for each of the KA cell subtypes also suggests that KA' and KA'' arise from different developmental origins. KA' originates from motoneuronal progenitors (pMN) domain. In contrast, KA'' has a different origin which is from the lateral floor plate (LFP) at about 16.5 hours postfertilization (hpf) (Cucun *et al*, 2024; Yang *et al*, 2020). The transcription factors and signaling pathways play crucial roles during KA'' differentiation. The participation of key transcription factors, including *nkx2.2a/b*, *nkx2.9*, *gata2a* and *tal2* is essential for the specification of KA'' cells. Shh and Notch signaling pathways are involved in regulating KA'' cell specification and play an important role in the transition from progenitor cells to differentiated KA'' cell (Cucun *et al*, 2024; Yang *et al*, 2020).

V3 interneurons are glutamatergic neurons that arise from the LFP during embryonic development in zebrafish and usually appear around 24 hpf. KA'' cells form before V3 interneurons in the LFP domain. V3 interneurons slowly spread from 24 hpf to 48 hpf along the anterior-posterior (AP) axis until they cover the maximum area. In contrast, the differentiation of KA'' cells happened quickly along AP axis from 16 hpf to 24 hpf (Jacobs *et al*, 2022). It also has specific marker genes such as early marker gene *sim1a* and late marker gene *slc17a6b*. The V3 interneuron differentiation is influenced by both the Shh and Notch signaling pathways (Cucun *et al*, 2024; Jacobs *et al*, 2022). The function of V3 interneurons include regulating locomotor patterns and coordinating muscle activity in the spinal cord. Their regulation of motor neurons can influence the timing and rhythm of movement. Additionally, It also contribute to the formation of neural circuits within the spinal cord (Danner *et al*, 2019; Zhang *et al*, 2022, 2008b).

2.1.10 The Intraspinal serotonergic neurons in zebrafish

Intraspinal serotonergic neurons (ISNs) are a unique group of neurons located within the spinal cord that synthesize and release serotonin (5-hydroxytryptamine or 5-HT). These neurons play an important role in regulating sensory, motor, and autonomic functions within the spinal cord, helping to integrate and process neural signals related to motor and sensory modulation (Montgomery *et al*, 2018).

The ISNs located mainly in the ventral region of the zebrafish spinal cord. The commonly used marker genes for intraspinal serotonergic neurons are tryptophan hydroxylase 2 (*tph2*) which is the rate-limiting enzyme involved in the synthesis of serotonin. *pet1* (*fev*) and *Imx1bb* (LIM homeobox transcription factor 1, beta b) are transcription factor involved in the development of serotonergic neurons (Zhao *et al*, 2006). Between 3- and 4-days post-fertilization (dpf), ISNs in zebrafish experience significant changes in neurite lengths and projection distances, indicating active morphological development. By 5 dpf, ISNs reach morphological maturity (Montgomery *et al*, 2018). These neurons synthesize and release serotonin, a neurotransmitter involved in the regulation of neuronal excitability, synaptic transmission, and circuit integration. Serotonergic neurons in the zebrafish spinal cord are known to play critical roles in various aspects of spinal cord physiology, including motor control, sensory processing, and neural development. Serotonin promotes the development and regeneration of motor neurons (Barreiro-Iglesias *et al*, 2015). Moreover, Serotonin plays an important role in both the regulation of axon regrowth and the generation of new neurons after nervous system injury (Sobrido-Cameán *et al*, 2018; Huang *et al*, 2021).

The primary source of ISNs in the spinal cord is the raphe nuclei located in the brainstem. However, research has shown that even after the spinal cord is complete and prolonged disconnected from the brain, serotonin is not entirely depleted from the spinal cord. This suggests that the spinal cord itself has the ability to locally synthesize serotonin (Montgomery *et al*, 2018). It is important to study the origin of serotonin in the spinal cord.

It provides valuable insights into the role of serotonin in spinal cord function, covering the regulation of sensory processing, motor control, locomotion, and autonomic regulation. Furthermore, studying the sources of serotonin may enhance our understanding of the functional organization of spinal cord circuits.

2.1.11The transcription factor Sox1

The transcription factor Sox1 is a member of the SRY (sex-determining region Y)-box (Sox) family of transcription factors, which is known for its crucial role in regulating gene expression during embryonic development, primarily in the development of central nervous system (Gerber *et al*, 2019). Sox1 has been implicated in neural determination and differentiation in vertebrate. It is expressed in most neural progenitor cells of the embryonic spinal cord (Panayi *et al*, 2010). *sox1a* is one of the two common orthologs of Sox1, alongside *sox1b* in zebrafish (Gerber *et al*, 2019; Gao *et al*, 2015).

sox1a and *sox1b* mRNA are expressed in various regions of the zebrafish central nervous system, including the ventral telencephalon, ventral diencephalon, spinal cord, and other tissues such as the lens and lateral line (Gerber *et al*, 2019; Lekk *et al*). *sox1a* and *sox1b* are involved in cell fate decisions during early nervous system development. Studies have shown that disruption of *sox1a* and *sox1b* function leads to a significant reduction in a specific subtype of V2 neurons, known as V2s (Gerber *et al*, 2019).

2.2 The aim

In this study, our aim was to explore the neuronal diversity of *sox1a* transgenic line in zebrafish ventral spinal cord using temporal single-cell transcriptome analysis. By capturing gene expression profiles at multiple developmental time points, we investigated the temporal regulation of gene expression patterns during spinal cord development. Specifically, we aim to identify genes that are dynamically regulated during the differentiation of neural progenitor cells and the specification of neuronal subtypes within the zebrafish spinal cord.

2.3 The Method and Material

2.3.1 Fish Stock Management and Embryo Collection

Wild-type (WT) zebrafish were obtained from the ABO line, an inbred line lineage originating from an intercross between the AB and OX lines (European Zebrafish Resource Center). Wild-type (WT) and transgenic zebrafish *sox1a:eGFP*, were kept under a 14-hour light/10-hour dark cycle at 28.5°C in a recirculation system (Schwarz). Their diet consisted of commercial food and in-house hatched brine shrimp, following established protocols.(A Guide for the Laboratory Use of Zebrafish *Danio (Brachydanio) rerio*). (performed by Mehmet Ilyas Cosacak).

For single-cell RNA sequencing (scRNA-seq), transgenic zebrafish *sox1a:eGFP* and WT-AB zebrafish were crossed, and embryos were promptly collected within 15 minutes to minimize age discrepancies. Any dead or unfertilized eggs were carefully removed at 4 hours post-fertilization (hpf). Batches of 40 embryos were gathered and placed in 30 mL of E3 zebrafish medium with methylene blue, maintaining a temperature of 28°C in a dark incubator until they reached the desired developmental stage. (performed by Mehmet Ilyas Cosacak).

2.3.2 Cell dissociation

Cell dissociation was performed following established protocols (Cosacak *et al*, 2020, 2019). A total of 60 embryos, representing developmental stages 1-, 2-, 3-, and 5-days post-fertilization (dpf), underwent a brief cooling step on ice for 10 minutes.

Subsequently, the trunk was separated from the head, and the yolk was carefully removed. The bodies were immersed directly in the dissociation buffer and kept on ice. Dissociation occurred at 28°C through continuous rotation and trituration every 10 minutes until complete dissociation. The cells underwent filtration using 40 mm filters (Falcon, Cat#431750), followed by centrifugation at 300 g in 2% BSA and resuspension

in 4% BSA. A total of 22,000 cells were sorted based on the GFP signal, and these cells were then processed for encapsulation. (performed by Mehmet Ilyas Cosacak).

2.3.3 Single-Cell RNA Sequencing and Generation of Count Matrix

Single-Cell RNAseq was performed in DRESDEN-Concept Genome Center deep sequencing facility based on the 10x Genomics Single-cell transcriptome workflow (Zheng *et al*, 2017). After sequencing, the raw sequencing data was processed by using the Cell Ranger software (v6.1.2) from 10X Genomics. For reference, the zebrafish genome (GRCz11) and gene annotation (Ensembl 104) were obtained from Ensembl. A custom sequence and GTF for GFP were created and appended to the zebrafish genome and gene annotation. The resulting combined data were then employed to build the reference using the 'mkref' command in CellRanger. The expression value matrix was generated by aligning sequencing reads from FASTQ files to a reference transcriptome using the 'count' command in CellRanger.

2.3.4 Preprocessing Single-Cell RNA Sequencing Data

We separately analyzed each dataset at 1, 2, 3 and 5dpf by using Seurat V4 (Hao *et al*, 2021). Firstly, we filtered the Low-quality cells or empty cells that are often caused by the failure during tissue dissociation or library preparation and are characterized by high mitochondrial proportions, low numbers of reads per cell, and low numbers of expressed genes. So, we removed the cells with more than 20% mitochondrial genes, those with fewer than 500 reads count, and cells with less than 200 feature counts. Then, the data was normalized by using the 'NormalizeData' function with default options. Effective data normalization guarantees that observed differences between different cell populations are attributed to biological factors rather than technical biases. The top 2000 variable genes were identified by using the 'FindVariableFeatures' function. The aim of feature selection is to retain genes containing important biological information while concurrently reducing the dataset's size. This process enhances computational efficiency and accelerates

downstream analysis. Prior to dimensionality reduction analysis, all genes were used to scale data by using 'ScaleData' function. Next, we used 'RunPCA' function to identify the principal component. The top 30 principal components (PCs) were selected for downstream analysis. We used the 'FindNeighbors' function to build a KNN graph using Euclidean distance in PCA space. The edge weights between any two cells were then refined based on the shared overlap in their local neighborhoods. Then we used 'FindClusters' function to cluster the cells with resolution = 0.5. Finally, we used the non-linear dimensional reduction method UMAP to visualize these data by using 'RunUMAP' function.

2.3.5 Preliminary quality control

In the initial analysis of 1 dpf embryos, a cell population exhibiting high expression of hbb genes, such as *hbbe1.1*, was identified. Utilizing scaled data from the 1 dpf embryos object, cells expressing the *hbbe1.1* gene more than 0.0 were labeled and subsequently removed. A subsequent round of analyses was conducted, similar to the one performed for 1 dpf embryos (with hbb cells removed). The remaining hbb cells, clustered together, were also excluded. After quality control and the removal of hbb cells, the 1 dpf datasets were converted to a Seurat object, applying the same settings as mentioned above. (performed by Mehmet Ilyas Cosacak).

2.3.6 Integration of Datasets

We integrated the Seurat objects across 4 time points (1, 2, 3 and 5 dpf) to cluster the same cell type from different timepoints. We used 'FindIntegrationAnchors' function to identify anchors, and subsequently used these anchors for dataset integration using the 'IntegrateData' function. Then we analyzed the integrated datasets by using the standard Seurat workflow for visualization and cell clustering.

2.3.7 Identification of main cell types

After integration and cell clustering, we identified 40 cell clusters by using the resolution 1.5. We used 'FindAllMarkers' function to identify marker genes responsible for driving the separation of the clusters. We annotated the main cell types using the top 20 marker genes and know markers from the literature. We employed *fabp7a/sox2/gfap* for neural progenitors (März *et al*, 2010; Lange *et al*, 2020; Delile *et al*, 2019), *elavl3/sv2a* for neurons (Delile *et al*, 2019; Lange *et al*, 2020), *cldnb/epcam/sox2* for lateral lines (Lush & Piotrowski, 2014; Behra *et al*, 2012), *meox1* for mesoderm (Row *et al*, 2018), *vcanb* for fin fold (Rabinowitz *et al*, 2017), *ccl25b* for the hematopoietic system (Wattrus & Zon, 2018), *pitx2* for cardiomyocytes (Franco *et al*, 2017), *mylpfa* for fast skeletal muscle (Zhang & Roy, 2016), and *fli1a* for the vascular system (Lawson & Weinstein, 2002). The remaining cells lacking clear markers were categorized as "unclassified" cells.

2.3.8 Iterative clustering refinement of neurons and neural progenitor cells

For a detailed examination of neurons and neural progenitors, we specifically isolated cells annotated as neural progenitors and neurons. This subset was used to create new Seurat objects, and integration was performed following the previously outlined methodology. we identified 32 cell clusters by using the resolution 2. We used 'FindAllMarkers' function to identify the top marker genes and transcription factors (TFs), obtained from AnimalTFDB(V3.0) (Hu *et al*, 2019), associated with cell clusters. We annotated the cells using the following markers: *sst1.1* (Kolmer-Agduhr', KA'), *urp1* (Kolmer-Agduhr", KA"), *slc6a5/nkx1.2lb* (V2s), *olig2/sox10* (oligodendrocyte precursor cells, OPC), *zic1/zic2a* (zic+ cells), *nkx2.9/sulf1* (lateral floor plate, LFP), *fev/lmx1bb/tph2* (intrapinal serotonergic neurons, ISN), *fabp7a/gfap/foxj1a* (ependymoradial glia, ERG), *neurod4/neurog1* (neural precursor cells, NPC), *slc17a6a/slc17a6b* (glutamatergic inhibitory neurons, GLU IN), *foxn4/vsx1* (V2 precursor cells, V2-pre), *mki67/sox2/sox19a/her4* (Neural Progenitors, NPro), *isl1/isl2a* (motor neurons, MN), *twist1b* (mesodermal cells, MC), and *epcam* (keratinocyte cells, KC).

2.3.9 Subclustering of ISN

We subsetting the ISN cluster from the integration dataset. We found that some neural precursor cells, motor neurons and glutamatergic inhibitory neurons were clustered with ISN cells within the integrated dataset. After checking genes characteristic, these cells were excluded from the ISN subset, and downstream analysis was performed on the adjusted ISN cells. Then We reanalyzed the ISN cells to create new Seurat objects and integration was performed following the previously outlined methodology. We identified 3 subclusters of ISN with resolution 0.3. We used 'FindAllMarkers' function to identify the variable genes between different subclusters. We designated Cluster 1 as ISN precursor cells (ISN-pre), while Clusters 0 and 2 were labeled as ISN cells. This classification was based on gene characteristics and developmental stages.

2.3.10 Trajectory inference and Pseudotime analysis of ISN cell lineage

We subsetting the LFP at 1dpf, ISN-pre and ISN cells from Seurat object to do trajectory and pseudotime analysis by using monocle(version 2.20.0) (Cao *et al*, 2019). We used 'estimateSizefactors' and 'estimateDispersions' function to estimate the size factors and dispersions for normalization of differences in mRNA recovered across cells and differential expression analysis. Then we filtered out the low-quality cells by using 'detectGenes' function and excluded the genes expressed in fewer than 10 cells. We subsequently used 'differentialGeneTest' function to identify a set of ordering genes by comparing the cell types throughout the ISN cell lineage. The differentially expressed genes ($qval < 0.01$) were set in the object using 'setOrderingFilter' function to use for ordering and trajectory reconstruction. We then employed 'reduceDimension' function to reduce the space into two dimensions for visualization and interpretation. We used 'orderCells' function to order cells along the trajectory and 'plot_cell_trajectory' function to visualize the trajectory.

After trajectory inference, we analyzed the cells to find genes that changes through ISN cell generation process. Monocle assigns each cell a “pseudotime” value, reflecting its

progression through the process by using 'differentialGeneTest' function. Then we selected the significantly differentially expressed genes ($qval < 0.00001$) and grouped these genes into 2 clusters based on similar trends by using 'plot_pseudotime_heatmap' function. The enrichment analysis of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms was performed on the clustered gene sets by using the clusterProfiler (Wu *et al*, 2021) R package. Terms with the $qvalue < 0.05$ were considered statistically significant.

2.3.11 Reanalysis of LFP, V3 precursors and V3 cells using Tg *olig2* scRNAseq data

We subsetting the LFP, Pre-V3 IN, and V3 IN cell populations from Tg *olig2:EGFP* (GSE173350) scRNA-seq data (Scott *et al*, 2021). We noticed the V3 IN cells mixed with other neurons. After subclustering V3 IN cell by using 'FindSubCluster' function in Seurat, We used *nkx2.2a* and *sim1a* positive V3 two subclusters to integrate with LFP and Pre-V3 IN by using the standard Seurat workflow for visualization and cell clustering, applying the same settings as mentioned above.

2.3.12 Integration of ISN and V3 cell lineage

We integrated the ISN and V3 cell lineage by using Seurat standard integration function (anchor.features = 1500, npcs = 25). We identified 18 cell clusters with resolution 3. We employed *nkx2.9* to label LFP cells, identifying *pcna/mki67* positive LFP as LFP-1, another subset as LFP-2, *neurod4* for NPC, *mnx1* for MN, *fev* for ISN-pre, *fev/tph2* for ISN, *dld/sim1a* for V3-pre, *nkx2.2a/sim1a* for V3. We excluded motor neurons and LFP-2 (late-stage LFP) and conducted trajectory inference analysis using the remaining cell types by using monocle, applying the same settings as mentioned above. The top 500 differentially expressed genes, determined by cell types using the Seurat 'FindAllMarkers' function, were employed for ordering and trajectory reconstruction.

Next, the monocle 'BEAM' function was used to detect genes that divide cells into different cell branches. We excluded batch effect genes and utilized the 'plot_genes_branched_heatmap' function to visualize the branch-related gene set with a q-value less than 0.01 and num_clusters = 4.

2.3.13 Data and code availability

The single-cell RNA-seq data for sox1a transgenic line has been deposited in the NCBI GEO database and is accessible by the number GSE226494.

The single-cell RNA-seq data for olig2 transgenic line is from the publication by Scott et al., 2021. It is accessible by the number GSE173350 (Scott *et al*, 2021).

The code for this work is available on Github at the following address:
https://github.com/Fushun-Chen/Sox1a_scRNA.git .

2.4 Result

2.4.1 Analysis of scRNA-seq data from the *sox1a* transgenic lines of zebrafish spinal cord

To gain a deeper understanding of the cell populations within the zebrafish spinal cord, specifically in the *sox1a* transgenic line, we performed droplet-based scRNA-seq (10x Genomics Chromium) in the body (trunk and tail) at 1, 2, 3, and 5 dpf (Figure 2.4.1A). We used Seurat (Hao *et al*, 2021) to individually analyze each dataset and subsequently integrated them. After applying quality filters, a dataset comprising 15,448 cells was retained for subsequent analysis, with distribution across developmental stages as follows: 3462 cells from 1 dpf, 2389 cells from 2 dpf, 4575 cells from 3 dpf, and 5022 cells from 5 dpf (Figure 2.4.1B). We used the UMAP dimensionality reduction to visualize the integrated dataset. We got 40 unique cell clusters using the resolutions 1.5 (Figure 2.4.1C). We examined the distribution at each developmental stage to assess the effectiveness of integration in removing batch effects. Integration of similar cell clusters from different stages emphasizes a successful integration, demonstrating a significant reduction in batch effects (Figure 2.4.1D).

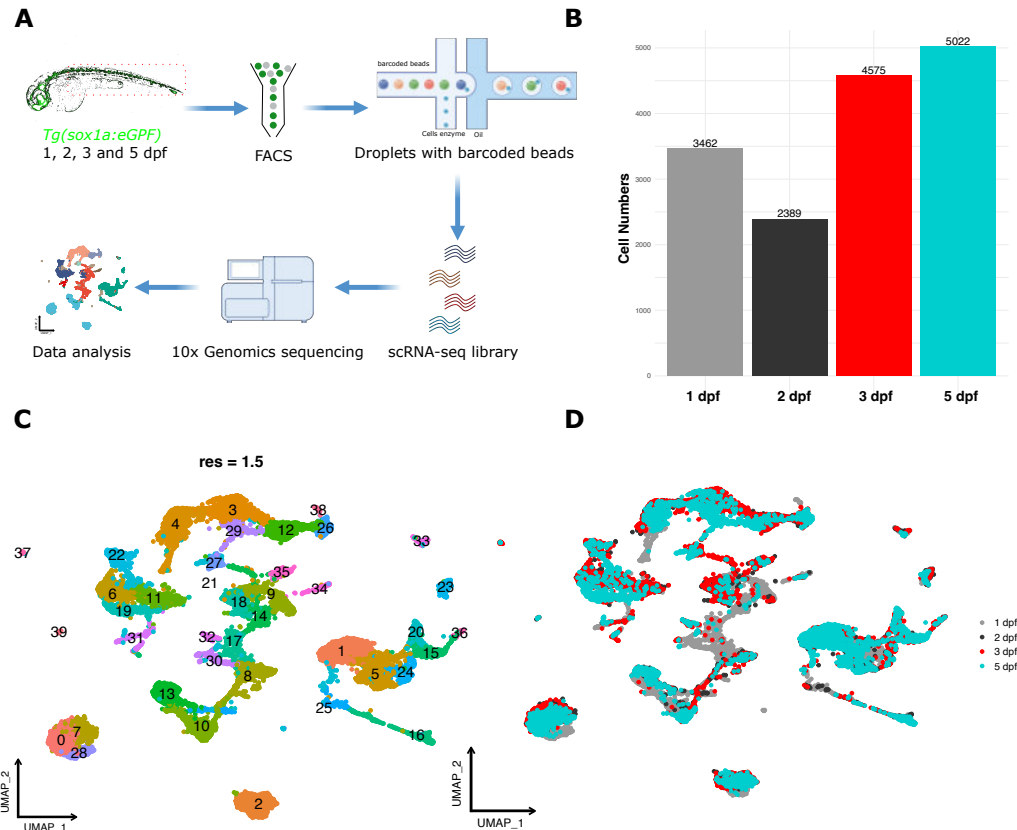


Figure 2.4.1 The workflow of scRNA-seq and preliminary data analysis.

- (A) Schematic Workflow of Single-Cell RNA Sequencing (scRNA-seq) (part of it was generation with BioRender.com).
- (B) Quantification of cell numbers at each developmental stage.
- (C) UMAP visualization shows the distinct cell clusters identified at a resolution of 1.5.
- (D) The distribution of cells across different developmental stages. Different colors represent distinct stages: grey for 1 dpf, black for 2 dpf, red for 3 dpf, and blue for 5 dpf.

2.4.2 The identification of cell clusters based on transcription markers

We then assigned the cells to different cell types based on combinatorial expression of a list of genes. We defined neural progenitors, neurons, lateral line, mesoderm, fins, hematopoietic system, cardiomyocytes, fast skeletal muscle, vascular cells and musculoskeletal cells (Figure 2.4.2A, C and D). The remaining cells didn't belong to any category, possibly because of these cells were from other tissues. We have classified these cells as Unclassified (Figure 2.4.2A, C and D). Subsequently, we examined the distribution of cells across developmental stages within each cell type. We observed that neural progenitors were predominantly present at 1 dpf, while neurons were primarily observed at 3 dpf and 5 dpf, aligning with the progression of neural development (Figure 2.4.2B). Then we checked the expression of GFP and *sox1a* at each cell type (Figure 2.4.2D). We found that GFP was highly expressed in the *sox1a* positive cells, especially in the lateral line and the neural progenitors and neural interneurons of the spinal cord (Figure 2.4.2D) (Gerber *et al*, 2019). The co-expression of GFP and *sox1a* validates the enrichment of *sox1a* cells in the zebrafish spinal cord.

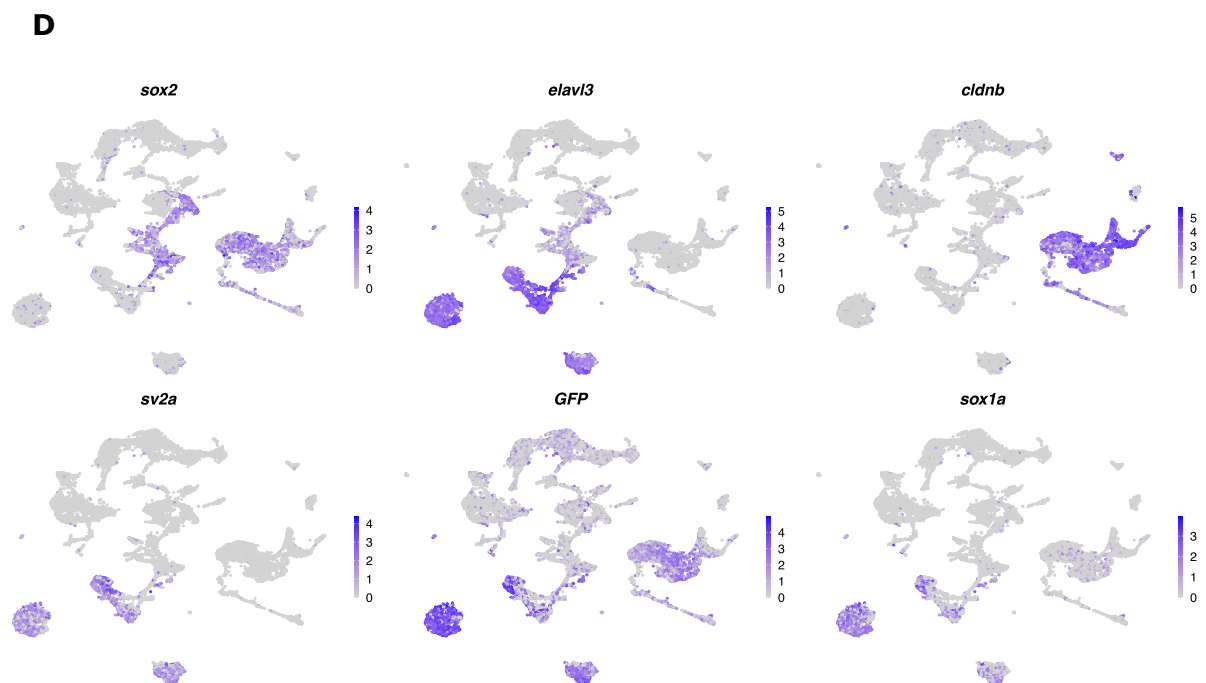
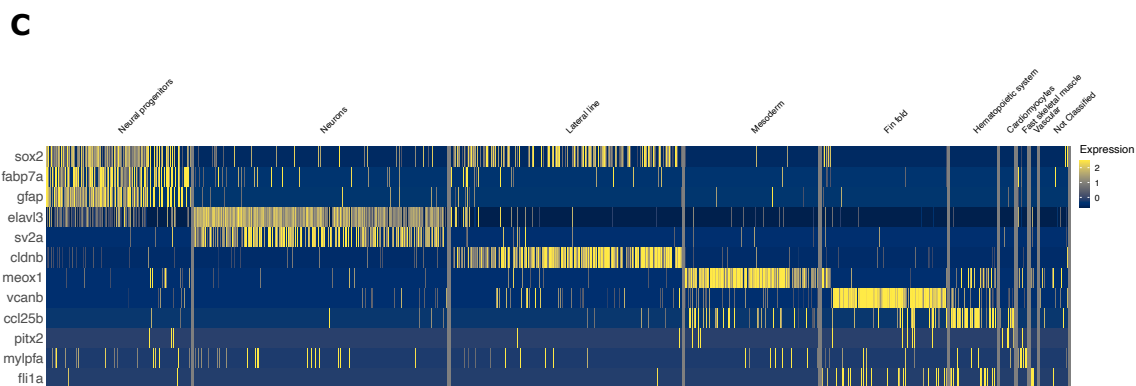
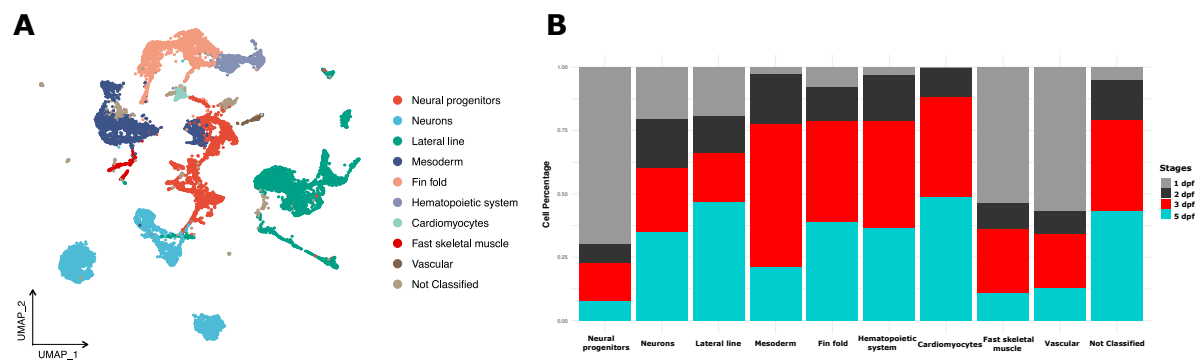


Figure 2.4.2 Comprehensive analysis of cell types and developmental dynamics in *sox1a:gfp* line.

- (A) UMAP shows the main cell types, delineated through the identification of key marker genes and previously recognized markers.
- (B) The distribution of cells across different developmental stages within each cell type.
- (C) A heatmap showing the expression of representative marker genes in main cell types.
- (D) Feature Plots showing marker genes utilized for sorting in-silico neural progenitor and neuronal cell populations. Note: *cldnb* is employed to exclude lateral line cells, which are *gfp*⁺ and *sox2*⁺.

2.4.3 Refined Analysis of Neurons and Neural Progenitor Cells

As our aim is to understand the gene regulatory network and the mechanisms of neural development and differentiation in the spinal cord. Therefore, we concentrated on the cells that were identified as spinal cord neural progenitors and neurons, excluding other cell types. After eliminating other cell types, we retained a dataset of 5963 cells for further analysis, comprising 2392 cells from 1 dpf, 929 cells from 2 dpf, 1203 cells from 3 dpf, and 1439 cells from 5 dpf (Figure 2.4.3A). Then we re-integrated neural progenitors and neurons of each timepoint. We got 33 cell clusters using the resolution 2 (Figure 2.4.3B). We evaluated the effectiveness of the integration in removing batch effects by examining the distribution at each developmental stage. The merging of similar cell clusters from different stages showed a successful integration, significantly reducing batch effects (Figure 2.4.3C).

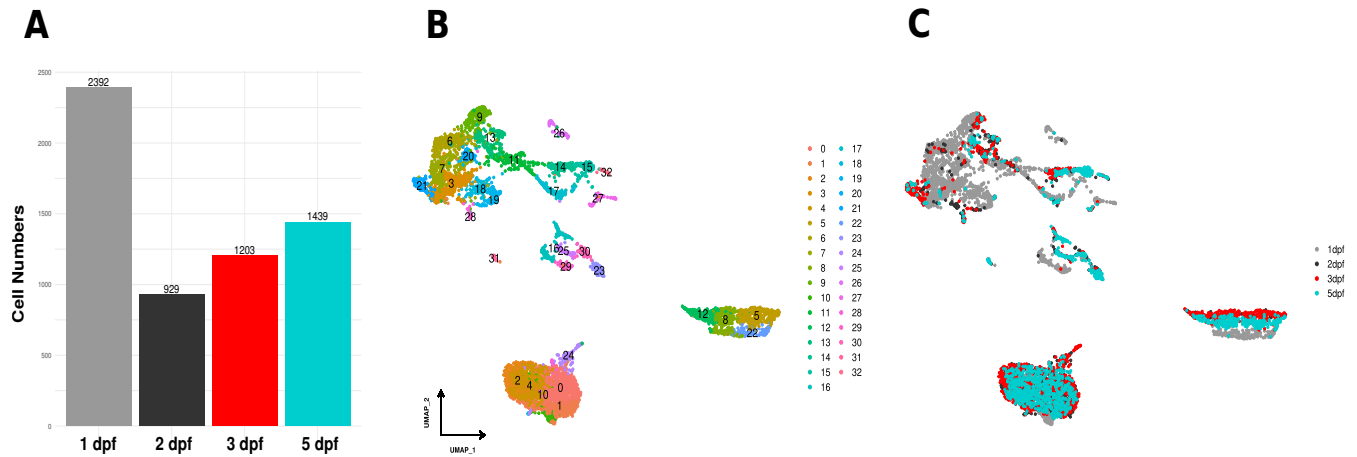


Figure 2.4.3 Quantification, clustering, and developmental distribution of neural progenitors and neurons.

(A) The cell numbers of neural progenitors and neurons at each developmental stage.

(B) UMAP visualization shows the distinct cell clusters of neural progenitors and neurons identified at a resolution of 2.

(C) The distribution of neural progenitors and neurons across different developmental stages.

2.4.4 The identification of cell subtypes within neural progenitors and neurons

Then we determined cell population identity based on the expression of markers genes. We identified distinct cell clusters representing various progenitor cells and neurons. In progenitor cells, we identified lateral floor plate (LFP, *nkx2.9/sulf1*), V2 precursor (V2-pre, *vsx1/foxn4*), neural precursor cells (NPC, *neurod4/neurog1*) and oligodendrocyte precursor cells (OPC, *olig2/sox10*). Additionally, we identified multiple neural progenitor groups (NPro) expressing markers of proliferation (*mki67*) and neural progenitors (*sox2* and *sox19a*). While we couldn't definitively link them to a known progenitor domain, these groups were designated as NPro1 to 4. Moreover, we discovered an additional neural

progenitor cell cluster (*mki67*, *sox2*, and *sox19a*) expressing distinct *zic* genes (*zic1* and *zic2a*), labeled as *zic*⁺ cells. Among the various clusters, we also identified a glial population representing ependymal radial glial cells (ERG, *gfap*, *foxj1a*, *fabp7*), known as neural stem cells in the spinal cord. In neuronal cells, we identified Kolmer-Agduhr' (KA', *sst1.1*), Kolmer-Agduhr" (KA", *urp1*), V2s (*slc6a5/nkx1.2lb*), intraspinal serotonergic neurons (ISN, *fev/lmx1bb/tph2*), glutamatergic inhibitory neurons (GLU IN, *slc17a6a/slc17a6b*) and motor neurons (MN, *isl1/isl2a*) (Figure 2.4.4A, B and C). In our analysis, we identified specific clusters of *sox1a*⁺ cells associated with KA', KA", V2-pre and V2s neurons (Figure 2.4.4A and B), aligning with previous research findings (Andrzejczuk *et al*, 2018; Gerber *et al*, 2019), thereby confirming the validity of our methodology and results. Additionally, our investigation also uncovered a previously unreported cell type: intraspinal serotonergic neurons (ISN) positive for *sox1a*, characterized by expression of *tph2* and *fev* (Montgomery *et al*, 2016, 2018) (Figure 2.4.4A and B).

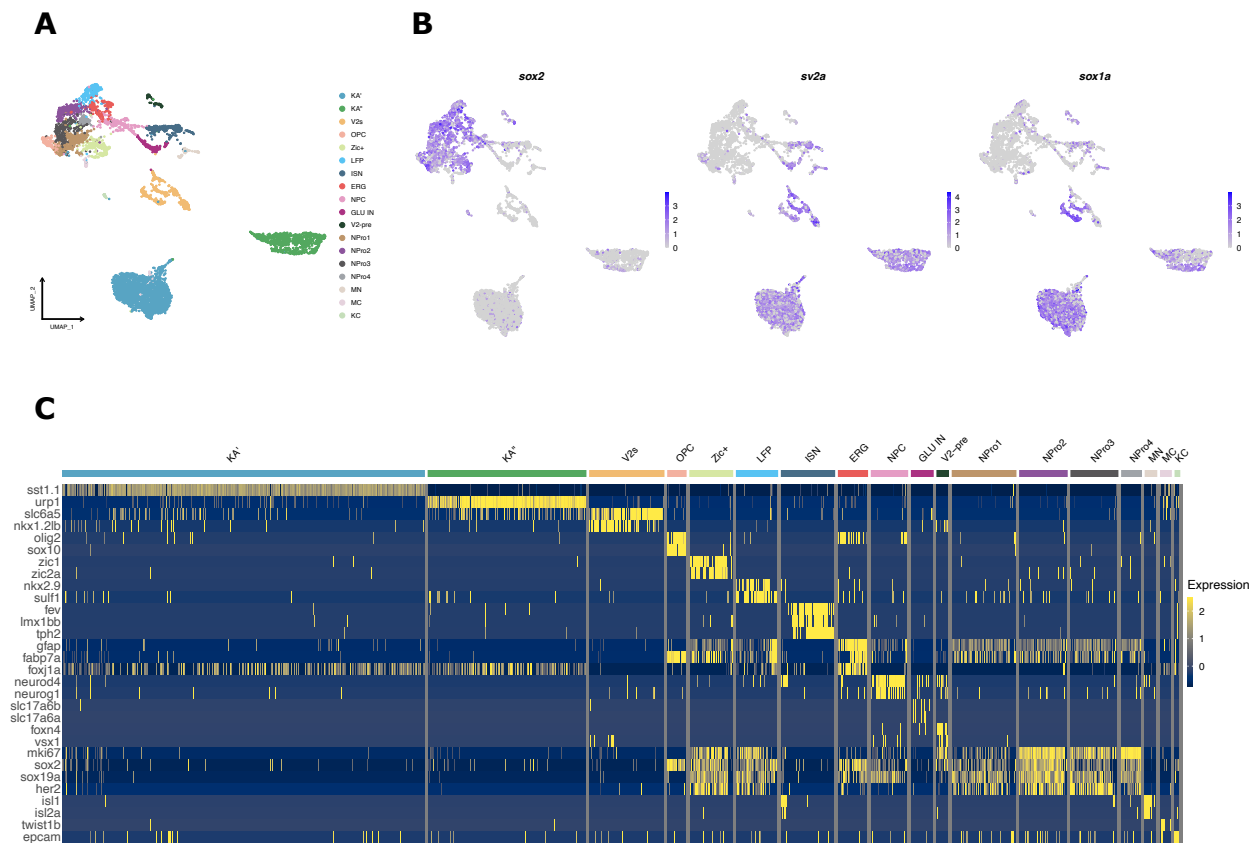


Figure 2.4.4 Annotation of cell subtypes within neural progenitors and neurons.

- (A) UMAP shows the neural progenitor and intraneuronal cell subtypes, delineated through the identification of key marker genes and previously recognized markers.
- (B) Feature Plots showing the expression of *sox2* (progenitor marker), *sv2a* (neuronal marker) and *sox1a*.
- (C) A heatmap showing the expression of representative marker genes in each neural progenitor and neuronal subtype.

2.4.5 The ISN cluster is made of precursor and mature neurons

The intraspinal serotonergic neurons (ISN) can synthesize the serotonin neurotransmitter which is critical for modulating sensory, motor, and autonomic functions (Huang *et al*, 2021). This cell cluster is characterised by the expression of the marker genes *fev*, *lmx1bb*, *gata3*, *gata2a*, *sox1a*, and *sox1b* as well as genes involved in the synthesis, transport, and reception of the neurotransmitter serotonin, namely *tph2*, *slc6a4a* and *htr1d* (5-hydroxytryptamine (serotonin) receptor 1D) (Figure 2.4.5A) (Spencer & Deneris, 2017). From the expression of *fev* and *tph2* at every timepoint, we can know it starts to appear in 2 to 3 days and some ISN cells is *fev* positive and *tph2* negative at 3dpf (Figure S1A and B). After checking marker genes expression, we also found that some neural precursor cells (*neurod4*), motor neurons (*isl1*) and glutamatergic inhibitory neurons (*mab21l1*) were clustered with ISN cells within the integrated dataset. These cells were subsequently excluded from the ISN subset, and subsequent downstream analyses were performed on the refined ISN cell population (Figure S1A and B).

To gain a deeper understanding of the ISN cell cluster, we performed subcluster analysis, dividing it into three distinct subclusters (Figure 2.4.5B). We then examined the cell distribution within each stage across these distinct subclusters. Sub-cluster 1 mainly consisted of cells from time points 2 and 3 dpf, while the remaining two sub-clusters were mainly composed of cells from time point 5 dpf (Figure 2.4.5C). We also identified the variable genes between different subclusters. We found a sub-cluster (sub-cluster 1) that was characterize by the positive expression of *nkx2.2a*, *nkx2.2b*, *fev*, *lmx1bb*, *gata2a* and *gata3*, but was negative for *tph2* and *ddc* (Deneris & Wyler, 2012; Deneris & Gaspar, 2018). The other two sub-clusters were positive for *fev*, *lmx1bb*, *gata2a*, *gata3*, *tph2* and *ddc*, but negative for *nkx2.2a* and *nkx2.2b*. Based on this information, we labeled the sub-cluster 1 as ISN-pre, and the other subclusters, which appeared more mature, as ISN (Figure 2.4.5D and E).

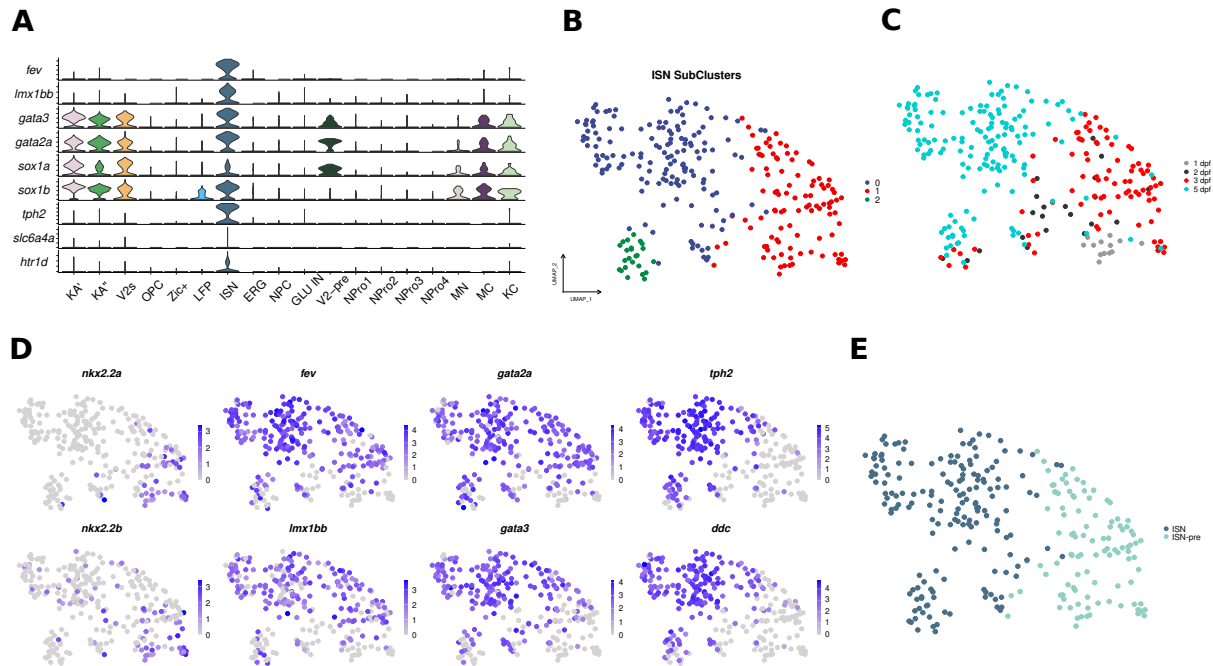


Figure 2.4.5 The sub-cluster analysis of ISNs reveals the existence of two distinct groups: ISN-pre and mature ISNs.

- (A) Violin plots showing the expression of key marker genes in ISNs co-expressed with *sox1a/b*.
- (B) UMAP shows the sub-clustering of the ISN cluster into three sub-clusters.
- (C) The distribution of ISN cells across different developmental stages.
- (D) Feature plots showing the expression of the variable genes between different subclusters.
- (E) The ISN cluster consists of ISN-pre (*fev*⁺ and *tph2*⁻ cells) and ISNs (*fev*⁺ and *tph2*⁺ cells).

2.4.6 ISN origins from LFP

Analysis of our data reveals that ISN generation occurs between 2 to 3 dpf. However, The exact origin and underlying mechanisms are still not fully elucidated. In the mouse, it has been reported that serotonin neurons are derivatives of the lateral floor plate of the hindbrain (Kittappa, 2017). Spinal cord development involves a complex process where

cells in different dorsal-ventral locations express different combinations of transcription factors (Lewis, 2006). This diversity of gene expression is critical for determining cell identity and functions along the dorsoventral axis of the spinal cord. We found ISN-pre shared some transcription factors with LFP like *foxa*, *nkx2.2a*, *nkx2.2b* (Figure 2.4.6A).

Previous studies using cell lineage tracing and genetic analysis have provided insights into the developmental sequences of lateral floor plate (LFP) progenitors. It is known that LFP first give rise to KA" at around 16 hours post-fertilization (hpf), followed by V3 cells at approximately 24 hpf (Jacobs *et al*, 2022). The fate decisions of LFP progenitor cells that differentiate into different interneuron types are closely related to the duration of Notch signaling. Those LFP cells that exhibit high levels of Notch signaling remain in a progenitor state, avoiding differentiation. In contrast, inhibition of the Notch pathway promotes the differentiation of progenitor cells into postmitotic interneurons, thereby contributing to spinal cord development (Jacobs & Huang, 2019; Huang *et al*, 2012; Jacobs *et al*, 2022; Schäfer *et al*, 2007; Yang *et al*, 2010).

Based on these data, we hypothesized that ISN originates from the lateral floor plate (LFP) in the zebrafish spinal cord. To understand the developmental trajectory and identify gene expression changes over this trajectory, we conducted trajectory inference and pseudotime analysis using 193 cells from LFP at 1 dpf, 103 cells from ISN-pre, and 172 cells from mature ISN, by using Monocle2 (Cao *et al*, 2019) (Figure 2.4.6B). Beginning with LFP, we predicted a single lineage trajectory through ISN-pre to reach mature ISN (Figure 2.4.6C). Analyzing the developmental trajectories and cell distribution, it clearly shows that LFP cells (originating from 1dpf) develop into ISN-pre cells (mainly from 2 and 3 dpf) and eventually mature into ISN (mainly from 3 and 5 dpf) (Figure 2.4.6C and D). Pseudotime increased follow the developmental trajectory. It can help us understand the continuum of cellular states and infer the progression of cells from one state to another. We observed *nkx2.9* is activated in the early stage of LFP, *nkx2.2a* and *foxa* exhibit activation in the LFP and ISN-Pre, *fev* and *lmx1bb* in ISN-pre and ISN, and *tph2* primarily expressed in the ISN (Figure 2.4.6E and F).

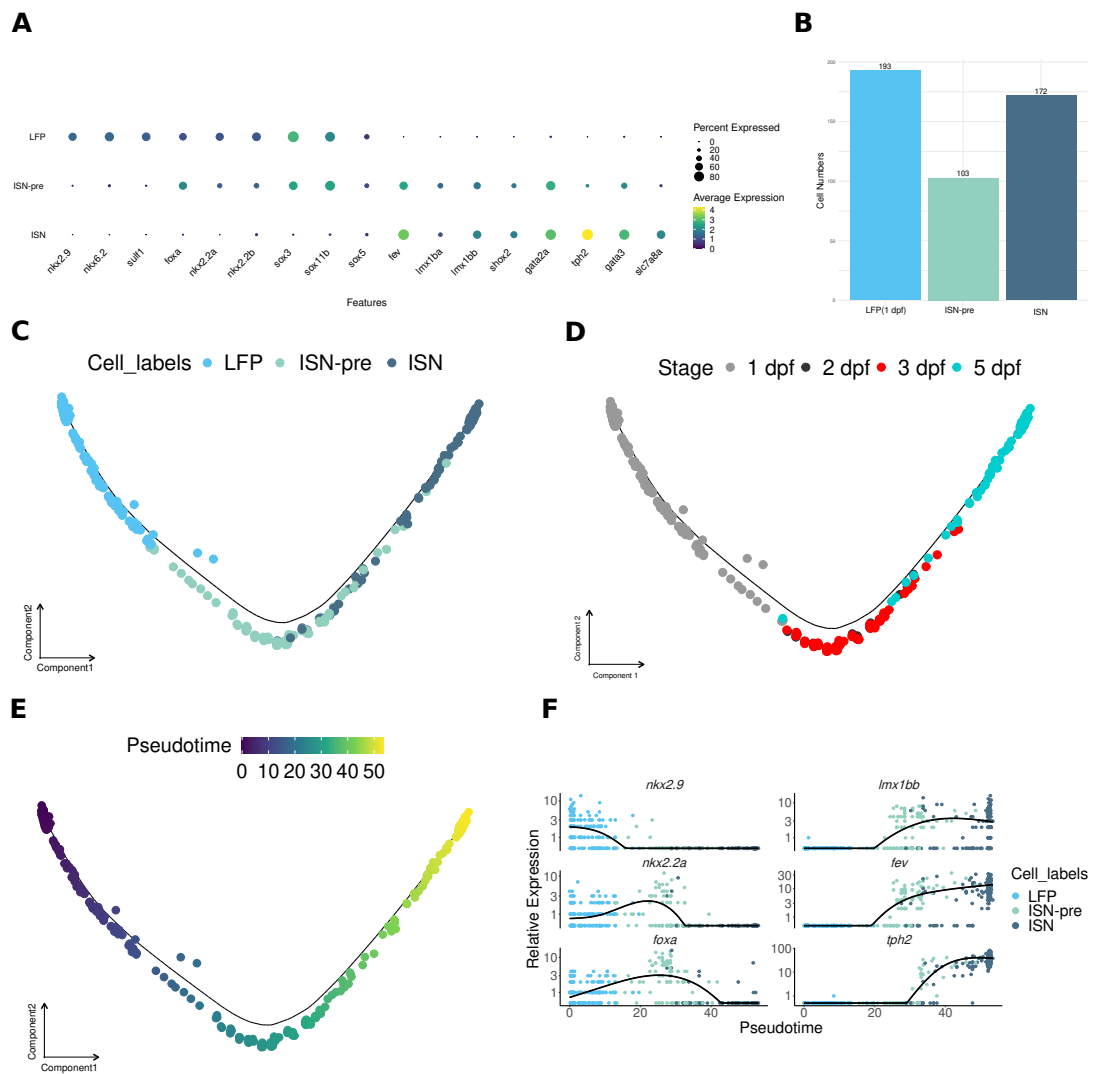


Figure 2.4.6 Analysis of transcription factor expression, trajectory, and pseudotime dynamics in LFP, ISN-pre, and ISN Cells.

- (A) The dot plot showing the expression of transcription factor (TF) genes in LFP, ISN-pre, and ISN cells. Varying dot sizes indicate the percentage of expression within each cell type, while different colors represent the level of gene expression in each cell type.
- (B) The number of cells used for the trajectory analysis for each cell type: LFP, ISN-pre, and ISN.
- (C) The developmental trajectory of LFP, ISN-pre, and ISN cells.

- (D) The distribution of cells from each developmental stage along the cell trajectory.
- (E) The Pseudotime is indicated along the cell trajectory.
- (F) The expression of essential transcription markers for LFP and ISN is displayed across pseudotime.

2.4.7 Notch signaling controls LFP cells differentiation into ISNs

To investigate the temporal changes in gene expression during the transition from LFP cells to ISN cells, we used monocle2 to identify two distinct gene clusters based on similar expression trends (Figure 2.4.7A). We subsequently performed gene ontology (biological process) and pathway (KEGG) analyses on the two identified gene clusters. We found the cluster 2 gene list (genes highly expressed in the LFP) are enriched in Notch signaling pathway, including genes such as *her12*, *her2*, *notch3*, and *her6* (Figure 2.4.7B and D). Conversely, the cluster 1 gene list (genes highly expressed in the ISN) exhibited enrichment in synaptic signaling and neurotransmitter transport, including genes like synaptotagmin Ia (*syt1a*), *scg2b*, *insm1a*, dopa decarboxylase (*ddc*), opiate receptor-like 1 (*oprl1*), GTP cyclohydrolase 1 (*gch1*), *gata3*, and *gata2a* (Figure 2.4.7C and D).

Previous studies have suggested that the differentiation of KA" and V3 interneurons from LFP progenitors requires downregulation of the Notch signaling pathway (Jacobs & Huang, 2019; Jacobs *et al*, 2022). To explore the impact of Notch signaling pathway downregulation on the development of ISNs from LFP cells, we treated wild-type (WT) embryos with the γ -secretase inhibitor LY411575 from 2 to 3 dpf. This inhibitor blocks Notch processing following ligand binding (Fauq *et al*, 2007; Jacobs & Huang, 2019). To evaluate the effects of LY411575 treatment, we analyzed the number of *tph2*⁺ neurons in the trunk of zebrafish embryos at 3 dpf. Our results revealed a significant increase in the number of *tph2*⁺ cells in embryos treated with inhibitor LY411575 compared to those treated with DMSO alone. This shows that blocking notch signaling leads to the differentiation of progenitors to ISN cells.

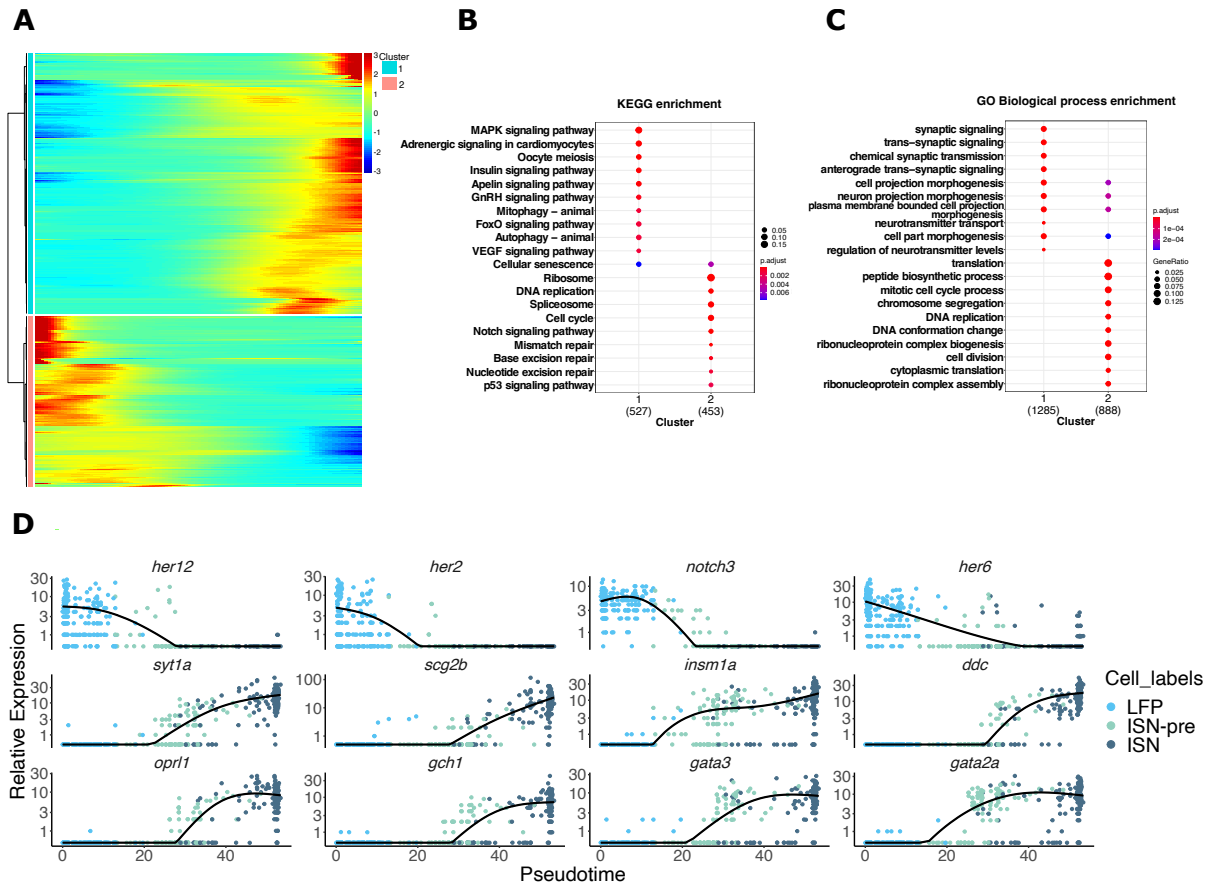


Figure 2.4.7 Notch signaling is involved in the differentiation of LFP into ISN.

- (A) Pseudotime heatmap showing pseudotime-dependent differentially expressed genes between clusters 1 and 2.
- (B) KEGG pathway analyses were performed using the two identified gene clusters.
- (C) GO analyses for Biological Processes were conducted using the two identified gene clusters.
- (D) The expression of key transcription markers for LFP and ISN is displayed across pseudotime.

2.4.8 ISN-pre cells were found in the Tg(*olig2:GFP*) scRNA-seq dataset

Previous studies have suggested that KA" and V3 interneurons develop from LFP progenitors (Jacobs *et al*, 2022). Our results suggest that ISNs also originate from LFP progenitors. To further confirm our results and provide evidence that both V3 and ISN originate from LFP cells, we conducted an analysis of scRNA-seq data as published by Scott *et al* (Scott *et al*, 2021). They used the *olig2* transgenic line to perform single-cell RNA sequencing on the zebrafish spinal cord at 24, 36, and 48 hours post-fertilization (hpf). In their sequencing data, they identified LFP, V3 precursors interneurons (Pre-V3 IN) and V3 interneurons (V3 IN) (Figure 2.4.8A). We initiated our analysis by examining the sequencing data of *olig2*⁺ cells to identify cells that might correspond to ISNs. Remarkably, within the group annotated as Pre-V3 IN, we found a small subset of cells that exhibited positive expression for *fev*, *lmx1bb*, and *gata2a* but lacked *tph2* expression. This observation strongly indicated that these cells represented ISN-pre cells (Figure 2.4.8A and B). When examining the expression of *sim1a* in the V3 IN cluster within the *olig2*⁺ sequencing data, we noticed that not all cells within this cluster expressed *sim1a*. Therefore, I conducted a subtype analysis of V3 IN and eliminated two subclusters that lacked *sim1a* expression but were positive for *lhx5* (Figure 2.4.8C). Subsequently, we combined all *sim1a*⁺ V3 cell clusters, and V3-pre cell and LFP cell, and then performed a re-analysis of the integrated data. This analysis uncovered six distinct clusters, aligning with LFP, MN (motor neurons), NPC (neural precursor cells), ISN-pre, V3-pre, and V3 cells (Figure 2.4.8D). The specific gene expression profiles for each of these clusters are visually represented in Figure 2.4.8D and E. We observed a distinct cell cluster labelled as ISN-pre expressing *fev* and *sox1a*, providing further confirmation that ISNs originate from LFP cells and express *sox1a*.

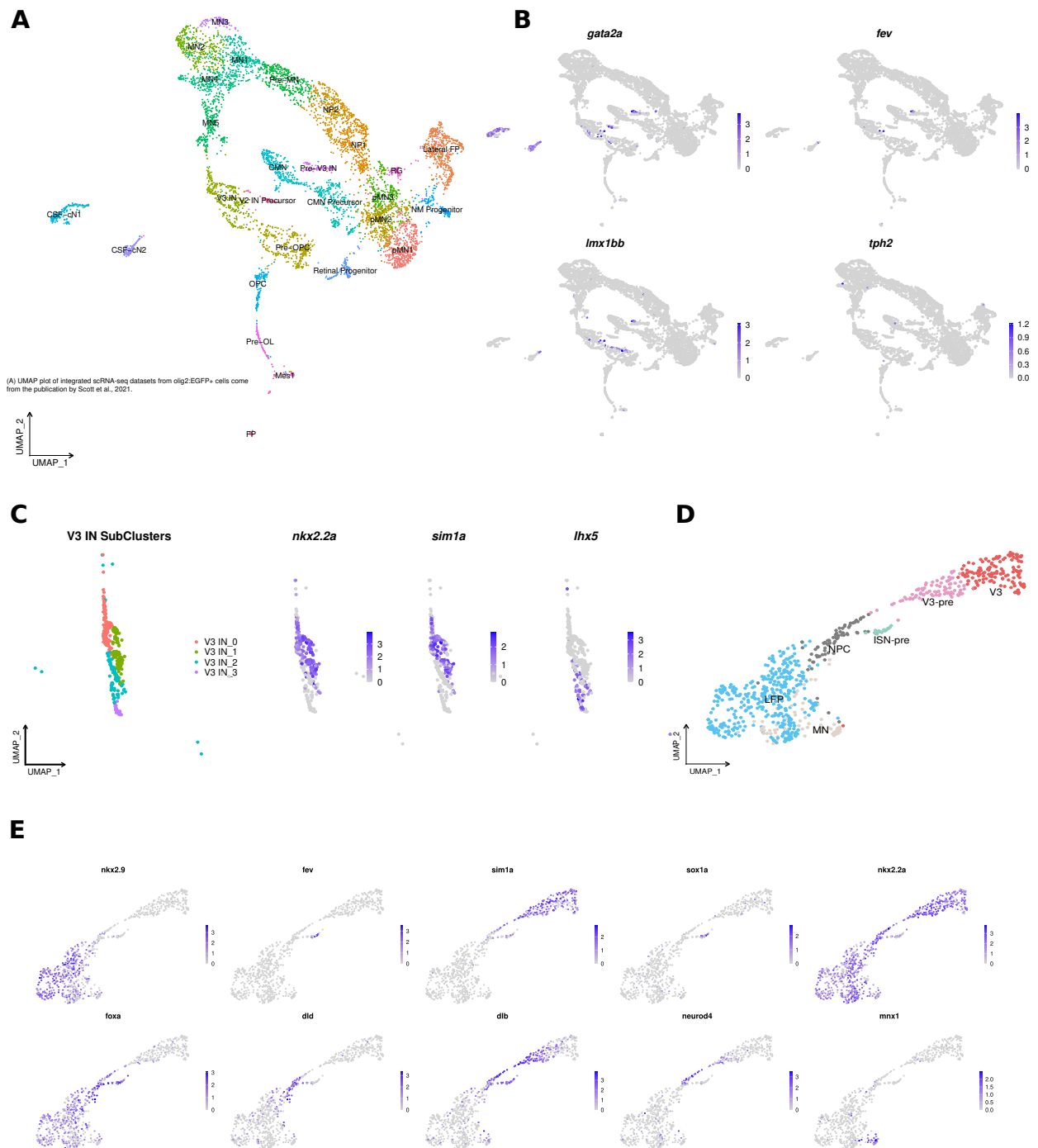


Figure 2.4.8 Confirmation of ISN origin from LFP using *olig2:gfp* scRNA-seq data.

(A) UMAP showing annotated cell types in the scRNA-seq data of the *olig2* transgenic line as published by Scott et al (Scott et al, 2021).

- (B) Feature Plot showing the expression of ISN-pre and ISN markers (*gata2a*, *fev*, *lmx1bb* and *tph2*).
- (C) The analyses of V3 interneurons and their subtypes (left), and the expression of selected marker genes shown on Feature Plots (right).
- (D) UMAP showing the reanalysis of LFP, V3, and ISN-pre cells in the scRNA-seq data of the *olig2* transgenic line.
- (E) Feature Plots showing marker genes related to the cell types in Figure 8D.
- Note 1: CSF-cN1 and CSF-cN2 denote KA' and KA" cells, respectively, which were used in this study (Gerber *et al*, 2019).

2.4.9 The comprehensive analysis of integrated scRNA-seq dataset from *olig2:eGFP+* and *sox1a:eGFP+* cells confirms the common origin of V3 and ISN

Subsequently, we integrated the LFP, V3-pre and V3 from *olig2+* sequencing results with LFP, ISN-pre and ISN from our own *sox1a+* sequencing results. This analysis uncovered eight distinct clusters, aligning with LFP-1, LFP-2, NPC, MN, ISN-pre, ISN, V3-pre and V3 (Figure 2.4.9A, B and C). Although both LFP-1 and LFP-2 cells express LFP markers *nkx2.2a* and *nkx2.9* (Figure 2.4.9D), the LFP-2 cluster represents a late-stage population, evidenced by the expression of the late LFP marker *shha* (Figure 2.4.9D), which is not present in LFP cells before 36 hpf (Oustah *et al*, 2014). Cells in the LFP-2 group mainly originate at 2, 3, and 5 dpf (Figure 2.4.9B). Moreover, these cells are no longer proliferative (*pcna* and *mki67* negative; Figure 2.4.9D) but express the radial glial cell marker (*gfap*; Figure 2.4.9D). As a result, the LFP-2 group is excluded from trajectory inference and pseudotime analysis. In contrast, the LFP-1 cluster represents a proliferative progenitor population, with cells emerging between 24 and 36 hpf.

The assessment of the developmental trajectory over pseudotime confirmed that LFP cells act as the common progenitor for both V3 and ISN neurons. This finding unveiled a bifurcated developmental path emanating from LFP cells, resulting in two distinct

branches. One trajectory led to the formation of V3 neurons, while the other trajectory ultimately leads to the development of ISN cells (Figure 2.4.9E, F, and G). This branching pattern marks differences in fate and gene expression profiles between V3 and ISN lineages, understanding of the molecular events guiding their differentiation.

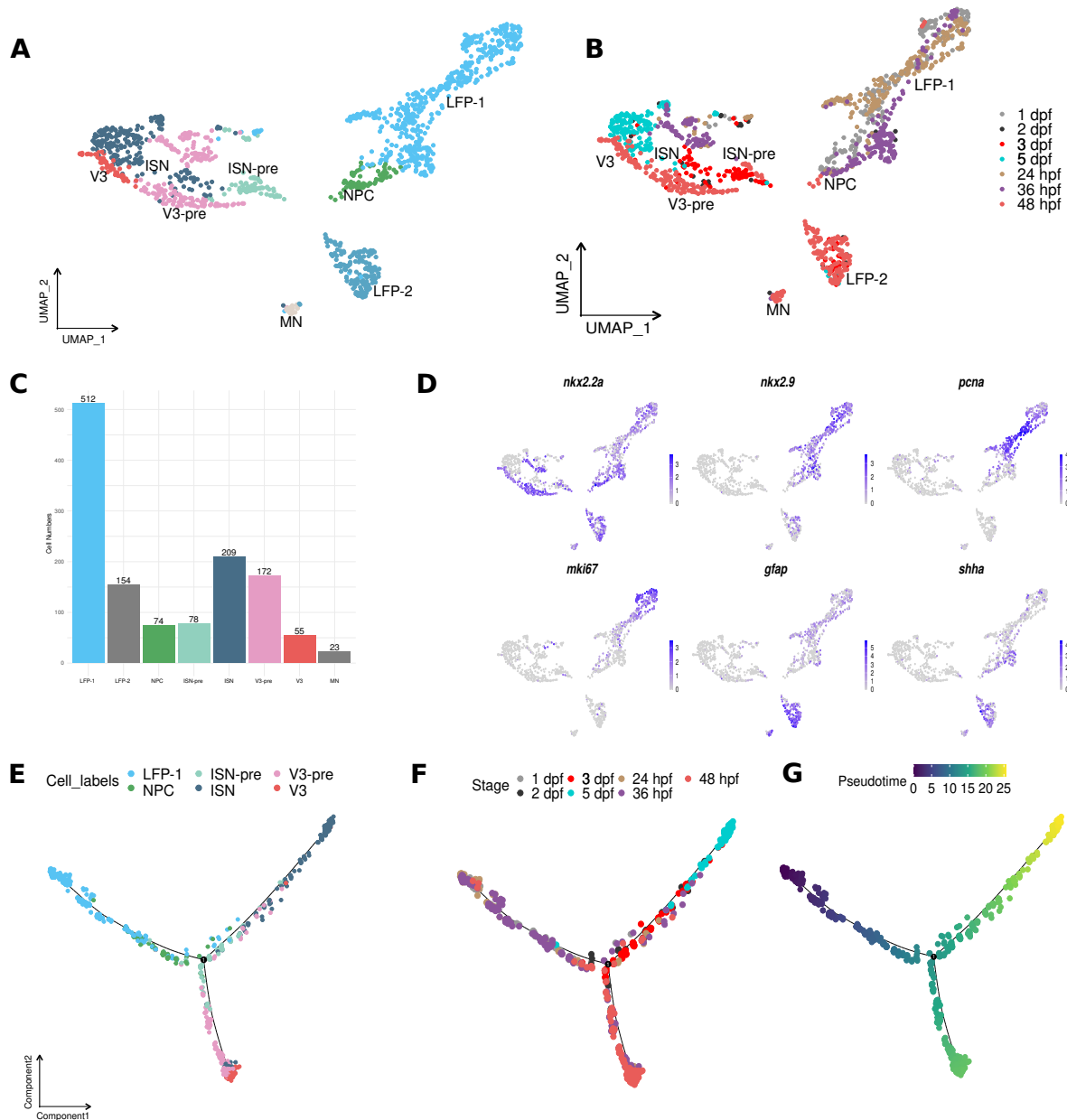


Figure 2.4.9 Establishing cell lineages of V3 and ISN through pseudo-time analysis using scRNA-seq data from *olig2:eGFP+* and *sox1a:eGFP+*.

- (A) UMAP showing the integration of LFP, V3, and ISN cells and their subtypes from *olig2:eGFP+* and *sox1a:eGFP+* scRNA-seq data.
- (B) The distribution of LFP, V3, and ISN cells and their subtypes across different developmental stages.
- (C) The number of cells for each cell type: LFP-1, LFP-2, NPC, ISN-pre, ISN, V3-pre, V3, and MN cells.
- (D) The feature plots in this part of the figure show marker genes that highlight the differences between LFP-1 and LFP-2 cells.
- (E) Cell trajectory showing LFP as the precursor of V3 and ISN.
- (F) The figure displays the distribution of cells from each developmental stage on the cell trajectory of LFP, ISN, and V3 cells.
- (G) Pseudotime is displayed on the cell trajectory of V3 and ISN.

Note 1: Sequencing data for 24 hpf, 36 hpf, and 48 hpf were obtained from *olig2:eGFP+* cells, while data for 1 dpf, 2 dpf, 3 dpf, and 5 dpf were obtained from *sox1a:eGFP+* cells in the current study.

We then generated a heatmap to visualize the changes in gene expression for all genes displaying significant branch dependencies. This analysis led to the identification of four distinct clusters (Figure 2.4.10A). Cluster 1 genes exhibited high expression in ISNs, with their cell type-specific markers being unique to fully differentiated ISNs, such as *fev* and *gata3*. Cluster 2 genes displayed significant expression in V3 cells, including *sim1a* and *nos1*. Cluster 3 genes were strongly expressed in LFP cells and ISNs during the transition to V3 and ISNs, with examples like *foxp4* and *fosab*. Lastly, Cluster 4 genes were specifically expressed in LFP and V3 cells during the transition to V3 and ISNs, featuring genes like *nkx2.2a* and *pcdh19* (Figure 2.4.10A and B).

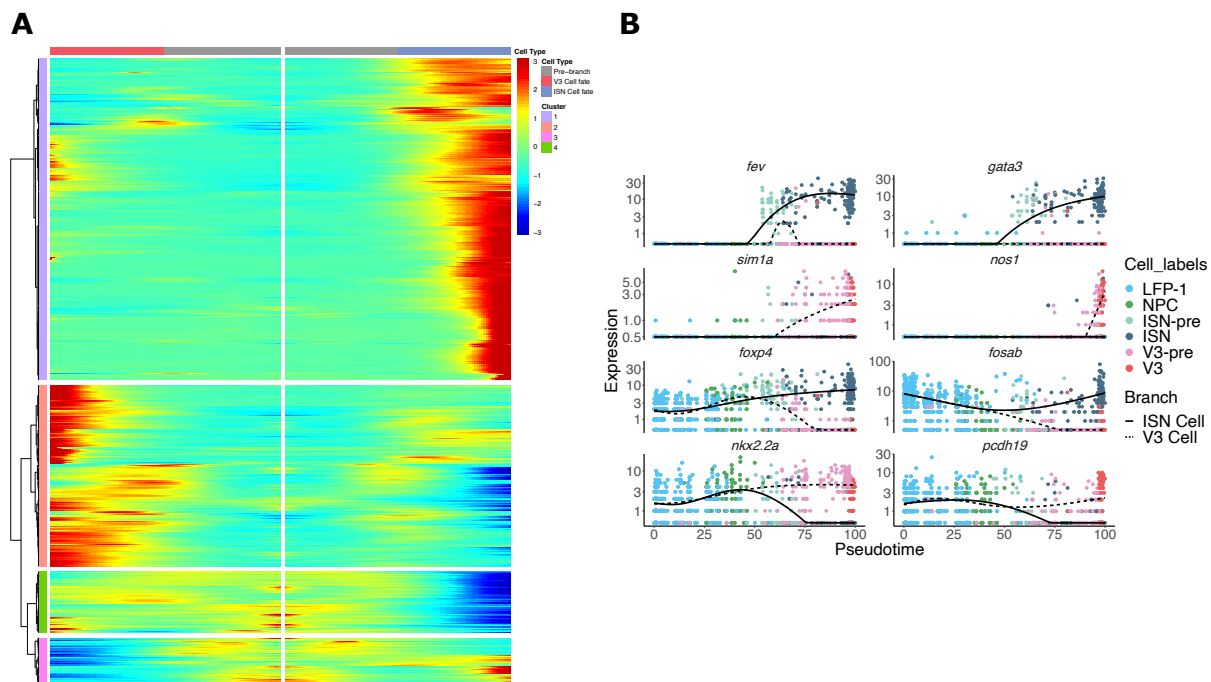


Figure 2.4.10 Comparison of pseudotime-dependent gene expression profiles between V3 and ISN cell lineages derived from LFP progenitors.

- (A) Heatmap showing pseudotime-dependent differentially expressed genes to reveal distinctions between the V3 and ISN cell lineages originating from LFP.
- (B) The expression of key marker genes for differentiation between the V3 and ISN cell lineages originating from LFP is displayed across pseudotime.

2.5 Discussion

Our results identified specific clusters of *sox1a* expressing cells, including KA', KA'', V2-pre, and V2s neurons, alongside the discovery of a new *sox1a*-positive cell type termed ISN. Our findings reveal the different ISN clusters, consisting of precursor cells (*fev*+, *tph2*-) and mature neurons (*fev*+, *tph2*+), suggesting a dynamic developmental trajectory from precursor to mature neuron within this population. Furthermore, we found that ISNs, like KA'' and V3 interneurons, originated from the lateral floor plate (LFP) around 48 hpf. The specification of ISNs is dependent on the downregulation of the Notch signaling pathway, aligning with the developmental process observed in KA'' and V3 interneurons.

2.5.1 LFP progenitor cells generate three distinct ventral neuronal populations with unique molecular profiles and functions

LFP progenitor cells give rise to three distinct ventral neuronal populations, each with unique molecular and functional characteristics. However, the mechanisms of differentiation of various neuronal types, including GABAergic KA'' cells, glutamatergic V3 interneurons, and serotonergic ISNs, arise from the same pool of LFP cells remain poorly understood. Previous studies have indicated a developmental sequence in lateral floor plate (LFP) progenitors, starting with the generation of KA'' cells around 16 hpf, followed by the emergence of V3 cells around 24 hpf (Jacobs *et al*, 2022; Cucun *et al*, 2024). Our studies further demonstrate that ISNs, similar to KA'' and V3 interneurons, originate from the LFP around 48hpf. The emergence of these neuronal populations at different stages suggests that there may be differences in the cellular environments or inducing signals that influence their progenitor pools. This may lead to the activation of different gene regulatory networks and ultimately to the formation of different types of interneurons. During spinal cord development, the amount and duration of the signaling protein Shh is a key factor in activating specific genes within neural progenitor cells along dorsal-ventral (DV) axis of the spinal cord (Sagner & Briscoe, 2017, 2019). Similarly, exposure of different progenitor cells within the lateral floor plate (LFP) to different concentrations and durations of Shh may contribute to the activation of different genes in these three

progenitor cell populations, ultimately driving the development of KA", V3, and ISN. This mechanism may contribute to the generation of these distinct neuronal populations from a shared pool of progenitor cells.

2.5.2 Diverse genetic programs for the development of KA", V3, and ISN from LFP progenitor cells

The developmental trajectories and functional roles of different cell populations within the spinal cord are complexly controlled by the expression and combinatorial action of specific transcription factors (TFs) (Sagner & Briscoe, 2019). When comparing the expression profiles of key TFs in KA", V3, and ISN, we observed only few common TFs primarily related to their shared spatial location within the spinal cord. For example these neurons express genes such as *insm1a* and several hox genes including *hoxb9a*, *hoxb10a*, *hoxc9a* and *hoxc10a*.

Although KA" develop at an earlier embryonic stage than ISNs, we observed more shared TFs between KA" and ISNs. Notably, TFs such as *gata2a*, *gata3*, *sox1a*, *sox1b*, *insm1a*, and *insm1b* were among those observed to be shared between these two cell types, along with other TFs. However, they also exhibit distinct TF expression patterns, for instance, *foxa2*, *nkx2.2a* and *nkx2.2b* are highly expressed in KA", whereas *fev* and *lmx1bb* are highly expressed in ISNs.

Further comparative analysis of the developmental trajectory of V3 and ISNs can provide valuable insights into the differences between these ventral neurons originating from LFP progenitor cells. For instance, our analysis revealed distinct expression patterns: *fev* and *gata3* exhibited high expression in ISNs, *sim1a* displayed significant expression in V3 cells, while *foxp4* and *fosab* were strongly expressed in LFP cells and ISNs. Additionally, *nkx2.2a* were specifically expressed in LFP and V3 cells.

So, it should be noted that the co-expression of a common set of TFs in two cell populations is not sufficient to conclude the existence of a shared gene regulatory

network. Therefore, further extensive research is required to fully understand the diverse genetic programs for the development of KA⁺, V3, and ISN from LFP progenitor cells.

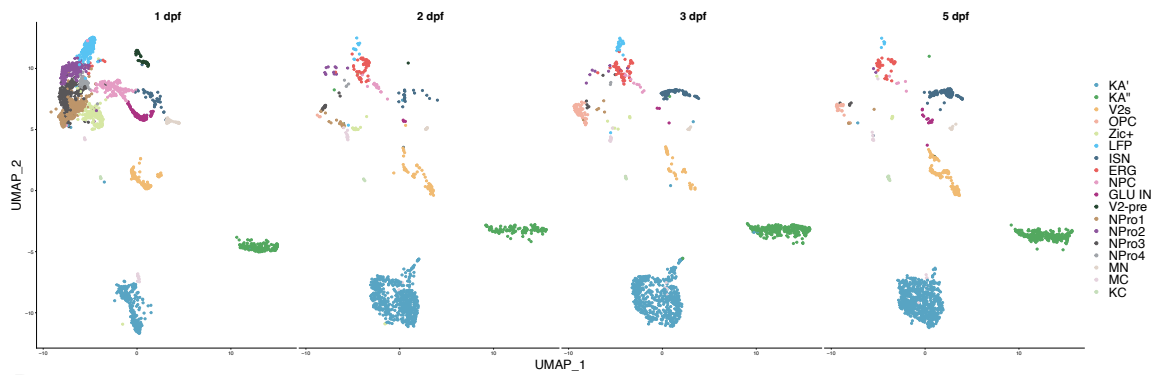
2.5.3 Limitation of the study

Firstly, this study only focused on sox1a expressing cells and does not represent all subtypes of corresponding neural cells. Additionally, our data collection was limited to four time points, lacking more precise time point information. Moreover, this study has shown that ISNs originate from the LFP and are regulated by Notch signaling. However, the precise molecular mechanism underlying this process remains unclear.

Furthermore, this study compared the developmental trajectories of V3 interneurons and ISNs but lacked information on the comparison with KA⁺ interneurons. And it did not explore the precise molecular interactions governing the sequential emergence of KA⁺ interneurons, V3 interneurons, and ISNs from LFP progenitor cells.

2.6 Supplementary Figure

A



B

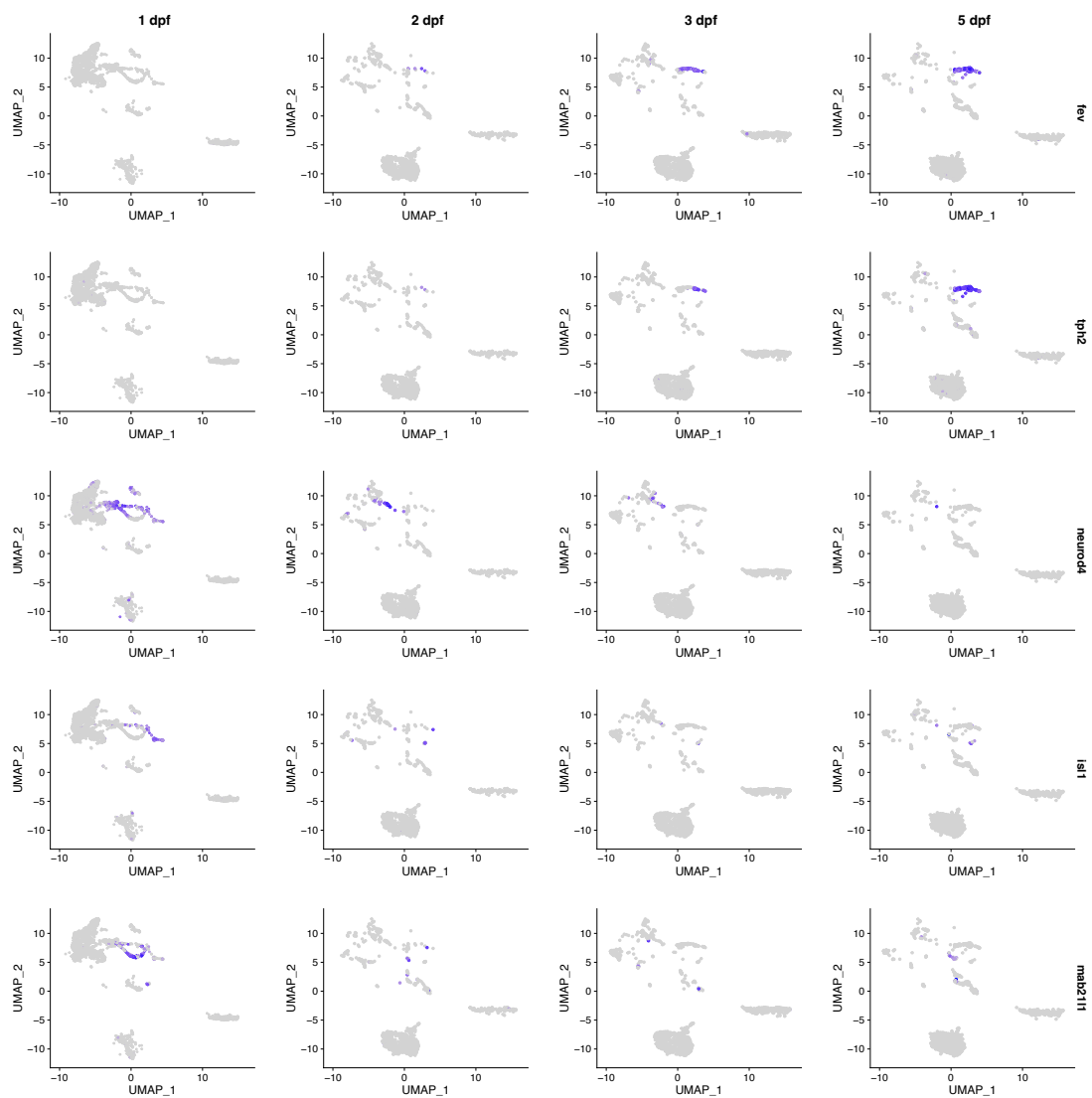


Figure S2.6.1 Cell type distribution and gene expression across developmental stages.

(A) UMAPs showing the distribution of cell types across developmental stages.

(B) Feature plots showing the expression of ISN markers (*fev* and *tph2*), NPC marker (*neurod4*), MN marker (*isl1*), and GLU IN marker (*man21l1*) across developmental stages. Some NPC cells, MN cells, and OPC-pre cells were initially clustered together with ISNs after integration at 1 dpf, but these cells were later excluded from the ISN subcluster and trajectory analysis.

Chapter 3: Integrative analysis of transcriptomes and chromatin accessibility reveals insights into gene regulatory networks across tissues during regeneration

3.1 Introduction

3.1.1 The importance of studying regeneration

In general, most mammals have very limited regenerative capabilities for their organs, such as the brain and heart, which makes recovery challenging after damage occurs. Understanding the cellular and molecular mechanisms underlying organ regeneration is of significant scientific and clinical importance (Shi *et al*, 2015).

The field of regenerative medicine is dedicated to identifying approaches for repairing tissues and organs that cannot recover naturally from injury or degenerative disease (Sánchez Alvarado, 2012). Two strategies have been considered for restoring organs. The first is to explore the potential of producing therapeutic cells from stem cells. However, the efficacy and underlying mechanisms of this treatment are still being evaluated. Another method is to study species capable of regeneration in order to gain a comprehensive understanding of the cellular and molecular mechanisms involved in developing effective organ repair therapies, and to apply the knowledge acquired in this regenerative model to humans in order to promote regeneration (Shi *et al*, 2015). In contrast to the limited regeneration capacity observed in mammals, some vertebrates, such as amphibians and certain fish species, have remarkable regeneration capacities (Khyeam *et al*, 2021). Among them, zebrafish and salamanders demonstrate exceptional regenerative prowess (Khyeam *et al*, 2021). Understanding how and why these animal models can regenerate their damaged or lost organs can guide the development of treatment strategies to promote healing and recovery in humans (Altyar *et al*, 2023). For

example, regeneration research is identifying the key genes and signalling pathways involved in tissue repair. Understanding these genes and pathways may help us to identify potential targets for clinical treatment. In this thesis, I used published omics data in the highly regenerative zebrafish model to identify the common and tissue-specific molecular signature of the regeneration process in the fin, heart, retina and spinal cord.

3.1.2 The mechanisms of tissue regeneration in zebrafish

The mechanism of tissue regeneration in zebrafish includes a variety of cellular and molecular processes, which allows zebrafish to efficiently regenerate various tissues and organs. Zebrafish regeneration usually begins with the activation of immune cell to promote wound healing, eliminate inflammation, and remove cellular debris. The mechanisms for generating new cells vary across different tissues. One such mechanism is activation and proliferation of resident stem cells, such as neural stem cells in the spinal cord and blastema cells in fin tissues. Additionally, tissues have latent stem cells, which are dormant or quiescent under normal conditions but become activated in response to injury. For example, Müller glia in the retina is activated to regenerate damaged retinal tissue. Furthermore, specialized cells can undergo dedifferentiation, transitioning to a stem cell-like state. For instance, dedifferentiation of cardiomyocytes occurs in the regeneration of the zebrafish heart.

Furthermore, a complex network of signaling pathways orchestrates the regenerative process. The key signaling pathways that orchestrate regeneration in various tissues including fin, heart, retina and spinal cord including Fibroblast Growth Factor (FGF), Hedgehog (Hh), Bone Morphogenetic Protein (BMP), Retinoic Acid (RA), Insulin-like Growth Factor (IGF), and Notch signaling (Campbell *et al*, 2021; Becker & Becker, 2020; Cardozo *et al*, 2017; Paredes *et al*, 2019; Gemberling *et al*, 2013; Sehring *et al*, 2016; Wehner & Weidinger, 2015; Wan *et al*, 2014; Lenkowski *et al*, 2013; Todd *et al*, 2018). These pathways regulate various aspects of regeneration, including cell proliferation, differentiation, and tissue remodeling, ensuring the coordinated and efficient repair of damaged tissues in zebrafish.

3.1.3 Fin regeneration in zebrafish

Zebrafish fin exhibits a remarkable ability to regenerate rapidly and stably after amputation (Wehner & Weidinger, 2015). Zebrafish fin regeneration begins with immediate wound healing following fin amputation. The wound quickly closes within 12 hours after the amputation as epithelial cells move to cover it (Gemberling *et al*, 2013; Sehring & Weidinger, 2020). Subsequently these cells aggregate to form a protective multilayered structure called the wound epidermis (Sehring & Weidinger, 2020). Next, the formation of blastema is an important step in fin regeneration. The blastema is comprising a mass of undifferentiated cells that accumulate at the site of injury (Sehring & Weidinger, 2020). It is created when fibroblasts and osteoblasts disorganize and form at the top of each ray through the migration of cells from the distal end (Gemberling *et al*, 2013). These cells possess remarkable plasticity and can give rise to various cell types needed for rebuilding the fin, including bone, skin, and blood vessels (Sehring & Weidinger, 2020). However, the termination signals and mechanisms responsible for this phase remain unknown (Shi *et al*, 2015).

3.1.4 Heart regeneration in zebrafish

In adult mammals, the capacity for cardiac muscle regeneration is significantly limited. Myocardial infarction (MI) is a frequent cause of heart damage in humans. It results in the permanent loss of cardiomyocytes and fibrotic scar tissue formation. This condition is severe in mammals with limited ability to regenerate the affected tissue (Beffagna, 2019). When the heart is injured, the inflammatory response is activated, recruiting immune cells to the damaged site (Simões & Riley, 2022). Dedifferentiated cardiomyocytes reenter the cell cycle and begin rapid proliferation to generate precursor cells (Jopling *et al*, 2010). The precursor cells differentiate into various cell types necessary for heart regeneration, including new cardiomyocytes and endothelial cells for blood vessel formation. After injury, the epicardium undergoes activation, characterized by epithelial-mesenchymal transition (EMT), which converts epicardial cells into migratory and multipotent

mesenchymal cells (Streef & Smits, 2021; Xia *et al*, 2022). Activated epicardial cells migrate toward the injury site and play an important role in initiating and coordinating the regeneration process. Concurrently with cell differentiation, the epicardium also stimulates extra cellular matrix (ECM) components able to create a supportive microenvironment for cell migration, proliferation, and tissue integration during regeneration (Xia *et al*, 2022). Endocardial cells are activated in response to injury, and they secrete factors that promote cardiomyocyte proliferation and differentiation (Lowe *et al*, 2021). Over time, damaged cardiac tissue undergoes remodeling, ultimately restoring structural and functional integrity (Gemberling *et al*, 2013).

3.1.5 Retina regeneration in zebrafish

In humans, neuronal cell loss within the retina can causes significant visual impairment. This loss of neurons may lead to irreversible damage to the visual system. This neuronal cell death often results in glial hypertrophy, restricted proliferation, and glial scarring, thereby preventing neuronal regeneration (Zhang *et al*, 2020). When the zebrafish retina is injured, The injury signals induce Müller glia to undergo dedifferentiation and re-enter the cell cycle, generating multipotent neuronal progenitor cells characterized by increased proliferative capacity and the ability to generate various retinal cell types (Lenkowski & Raymond, 2014; Gemberling *et al*, 2013). These retinal progenitor cells migrate to the area of injury and differentiate into appropriate retinal neuronal cell types, including photoreceptors, bipolar cells, and ganglion cells (Lenkowski & Raymond, 2014). In addition to producing new retinal neuronal cells, zebrafish also display a strong ability to regenerate axons in response to retinal damage (Harvey *et al*, 2019). As regenerating retinal cells mature and make connections, synaptic remodeling occurs to improve neural circuitry and restore visual function (Barrett *et al*, 2022). Over time, the regenerated retina fuses with existing retinal tissue, ultimately restoring visual function to the zebrafish (Gemberling *et al*, 2013).

3.1.6 Spinal cord regeneration in zebrafish

After injury, the specialized progenitor radial glial cells are activated to differentiate into the various cell types to repair the damaged tissue (Briona & Dorsky, 2014). These progenitor radial glial cells are a source of new neurons and glial cells at the site of injury, contributing to the recovery of the neural circuit. An important step in the regeneration of the zebrafish spinal cord is the formation of a "glial bridge" at the site of injury (Goldshmit *et al*, 2012). The glial cells migrate into the damaged site and connect the two sides of the lesion, forming a bridge that serves as a scaffold for axon regeneration (Gemberling *et al*, 2013). Regenerating axons restores communication between neurons and promotes recovery of motor and sensory functions, thereby promoting functional recovery (Tsata & Wehner, 2021). Similar to developmental processes, the types of neurons regenerated in the zebrafish spinal cord correlate with the dorsoventral location of their respective progenitor cells within the spinal cord (Gemberling *et al*, 2013).

3.1.7 Genomics research in zebrafish regeneration

Zebrafish regeneration research has used a variety of high-throughput genomic techniques, including bulk RNA-seq, scRNA-seq, Chip-seq, and ATAC-seq to investigate the gene expression patterns, epigenetic modifications, and regulatory elements involved in the regeneration process (Cao *et al*, 2022; Thompson *et al*, 2020; Hoang *et al*, 2020; Cigliola *et al*, 2023). By comparing the transcriptome and epigenome of injured uninjured tissues, it can help us to identify key genes and regulatory elements specifically activated during regeneration process that encode factors involved in cell proliferation, differentiation, and immune responses. For instance, Thompson JD *et al*., used ATAC-seq and bulk RNA-seq from bulk tissue and purified fibroblasts of uninjured and regenerating caudal fins to identify candidate tissue regeneration enhancer elements for zebrafish fin regeneration. Moreover, they used transgenic and mutant lines to confirm novel non-coding regulatory sequences near key genes involved in fin regeneration, such as *fgf20a*, *mdka*, and *cx43* (Thompson *et al*, 2020). Cao Y *et al*., identified tissue

regeneration enhancer elements responsible for regeneration of injury-induced gene expression by profiling transcriptomes (bulk RNA-seq) and chromatin accessibility (ATAC-seq) in epicardial cells during heart regeneration. And they validated several candidate tissue regeneration enhancer elements for *gnai3* and *ncam1a* genes using stable transgenic reporter lines (Cao *et al*, 2022). Cigliola V *et al.* employed bulk RNA-seq, ATAC-seq, and scRNA-seq techniques to uncover that ependymal cells produce the neurogenic factor *hb-egfa* in response to transection injury. Furthermore, they identified a tissue regeneration enhancer element associated with *hb-egfa*, which regulates gene expression in spinal cord injuries in zebrafish (Cigliola *et al*, 2023). This comprehensive genomic analysis helps to understand the molecular events behind regeneration more quickly and easily.

Another genomic research direction in regeneration is comparative genomic studies in different species. It can provide valuable insights into evolutionarily conserved and species specificity of regeneration. Comparative analyzes can reveal common genetic pathways and regulatory elements of regeneration, as well as unique properties of regenerating organisms. These findings reveal fundamental principles of regeneration and may inspire new regeneration strategies in non-regenerating species, including humans. For example, Wang W *et al.*, used Chip-seq, bulk RNA-seq and scRNA-seq to identify conserved regeneration responsive enhancers between zebrafish and african killifish. The study reveals significant differences in the genomic responses to amputation between killifish and zebrafish, alongside an evolutionarily conserved regeneration response program, predominantly activated by regeneration-specific blastema cells (Wang *et al*, 2020). Hoang T *et al.* performed bulk RNA-seq and ATAC-seq to separately profile gene expression and chromatin accessibility before and after injury in both mouse and zebrafish retina. Additionally, they utilized scRNA-seq to examine changes in gene expression at cell level before and after injury in retinas of mouse, zebrafish, and chick. Cross-species analysis showed that there are similar and different gene networks controlling how Müller glia change when they're stimulated to become active or start growing again. Specifically, in mice, a specialized gene network guided Müller glia to a quiescent state, while in zebrafish and chick, selective genes expression in reactive

Müller glia contributes the transition to a proliferative and neurogenic progenitor state (Hoang *et al*, 2020). These studies can enhance our understanding of conservation and unique features of regeneration across species, potentially providing insights for species with limited regenerative capacity, including human.

Insights gained from zebrafish regenerative genomics studies hold great promise for regenerative medicine. By understanding the gene expression patterns, epigenetic modifications, and regulatory elements that drive tissue repair in zebrafish, researchers can develop new therapies to treat human disease and injury.

3.2 The Aim

Although there are significant differences in the regeneration processes of different tissues due to differences in the cell types and gene regulatory networks, some commonalities can be observed. For example, one study found that leptin b (*lepb*) was significantly upregulated during heart and fin regeneration in zebrafish. Additionally, this study identified a short DNA sequence element upstream and distal to *lepb* that is marked by open chromatin during regeneration (Kang *et al*, 2016). To comprehensively understand the regeneration process across various tissues, we collected the transcriptional and epigenomic sequencing data from tissues involved in regeneration, such as the fin, heart, retina, and spinal cord. By conducting comparative analyses, we aim to unveil both shared gene regulatory networks and tissue-specific factors that drive regeneration of different organs. This work not only deepens our understanding of the intricate process of regeneration, but also provides valuable insights into the development of novel regenerative therapies.

3.3 The Method

3.3.1 The data collection on regenerating tissues including fin, heart, retina, and spinal cord

To understand the core gene regulatory networks involved in the regeneration process, The bulk RNA-seq, scRNA-seq and ATAC-seq datasets were collected from regenerating tissues, including the fin, heart, retina, and spinal cord.

In the case of fin, the GSE146960 dataset was collected (Thompson *et al*, 2020), which provides bulk RNA-seq and ATAC-seq datasets at different time points, ranging from freshly amputated tissue (0 days post-amputation, dpa) and regenerating tissues (1 dpa and 4 dpa). Additionally, for the fin regenerating scRNA-seq data, the GSE137971 was collected which includes single-cell transcriptome profiles of both before injury and at 1/2/4 dpa during regeneration process (Hou *et al*, 2020).

For heart regeneration, the GSE89444 dataset was collected, which provides bulk RNA-seq and ATAC-seq datasets. This dataset includes sample from uninjured ventricles and those injured ventricles at 3/7 dpa from *tcf21:nucEGFP* animals that were performed ventricular resection injuries (Cao *et al*, 2022). scRNA-seq data were obtained from GSE159032 and GSE158919 at uninjured and 3/7/30 dpa (Hu *et al*, 2022).

Retina regeneration data constituted the bulk RNA-seq, ATAC-seq and scRNA-seq conducted at uninjured and 4/10/20/36 hpa (hours post-amputation) following N-methyl-D-aspartate (NMDA) treatment which damage inner retinal neurons. The bulk RNA-seq and ATAC-seq data source is in GSE135406, and scRNA-seq is available through the provided link (<http://bioinfo.wilmer.jhu.edu/jiewang/scRNAseq/>) (Hoang *et al*, 2020).

Lastly, for spinal cord, the GSE193503 was collected for bulk RNA-seq and ATAC-seq and GSE213435 for scRNA-seq which at uninjured and 1 wpi (weeks post transection injury) in wild-type zebrafish (Cigliola *et al*, 2023).

Table 3.1 Sequence Data Sources

Tissues Types	Fin		Heart		Retina		Spinal Cord	
	dpa	source	dpa	source	hpa	source	wpi	source
bulk RNA-seq	0/1/4	GSE146960	0/3/7	GSE89444	0/4/10/20/36	GSE135406	0/1	GSE193503
ATAC-seq								
scRNA-seq	0/1/2/4	GSE137971	0/3/7/30	GSE159032 GSE158919		Below link		GSE213435

The link of scRNA-seq in Retina: <http://bioinfo.wilmer.jhu.edu/jiewang/scRNAseq/>

3.3.2 Bulk RNA-seq analysis on datasets from fin, heart, retina, and spinal cord

The raw sequencing datasets were download by SRA Toolkit (with prefetch --type sra). The fastq files were extracted by fastq-dump in SRA Toolkit (with --gzip --split-3). The adapter in reads were trimmed by using Trim Galore (Krueger) (with -q 25 --phred33 --length 36 -e 0.1 --stringency 3). Low-quality ends in the reads were also trimmed. Shorter reads less than 36 were discarded after adapter trimming. Trimmed reads were aligned to zebrafish genome reference (danRer10) using HISAT2 (Kim *et al*, 2019). The sam files were converted to bam files using samtools (Danecek *et al*, 2021). The read counts in gene level were summarized using featureCounts (Liao *et al*, 2014) (with -p -t exon -g gene_id) with danRer10 gene annotation. The differentially expressed genes were detected using DESeq2 (Love *et al*, 2014). The differentially expressed genes were filtered with log2FoldChange > 1 or log2FoldChange < -1, and adjusted p-value < 0.05. Enrichment analysis of commonly upregulated genes was performed using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) using the

clusterProfiler (Wu *et al*, 2021) R package. The significant GO and KEGG terms were filtered based on the condition of q-value < 0.05. The Venn diagram was generated using the VennDiagram (Chen & Boutros, 2011) R package. The heatmap plots were created using pheatmap and complexHeatmap (Gu *et al*, 2016) R package.

3.3.3 scRNA-seq analysis for fin, heart, retina, and spinal cord

For the fin scRNA-seq data, the count matrixes from preinjury and 1/2/4 dpa samples in GSE137971 were downloaded (Hou *et al*, 2020). These count matrixes were individually analyzed using Seurat V4 (Hao *et al*, 2021). The low-quality, empty, and dead cells exhibit high mitochondrial proportions, low read counts and low numbers of expressed genes. Specifically, cells with more than 20% mitochondrial genes, fewer than 500 reads, or less than 200 features were excluded. Then, the filtered matrixes were created Seurat objects using the 'CreateSeuratObject' function and normalized using the 'NormalizeData' function. The highly variable features were identified using the 'FindVariableFeatures' function (with selection.method = "vst", nfeatures = 2000). All genes were scaled using 'ScaleData' function before dimensionality reduction analysis. The scaled data of variable features were performed PCA using 'RunPCA' function. The top 30 principal components (PCs) were selected for subsequent analysis. The cells were clustered using the 'FindNeighbors' and 'FindClusters' functions (with resolution = 0.5).

The Seurat objects of every sample were integrated using 'FindIntegrationAnchors' and 'IntegrateData' function. The integrated object was dimensionality reduction analysis using 'ScaleData' and 'RunPCA' functions (with npcs = 30). Cell clustering was performed using the 'RunUMAP', 'FindNeighbors' and 'FindClusters' functions (with resolution = 1).

The 23 cell clusters were identified with the resolution 1. The marker genes were identified using 'FindAllMarkers' function. Based on cell type information from Hou *et al*. (Hou *et al*, 2020), The cell clusters were labeled as follows: epithelial cells (*epcam* and *cdh1*), hematopoietic cells (*mpeg1.1* and *cxc3.2*), mesenchymal cells (*msx1b* and *twist1a*), and Mucosal-like cells (*foxp1b* and *arg2*). Moreover, epithelial cells exhibited transcriptional

diversity across three distinct subgroups, comprising the superficial layer (*krt4*), intermediate layer (*tp63*), and basal layers (*tp63* and *krtt1c19e*) of the epithelium.

For the heart scRNA-seq data, the count matrixes from before injury and 3/7/30 dpa were downloaded in GSE159032 and GSE158919 (Hu *et al*, 2022). Hu *et al*. provided the custome code and cell types of information for this data (https://github.com/Bastiaanspanjaard/Heart_regeneration). The matrixes were readed using 'Read10X_h5' function and created Seurat objects by 'CreateSeuratObject' function (with min.cells = 3, min.features = 150) in Seurat (Hao *et al*, 2021). The cells were kept with less than more than 25% mitochondrial genes and fewer than 4100 features. All objects from different time points were merged. The merged object was labeled based on their cell type information (cell.id.table.csv from github link above). Then the merged object was normalized using the 'NormalizeData' function. The highly variable features were identified using the 'FindVariableFeatures' function (with selection.method = "vst", nfeatures = 3500). The data were scaled and dimensionality reduction analysis using 'ScaleData' function and 'RunPCA' function (with npcs = 100). The cells were clustered using the 'FindNeighbors' function (with dims = 1:50) and 'FindClusters' function (with resolution = 4). The data were analyzed using 'RunUMAP' function for visualization.

For the retina scRNA-seq data, the Seurat object was downloaded from the provided link (<http://bioinfo.wilmer.jhu.edu/jiewang/scRNAseq/>) (Hoang *et al*, 2020). This object included the retinal sample at uninjured and 4/10/20/36 hpa following N-methyl-D-aspartate (NMDA) treatment which damage inner retinal neurons. And it also included the cell type information.

For the spinal cord scRNA-seq data, the matrixes from wild-type samples in both sham-injured and 1 week post-injury were downloaded from GSE213435 (Cigliola *et al*, 2023). The matrixes were readed using 'Read10X_h5' function and created Seurat objects by 'CreateSeuratObject' function in Seurat4 (Hao *et al*, 2021). The objects were filtered and integrated using the same analysis steps as in the fin scRNA-seq analysis. The 20 cell

clusters were identified with the resolution 0.5. The marker genes for each cell clusters were identified using 'FindAllMarkers' function. The cell clusters were labeled as following: neural crest cells (*scn4ab*, *foxd3*), dendritic cells (*ccr7*, *ccl38.6*), Oligodendrocytes (*sox10*, *olig2*), meninges cell (*dcn*, *col1a2*), T-cells (*cd83*, *ccl19a.1*), GABAergic neurons (*gad2*, *snap25a*), ependymo-radial glial cells (ERGs: *foxj1a*, *gfap*), endothelial cells (*cd248a*, *kdrl*), neutrophils (*lect2l*, *npsn*, *mpx*), Erythrocytes (*hbba1*, *hemgn*), and macrophages (*mpeg1.1*, *mfap4*). The cluster16 and cluster18 lacking clear markers were not labeled.

3.3.4 ATAC-seq analysis for fin, heart, retina, and spinal cord

The preprocessing of ATAC-seq reads included adapter trimming using Trim Galore (Krueger) (with -q 25 --phred33 --length 35 -e 0.1 --stringency 4). The trimmed reads were aligned to zebrafish genome reference (danRer10) using bowtie2 (Langmead & Salzberg, 2012) (with -X 5000 --no-discordant for paired-end reads), followed by the conversion of sam files to bam files using samtools (Danecek *et al*, 2021). The duplicate reads were marked and removed in the bam files using markdup function in sambamba (Tarasov *et al*, 2015) (with --overflow-list-size 600000 -r). The ATACseq signal peaks were identified using MACS2 (Zhang *et al*, 2008a) (with -f BAM -g 1.4e9 --keep-dup all --nomodel --shift -100 --extsize 200). The bigwig files were generated by using bamCoverage in deeptools (Ramírez *et al*, 2016) (with --normalizeUsing RPKM). The differentially expressed regions were detected using edgeR function in DiffBind (Ross-Innes *et al*, 2012) R package. The significant differentially expressed regions were filtered with FDR < 0.05. These regions were annotated with danRer10 gene annotation information using ChIPseeker (Yu *et al*, 2015).

The upregulated regions were merged together across different time points within each tissue using merge function in bedtools (Quinlan & Hall, 2010). The Intervene tool (Khan & Mathelier, 2017) was used for intersecting and visualizing the merged upregulated regions across different tissues. The heatmap visualizing the ATAC-seq signal of commonly upregulated regions was created using computeMatrix and plotHeatmap

functions in deeptools (Ramírez *et al*, 2016). The enriched motifs within the commonly upregulated regions were analyzed using HOMER (Heinz *et al*, 2010).

3.4 Result

3.4.1 Bulk RNA-seq profiling of regenerating tissues including fin, heart, retina, and spinal cord

To investigate the dynamic changes in gene expression patterns during the regeneration process, thorough analysis of these large bulk RNA-seq datasets was conducted to identify the differentially expressed genes between injured with uninjured at each time point across various tissues. The same method was employed to analyze the data and filter the differentially expressed genes based on the condition with a $\log_2\text{FoldChange} > 1$ or $\log_2\text{FoldChange} < -1$, and adjusted p-value < 0.05 . Genes with a $\log_2\text{FoldChange} > 1$ were classified as upregulated genes, while those with a $\log_2\text{FoldChange} < -1$ were classified as downregulated genes. A list of differentially expressed genes was obtained at each time point across various tissues (Table2).

Table 3.2 The number of differentially expressed genes

Tissues	Group	Up	Down
Fin	0vs1 dpa	4403	5488
Fin	0vs4 dpa	3482	3656
Heart	0vs3 dpa	3934	3475
Heart	0vs7 dpa	1161	1035
Retina	0vs4 hpa	541	480
Retina	0vs10 hpa	1051	1085
Retina	0vs20 hpa	1342	1098
Retina	0vs36 hpa	1068	867
Spinal Cord	0vs1 wpi	2133	1158

To compare the differences in gene expression and similarities across various tissues during the regeneration process, The differentially expressed genes that were

To understand the function of these genes, an enrichment analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) was performed. The results revealed significant enrichment of pathways related to cell cycle and DNA replication and biological processes related cell division and nuclear division. This suggests that these genes are involved in key signaling pathways that control cellular processes, such as cell cycle regulation and DNA replication during regeneration (Figure 3.4.2A and B). Further exploration of the significant enrichment of biological processes revealed three specific items directly related to regeneration, including regeneration, tissues regeneration and cardiac muscle tissue regeneration. The corresponding genes involved in these regenerative biological processes are *tubb5*, *cdh2*, *uhrf1*, *esco2*, *ttk*, *hmmr*, *top2a*, *runx1* and *dkk1a* (Table 3).

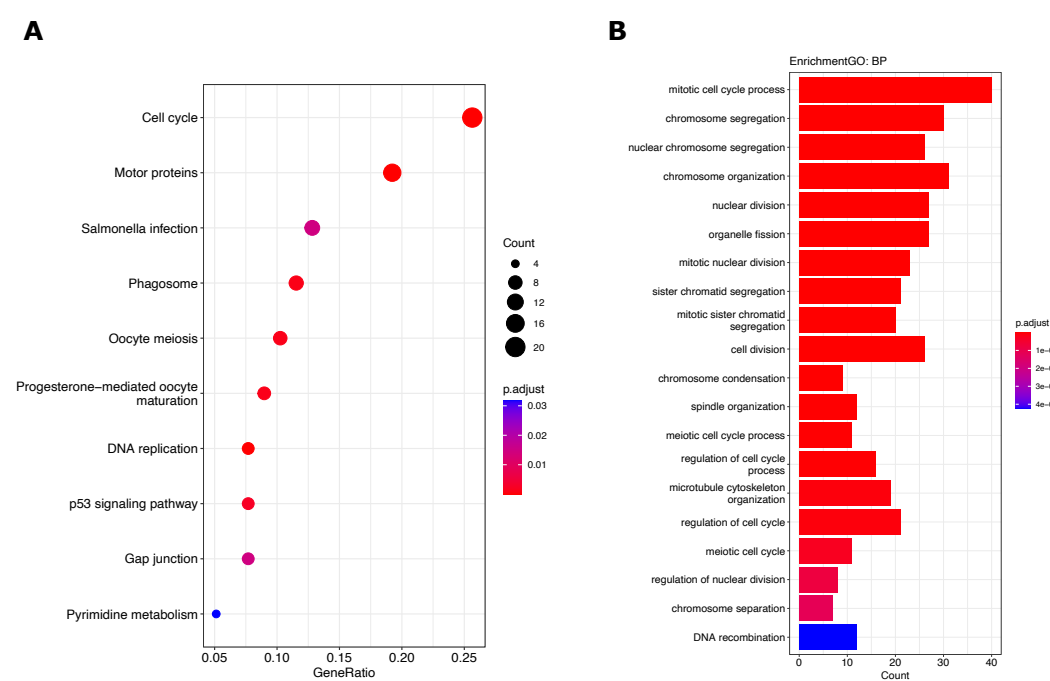


Figure 3.4.2 Functional analysis of 159 commonly upregulated genes.

- (A) KEGG pathway analyses were performed using the 159 commonly upregulated genes.
- (B) GO analyses for Biological Processes were conducted the 159 commonly upregulated genes.

Table 3.3 The GO: BP items related the regeneration

ONTOLOGY	ID	Description	p.adjust	geneID	Count
BP	GO:0031099	regeneration	0.00046	<i>tubb5/cdh2/uhrf1/esco2/ttk/hmmr/top2a/runx1/dkk1a</i>	9
BP	GO:0042246	tissue regeneration	0.00134	<i>tubb5/cdh2/esco2/ttk/hmmr/runx1/dkk1a</i>	7
BP	GO:0061026	cardiac muscle tissue regeneration	0.04884	<i>hmmr/runx1</i>	2

3.4.2 Tissue specific patterns in zebrafish regeneration

In addition to commonly upregulated genes across all four tissues, each tissue exhibits its own set of specific marker genes during the regeneration process. For instance, a total of 2490 upregulated genes were identified in the fin regeneration process, 1288 upregulated genes in heart regeneration process, 890 upregulated genes in retina regeneration and 922 upregulated genes in spinal cord regeneration. This diversity underscores the presence of tissue specific molecular signatures that are likely to be closely related to the different functions and cell types during zebrafish regeneration. In the zebrafish fin regeneration, keratins such as *krt93* and *krt94* play an important role (Figure 3.4.3A). These keratins are structural proteins associated with the cytoskeleton of epithelial cells and help provide necessary structural support during fin regeneration (Doganyigit *et al*, 2023). These specific upregulated genes were enriched in axon development and Wnt signaling pathway during fin regeneration (Figure 3.4.3B). As reported Wnt signaling pathway involved blastema cell proliferation during fin regeneration (Wehner *et al*, 2014). In the zebrafish heart regeneration process, *cfl1l* is involved in actin filament binding activity (Figure 3.4.4A). Additionally, the expression pattern of these genes shows their roles as early induced gene and later induced genes across different the injured stages. For instance, genes like *cfl1l*, *ugt5g1* and *scxa* were highly expressed at 3 dpa, while *rsad2* and *mmrn2a* were highly expressed at 7 dpa. In the zebrafish spinal cord regeneration process, *tbata* is related to the central nervous system (Figure 3.4.4B). In the process of zebrafish retina regeneration, *il34* is involved in immune regulation and tissue homeostasis (Figure 3.4.4C). Additionally *clcn5b* and *tdh2*

showed high expression levels at 4 hpa, while *myo1ea*, *il34*, and *derl3* showed high expression levels at 10 hpa. These unique molecular signatures highlight the diversity of regenerative processes in different tissues at different time points in zebrafish, revealing the complex cellular mechanisms underlying each tissue's unique regeneration.

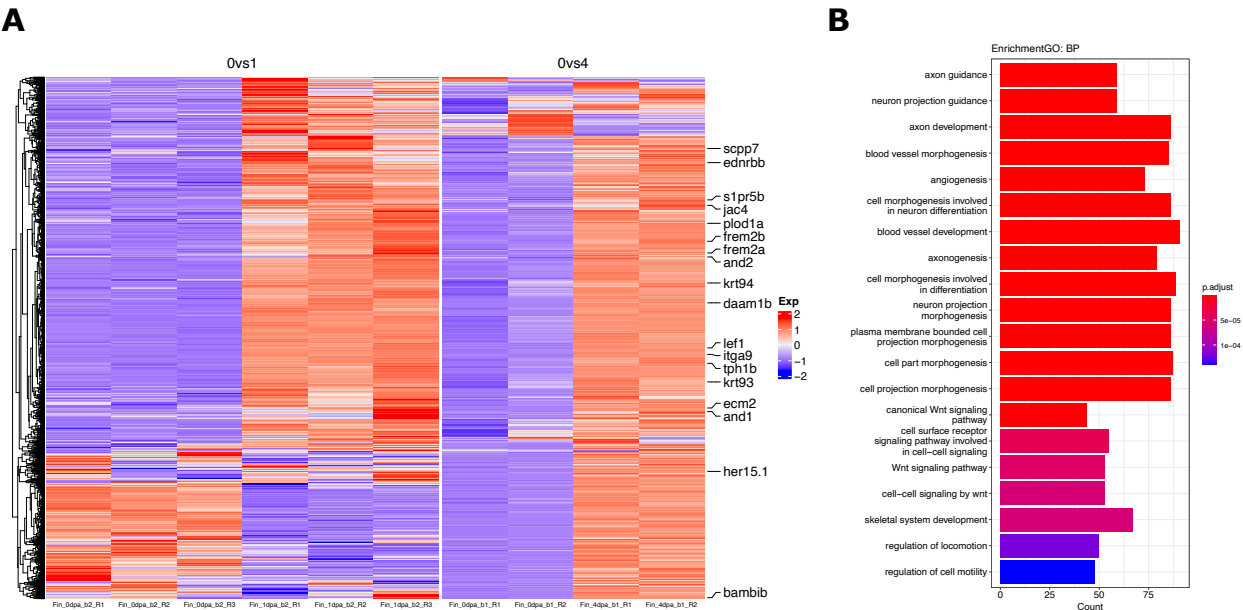


Figure 3.4.3 Specific upregulated genes in zebrafish fin regeneration.

- (A) Heatmap showing the expression profiles of the specific upregulated genes at 1 and 4 dpa during fin regeneration. Notable genes are labeled for reference.
- (B) GO analyses for Biological Processes were conducted the specific upregulated genes at 1 and 4 dpa during fin regeneration.

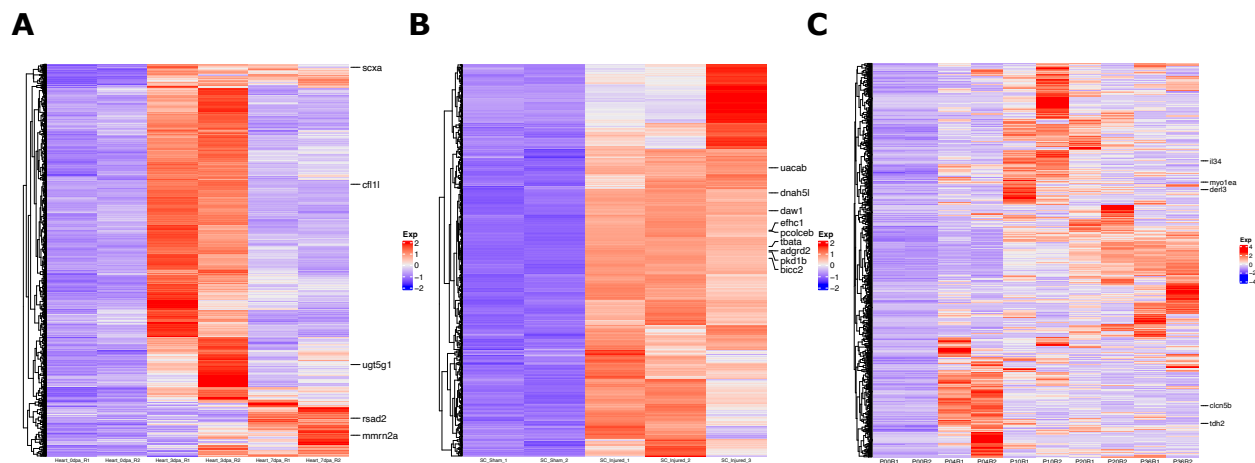


Figure 3.4.4 Specific upregulated genes in zebrafish heart, spinal cord, and retina regeneration.

- (A) Heatmap showing the expression profiles of the specific upregulated genes at 3 and 7 dpa compared with uninjured sample during heart regeneration. Notable genes are labeled for reference.
- (B) Heatmap showing the expression profiles of the specific upregulated genes at 1 wpi compared with uninjured sample during spinal cord regeneration. Notable genes are labeled for reference.
- (C) Heatmap showing the expression profiles of the specific upregulated genes at 4, 10, 20 and 36 hpa compared with uninjured sample during retina regeneration. Notable genes are labeled for reference.

3.4.3 scRNA-seq profiling of regenerating tissues including fin, heart, retina, and spinal cord

Now, insight into the gene expression changes during the regeneration of each tissue has been gained. However, the traditional bulk RNA-seq only provides average expression levels and can't discern expression changes at the individual cell level from whole sample sequencing. the regeneration process involved complex biologic mechanisms. Different

cell types play an important role in different tissue regeneration process. To understand this complex biological mechanism at the cell level, scRNA-seq analysis was performed.

For the fin regenerating scRNA-seq data, Hou et al. collected the regenerating fin samples before injury and at 1-, 2-, and 4-days post-amputation (dpa) to investigate the stages of blastema formation, outgrowth, and maintenance. They identified epithelial cells (*epcam* and *cdh1*), hematopoietic cells (*mpeg1.1* and *cxcr3.2*), mesenchymal cells (*msx1b* and *twist1a*), and Mucosal-like cells (*foxp1b* and *arg2*). Epithelial cells are transcriptional diversity across three distinct subgroups, including the superficial layer (*krt4*), intermediate layer (*tp63*), and basal layers (*tp63* and *krtt1c19e*) of the epithelium (Hou et al, 2020). To further investigate the fin regeneration process at the cell level, the data was downloaded for reanalysis, successfully reproducing the results (Figure 3.4.5A, B and C).

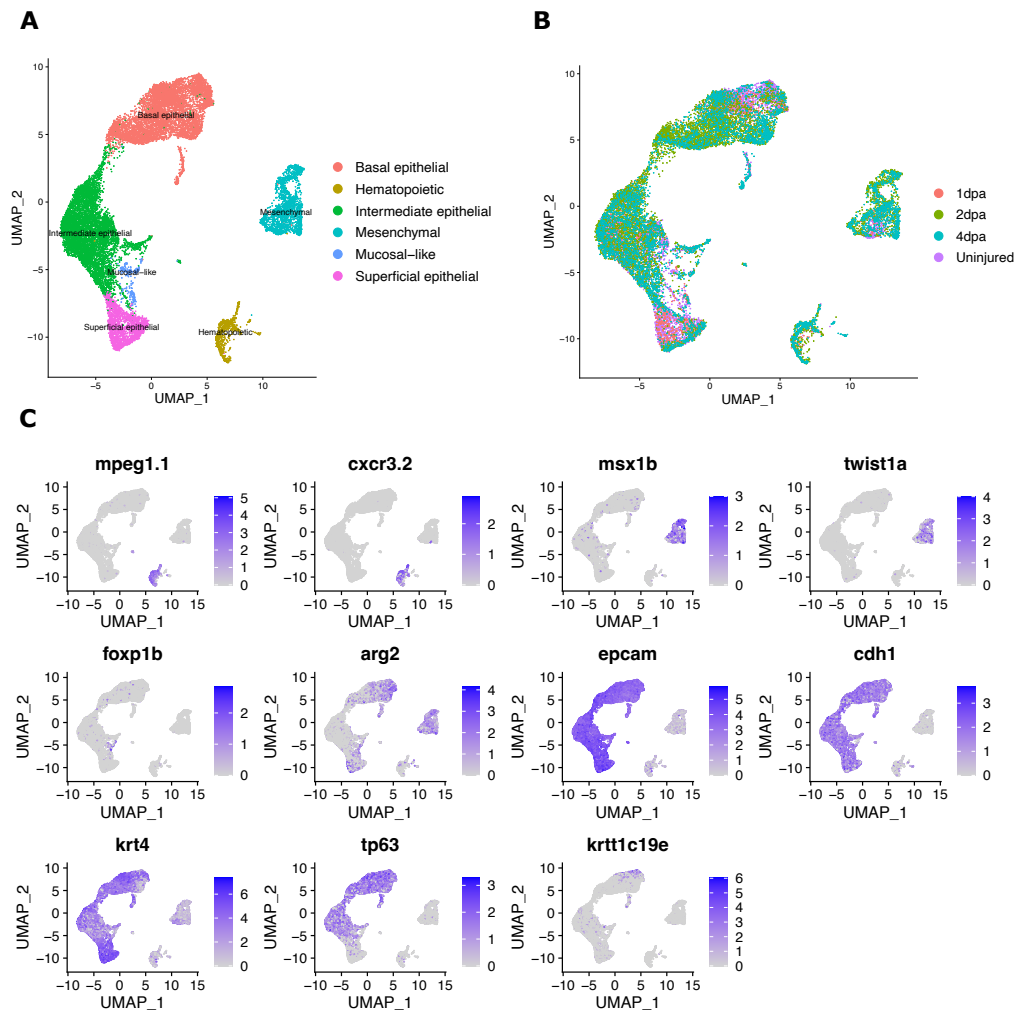


Figure 3.4.5 The identification of cell type in zebrafish fin.

- (A) Cluster assignments for caudal fin cells obtained from both pre-injury and regenerating stages were visualized on UMAP axes from (Hou *et al*, 2020) .
- (B) The distribution of cell across different stages.
- (C) The expressions of key marker genes associated with the major cell types.

Hu *et al.* performed scRNA-seq on approximately 200,000 dissociated cells for the heart regeneration, collecting samples at various stages both before injury and 3/7/30 dpa (Hu *et al*, 2022). They found a strong increase in fibroblasts and immune cells following injury. Furthermore, they characterized the cell types within the three primary layers of the heart:

epicardium, myocardium, and endocardium in both the atrium and ventricle (Figure 3.4.6A and B).

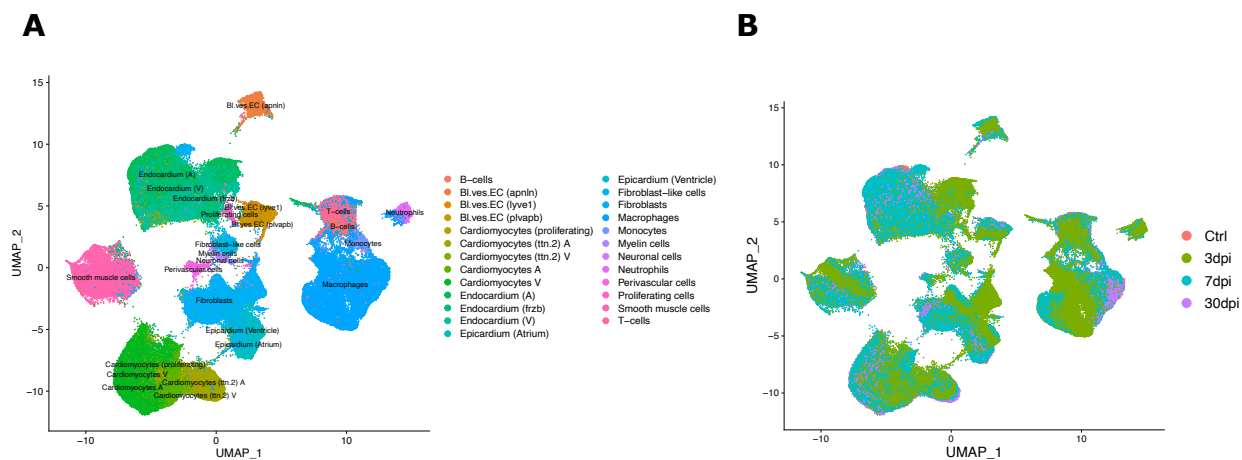


Figure 3.4.6 The identification of cell type in zebrafish heart.

- (A) Cluster assignments for heart cells obtained from both pre-injury and regenerating stages were visualized on UMAP axes from (Hu *et al*, 2022).
- (B) The distribution of cell across different stages.

Hoang et al. performed the scRNA-seq for the retinal sample at uninjured and 4/10/20/36 hpa (hours post-amputation) following N-methyl-D-aspartate (NMDA) treatment which damage inner retinal neurons (Hoang *et al*, 2020). They identified the activated Müller glia cells which were mainly from injured samples (Figure 3.4.7A and B).

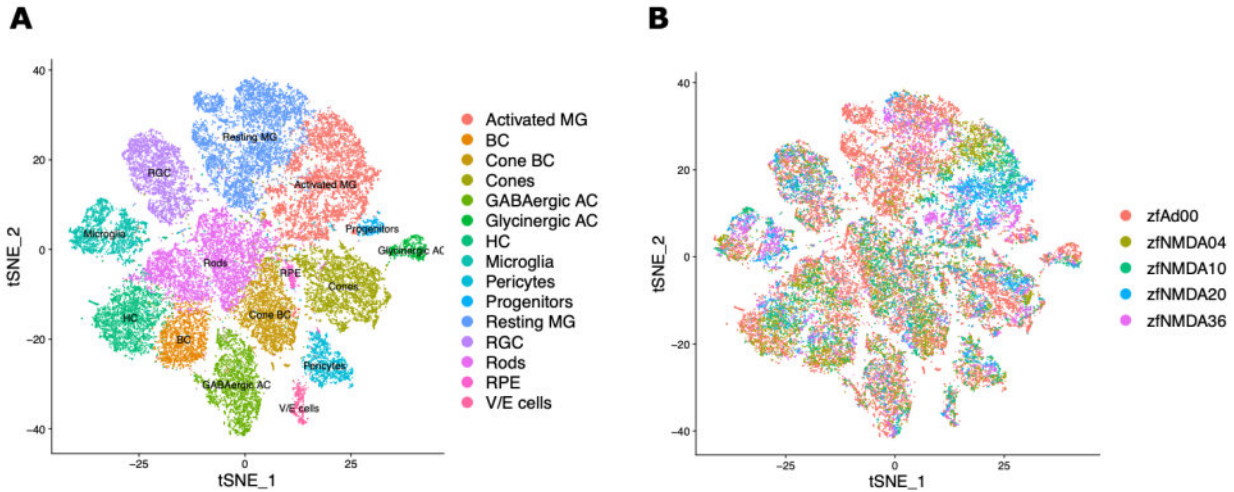


Figure 3.4.7 The identification of cell type in zebrafish retina.

- (A) Cluster assignments for retina cells obtained from both pre-injury and regenerating stages were visualized on tSNE axes from (Hoang *et al*, 2020).
- (B) The distribution of cell across different stages.

Cigliola *et al.* performed scRNA-seq profiling on spinal cords from both wild-type and *hbegfa*KO, including sham-injured and 1 week post-injury (wpi) (Cigliola *et al*, 2023). The wild-type data was downloaded reanalysis to investigate the spinal cord regeneration. The neural crest cells (*scn4ab*, *foxd3*), dendritic cells (*ccr7*, *ccl38.6*), Oligodendrocytes (*sox10*, *olig2*), meninges cell (*dcn*, *col1a2*), T-cells (*cd83*, *ccl19a.1*), GABAergic neurons (*gad2*, *snap25a*), ependymo-radial glial cells (ERGs: *foxj1a*, *gfap*), endothelial cells (*cd248a*, *kdrl*), neutrophils (*lect2l*, *npsn*, *mpx*), Erythrocytes (*hbba1*, *hemgn*), macrophages (*mpeg1.1*, *mfap4*) were identified (Figure 3.4.8A, B and C).

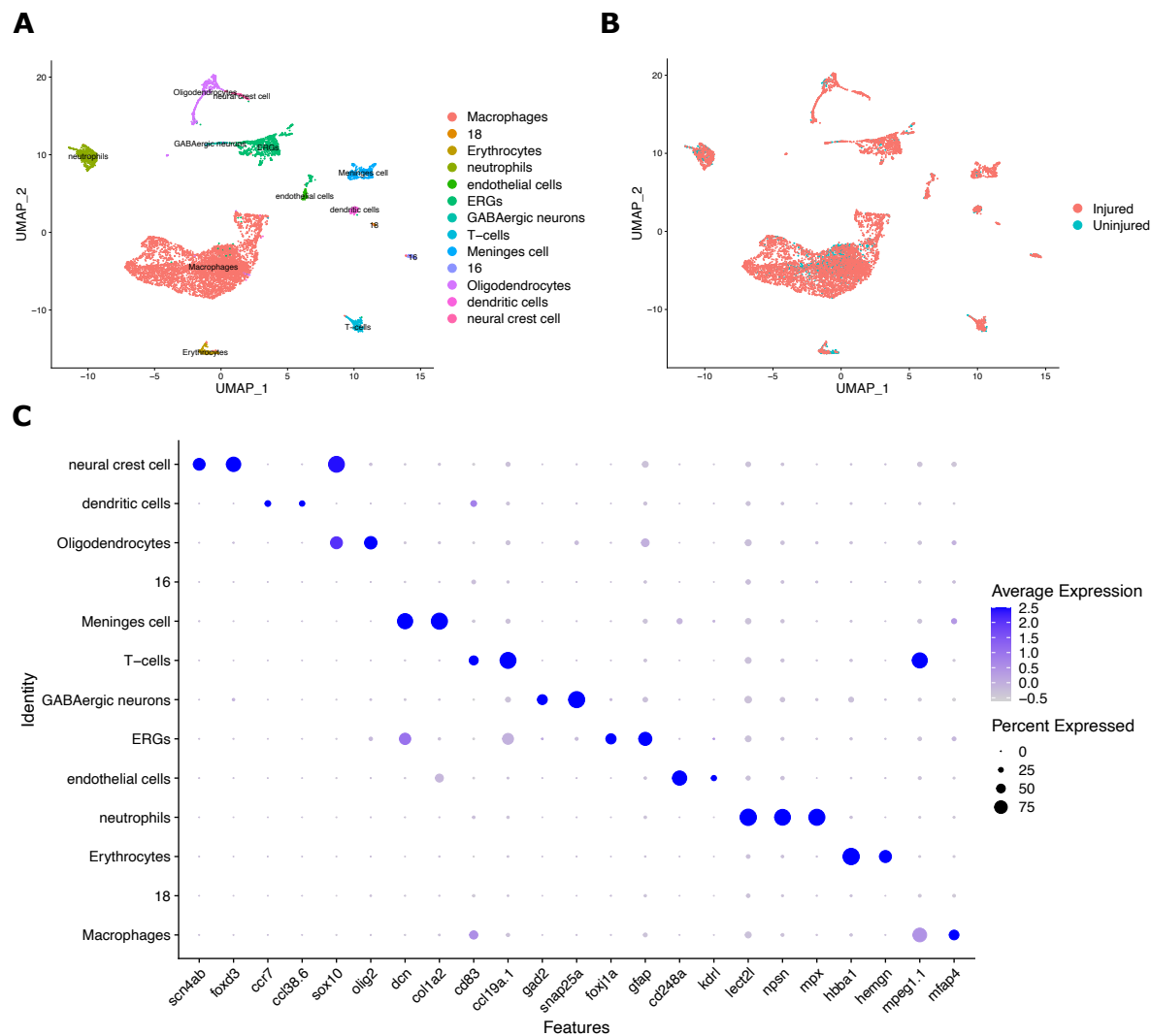


Figure 3.4.8 The identification of cell type in zebrafish spinal cord.

- (A) Cluster assignments for spinal cord cells obtained from both pre-injury and regenerating stages were visualized on UMAP axes from (Cigliola *et al*, 2023).
- (B) The distribution of cell of uninjured and injured samples.
- (C) The expressions of key marker genes associated with the major cell types.

We have defined 195 commonly upregulated genes in these tissues during regeneration process by Bulk RNA-seq data analysis. To know in which cell types this gene set is enriched, the AUCCell software was used, which can estimate the proportion of genes within the set that exhibit high expression levels in each cell (Aibar *et al*, 2017). After

analysis, it was observed that this gene set is enriched in various cell types: hematopoietic, mesenchymal, intermediate, and basal epithelial cells in the fin; fibroblasts, epicardium, endocardium, and macrophages in the heart; activated Müller glia, microglia and progenitors in retina; ependymo-radial glial cells, macrophages and meninges cell in spinal cord (Figure 3.4.9). This observation suggests that commonly upregulated genes are enriched in cells with regenerative potential and immune related functions in different tissues. It highlights the potential universality of commonly upregulated genes contributing to the regeneration process in different tissues.



Figure 3.4.9 AUCCell analysis of commonly upregulated genes across cell types in fin, heart, retina, and spinal cord scRNA-seq datasets.

UMAP/tSNE showing the AUCCell scores for the activities of commonly upregulated genes set in each cell of fin (A), heart (B), retina (C) and spinal cord (D).

lepb gene is a paralog of the leptin gene in zebrafish. It plays an important role in energy homeostasis. Notably, It has been reported to exhibit high expression levels during fin and heart regeneration processes (Kang *et al*, 2016). Moreover, *lepb* can be specific marker to identify the injured activated endocardium cells (Shin *et al*, 2022). Analysis of scRNA-seq reveals the expression patterns for *lepb* in various tissues: highly expressed in mesenchymal cells after fin injury (Figure 3.4.10A) and predominantly in edocardium cell after heart injury, consistent with prior report (Figure 3.4.11A). In the retina, activated müller glia cells exhibited this expression pattern (Figure 3.4.12A). In spinal cord, meninges showed similar expression profiles (Figure 3.4.13A).

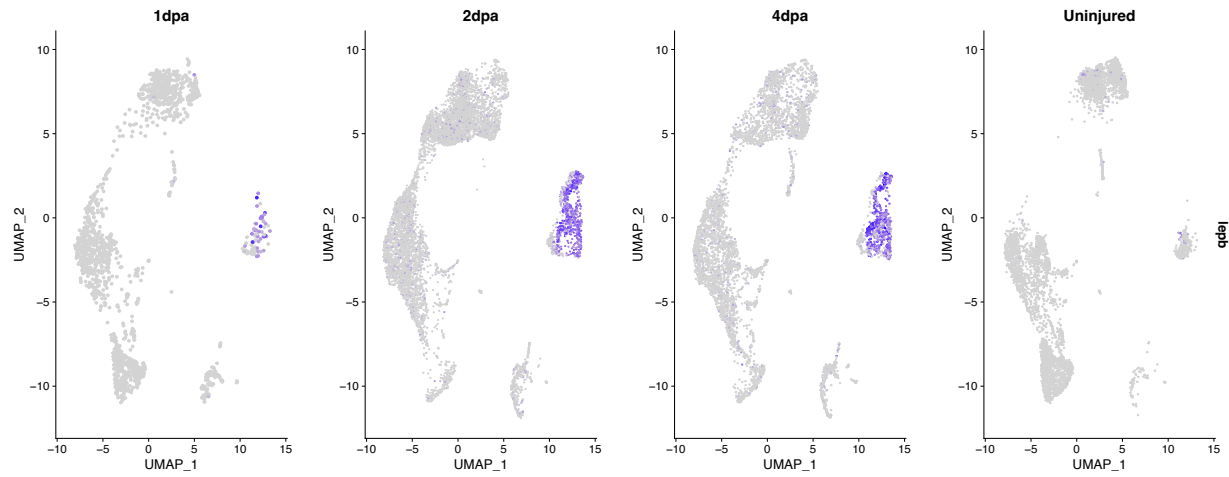
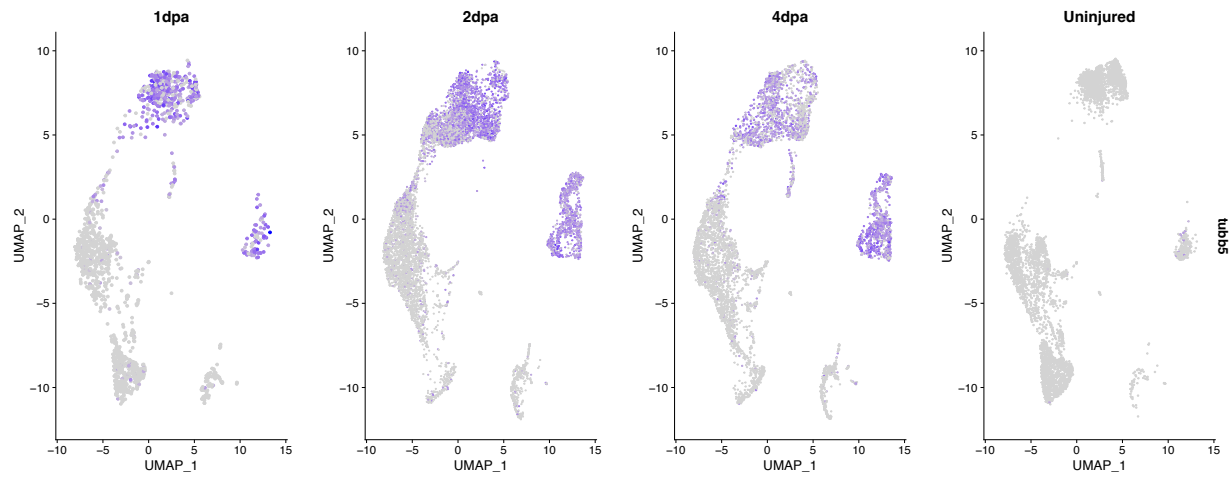
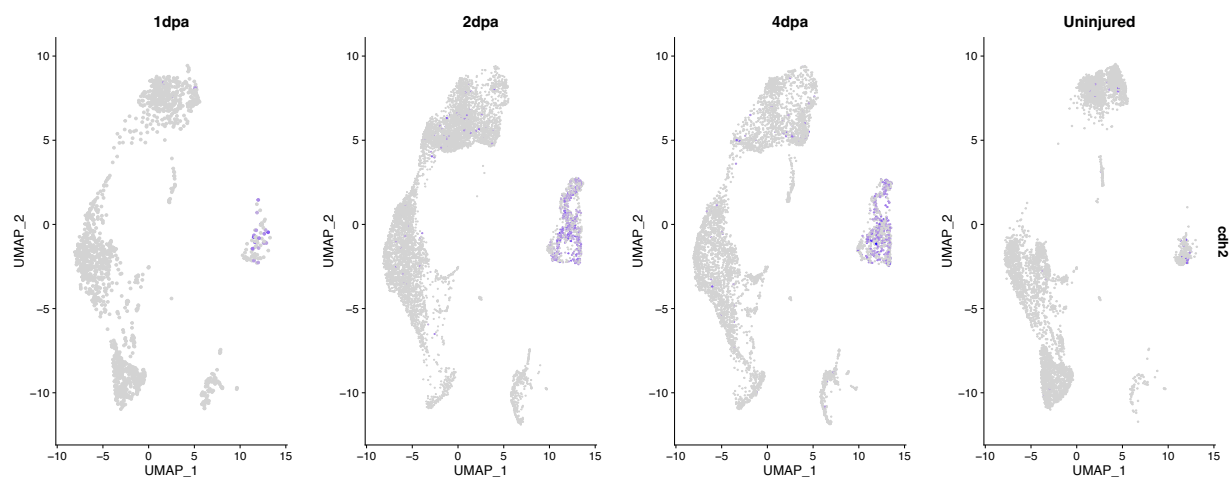
A**B****C**

Figure 3.4.10 The expression of *lepb*, *tubb5* and *cdh2* in zebrafish fin.

FeaturePlot showing the expression of *lepb* in zebrafish fin (A), *tubb5* in zebrafish fin (B) and *cdh2* in zebrafish fin (C) at different stages during regeneration.

At the same time, these commonly upregulated gene sets were enriched in regeneration-related items, and the corresponding genes *tubb5* and *cdh2* were also found. Then these genes were checked in scRNA-seq data. *tubb5* exhibited high expression levels in basal epithelial and Mesenchymal cells after fin injury (Figure 3.4.10B). And *cdh2* exhibited high expression levels in Mesenchymal cells after fin injury (Figure 3.4.10C). Moreover, *tubb5* and *cdh2* were highly expressed in Fibroblasts and Epicardium cell after heart injury (Figure 3.4.11B and Figure 3.4.11C). In the retina, activated Müller glia cells exhibited this expression pattern (Figure 3.4.12B and Figure 3.4.12C). In spinal cord, ERGs showed similar expression profiles (Figure 3.4.13B and Figure 3.4.13C).

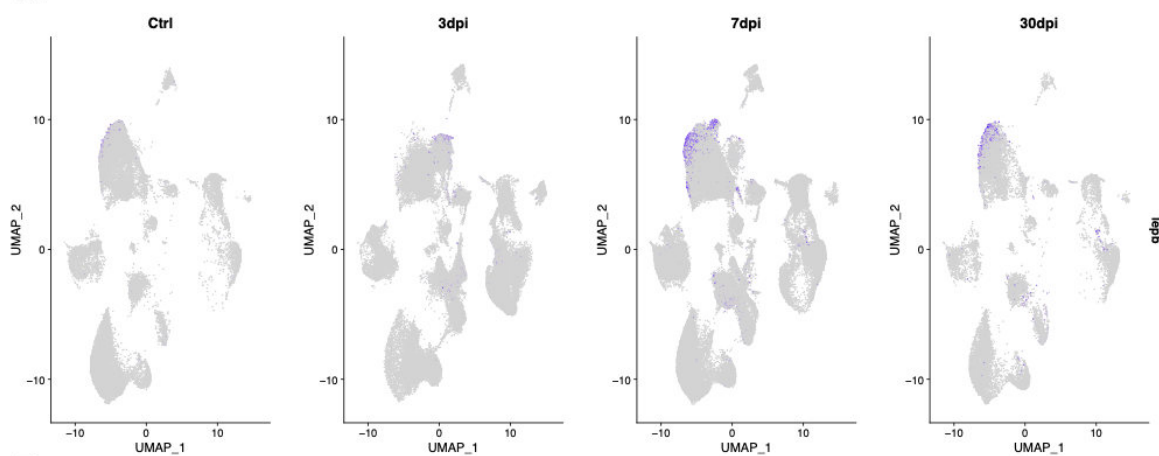
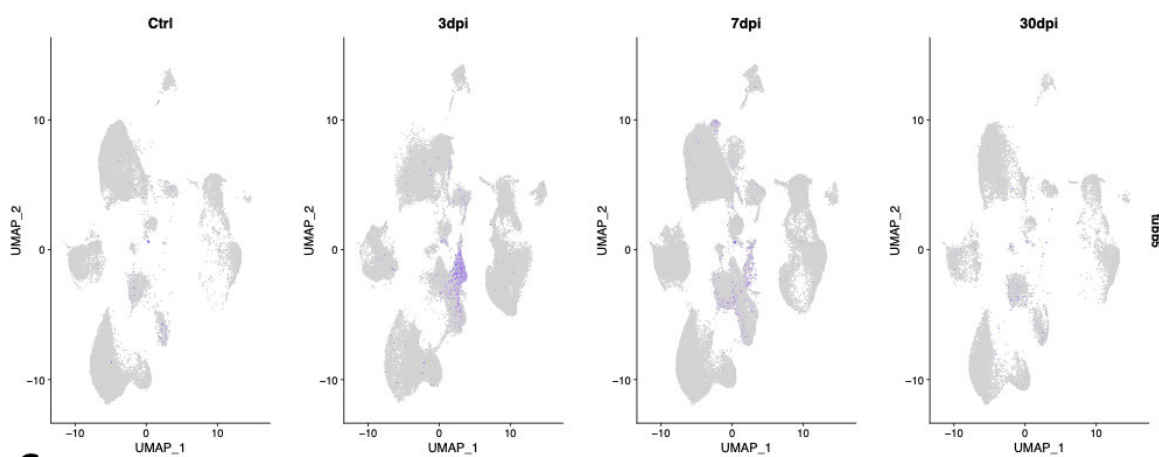
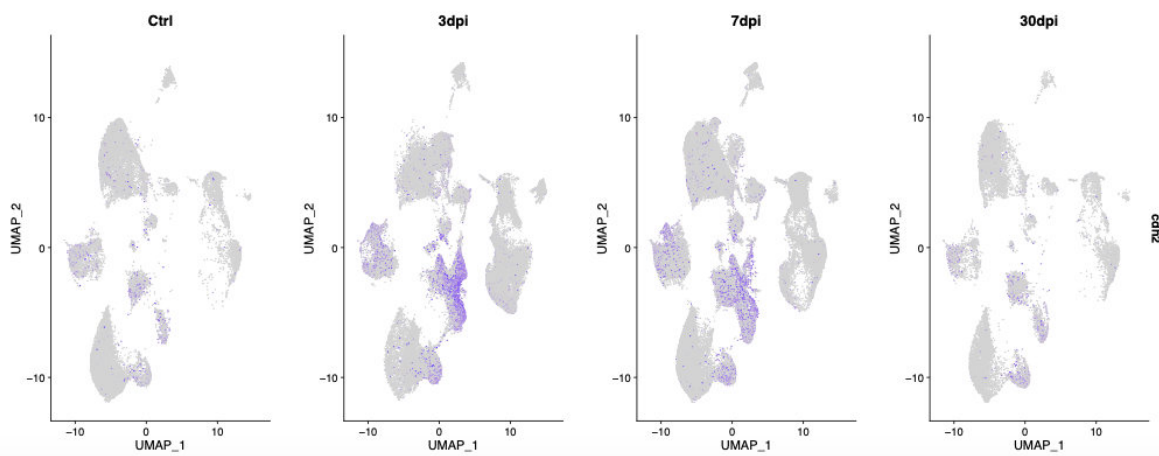
A**B****C**

Figure 3.4.11 The expression of *lepb*, *tubb5* and *cdh2* in zebrafish heart.

FeaturePlot showing the expression of *lepb* in zebrafish heart (A), *tubb5* in zebrafish heart (B) and *cdh2* in zebrafish heart (C) at different stages during regeneration.

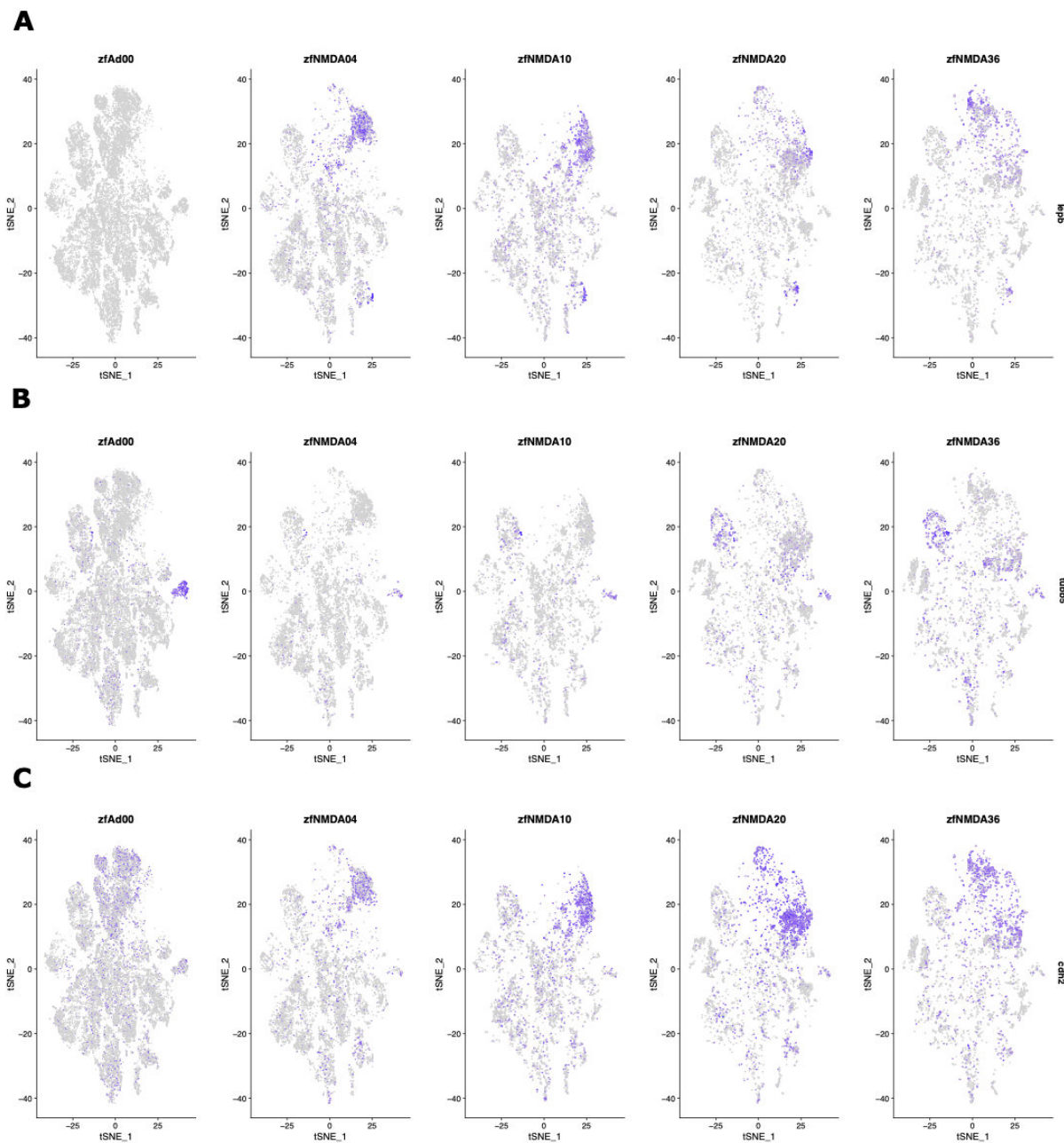


Figure 3.4.12 The expression of *lepb*, *tubb5* and *cdh2* in zebrafish retina.

FeaturePlot showing the expression of *lepb* in zebrafish retina (A), *tubb5* in zebrafish retina (B) and *cdh2* in zebrafish retina (C) at different stages during regeneration.

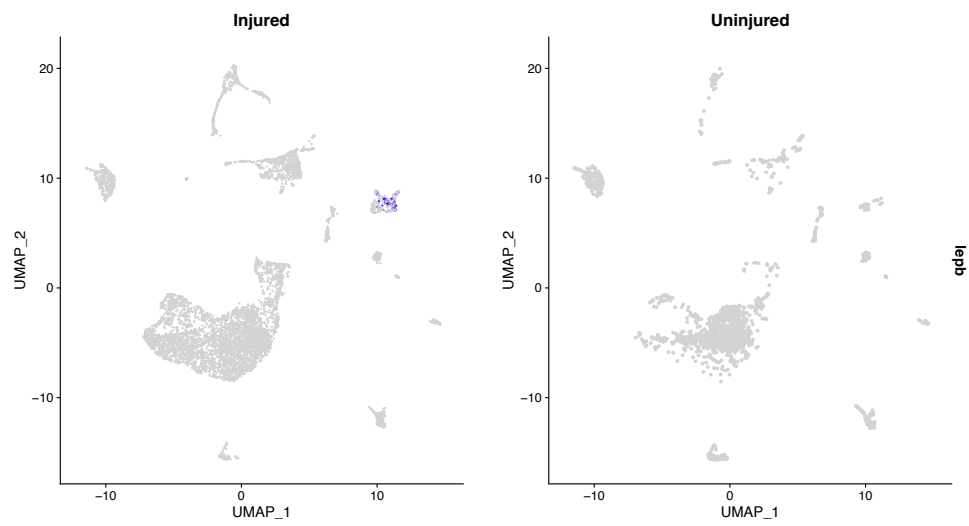
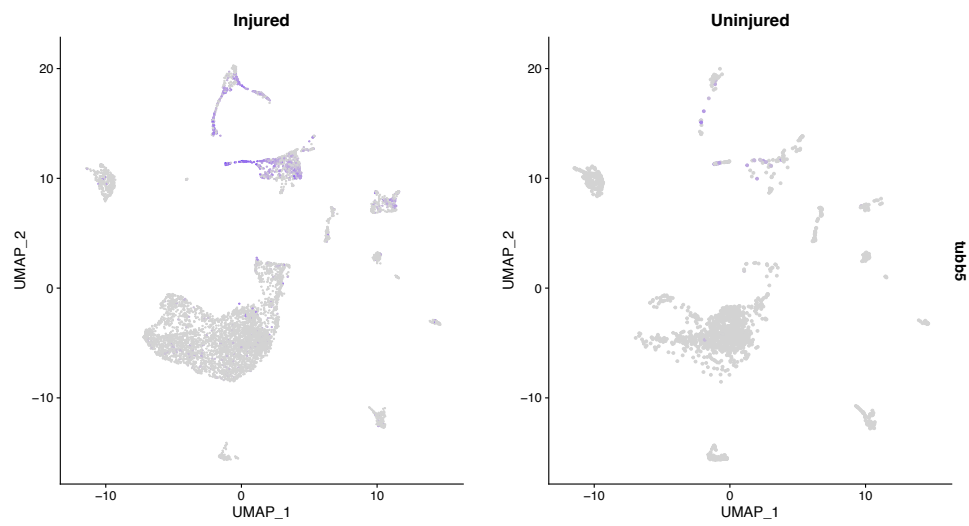
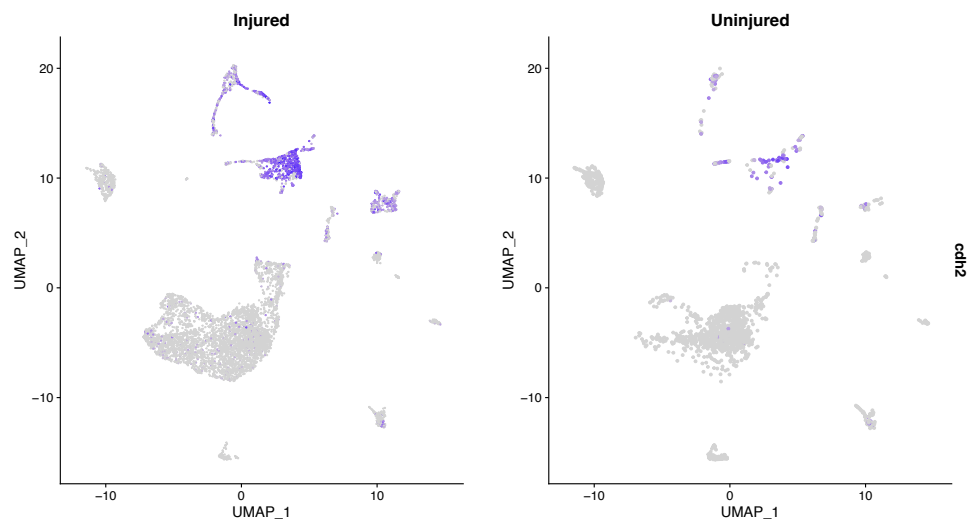
A**B****C**

Figure 3.4.13 The expression of *lepb*, *tubb5* and *cdh2* in zebrafish spinal cord.

FeaturePlot showing the expression of *lepb* in zebrafish spinal cord (A), *tubb5* in zebrafish spinal cord (B) and *cdh2* in zebrafish spinal cord (C) at different stages during regeneration.

Additionally, *krt93* and *krt94* emerged as uniquely upregulated genes during fin regeneration, showing specificity in their expression that was not observed in the regeneration of other tissues (Figure 3.4.14). They were highly expressed in intermediate epithelial cells during the fin regeneration process. The unique upregulation of *krt93* and *krt94* during fin regeneration not only emphasizes their specific involvement in this specific regeneration process, but also suggests that different tissues exhibit different functional requirements.

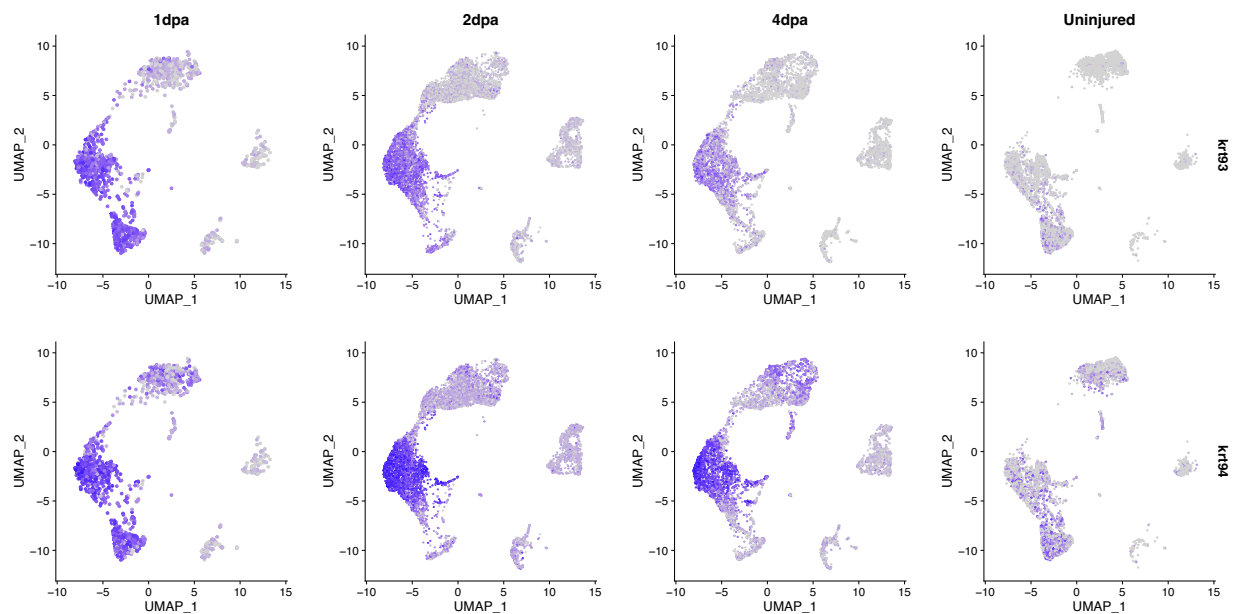


Figure 3.4.14 The expression of *krt93* and *krt94* in zebrafish fin.

FeaturePlot showing the expression of *krt93* and *krt94* in zebrafish fin at different stages during regeneration.

3.4.4 ATAC-seq profiling of regenerating tissues including fin, heart, and retina

To study the dynamic changes of chromatin during the regeneration process, the ATAC-seq datasets were examined to identify the differentially expressed regions between injured with uninjured at each time point in different tissues. ATAC-seq datasets were obtained for the fin, heart, retina, and spinal cord. However, it was observed that the mapping rate for most sample in spinal cord ATAC-seq sequencing data (GSE193503) (Cigliola *et al*, 2023) were very low, suggesting contamination by bacteria. Therefore, for subsequent analysis, focused was only directly towards the ATAC-seq analysis of the fin, retina, and heart.

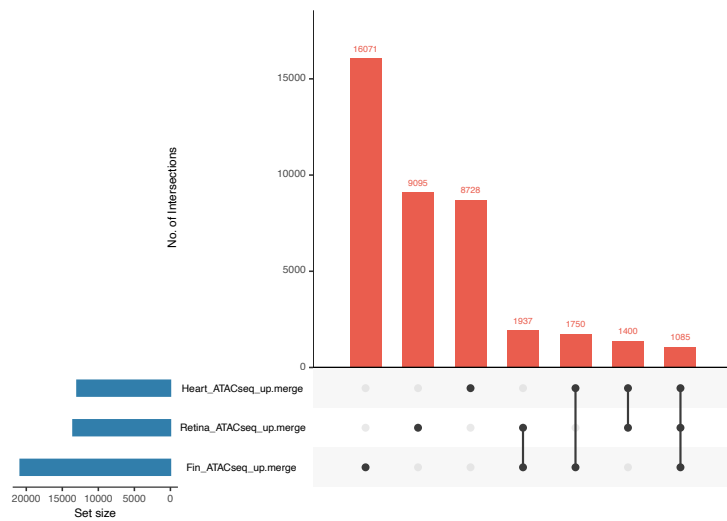
The same method was used to analyze the data, and the differentially expressed regions were filtered based on the condition with $FDR < 0.05$. Subsequently, the differentially expressed regions exhibiting increased accessibility after injury were classified as upregulated peak regions, while those exhibiting reduced accessibility after injury were labeled as downregulated peak regions. A list of differentially expressed regions at each time point in different tissues was then generated (Table 4).

Table 3.4 The number of differentially expressed regions

Tissues	Group	Up	Down
Fin	0vs1 dpa	19529	52
Fin	0vs4 dpa	2506	119
Heart	0vs3 dpa	12649	5116
Heart	0vs7 dpa	2746	76
Retina	0vs4 hpa	11438	689
Retina	0vs10 hpa	919	2053
Retina	0vs20 hpa	997	21509
Retina	0vs36 hpa	3700	4997

To compare the differences and similarities in chromatin dynamics across various tissues during the regeneration process, the upregulated regions from different time points in each tissue were merged, followed by comparative analysis between these regions across different tissues. A common set of 1085 upregulated peak regions, showing increased activity during the regeneration processes of fin, heart, and retina, was identified (Figure 3.4.15A and B). These regions show a significant increase in ATAC-seq signal in the injured samples compared to the uninjured samples in fin, heart, and retina (Figure 3.4.15B).

A



B

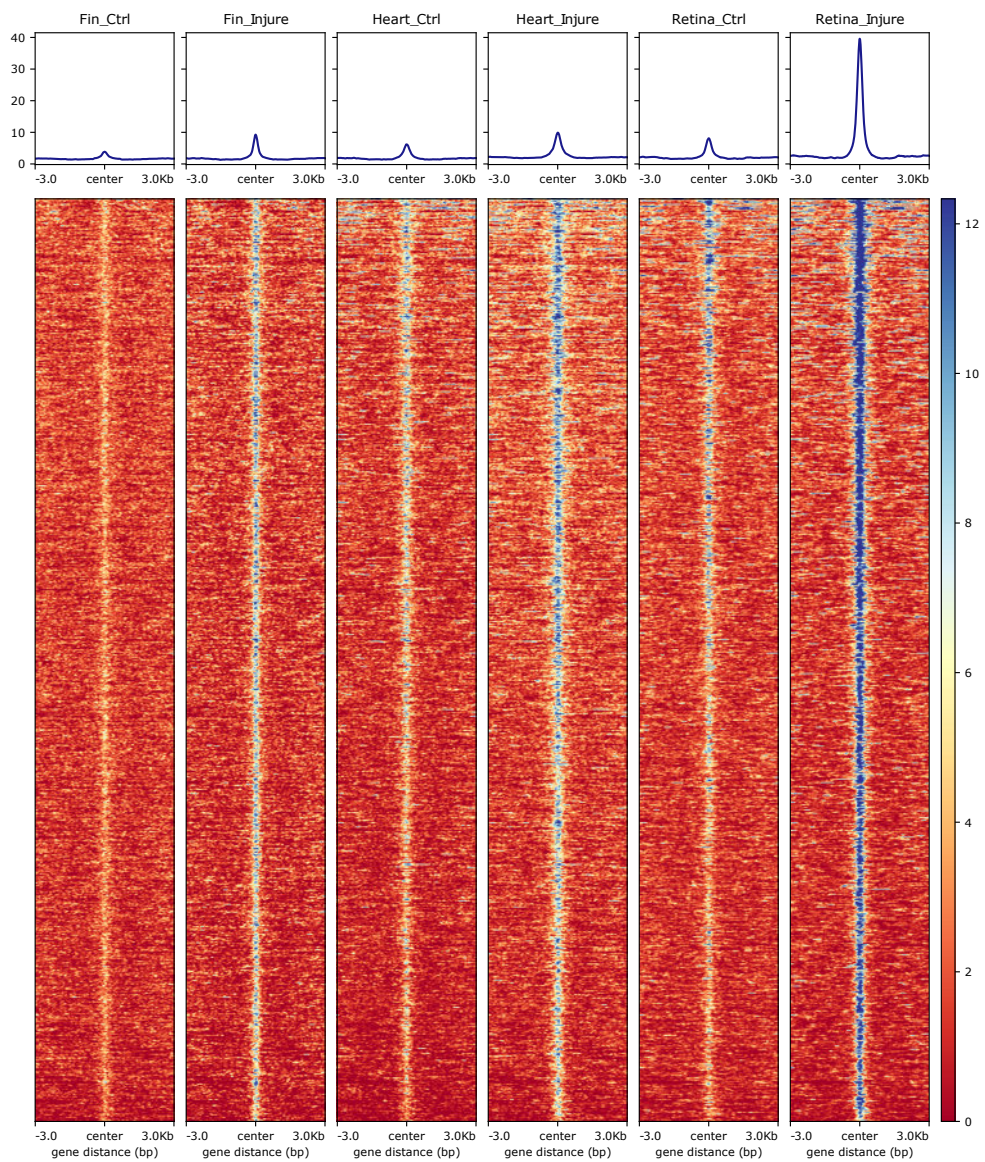


Figure 3.4.15 Comparative analysis of upregulated regions in fin, heart, and retina.

- (A) The UpSet plot illustrates the intersection of upregulated regions in ATAC-seq, providing the number of upregulated regions across comparisons involving fin, heart, and retina datasets.
- (B) The Heatmap shows the ATAC-seq signal of 1085 commonly upregulated regions at uninjured and injured samples in fin, heart, and retina.

To identify the candidate binding sites of transcriptional factors, the motif analysis for these 1085 commonly upregulated regions was conducted by using HOMER (Heinz *et al*, 2010). The motifs related to regeneration, such as Fos, AP-1, Fosl2, Jun-AP, NF-E2 and Nrf2, were enriched in these regions (Figure 3.4.16). The previous report suggests that the AP-1 motif activates gene expression after injured during zebrafish heart regeneration (Beisaw *et al*, 2020). NF-E2 is essential for thrombocyte formation in adult zebrafish (Rost *et al*, 2018). Nrf2 is important for liver regeneration (Bellanti *et al*, 2023). This result suggests the existence of conserved chromatin features during different tissue regeneration processes. The enrichment of specific motifs such as Fos, AP-1, Fosl2, Jun-AP, NF-E2 and Nrf2 suggests the existence of common regulatory elements in the regeneration of different tissues.

Rank	Motif	Name	P-value
1		Atf3(bZIP)/GBM-ATF3-ChIP-Seq(GSE33912)/Homer	1e-074
2		Fos(bZIP)/TSC-Fos-ChIP-Seq(GSE110950)/Homer	1e-073
3		Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46160)/Homer	1e-065
4		BATF(bZIP)/Th17-BATF-ChIP-Seq(GSE39756)/Homer	1e-062
5		Fra2(bZIP)/Striatum-Fra2-ChIP-Seq(GSE43429)/Homer	1e-046
6		AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-037
7		JunB(bZIP)/Dendritic Cells-Junb-ChIP-Seq(GSE36099)/Homer	1e-023
8		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)/Homer	1e-058
9		Jun-AP1(bZIP)/K562-cJun-ChIP-Seq(GSE31477)/Homer	1e-490
10		Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	1e-240
11		Bach1(bZIP)/K562-Bach1-ChIP-Seq(GSE31477)/Homer	1e-108
12		NF-E2(bZIP)/K562-NFE2-ChIP-Seq(GSE31477)/Homer	1e-95
13		MafK(bZIP)/C2C12-MafK-ChIP-Seq(GSE36030)/Homer	1e-93
14		Nrf2(bZIP)/Lymphoblast-Nrf2-ChIP-Seq(GSE37589)/Homer	1e-81
15		NFE2L2(bZIP)/HepG2-NFE2L2-ChIP-Seq(Encode)/Homer	1e-64

Figure 3.4.16 The enriched motifs in 1085 commonly upregulated regions.

The figure displays the top 15 enriched motifs identified through analysis of motif enrichment in 1085 commonly upregulated regions.

3.5 Discussion

Our results identified a set of 159 genes that are consistently upregulated in all four tissues during regeneration, enriched in cell cycle and DNA replication pathways. Moreover, this gene set was associated with cell types possessing regenerative potential and immune-related functions in different tissues. For instance, *tubb5* and *cdh2*, both important marker genes of the regeneration process. They exhibited high expression levels in fin Mesenchymal cells, heart Fibroblasts and Epicardium cell, activated Müller glia cells in the retina and ependymo-radial glial cells in the spinal cord following injury. Our study also identified tissue-specific genes involved in the regeneration of different tissues. For example, *krt93* and *krt94* were uniquely upregulated during fin regeneration, particularly in the intermediate epithelial cells.

3.5.1 Zebrafish employ varied mechanisms to regenerate distinct tissues

Zebrafish used a diverse array of mechanisms to regenerate various tissues (Khyeam *et al*, 2021). Each tissue utilizes a unique combination of cellular and molecular processes to facilitate tissue repair and regeneration, reflecting the complex biology of zebrafish regeneration (Shi *et al*, 2015). For instance, the ependymo-radial glia cell in the spinal cord and blastema cells in fin serve as resident stem cells, activating and proliferating following injury (Sehring & Weidinger, 2020; Goldshmit *et al*, 2012). Müller glia cells in the retina act as latent stem cells, remaining quiescent until activated by injury (Lenkowski & Raymond, 2014). While cardiomyocytes in the heart undergo dedifferentiation post-injury (Xia *et al*, 2022; Streef & Smits, 2021). Zebrafish coordinates these regenerative mechanisms through complex signaling pathways and coordinated cellular responses.

Our observations reveal the commonalities across different tissue regeneration processes, such as the involvement of genes like *tubb5* and *cdh2*. However, the precise

role of these shared elements in the complex zebrafish regeneration mechanism remains unclear. Further investigation is needed to explain how these similarities contribute to the regulation and coordination of regeneration across diverse tissues and cell types.

3.5.2 Limitation of the study

Firstly, the diversity of data collection across organizations may lead to differences in our analyses. Furthermore, the lack of data on regeneration of some other tissues such as the brain limits the comprehensiveness of our findings. Moreover, Data quality is a challenge, especially in the absence of high-quality spinal cord ATAC-seq data. Furthermore, Current data analysis and technology lack reliable methods to determine which genes are specifically regulated by regulatory elements. Additionally, although our results provide valuable insights, further experimental validation is necessary to confirm their accuracy and reliability. Future studies should aim to address these limitations and delve into the functional significance of the identified marker genes and regulatory elements in tissue regeneration.

References

A Guide for the Laboratory Use of Zebrafish *Danio* (*Brachydanio*) *rerio*

Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine J-C, Geurts P, Aerts J, *et al* (2017) SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 14: 1083–1086

Andersson ER, Sandberg R & Lendahl U (2011) Notch signaling: simplicity in design, versatility in function. *Development* 138: 3593–3612

Andrews MG, Kong J, Novitch BG & Butler SJ (2019) New perspectives on the mechanisms establishing the dorsal-ventral axis of the spinal cord. In *Current Topics in Developmental Biology* pp 417–450. Elsevier

Andrzejczuk LA, Banerjee S, England SJ, Voufo C, Kamara K & Lewis KE (2018) *Tall1*, *Gata2a*, and *Gata3* Have Distinct Functions in the Development of V2b and Cerebrospinal Fluid-Contacting KA Spinal Neurons. *Front Neurosci* 12

Barreiro-Iglesias A, Mysiak KS, Scott AL, Reimer MM, Yang (杨宇婕) Y, Becker CG & Becker T (2015) Serotonin Promotes Development and Regeneration of Spinal Motor Neurons in Zebrafish. *Cell Rep* 13: 924–932

Barrett LM, Mitchell DM, Meighan PC, Varnum MD & Stenkamp DL (2022) Dynamic functional and structural remodeling during retinal regeneration in zebrafish. *Front Mol Neurosci* 15: 1070509

Batista MF, Jacobstein J & Lewis KE (2008) Zebrafish V2 cells develop into excitatory CiD and Notch signalling dependent inhibitory VeLD interneurons. *Dev Biol* 322: 263–275

Becker T & Becker CG (2020) Dynamic cell interactions allow spinal cord regeneration in zebrafish. *Curr Opin Physiol* 14: 64–69

Beffagna G (2019) Zebrafish as a Smart Model to Understand Regeneration After Heart Injury: How Fish Could Help Humans. *Front Cardiovasc Med* 6

Behra M, Gallardo VE, Bradsher J, Torrado A, Elkahloun A, Idol J, Sheehy J, Zonies S, Xu L, Shaw KM, *et al* (2012) Transcriptional signature of accessory cells in the lateral line, using the *Tnk1bp1:EGFP* transgenic zebrafish line. *BMC Dev Biol* 12: 6

Beisaw A, Kuenne C, Guenther S, Dallmann J, Wu C-C, Bentsen M, Looso M & Stainier DYR (2020) AP-1 Contributes to Chromatin Accessibility to Promote Sarcomere Disassembly and Cardiomyocyte Protrusion During Zebrafish Heart Regeneration. *Circ Res* 126: 1760–1778

- Bellanti F, Serviddio G & Vendemiale G (2023) Modulation of liver regeneration by the nuclear factor erythroid 2-related factor 2. *Adv Redox Res* 7: 100066
- Bello-Rojas S & Bagnall MW (2022) Clonally related, Notch-differentiated spinal neurons integrate into distinct circuits. *eLife* 11: e83680
- Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, Schellen P, Verschueren H, Post E, Koster J, *et al* (2015) RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell* 28: 666–676
- Blader P (2000) Zebrafish developmental genetics and central nervous system development. *Hum Mol Genet* 9: 945–951
- Boareto M, Iber D & Taylor V (2017) Differential interactions between Notch and ID factors control neurogenesis by modulating Hes factor autoregulation. *Dev Camb Engl* 144: 3465–3474
- Briona LK & Dorsky RI (2014) Radial glial progenitors repair the zebrafish spinal cord following transection. *Exp Neurol* 256: 81–92
- Campbell LJ, Levendusky JL, Steines SA & Hyde DR (2021) Retinal regeneration requires dynamic Notch signaling. *Neural Regen Res* 17: 1199–1209
- Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, *et al* (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 566: 496–502
- Cao Y, Xia Y, Balowski JJ, Ou J, Song L, Safi A, Curtis T, Crawford GE, Poss KD & Cao J (2022) Identification of enhancer regulatory elements that direct epicardial gene expression during zebrafish heart regeneration. *Development* 149: dev200133
- Cardozo MJ, Mysiak KS, Becker T & Becker CG (2017) Reduce, reuse, recycle – Developmental signals in spinal cord regeneration. *Dev Biol* 432: 53–62
- Catala M & Kubis N (2013) Gross anatomy and development of the peripheral nervous system. In *Handbook of Clinical Neurology*, Said G & Krarup C (eds) pp 29–41. Elsevier
- Cearns MD, Escuin S, Alexandre P, Greene NDE & Copp AJ (2016) Microtubules, polarity and vertebrate neural tube morphogenesis. *J Anat* 229: 63–74
- Chen F, Köhler M, Cucun G, Takamiya M, Kizil C, Cosacak MI & Rastegar S (2023) sox1a:eGFP transgenic line and single-cell transcriptomics reveal the origin of zebrafish intraspinal serotonergic neurons. *iScience* 26: 107342
- Chen H & Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12: 35

- Cigliola V, Shoffner A, Lee N, Ou J, Gonzalez TJ, Hoque J, Becker CJ, Han Y, Shen G, Faw TD, *et al* (2023) Spinal cord repair is modulated by the neurogenic factor Hb-egf under direction of a regeneration-associated enhancer. *Nat Commun* 14: 4857
- Cosacak MI, Bhattarai P & Kizil C (2020) Protocol for Dissection and Dissociation of Zebrafish Telencephalon for Single-Cell Sequencing. *STAR Protoc* 1: 100042
- Cosacak MI, Bhattarai P, Reinhardt S, Petzold A, Dahl A, Zhang Y & Kizil C (2019) Single-Cell Transcriptomics Analyses of Neural Stem Cell Heterogeneity and Contextual Plasticity in a Zebrafish Brain Model of Amyloid Toxicity. *Cell Rep* 27: 1307-1318.e3
- Cucun G, Köhler M, Pfitsch S & Rastegar S (2024) Insights into the mechanisms of neuron generation and specification in the zebrafish ventral spinal cord. *FEBS J* 291
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, *et al* (2021) Twelve years of SAMtools and BCFtools. *GigaScience* 10: giab008
- Danner SM, Zhang H, Shevtsova NA, Borowska-Fielding J, Deska-Gauthier D, Rybak IA & Zhang Y (2019) Spinal V3 Interneurons and Left–Right Coordination in Mammalian Locomotion. *Front Cell Neurosci* 13
- Delile J, Rayon T, Melchionda M, Edwards A, Briscoe J & Sagner A (2019) Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development* 146: dev173807
- Deneris E & Gaspar P (2018) Serotonin neuron development: shaping molecular and structural identities. *WIREs Dev Biol* 7: e301
- Deneris ES & Wyler SC (2012) Serotonergic transcriptional networks and potential importance to mental health. *Nat Neurosci* 15: 519–527
- Diotel N, Lübke L, Strähle U & Rastegar S (2020) Common and Distinct Features of Adult Neurogenesis and Regeneration in the Telencephalon of Zebrafish and Mammals. *Front Neurosci* 14
- Doganyigit Z, Eroglu E & Okan A (2023) Intermediate filament proteins are reliable immunohistological biomarkers to help diagnose multiple tissue-specific diseases. *Anat Histol Embryol* 52: 655–672
- Farnsworth DR, Saunders LM & Miller AC (2020) A single-cell transcriptome atlas for zebrafish development. *Dev Biol* 459: 100–108
- Fauq AH, Simpson K, Maharvi GM, Golde T & Das P (2007) A multigram chemical synthesis of the gamma-secretase inhibitor LY411575 and its diastereoisomers. *Bioorg Med Chem Lett* 17: 6392–6395

- Franco D, Sedmera D & Lozano-Velasco E (2017) Multiple Roles of Pitx2 in Cardiac Development and Disease. *J Cardiovasc Dev Dis* 4: 16
- Gao J, Zhang W, Li P, Liu J, Song H, Wang X & Zhang Q (2015) Identification, molecular characterization and gene expression analysis of *sox1a* and *sox1b* genes in Japanese flounder, *Paralichthys olivaceus*. *Gene* 574: 225–234
- Gemberling M, Bailey TJ, Hyde DR & Poss KD (2013) The zebrafish as a model for complex tissue regeneration. *Trends Genet TIG* 29: 10.1016/j.tig.2013.07.003
- Gerber V, Yang L, Takamiya M, Ribes V, Gourain V, Peravali R, Stegmaier J, Mikut R, Reischl M, Ferg M, *et al* (2019) The HMG box transcription factors Sox1a and Sox1b specify a new class of glycinergic interneuron in the spinal cord of zebrafish embryos. *Development* 146: dev172510
- Gilbert SF (2000) Differentiation of the Neural Tube. In *Developmental Biology. 6th edition* Sinauer Associates
- Goldman JA & Poss KD (2020) Gene regulatory programmes of tissue regeneration. *Nat Rev Genet* 21: 511–525
- Goldshmit Y, Sztal TE, Jusuf PR, Hall TE, Nguyen-Chi M & Currie PD (2012) Fgf-Dependent Glial Cell Bridges Facilitate Spinal Cord Regeneration in Zebrafish. *J Neurosci* 32: 7477–7492
- Gu Z, Eils R & Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32: 2847–2849
- Hajam YA, Rani R, Sharma P, Kumar R & Verma SK (2021) Zebrafish (*Danio rerio*): A Versatile Model for Reproductive Biology. In *Recent updates in molecular Endocrinology and Reproductive Physiology of Fish: An Imperative step in Aquaculture*, Sundaray JK Rather MA Kumar S & Agarwal D (eds) pp 105–120. Singapore: Springer
- Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, *et al* (2021) Integrated analysis of multimodal single-cell data. *Cell* 184: 3573–3587.e29
- Haque A, Engel J, Teichmann SA & Lönnberg T (2017) A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med* 9: 75
- Harvey BM, Baxter M & Granato M (2019) Optic nerve regeneration in larval zebrafish exhibits spontaneous capacity for retinotopic but not tectum specific axon targeting. *PLoS ONE* 14: e0218667
- Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR & Ordoukhanian P (2014) Library construction for next-generation sequencing: Overviews and challenges. *BioTechniques* 56: 61-passim

- Hegenbarth J-C, Lezsoche G, De Windt LJ & Stoll M (2022) Perspectives on Bulk-Tissue RNA Sequencing and Single-Cell RNA Sequencing for Cardiac Transcriptomics. *Front Mol Med* 2: 839338
- Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H & Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38: 576–589
- Ho JWK (2012) Application of a systems approach to study developmental gene regulation. *Biophys Rev* 4: 245–253
- Hoang T, Wang J, Boyd P, Wang F, Santiago C, Jiang L, Yoo S, Lahne M, Todd LJ, Jia M, *et al* (2020) Gene regulatory networks controlling vertebrate retinal regeneration. *Science* 370: eabb8598
- Hou Y, Lee HJ, Chen Y, Ge J, Osman FOI, McAdow AR, Mokalled MH, Johnson SL, Zhao G & Wang T (2020) Cellular diversity of the regenerating caudal fin. *Sci Adv* 6: eaba2084
- Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, *et al* (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496: 498–503
- Hu B, Lelek S, Spanjaard B, El-Sammak H, Simões MG, Mintcheva J, Aliee H, Schäfer R, Meyer AM, Theis F, *et al* (2022) Origin and function of activated fibroblast states during zebrafish heart regeneration. *Nat Genet* 54: 1227–1237
- Hu H, Miao Y-R, Jia L-H, Yu Q-Y, Zhang Q & Guo A-Y (2019) AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res* 47: D33–D38
- Huang C-X, Zhao Y, Mao J, Wang Z, Xu L, Cheng J, Guan NN & Song J (2021) An injury-induced serotonergic neuron subpopulation contributes to axon regrowth and function restoration after spinal cord injury in zebrafish. *Nat Commun* 12: 7093
- Huang P, Xiong F, Megason SG & Schier AF (2012) Attenuation of Notch and Hedgehog Signaling Is Required for Fate Specification in the Spinal Cord. *PLOS Genet* 8: e1002762
- Jacobs CT & Huang P (2019) Notch signalling maintains Hedgehog responsiveness via a Gli-dependent mechanism during spinal cord patterning in zebrafish. *eLife* 8: e49252
- Jacobs CT & Huang P (2021) Complex crosstalk of Notch and Hedgehog signalling during the development of the central nervous system. *Cell Mol Life Sci* 78: 635–644
- Jacobs CT, Kejriwal A, Kocha KM, Jin KY & Huang P (2022) Temporal cell fate determination in the spinal cord is mediated by the duration of Notch signalling. *Dev Biol* 489: 1–13

- Johansson CB, Momma S, Clarke DL, Risling M, Lendahl U & Frisén J (1999) Identification of a Neural Stem Cell in the Adult Mammalian Central Nervous System. *Cell* 96: 25–34
- Jopling C, Sleep E, Raya M, Martí M, Raya A & Belmonte JCI (2010) Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* 464: 606–609
- Jovic D, Liang X, Zeng H, Lin L, Xu F & Luo Y (2022) Single-cell RNA sequencing technologies and applications: A brief overview. *Clin Transl Med* 12: e694
- Kageyama R, Ohtsuka T, Shimojo H & Imayoshi I (2008) Dynamic Notch signaling in neural progenitor cells and a revised view of lateral inhibition. *Nat Neurosci* 11: 1247–1251
- Kang J, Hu J, Karra R, Dickson AL, Tornini VA, Nachtrab G, Gemberling M, Goldman JA, Black BL & Poss KD (2016) Modulation of tissue repair by regeneration enhancer elements. *Nature* 532: 201–206
- Khan A & Mathelier A (2017) Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC Bioinformatics* 18: 287
- Khyeam S, Lee S & Huang GN (2021) Genetic, epigenetic, and post-transcriptional basis of divergent tissue regenerative capacities among vertebrates. *Adv Genet* 2
- Kim D, Paggi JM, Park C, Bennett C & Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37: 907–915
- Kimura Y, Satou C & Higashijima S (2008) V2a and V2b neurons are generated by the final divisions of pair-producing progenitors in the zebrafish spinal cord. *Development* 135: 3001–3005
- Kittappa R (2017) Regulation of the Serotonin Neuron Fate in Stem Cells by *foxa2* and *shh*. *J Stem Cell Res Ther* 3
- Krueger F Trim galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. 2015;516(517).
- Kutejova E, Sasai N, Shah A, Gouti M & Briscoe J (2016) Neural Progenitors Adopt Specific Identities by Directly Repressing All Alternative Progenitor Transcriptional Programs. *Dev Cell* 36: 639–653
- Kuzmicz-Kowalska K & Kicheva A (2021) Regulation of size and scale in vertebrate spinal cord development. *Wiley Interdiscip Rev Dev Biol* 10: e383
- Lafzi A, Moutinho C, Picelli S & Heyn H (2018) Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Nat Protoc* 13: 2742–2757

- Lange C, Rost F, Machate A, Reinhardt S, Lesche M, Weber A, Kuscha V, Dahl A, Rulands S & Brand M (2020) Single cell sequencing of radial glia progeny reveals the diversity of newborn neurons in the adult zebrafish brain.
- Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359
- Lanigan LG, Russell DS, Woolard KD, Pardo ID, Godfrey V, Jortner BS, Butt MT & Bolon B (2021) Comparative Pathology of the Peripheral Nervous System. *Vet Pathol* 58: 10–33
- Lawson ND & Weinstein BM (2002) In Vivo Imaging of Embryonic Vascular Development Using Transgenic Zebrafish. *Dev Biol* 248: 307–318
- Lekk I, Duboc V, Faro A, Nicolaou S, Blader P & Wilson SW Sox1a mediates the ability of the parapineal to impart habenular left-right asymmetry. *eLife* 8: e47376
- Lenkowski JR, Qin Z, Sifuentes CJ, Thummel R, Soto CM, Moens CB & Raymond PA (2013) Retinal Regeneration in Adult Zebrafish Requires Regulation of TGF β Signaling. *Glia* 61: 1687–1697
- Lenkowski JR & Raymond PA (2014) Müller glia: Stem cells for generation and regeneration of retinal neurons in teleost fish. *Prog Retin Eye Res* 40: 94–123
- Lewis KE (2006) How do genes regulate simple behaviours? Understanding how different neurons in the vertebrate spinal cord are genetically specified. *Philos Trans R Soc B Biol Sci* 361: 45–66
- Li X & Wang C-Y (2021) From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci* 13: 36
- Liao Y, Smyth GK & Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma Oxf Engl* 30: 923–930
- Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550
- Lowe V, Wisniewski L & Pellet-Many C (2021) The Zebrafish Cardiac Endothelial Cell—Roles in Development and Regeneration. *J Cardiovasc Dev Dis* 8: 49
- Ludwig PE, Reddy V & Varacallo M (2023) Neuroanatomy, Central Nervous System (CNS). In *StatPearls* Treasure Island (FL): StatPearls Publishing
- Lupo G, Harris WA & Lewis KE (2006) Mechanisms of ventral patterning in the vertebrate nervous system. *Nat Rev Neurosci* 7: 103–114
- Lush ME & Piotrowski T (2014) Sensory hair cell regeneration in the zebrafish lateral line. *Dev Dyn* 243: 1187–1202

- Manzari-Tavakoli A, Babajani A, Farjoo MH, Hajinasrollah M, Bahrami S & Niknejad H (2022) The Cross-Talks Among Bone Morphogenetic Protein (BMP) Signaling and Other Prominent Pathways Involved in Neural Differentiation. *Front Mol Neurosci* 15: 827275
- März M, Chapouton P, Diotel N, Vaillant C, Hesl B, Takamiya M, Lam CS, Kah O, Bally-Cuif L & Strähle U (2010) Heterogeneity in progenitor cell subtypes in the ventricular zone of the zebrafish adult telencephalon. *Glia* 58: 870–888
- Mendoza-Parra M-A, Malysheva V, Mohamed Saleem MA, Lieb M, Godel A & Gronemeyer H (2016) Reconstructed cell fate–regulatory programs in stem cells reveal hierarchies and key factors of neurogenesis. *Genome Res* 26: 1505–1519
- Montgomery JE, Wahlstrom-Helgren S, Wiggin TD, Corwin BM, Lillesaar C & Masino MA (2018) Intraspinal serotonergic signaling suppresses locomotor activity in larval zebrafish. *Dev Neurobiol*
- Montgomery JE, Wiggin TD, Rivera-Perez LM, Lillesaar C & Masino MA (2016) Intraspinal serotonergic neurons consist of two, temporally distinct populations in developing zebrafish. *Dev Neurobiol* 76: 673–687
- Morimoto K & Nakajima K (2019) Role of the Immune System in the Development of the Central Nervous System. *Front Neurosci* 13
- Nguyen QH, Pervolarakis N, Nee K & Kessenbrock K (2018) Experimental Considerations for Single-Cell RNA Sequencing Approaches. *Front Cell Dev Biol* 6
- Norton WH, Mangoli M, Lele Z, Pogoda H-M, Diamond B, Mercurio S, Russell C, Teraoka H, Stickney HL, Rauch G-J, *et al* (2005) Monorail/Foxa2 regulates floorplate differentiation and specification of oligodendrocytes, serotonergic raphé neurones and cranial motoneurons. *Dev Camb Engl* 132: 645–658
- Odenthal J, Van Eeden FJM, Haffter P, Ingham PW & Nüsslein-Volhard C (2000) Two Distinct Cell Populations in the Floor Plate of the Zebrafish Are Induced by Different Pathways. *Dev Biol* 219: 350–363
- Ohtaki H & Shioda S (2015) Anatomy of Adult Central Nervous System: Structure and Function of the Brain and Spinal Cord. In *Neuroanesthesia and Cerebrospinal Protection*, Uchino H Ushijima K & Ikeda Y (eds) pp 3–22. Tokyo: Springer Japan
- Oustah AA, Danesin C, Khouri-Farah N, Farreny M-A, Escalas N, Cochard P, Glise B & Soula C (2014) Dynamics of Sonic hedgehog signaling in the ventral spinal cord are controlled by intrinsic changes in source cells requiring Sulfatase 1. *Development* 141: 1392–1403
- Panayi H, Panayiotou E, Orford M, Genethliou N, Mean R, Lapathitis G, Li S, Xiang M, Kessaris N, Richardson WD, *et al* (2010) Sox1 Is Required for the Specification of a Novel p2-Derived Interneuron Subtype in the Mouse Ventral Spinal Cord. *J Neurosci* 30: 12274–12280

- Paredes LC, Olsen Saraiva Camara N & Braga TT (2019) Understanding the Metabolic Profile of Macrophages During the Regenerative Process in Zebrafish. *Front Physiol* 10: 617
- Peymani F, Farzeen A & Prokisch H (2022) RNA sequencing role and application in clinical diagnostic. *Pediatr Investig* 6: 29–35
- Qiu C, Cao J, Martin BK, Li T, Welsh IC, Srivatsan S, Huang X, Calderon D, Noble WS, Disteche CM, *et al* (2022) Systematic reconstruction of cellular trajectories across mouse embryogenesis. *Nat Genet* 54: 328–341
- Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842
- Rabinowitz JS, Robitaille AM, Wang Y, Ray CA, Thummel R, Gu H, Djukovic D, Raftery D, Berndt JD & Moon RT (2017) Transcriptomic, proteomic, and metabolomic landscape of positional memory in the caudal fin of zebrafish. *Proc Natl Acad Sci* 114: E717–E726
- Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F & Manke T (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44: W160–W165
- Rayon T, Maizels RJ, Barrington C & Briscoe J (2021) Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. *Development* 148: dev199711
- Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, *et al* (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481: 389–393
- Rost MS, Shestopalov I, Liu Y, Vo AH, Richter CE, Emly SM, Barrett FG, Stachura DL, Holinstat M, Zon LI, *et al* (2018) Nfe2 is dispensable for early but required for adult thrombocyte formation and function in zebrafish. *Blood Adv* 2: 3418–3427
- Row RH, Pegg A, Kinney BA, Farr GH 3rd, Maves L, Lowell S, Wilson V & Martin BL (2018) BMP and FGF signaling interact to pattern mesoderm by controlling basic helix-loop-helix transcription factor activity. *eLife* 7: e31018
- Sagner A & Briscoe J (2017) Morphogen interpretation: concentration, time, competence, and signaling dynamics. *WIREs Dev Biol* 6: e271
- Sagner A & Briscoe J (2019) Establishing neuronal diversity in the spinal cord: a time and a place. *Development* 146: dev182154
- Sánchez Alvarado A (2012) Q&A: What is regeneration, and why look to planarians for answers? *BMC Biol* 10: 88

- Schäfer M, Kinzel D & Winkler C (2007) Discontinuous organization and specification of the lateral floor plate in zebrafish. *Dev Biol* 301: 117–129
- Schäfer M, Rembold M, Wittbrodt J, Schartl M & Winkler C (2005) Medial floor plate formation in zebrafish consists of two phases and requires trunk-derived Midkine-a. *Genes Dev* 19: 897–902
- Schmidt R, Strähle U & Scholpp S (2013) Neurogenesis in zebrafish – from embryo to adult. *Neural Develop* 8: 3
- Scholpp S, Delogu A, Gilthorpe J, Peukert D, Schindler S & Lumsden A (2009) Her6 regulates the neurogenetic gradient and neuronal identity in the thalamus. *Proc Natl Acad Sci U S A* 106: 19895–19900
- Scott K, O'Rourke R, Winkler CC, Kearns CA & Appel B (2021) Temporal single-cell transcriptomes of zebrafish spinal cord pMN progenitors reveal distinct neuronal and glial progenitor populations. *Dev Biol* 479: 37–50
- Sehring IM, Jahn C & Weidinger G (2016) Zebrafish fin and heart: what's special about regeneration? *Curr Opin Genet Dev* 40: 48–56
- Sehring IM & Weidinger G (2020) Recent advancements in understanding fin regeneration in zebrafish. *WIREs Dev Biol* 9: e367
- Shen W-Y, Fu X-H, Cai J, Li W-C, Fan B-Y, Pang Y-L, Zhao C-X, Abula M, Kong X-H, Yao X, *et al* (2021) Identification of key genes involved in recovery from spinal cord injury in adult zebrafish. *Neural Regen Res* 17: 1334–1342
- Shi W, Fang Z, Li L & Luo L (2015) Using zebrafish as the model organism to understand organ regeneration. *Sci China Life Sci* 58: 343–351
- Shin J, Poling J, Park H-C & Appel B (2007) Notch signaling regulates neural precursor allocation and binary neuronal fate decisions in zebrafish. *Development* 134: 1911–1920
- Shin K, Begeman IJ, Cao J & Kang J (2022) leptin b and its regeneration enhancer illustrate the regenerative features of zebrafish hearts. *Dev Dyn* 253: 91–106
- Silbereis JC, Pochareddy S, Zhu Y, Li M & Sestan N (2016) The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron* 89: 248–268
- Simões FC & Riley PR (2022) Immune cells in cardiac repair and regeneration. *Dev Camb Engl* 149: dev199906
- Simons M & Nave K-A (2016) Oligodendrocytes: Myelination and Axonal Support. *Cold Spring Harb Perspect Biol* 8: a020479

- Sobrido-Cameán D, Rodicio MC & Barreiro-Iglesias A (2018) Serotonin controls axon and neuronal regeneration in the nervous system: lessons from regenerating animal models. *Neural Regen Res* 13: 237
- Soerensen J, Jakupoglu C, Beck H, Förster H, Schmidt J, Schmahl W, Schweizer U, Conrad M & Brielmeier M (2008) The Role of Thioredoxin Reductases in Brain Development. *PLoS ONE* 3: e1813
- Spencer WC & Deneris ES (2017) Regulatory Mechanisms Controlling Maturation of Serotonin Neuron Identity and Function. *Front Cell Neurosci* 11: 215
- Strähle U, Lam CS, Ertzer R & Rastegar S (2004) Vertebrate floor-plate specification: variations on common themes. *Trends Genet TIG* 20: 155–162
- Streef TJ & Smits AM (2021) Epicardial Contribution to the Developing and Injured Heart: Exploring the Cellular Composition of the Epicardium. *Front Cardiovasc Med* 8: 750243
- Tarasov A, Vilella AJ, Cuppen E, Nijman IJ & Prins P (2015) Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 31: 2032–2034
- van Tetering G & Vooijs M (2011) Proteolytic Cleavage of Notch: “HIT and RUN”. *Curr Mol Med* 11: 255–269
- Thau L, Reddy V & Singh P (2022) Anatomy, Central Nervous System. In *StatPearls* Treasure Island (FL): StatPearls Publishing
- Thompson JD, Ou J, Lee N, Shin K, Cigliola V, Song L, Crawford GE, Kang J & Poss KD (2020) Identification and requirements of enhancers that direct gene expression during zebrafish fin regeneration. *Development* 147: dev191262
- Todd L, Suarez L, Quinn C & Fischer AJ (2018) Retinoic Acid-Signaling Regulates the Proliferative and Neurogenic Capacity of Müller Glia-Derived Progenitor Cells in the Avian Retina. *Stem Cells* 36: 392–405
- Tokunaga A, Kohyama J, Yoshida T, Nakao K, Sawamoto K & Okano H (2004) Mapping spatio-temporal activation of Notch signaling during neurogenesis and gliogenesis in the developing mouse brain. *J Neurochem* 90: 142–154
- Tsata V & Wehner D (2021) Know How to Regrow—Axon Regeneration in the Zebrafish Spinal Cord. *Cells* 10: 1404
- Van den Berge K, Hembach KM, Soneson C, Tiberi S, Clement L, Love MI, Patro R & Robinson MD (2019) RNA Sequencing Data: Hitchhiker’s Guide to Expression Analysis. *Annu Rev Biomed Data Sci* 2: 139–173

- Wan J, Zhao X-F, Vojtek A & Goldman D (2014) Retinal Injury, Growth Factors, and Cytokines Converge on β -Catenin and pStat3 Signaling to Stimulate Retina Regeneration. *Cell Rep* 9: 285–297
- Wang W, Hu C-K, Zeng A, Alegre D, Hu D, Gotting K, Ortega Granillo A, Wang Y, Robb S, Schnittker R, *et al* (2020) Changes in regeneration-responsive enhancers shape regenerative capacities in vertebrates. *Science* 369: eaaz3090
- Wattrus SJ & Zon LI (2018) Stem cell safe harbor: the hematopoietic stem cell niche in zebrafish. *Blood Adv* 2: 3063–3069
- Wehner D, Cizelsky W, Vasudevaro MD, Ozhan G, Haase C, Kagermeier-Schenk B, Röder A, Dorsky RI, Moro E, Argenton F, *et al* (2014) Wnt/ β -catenin signaling defines organizing centers that orchestrate growth and differentiation of the regenerating zebrafish caudal fin. *Cell Rep* 6: 467–481
- Wehner D & Weidinger G (2015) Signaling networks organizing regenerative growth of the zebrafish fin. *Trends Genet* 31: 336–343
- White RJ, Collins JE, Sealy IM, Wali N, Dooley CM, Digby Z, Stemple DL, Murphy DN, Billis K, Hourlier T, *et al* (2017) A high-resolution mRNA expression time course of embryonic development in zebrafish. *eLife* 6: e30860
- Wilson DM & McNeill DR (2007) Base excision repair and the central nervous system. *Neuroscience* 145: 1187–1200
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, *et al* (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* 2: 100141
- Xia Y, Duca S, Perder B, Dündar F, Zumbo P, Qiu M, Yao J, Cao Y, Harrison MRM, Zangi L, *et al* (2022) Activation of a transient progenitor state in the epicardium is required for zebrafish heart regeneration. *Nat Commun* 13: 7704
- Xu X, Mee T & Jia X (2020) New era of optogenetics: from the central to peripheral nervous system. *Crit Rev Biochem Mol Biol* 55: 1–16
- Yang L, Rastegar S & Strähle U (2010) Regulatory interactions specifying Kolmer-Agduhr interneurons. *Development* 137: 2713–2722
- Yang L, Wang F & Strähle U (2020) The Genetic Programs Specifying Kolmer–Agduhr Interneurons. *Front Neurosci* 14
- Yeo S-Y, Kim M, Kim H-S, Huh T-L & Chitnis AB (2007) Fluorescent protein expression driven by her4 regulatory elements reveals the spatiotemporal pattern of Notch signaling in the nervous system of zebrafish embryos. *Dev Biol* 301: 555–567

- Yu G, Wang L-G & He Q-Y (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinforma Oxf Engl* 31: 2382–2383
- Zhang H, Shevtsova NA, Deska-Gauthier D, Mackay C, Dougherty KJ, Danner SM, Zhang Y & Rybak IA (2022) The role of V3 neurons in speed-dependent interlimb coordination during locomotion in mice. *eLife* 11: e73424
- Zhang L, Lei Z, Guo Z, Pei Z, Chen Y, Zhang F, Cai A, Mok G, Lee G, Swaminathan V, *et al* (2020) Development of Neuroregenerative Gene Therapy to Reverse Glial Scar Tissue Back to Neuron-Enriched Tissue. *Front Cell Neurosci* 14: 594170
- Zhang W & Roy S (2016) The zebrafish fast myosin light chain mylpfa:H2B-GFP transgene is a useful tool for in vivo imaging of myocyte fusion in the vertebrate embryo. *Gene Expr Patterns* 20: 106–110
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, *et al* (2008a) Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* 9: R137
- Zhang Y, Narayan S, Geiman E, Lanuza GM, Velasquez T, Shanks B, Akay T, Dyck J, Pearson K, Gosgnach S, *et al* (2008b) V3 spinal neurons establish a robust and balanced locomotor rhythm during walking. *Neuron* 60: 84–96
- Zhao Z-Q, Scott M, Chiechio S, Wang J-S, Renner KJ, Gereau RW, Johnson RL, Deneris ES & Chen Z-F (2006) Lmx1b Is Required for Maintenance of Central Serotonergic Neurons and Mice Lacking Central Serotonergic System Exhibit Normal Locomotor Activity. *J Neurosci* 26: 12781–12788
- Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, *et al* (2017) Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 8: 14049
- Zheng Y, Zhang F, Xu S & Wu L (2021) Advances in neural organoid systems and their application in neurotoxicity testing of environmental chemicals. *Genes Environ* 43: 39
- Zhou B, Lin W, Long Y, Yang Y, Zhang H, Wu K & Chu Q (2022) Notch signaling pathway: architecture, disease, and therapeutics. *Signal Transduct Target Ther* 7: 95
- Zhou W, Yan Y, Guo Q, Ji H, Wang H, Xu T, Makabel B, Pilarsky C, He G, Yu X, *et al* (2021) Microfluidics applications for high-throughput single cell sequencing. *J Nanobiotechnology* 19: 312
- Zuñiga NR & Stoeckli ET (2017) Sonic Hedgehog—‘Jack-of-All-Trades’ in Neural Circuit Formation. *J Dev Biol* 5: 2