Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors

Anneke Dixie Kakebeen

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Andrea Wills, Chair

David Kimelman

Thomas Reh

Program Authorized to Offer Degree:
Department of Biochemistry

University of Washington

Abstract

Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors

Anneke Dixie Kakebeen

Chair of the Supervisory Committee:
Professor Andrea Wills, PhD
Biochemistry

Vertebrate appendage regeneration requires precisely coordinated remodeling of the transcriptional landscape to enable the growth and differentiation of new tissue, a process executed over multiple days and across dozens of cell types. The heterogeneity of tissues and temporally-sensitive fate decisions involved has made it difficult to articulate the gene regulatory programs enabling regeneration of individual cell types. To better understand how a regenerative program is fulfilled by neural progenitor cells (NPCs) of the spinal cord, we analyzed pax6-expressing NPCs isolated from regenerating *Xenopus tropicalis* tails. By intersecting chromatin accessibility data with single-cell transcriptomics, we find that NPCs place an early priority on neuronal differentiation. Late in regeneration, the priority returns to proliferation. Our analyses identify Pbx3 and Meis1 as critical regulators of tail regeneration and axon organization. Overall, we use

transcriptional regulatory dynamics to present a new model for cell fate decisions and their regulators in NPCs during regeneration.

TABLE OF CONTENTS

List of figures v

List of Tables vii

Chapter 1. Introduction 11

1.1 *Xenopus tropicalis*: model for regeneration 12

1.2 Progression of tail regeneration 15

1.3 Neural progenitor cells in spinal cord regeneration 16

Chapter 2. Tissue disaggregation and isolation of specific cell types from transgenic Xenopus appendages for transcriptional analysis by FACS 20

2.1 Abstract 20

2.2 Introduction 21

2.3 Results and Experimental Procedures 25

2.3.1 Transgenic lines used 25

2.3.2 Tadpole tail and limb isolation 27

2.3.3 Tissue dissociation to single cell suspension 27

2.3.4 Viability analysis with Trypan blue 30

2.3.5 Isolation of GFP+ single cells using FACS 31

2.4 Downstream library preparation and analysis 34

2.5 Strategies for validating cell type recovery and quality 34

2.6 Conclusions 40

2.7 Acknowledgements 41

Chapter 3. Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors 42

3.1 Introduction 42

3.2 Results 45

3.2.1 Transgenic pax6:GFP is expressed in NPCs during regeneration 45

3.2.2 The chromatin accessibility profile of FACS-sorted NPCs is highly enriched for neural-specific regulatory regions 48

3.2.3 Analysis of differentially accessible peaks between pax6 libraries reveals an early prioritization of neuronal differentiation in regeneration 51

3.2.4 Comparative analysis of uninjured and 24hpa tail single cell RNA-Seq data sets reveals expansion in differentiated neuronal clusters at 24hpa 54

3.2.5 Gene regulatory network prediction reveals Pbx3 and Meis1 as candidate regulators of neuronal regeneration 59

3.2.6 Meis1 and Pbx3 are necessary for successful spinal cord and tail regeneration 64

3.3 Discussion 69

3.3.1 Neural lineage-specific analysis of regeneration identifies new regulatory factors and target genes 69

3.3.2 Evidence for a temporal uncoupling of differentiation and proliferation in the regenerating neural lineage 71

3.3.3 Neural cell types preserve their identities in the regenerating Xenopus tail 73

3.3.4 Pbx3 and Meis1 are required for axon organization and for tail regeneration 74

3.4 Conclusion 76

3.5 Methods 77

3.5.1 X. tropicalis husbandry and use 77
3.5.2 Sectioning and imaging 77
3.5.3 X. tropicalis amputation assay 78
3.5.4 Cell dissociation and fluorescent activated cell sorting (FACS) 78
3.5.5 Quantitative RT-PCR 79
3.5.6 ATAC-Seq library preparation 80
3.5.7 ATAC-Seq analysis pipeline 81
3.5.8 Gene Ontology Analysis 82
3.5.9 Single cell RNA-Seq library preparation 83
3.5.10 Single cell RNA-Seq data processing 84
3.5.11 Whole-Mount In situ hybridization 87
3.5.12 Morpholino microinjections 87
3.5.13 CRISPR-mediated mutation 88
3.5.14 Immunohistochemistry 88
3.5.15 Regenerated tissue measurement and analysis 90
3.5.16 PH3 cell quantification 90
3.6 Data Availability 91
3.7 Acknowledgements 91
Chapter 4. Concluding remarks and future directions 103
4.1 Cell type-specific discovery approach 103
4.2 Regeneration uniquely utilizes developmental programs 105
4.3 Identification of novel transcription factors in regeneration 107
4.4 Future use of generated data-sets 109

4.5 Mile high view 110

Chapter 5. References 113

LIST OF FIGURES

Figure 1-1 *Cellular processes activated by injury in Xenopus tail regeneration.* 14
Figure 1-2 Regeneration time-course of *X. tropicalis* tail regeneration 15
Figure 2-1 At a glance: Tissue disaggregation and FACS purification of transgenic *Xenopus* tissue 24
Figure 2-2 Disaggregation of *Xenopus tropicalis* tails and limbs..... 26
Figure 2-3 FACS gating settings for GFP+ cell isolation..... 33
Figure 2-4 Sorting by FACS does not significantly change gene expression or chromatin quality. 36
Figure 2-5 Sorting cells by FACS enriches specific cell types..... 39
Figure 3-1 Expression dynamics of pax6:GFP in the regenerating spinal cord. 47
Figure 3-2 Pax6 ATAC-Seq libraries resolve accessible neural specific regions that were not identified in all-tissue libraries..... 49
Figure 3-3 Differential accessibility analysis of pax6 libraries over regenerative time reveals chromatin accessibility prioritizes first tubule morphogenesis, followed by neural differentiation, and later proliferation..... 52
Figure 3-4 scRNA-Seq of uninjured and 24hpa tails reveals a transcriptional shift to differentiated neuronal types at 24hpa. 57
Figure 3-5 Markers of neuronal differentiation are increased at 24hpa and proliferation increases at 72hpa. 61
Figure 3-6 Gene regulatory network prediction identifies Meis1 and Pbx3 as key regulators of neural regeneration..... 63
Figure 3-7 Pbx3 and Meis1 are independently required for successful regeneration of neural tissues and tails in response to injury..... 67
Figure 3-7 Pbx3 and Meis1 are independently required for successful regeneration of neural tissues and tails in response to injury..... 68
Figure 3-8 Model for neural progenitor regeneration. 75
Figure 3-9 FACS gating was set using wild-type tadpoles to avoid false-positive GFP+ cells. 92

Figure 3-10 Low input ATAC-Seq libraries meet sufficient quality for downstream analysis. 93
Figure 3-11 scRNA-Seq analysis of all cells sequenced in uninjured and 24hpa tails. 94
Figure 3-12 Embryonic Meis1 and Pbx3 morphants are characterized by small heads and missing or small eyes. 95
Figure 3-13 Meis1 and Pbx3 are independently necessary for successful regeneration of the axial tissue and neuronal patterning. 96

LIST OF TABLES

Table 1-1 Comparative overview of spinal cord regeneration..... 18
Table 3-1 Key Resources Table..... 97

ACKNOWLEDGEMENTS

I am incredibly appreciative for the endless support I have received throughout my thesis research. Thank you to members of the Department of Biochemistry and the developmental biology community for creating an inspiring and welcoming community to spend five years.

I thank my mentor, Andrea Wills for introducing this wandering biochemist to the wonders of developmental and regenerative biology. I am truly thankful for the ever-present support in my project, scientific career, and personal growth over the last 5 years. Thank you for allowing me to take this project where it interested me and develop techniques to make it happen. Last, thank you for your willingness to be first and foremost a mentor, I wish to emulate you when I have a lab of my own.

I thank my committee members, David Kimelman, Tom Reh, Nancy Maizels, and Jim Hurley for their time, constructive feedback, and career guidance. I specifically want to thank my members for their direction during my general exam. Suggesting I go after more exciting experiments because I was capable of doing something better is one of the most confidence-building pieces of feedback I have received as a scientist. Last, I would like to thank my committee for suggesting I use scRNA-Seq in my work.

I thank my past scientific mentors, especially Tina Singh, Will Martin, Elizabeth Komives, Jonathan Parnell, and Jimmy Marion, for how they inspired the love for science in me and gave me a base set of skills I have carried into how I do science today.

While it is my name that adorns this thesis, each section has been informed by support and discussions of my community. To all iterations of the Wills lab, especially Alex Chitsazan, Hannah Arbach, Lauren Loh, Jeet Patel, and Madison Williams, I thank you for countless conversations which have enriched my work and scientific perspective. I thank Arianne Caudal

for providing perspective both in my project and in my life and being a real one. I thank honorary Wills lab older graduate students, Lauren Saunders and Vic Lewis for always being willing to share their experiences and guide our graduate career. Science can be an isolating place, but the conversations, laughter, and shenanigans always made it feel like I was part of something bigger with people I consider family.

To the friends that have become family to me in Seattle, thank you for sharing in meals and always making time for me to check the frogs. The Seattle grey could not compete with the joy I felt spending time with y'all. To my friends I have kept through many years, thank you for always asking how the frogs are, but also keeping me grounded and reminding me of the world outside of science. To my family, thank you for instilling core values that drive how I have completed this research and for your love and support. Thank you for always letting me know you are proud even if you didn't know what I was doing. To Antoinette, thank you for all you do to keep us steady while I chase down my dreams.

To everyone I have encountered since September 2015, thank you for listening to me go on about tadpoles and regeneration for hours on end.

DEDICATION

For Grandma and Grandpa

Chapter 1. INTRODUCTION

Humans have long been fascinated by regeneration and the elegance it presents as a treatment for traumatic injury. Notably, the ancient Greeks described the ability of the fire-stealing Prometheus to regrow his liver after suffering the punishment of daily eating of his liver by an eagle (Power and Rasko, 2008). Although the story of Prometheus captures the fact that humans can regenerate their livers (albeit in an extreme way), regenerative capacity in other tissues is quite restricted (Koniaris et al., 2003). With this knowledge, we find ourselves in admiration of contemporary characters like Wolverine, who has a seemingly limitless regenerative capacity, being able to regenerate heavily injured tissues by the use of his healing factor. While the extreme capacity to regenerate a lost limb may seem to only exist in science-fiction, some non-mammalian vertebrate organisms do possess such regenerative capacities. Regenerating vertebrate organisms present scientists and clinicians with a unique opportunity to study the process they want to emulate. Rather than re-inventing the wheel, researchers can study and articulate the molecular mechanisms facilitating naturally occurring regeneration in salamanders, frogs, and zebrafish, then introduce them to mammals in hopes of promoting a regenerative response.

For my thesis research, I sought to uncover the molecular mechanisms underlying how neural progenitor cells are regulated during naturally occurring tail regeneration in *Xenopus tropicalis*. In this chapter, I introduce (1) *Xenopus tropicalis* as a model system for regeneration and (2) the impetus for studying neural progenitor cells in spinal cord regeneration.

1.1 *XENOPUS TROPICALIS*: MODEL FOR REGENERATION

Section 1.1 and Figure 1.1 of Chapter 1 have been adapted from:

Kakebeen AD, Wills AE. More Than Just a Bandage: Closing the Gap Between Injury and Appendage Regeneration. Front Physiol. 2019 Feb 8;10:81. doi: 10.3389/fphys.2019.00081. PMID: 30800076; PMCID: PMC6376490.

Injuries that sever tissues such as the limb or spinal cord are met with radically different outcomes among vertebrates. In mammals, a limb amputation or spinal cord transection is followed by inflammation and fibrotic scarring that leaves the animal with a permanent disability. In urodele amphibians such as axolotls and newts, the same injury is followed by scarless regenerative healing that can fully restore both the lost tissue and its function (reviewed in (Tanaka, 2016)). The anuran frogs *Xenopus laevis* and *Xenopus tropicalis* represent a middle ground: injuries to the tadpole tail, limb bud, or spinal cord are readily repaired through regeneration, but this ability declines during metamorphosis (Cannata et al., 2001; Phipps et al., 2020). As adults, *Xenopus* can no longer functionally recover from a spinal cord transection (Filoni and Bosco), while amputation of the hindlimb results in regeneration of a single digit, rather than the whole limb(Dent, 1962; Suzuki et al., 2006). This temporally-restricted regenerative competence therefore makes *Xenopus* an appealing model for defining the features that enable or inhibit regenerative healing. In addition to the loss of regenerative competence undergone during metamorphosis, *Xenopus* tadpoles also experience a transient loss of regenerative competence called the refractory period at Nieuwkoop and Faber stages 45-47, shortly after the onset of independent feeding (Beck et al., 2003). Appendage regeneration, particularly of the tadpole tail, has been widely studied before, during and after this period. As a complement to the limb or limb bud, the tail is an excellent model for appendage regeneration because it comprises multiple cell types from epidermal, neural,

mesodermal, and neural crest lineages, is easily accessible experimentally, and regenerates fully in a matter of days (Beck et al., 2009a; Chen et al., 2014).

The regeneration of a tissue intuitively recapitulates aspects of its embryonic development. In both processes, rapid proliferation gives rise to new tissue, cell fate has to be specified within that tissue, and distinct positional identities have to be established to generate a properly patterned structure. Molecular evidence has validated multiple aspects of this parallel. Experimental perturbations using small molecule inhibitors and heat-shock inducible inhibitory proteins have established that BMP, FGF, Wnt, Notch, Shh, and Nodal/TGF-b signaling pathways are required for proper formation of the regenerated tail, paralleling their requirements in early embryonic patterning ((Beck et al., 2003a; Lin and Slack, 2008)(Taniguchi et al., 2014)(Ho and Whitman, 2008)). Elegant experiments using heat-shock inducible expression of inhibitory proteins have further refined these observations to establish epistatic relationships, in which BMP acts upstream of Wnt, which in turn acts upstream of FGF during regeneration of the limb bud and tail (Lin and Slack, 2008). As during development, the establishment of positional identity appears to rely on the action of posterior Hox transcription factors (Christen et al., 2003). Numerous genes expressed in the developing limb and tail buds are re-expressed during tail regeneration, suggesting that many factors used to form these structures during embryogenesis are recapitulated during regeneration (Love et al., 2011).

More recently, next-generation sequencing approaches have endeavored to comprehensively catalog the transcriptional responses undergone by regenerating tissues. Microarray and RNA-Seq studies of the whole regenerating tail (Love et al., 2011)(Chang et al., 2017)(Aztekin et al., 2019), proliferating blastemal cells (Tsujioka et al., 2015), and spinal cord (Lee-Liu et al., 2014), have highlighted that embryonic patterning and developmental processes

are indeed highly prioritized beginning at one day after amputation. However, these studies show that the initial transcriptional responses triggered by injury include a distinct set of target genes that characterize regeneration independent from development, and also hold clues to how regenerative healing may be differentiated from other forms of wounding response. These include changes in cell metabolic enzymes, factors used to generate reactive oxygen species (ROS), ion channels, innate immune cell factors, and epigenetic modifiers (Figure 1). Functional interrogation of many of these cell physiological mechanisms has begun to confirm that they are not only upregulated but are also necessary for regeneration of the tail, limb, or spinal cord.

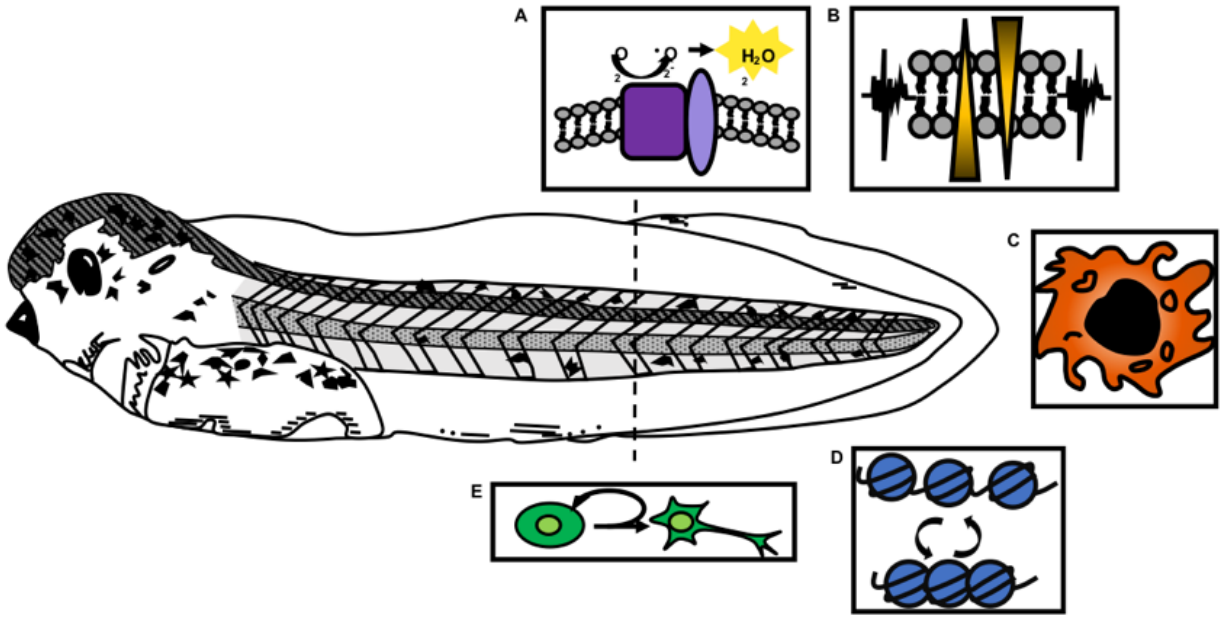

Figure 1-1 ***Cellular processes activated by injury in Xenopus tail regeneration.***
A regenerative stage 41 tadpole is shown, prior to the onset of independent feeding and the refractory period. Responses to injury that are critical for regeneration include (A) formation of reactive oxygen species such as H_2O_2 through the action of NOX complexes (purple) and p22-phox/cyba (light purple); (B) bioelectrical signaling mediated by ion channel activation; (C) recruitment of innate immune cell types such as macrophages; (D) epigenetic modifications that affect chromatin accessibility and transcription, and (E) activation of proliferation of blastemal cells and tissue-specific progenitors.

1.2 PROGRESSION OF TAIL REGENERATION

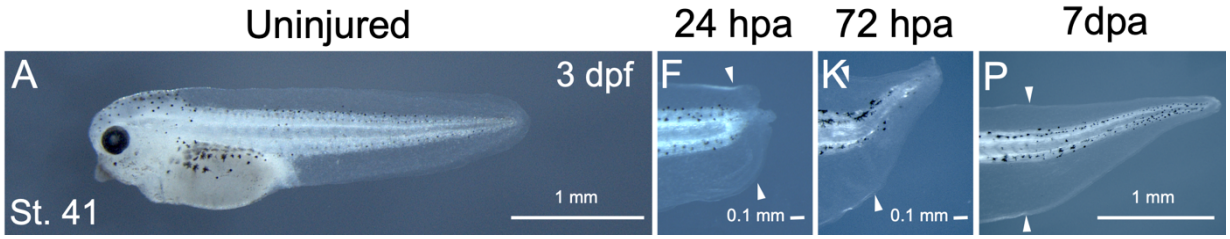

Figure 1-2 Regeneration time-course of *X. tropicalis* tail regeneration
(adapted from Williams and Patel, et al., *in submission*)

In this thesis, I will restrict the scope of discussion to regeneration of the tail and spinal cord. At NF stage 41, complete tail regeneration takes place over the course of 7 days, where the 7 days post amputation (dpa) tadpole is no longer distinguishable from its uninjured clutch mates. Generally, the first several hours of whole-tail regeneration are characterized by an inflammatory response which facilitates wound healing and initiates regeneration. Following tail amputation, reactive oxygen signaling, bioelectrical signaling, and innate immune cell recruitment are the earliest signals to be detected and investigated in regeneration (reviewed in (Kakebeen and Wills, 2019)). Wound closure is concluded when a wound epithelium covers the posterior tip of the tail. After wound closure, the main objective of the tail is to establish a regeneration bud over the next 24-48 hours. The *Xenopus* regeneration bud does not comprise a mass of undifferentiated cells which makes it unique from the regeneration blastemas found in axolotls and zebrafish (Gemberling et al., 2013; Gerber et al., 2018). Lineage tracing experiments in *Xenopus* and axolotl tail amputation have shown the cellular source for the regeneration bud originates from the 500µm zone anterior to the cut site (Lin et al., 2007; Mchedlishvili et al., 2007). Furthermore, lineage tracing experiments in *Xenopus* and axolotls also made it clear that cells in the regeneration bud and blastema retain a cellular memory of tissue of origin. In Kakebeen and Wills 2019, we described the difference in regeneration bud cellular memory:

> "In both species, this was demonstrated by grafting specific tissues from a GFP-expressing donor into an unlabeled host, letting a full limb or tail develop, and then amputating the resulting, tissue-specifically labeled appendage. In axolotls, GFP-labeled muscle gave rise to muscle in the regenerated limb (but not skin or cartilage), labeled skin gave rise to skin (but not muscle or cartilage), and cartilage to cartilage (but not skin or muscle) (Kragl et al., 2009). It is worth noting that in axolotls, cartilage, and bone from the truncated limb do not contribute to the final regenerated skeleton, suggesting that other connective tissue cells are responsible for skeletogenesis, and therefore exhibit plasticity of fate (McCusker et al., 2016). Similar lineage tracing experiments in *Xenopus* examined animals where either the spinal cord, notochord, or muscle had been labeled with GFP, and found that only the same tissue was labeled following tail amputation and regeneration (Gargioli and Slack, 2004)."

Following the establishment of the regeneration bud, a phase of increasing proliferative activity and tail elongation ensues to complete regeneration (Lee-Liu et al., 2014; Tsujioka et al., 2015).

1.3 NEURAL PROGENITOR CELLS IN SPINAL CORD REGENERATION

Neural progenitor cells have been a popular therapeutic target in efforts to repair and regenerate neurons in response to spinal cord injury in humans (Khazaei et al., 2019; Levi et al., 2018; Shin et al., 2015). The utility of neural progenitor cells in regeneration has been discussed across species with regenerating nervous systems as well (Alunni and Bally-Cuif, 2016). In this thesis, I define neural progenitor cells as self-renewing, multipotent progenitors that may give rise to neurons and glial cells (Martínez-Cerdeño and Noctor 2018). These cells are an attractive target due to their ability to either self-renew or differentiate into neurons under the correct conditions, supplying both a cell type that can expand a population of cells as well as produce cells that will bring function to the regenerated tissue. The dual roles of proliferation and differentiation have been well-studied and appreciated in development, where neural progenitor cells necessarily self-renew to expand the population of neural cells and differentiate to develop the nervous system (Thuret et al., 2015). In *Xenopus,* two main periods of rapid neural progenitor cell proliferation

coincide with neurogenesis, followed by a decline to quiescence (Thuret et al., 2015). In addition to rapidly expanding the progenitor pool in development, progenitor cells are also being specified to distinct lineage-restricted cell fates. Opposing gradients of Shh signaling (ventral) and BMP signaling (dorsal) elicit complex gene regulatory networks specifying neural progenitor domains dorsoventrally (Alaynick et al., 2011; Cardozo et al., 2017; Lai et al., 2016). After propagation of lineage-restricted progenitors, the decision to exit the cell cycle and undergo neurogenesis is made by extrinsic signaling by the Notch signaling pathway, specifically the repression of Notch signaling (Lara-Ramirez et al., 2019).

In a spinal cord injury context, neural progenitor cells are tasked with healing the opened tube, proliferating to contribute to growth of the tissue, and differentiating to neurons to make the new tissue functional. A comparative analysis of these processes in regenerative animals is reviewed in Diaz Quiroz and Echeverri 2013 and summarized here in (**Table 1-1**). In animals that can regenerate their spinal cord, such as *Xenopus*, axolotls, and zebrafish, neural progenitors must first repair the wounded spinal cord. The spinal cord self-repairs over the first few hours by reorganizing the ependymal cells (progenitor cells that line the central canal) into a bulb-like structure referred to as the neural ampulla (Beck et al., 2009; Gaete et al., 2012; Stefanelli, 1951; Tanaka and Ferretti, 2009). The neural ampulla is populated with *sox2*$^{+}$ neural progenitor cells which have been a concentrated area of focus in spinal cord regeneration in *Xenopus* as well as in axolotls and zebrafish. *Sox2* is a transcription factor restricted to the neural lineage, and used as a pan-neural progenitor maker (Graham et al., 2003); thus, it has served as a way to track and manipulate neural progenitor cells in regeneration. For this reason, what is known about how neural progenitors function in *Xenopus* spinal cord regeneration is mainly known through the lens of *sox2* manipulation. Studies in *Xenopus,* axolotls, and zebrafish have shown that spinal cord

amputation or transection stimulates proliferation of *sox2*$^{+}$ cells at the wound site (Fei et al., 2014, 2014; Gaete et al., 2012; Muñoz et al., 2015; Ogai et al., 2014). Additionally, in response to injury, *Xenopus sox2/3*$^{+}$ cells have been shown to give rise to new neurons by lineage tracing experiments (Muñoz et al., 2015). In *Xenopus* and axolotls, *sox2* loss of function inhibits spinal cord regeneration (Fei et al., 2014; Muñoz et al., 2015) whereas in zebrafish, *sox2* loss of function has only been shown to result in decrease of proliferating cells (Ogai et al., 2014). Together these studies present a picture where neural progenitor cells both proliferate and differentiate in regeneration and the expression of *sox2* is necessary. However, the temporal ordering of these processes has not been established and the transcriptional regulation of these processes have yet to be interrogated.

Table 1-1 Comparative overview of spinal cord regeneration
(Adapted from Diaz Quiroz and Echeverri 2013, see review for references)

Process	**Fish**	**Frogs**	**Axolotls & newts**	**Mammals**
Glial response	Proliferation, migration, bridge formation	Proliferation, migration, tube formation	Proliferation, migration, tube formation	Proliferation, migration, up-regulation of scar proteins
Glial scar	No	Not in larvae, unknown in adult	No	Yes
***Sox2*$^{+}$ cell proliferation**	Yes	Yes in larvae, no in adults	Yes	
Axon degeneration	Some	In adults	Unknown	Yes
Axon regrowth	Yes	Yes in larvae, no in adults	Yes	No (limited regeneration)
Functional recovery	Partial	Yes in larvae, no in adults	Yes	No (limited recovery)

To gain gene regulatory insight into how neural progenitors make cell fate decisions in regeneration, I analyzed the genomic and transcriptomic profiles of neural progenitor cells purified from the *pax6:*GFP transgenic line. ATAC-Seq was utilized to profile genome-wide changes in chromatin accessibility and identify transcription factor binding sites in the uninjured tail, 0 hours post amputation (hpa), 6hpa, 24hpa, and 72hpa. To complement the chromatin accessibility data, I used scRNA-Seq of the uninjured tail and 24hpa tail to identify gene expression changes and the heterogeneity of neural cell types at these timepoints. By integrating the two data sets I made the new observation that neural progenitor cells are prioritizing tube formation in early regeneration, differentiation at 24hpa, and proliferation at 72hpa. This hypothesis comes in contrast to the intuitive idea that neural progenitor cells would have to first proliferate to expand the neural population prior to differentiation. Further, I used the integration of the data sets to predict gene regulatory networks and identified two novel candidate regulators for neural differentiation and axonogenesis in regeneration. Together, the conclusions drawn from my thesis work advance our knowledge of transcriptional regulation in neural progenitor cells in *Xenopus* regeneration.

In the following chapters I detail my completed work as well as future directions suggested by this work. Chapter 2 describes the optimization of a method to purify neural progenitor cells from transgenic *Xenopus* tadpole. Chapter 3 presents the main experimental work and conclusions drawn from my thesis work. Chapter 4 explains where I see this work fitting into the broader framework of regeneration and the future directions it evokes.

Chapter 2. TISSUE DISAGGREGATION AND ISOLATION OF SPECIFIC CELL TYPES FROM TRANSGENIC XENOPUS APPENDAGES FOR TRANSCRIPTIONAL ANALYSIS BY FACS

Chapter 2 is adapted with minimal modification from:

Kakebeen AD, Chitsazan AD, Wills AE. Tissue disaggregation and isolation of specific cell types from transgenic *Xenopus* appendages for transcriptional analysis by FACS. Dev Dyn. 2020 Nov 2. doi: 10.1002/dvdy.268. PMID: 33137227.

2.1 ABSTRACT

Xenopus embryos and tadpoles are versatile models for embryological, cell biological, and regenerative studies. Genomic and transcriptomic approaches have been increasingly employed in these frogs. Most of these genome-wide analyses have profiled tissues in bulk, but there are many scenarios where isolation of single cells may be advantageous, including isolation of a preferred cell type, or generation of a single-cell suspension for applications such as scRNA-Seq. Here we present a protocol for the disaggregation of complex tail and limb bud tissue, and use cell type-specific fluorescence in transgenic *X. tropicalis* appendages to isolate specific cell populations using fluorescence activated cell sorting (FACS). Our protocol addresses a specific challenge in *Xenopus* embryos and tadpoles: the storage of maternal yolk platelets in each cell, which can introduce light scatter and thereby false positives into FACS analysis. Here we gate against both non-transgenic and ubiquitously transgenic animals to reduce both false positives and false negatives. We use the Xtr.Tg(pax6:*GFP*;cryga:*RFP*;actc1:*RFP*)^{Papal} transgenic line as a test case to demonstrate that nucleic acid preparations made from sorted cells are high quality and specific. We anticipate this method will be adaptable to study various cell types that have transgenic reporter lines to better profile cell types of interest.

2.2 INTRODUCTION

Xenopus tropicalis and *Xenopus laevis* are leading models for embryonic development, appendage regeneration, and human disease modeling (reviewed in (Duncan and Khokha, 2016; Harland and Grainger, 2011; Nenni et al., 2019; Phipps et al., 2020). Next generation sequencing approaches are used routinely in both these species to profile changes in gene expression, transcription factor occupancy, chromatin conformation and chromatin accessibility (reviewed in (Grainger, 2012; Harland and Grainger, 2011; Kakebeen and Wills, 2019a)). More recently, scRNA-Seq has also allowed interrogation of transcriptional changes in specific cell types (Fei et al., 2014; Gaete et al., 2012; Muñoz et al., 2015)(Aztekin et al., 2019; Briggs et al., 2018; Kakebeen et al., 2020). However, there remain many scenarios where it may be advantageous to isolate a cell population of interest prior to sequencing or other downstream applications. Among other reasons, in situations where one seeks to identify transcriptional changes in a specific or rare cell type, profiling gene expression in all the other cell types within that tissue may be both superfluous and cost-prohibitive. Other applications where capture of defined cell types may be useful include assaying simple abundance of a target cell type, capturing a cell based on a reportable state transition (such as cell cycle state or signaling pathway reporter), or capturing a subset of cell types with shared expression of a single marker. For most of these applications, fluorescent transgenic reporters can be used to identify the target cell population, and in many other species, fluorescence activated cell sorting (FACS) based on transgenic fluorescent reporters or fluorescent cell surface markers is routinely used to capture the cell population of interest (Bonner et al., 1972; Cerda et al., 2009; Gallardo and Behra, 2013; Gerber et al., 2018; Hadjantonakis and Nagy, 2000; Manoli and Driever, 2012).

A specific challenge to FACS in *Xenopus* species is that for the first seven days after fertilization each cell has its own maternal yolk store (Selman and Pawsey, 1965). The yolk is light scattering, thus potentially making it more difficult to confidently recover true positive fluorescent cells. Several groups have successfully used FACS in *Xenopus* embryos or tadpoles, demonstrating that the technique can be leveraged in these species to recover diverse cell types ranging from blood cells to neurons (Gilbert, 1986; Kakebeen et al., 2020; Noelanders and Vleminckx, 2018; Ny et al., 2013; Sato et al., 2018; Tsujioka et al., 2015). However, there is not yet a consensus protocol available that specifically addresses yolk light scatter and how to recover true fluorescence from a yolky background. As the library of cell type-specific fluorescent reporter lines continues to expand in both *Xenopus* species, a standardized protocol for FACS will be advantageous. The tadpole tail and limb bud represent useful starting tissues for this analysis. In addition to their widespread value as models for regeneration (Beck et al., 2009b; Kakebeen and Wills, 2019b; Phipps et al., 2020b; Suzuki et al., 2006), both appendage types comprise a variety of cell types from ectodermal, mesodermal, and neural crest lineages. Moreover, a prerequisite for FACS as well as other techniques that require cell disaggregation, like scRNA-Seq, is generation of a single-cell suspension. At younger embryonic stages, this can be achieved by using CMFM or PhoNaK media (Godsave and Slack, 1989; Sargent et al., 1986), but as the tadpole matures, the extracellular matrix becomes denser, and enzymatic methods are required. There are a wide array of enzymatic approaches for tissue disaggregation, many of which have been optimized for cell culture or mammalian tissues at 37 degrees celsius. These include trypsin, papain, and collagenase mixtures (Freshney, 2015; Reichard and Asosingh, 2019). Because over-aggressive disaggregation strategies may lead to cell damage or death, it is valuable to assay the performance of enzymatic cell disaggregation strategies in new species or

tissues. The viability and recovery of *Xenopus* cells in these reagents has not been thoroughly compared and reported.

Here we introduce a FACS strategy for *Xenopus* that accommodates light scatter due to yolk by gating both against non-fluorescent wild-type cells and against cells expressing GFP under the control of a ubiquitous promoter (Xtr.Tg(*eef1a1*:GFP)^{Krieg} (Johnson and Krieg, 1994). We previously reported a specific application of FACS for isolating neural progenitors during tail regeneration (Kakebeen et al., 2020). Here we expand on that strategy to assay multiple enzymatic approaches, and efficacy in the limb bud as well as tail. Our cell disaggregation and FACS method (**Figure 2-1 At a glance: Tissue disaggregation and FACS purification of transgenic *Xenopus* tissueFigure 2-1**) is optimized to retrieve only true positive cells while excluding light scattering cells. Using neural progenitor cells with the Xtr.Tg(*pax6*:GFP;*cryga*:RFP;*actc1*:RFP)^{Papal} (Hartley et al., 2001; Hirsch et al., 2002) transgenic line as a test case, we evaluated the quality of RNA or chromatin extracted from cells that have gone through the sorting process by comparing gene expression and chromatin accessibility to cells that were not sorted. We also used expression analysis to validate the neural progenitor identity of sorted cells. These assay measures present a potential guideline for assaying cell type fidelity and transcriptional stress response signatures in *Xenopus* sorted cells. Our hope is that this optimized protocol will be applicable to a variety of cell types represented by new and established transgenic lines, and be adjustable for several forms of downstream analysis.

At a Glance: Disaggregation and FACS purification of transgenic *Xenopus* tissue

Materials

Liberase	5mg (working: 0.035 mg/ml)	Roche 5401119001
PBS	1x	in house
Cell Strainer	70um	Fisher 22-363-548
Polystyrene round bottom tube	12x75mm, 5 mL	Corning 352008
Trypan Blue	0.4%	Thermo 15250061

Cell Disaggregation (For 20xNF stage 41 tails or 1-2x NF stage 53 limb buds)

** Prepare an extra wild-type and transgenic sample for gate setting*

1. Isolate tissue of interest into a 1.5ml Eppendorf tube.
2. Spin down for 1 minute at 500-1000g to pellet tissue.
3. Take off media and replace with 200µl of 0.035mg/ml (1x Liberase). Gently Vortex and incubate at room temperature for 20 minutes. Keep samples in dark while incubating to retain reporter fluorescence.
4. Gently pipette sample up and down using a p200 until there are no visible tissue chunks. Avoid vigorous pipetting and introducing bubbles.
5. When tissue is entirely disaggregated, spin down cells for 3 minutes at 1000xg.
6. Remove supernatant and wash cells with 180µl of 1xPBS.
7. Spin down cells for 1 minute at 1000xg.
8. Remove supernatant, add final 1xPBS volume for sorting (100-250µl), and mix until cells are in suspension.
9. Pipette cell suspension through a 70um cell strainer into a polystyrene, round bottomed tube. Place cells on ice and in the dark until sorting.

Cell Viability

1. Mix 10µl of cell suspension with 10µl of 0.4% Trypan Blue (or dilution of choice).
2. Load 10µl of solution to each side of hemocytometer.
3. On side one: Count the total number of cells in a fixed grid (non-blue cells).
4. Count the number of dead cells in the same fixed grid (blue cells).
5. Repeat for side two
6. Average the total number of cells and number of dead cells from the two sides, then calculate cell viability: (Total cells - dead cells)/ Total cells * 100

Sorting and collecting by FACS on BD FACSAria III Cell Sorter

1. First run the wild-type gating sample through the flow cytometer and set gates:
    a. All cells: FSC-A v. SSC-A (to select against debris) → p1
    b. P1: FSC-H v. FSC-W (to select for singlets) → p2
    c. P2: GFP cells v. Count (histogram to select for GFP cells) → p3/p4
    d. Alt. Pop 2: GFP cells v. SSC-A (scatter plot for GFP cells) → p3/p4
        * p3=GFP$^-$ cells, p4=GFP$^+$ cells
        ** All cells in wild-type gating sample should be in p3
2. Next run the fluorescent gating sample through the flow cytometer and adjust GFP gate (p4) to be more or less stringent. **Do not adjust the p3 gate.**
3. Apply samples and collect GFP+ cells from p4 into chilled buffer.

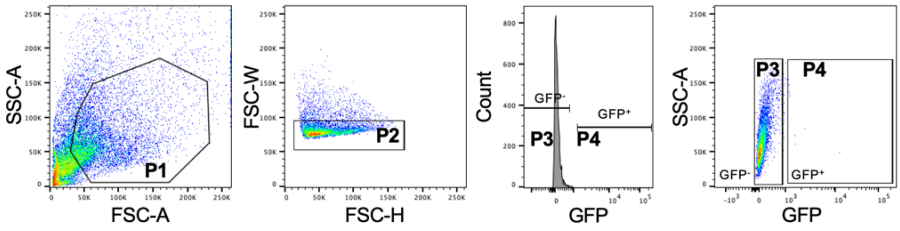

Figure 2-1 At a glance: Tissue disaggregation and FACS purification of transgenic *Xenopus* tissue

2.3 Results and Experimental Procedures

2.3.1 *Transgenic lines used*

Our primary goal was to articulate FACS conditions that would use transgenic expression of fluorescent proteins to let us confidently recover specific cell types from the complex tissues of the tadpole tail and limb. We therefore selected comparable *Xenopus tropicalis* lines that were either non-transgenic (wildtype), ubiquitously expressed GFP [Xtr.Tg(eef1a1:GFP)^{Krieg} , hereafter *eef1a:GFP*], or had tissue-specific expression of GFP [Xtr.Tg(pax6:*GFP*;cryga:*RFP*;actc1:*RFP*)^{Papal}, hereafter *pax6:GFP*](Hartley et al., 2001; Hirsch et al., 2002; Johnson and Krieg, 1994). Wildtype parent frogs were purchased from Nasco, and transgenic parent frogs from the National *Xenopus* Resource (https://www.mbl.edu/xenopus/)(Horb et al., 2019). Information on the *eef1a:*GFP line and *pax6:*GFP line can be found with the reference ids RRID:NXR_1.0008 and RRID:NXR_1.0021, respectively. The *eef1a:GFP* line was used as a positive control for sorting, as *eef1a* is stably expressed during development and used in *Xenopus* as a reference gene for RT-PCR studies (Dhorne-Pollet et al., 2013). In an *eef1a:*GFP transgenic animal, we observed robust reporter expression in most tissues in the head, the gut, and the axial tissues (**Figure 2-2**A). This transgenic line was used to evaluate how well our gating returned true positive reporter cells. Transgenic *pax6:GFP* tadpoles were utilized to isolate a rare cell type (neural progenitor cells) from the tail. The GFP reporter is expressed specifically in the forebrain, hindbrain, eye, and spinal cord of these tadpoles (**Figure 2-2**B). Transgenic matings were performed by crossing a heterozygous transgenic frog to a wild-type frog. These matings yielded clutches with 50/50 wt/*pax6*:GFP+ populations.

All of the following work was carried out under the supervision of IACUC at the University of Washington, an AAALAC-accredited institution (protocol 4374-01).

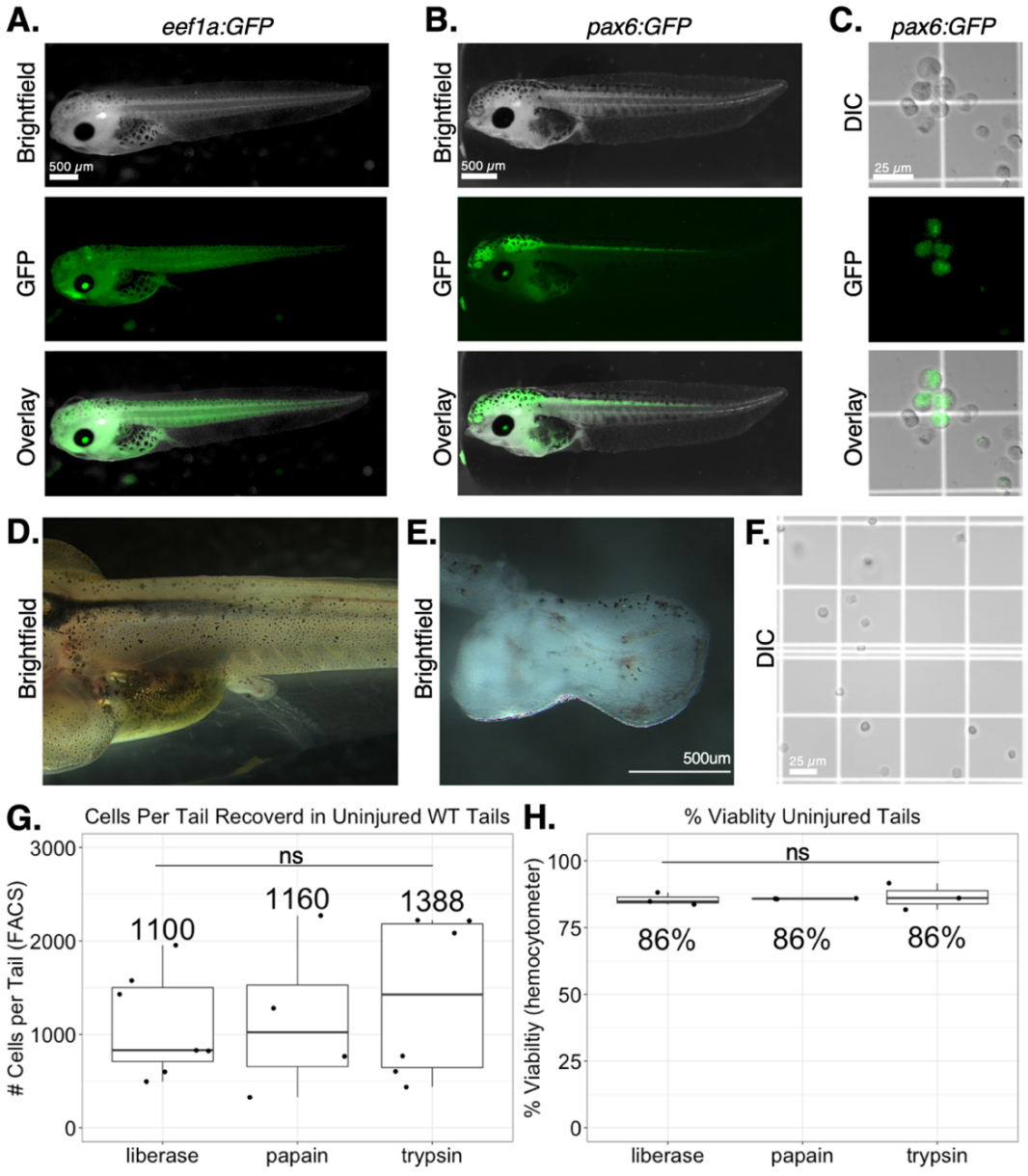

Figure 2-2 Disaggregation of *Xenopus tropicalis* tails and limbs.
(A-B) NF stage 41/42 live tadpoles imaged for reporter fluorescence. (A) *Eef1a:GFP* transgenic tadpole. (B) *Pax6:GFP* transgenic tadpole. (C) Dissociated cells from a *pax6:GFP* animal. Images were taken in DIC and in the GFP channel. (D) Wild-type NF stage 53 tadpole trunk and (E) the isolated limb bud from that tadpole. (F) Cells from dissociated limb bud in (E). (G) Boxplot showing cells per tail for different enzyme preparations measured from FACS treated wild-type tail tissue. Liberase, n=7; Papain, n=4; Trypsin, n=6. A pairwise t-test was performed between any two enzymes. ns = not significant. (H) Boxplots showing the % cell viability from disaggregated tissue prior to FACS for different enzyme types (n=3). The number reported above the plot is the mean viability for n=3.

2.3.2 *Tadpole tail and limb isolation*

Tadpoles were reared in 1/9 modified Ringer's solution (1/9 MR)(Kay and Peng, 1991) to 3 days post fertilization or stage 41(Nieuwkoop and Faber, 1956, 1994), at which time tail collections were performed. For limb bud isolation, tadpoles were reared to stage 53/54. Tadpoles were anesthetized using 0.5% MS222 in 1/9 MR, and tested for the loss of the escape reflex and other touch responses prior to amputation surgery. Once fully anesthetized, a sterilized scalpel was used to amputate the posterior third of the tail by cutting the tail from dorsal to ventral. Amputated tails were collected by pipetting with a p200 from the collection dish to an Eppendorf tube. Numbers of tails collected were noted. Following surgery, tadpoles were either removed from anesthetic media within 10 minutes of amputation into new 1/9x MR for regeneration studies or left in anesthetic for euthanasia.

For limb bud collection, NF stage 53 tadpoles were euthanized by anesthetic immersion in MS-222, and limb buds were removed using forceps to pinch off limb buds where the limb connects to the body (**Figure 2-2**D-E). Single limbs were collected into their own eppendorf tube.

2.3.3 *Tissue dissociation to single cell suspension*

Successfully dissociating tissue to a single cell suspension full of intact and viable cells is the first major check-point in the cell sorting process. To this end, we compared three different enzyme systems for tissue dissociation. Our goal was to break apart the tissue while preserving reporter expression and cell integrity. We optimized our protocol for speed and gentleness, to minimize disruptions to the gene expression or chromatin accessibility profile through time and stress. We interrogated three commonly-used enzymes for our dissociation protocol: liberase,

trypsin, and papain. We describe the performance of each below, but selected liberase for downstream experiments.

Liberase

Liberase (a proprietary collagenase mixture from Roche, cat 5401020001) was resuspended from a 5mg lyophilized pellet to 0.35mg/ml solution in 1xPBS. The working concentration was made by diluting the stock ten fold in 1xPBS to a concentration of 0.035mg/ml. Once tails were collected into Eppendorf tubes, tails were spun down for 1 minute at 500g to pellet the tails. The supernatant was discarded and replaced with 200µl of 0.035 mg/ml Liberase solution. Tails were left to incubate for 20 minutes at room temperature. To preserve reporter fluorescence, incubations were performed in dark conditions. Following the incubation period, tails were manually dissociated using a p200 pipette and gently pipetted up and down until no visible tissue chunks were apparent (~2-5 minutes). Once cells were dissociated, the suspension was spun down for 5 minutes at 1000g. The supernatant was discarded, and the pellet was then washed in 180µl of 1xPBS. The washed pellet was then spun down for 1 minute at 1000g. The supernatant was discarded, and the pellet was resuspended in 250ul of 1xPBS (or volume suitable for tail amount and FACS analysis). The cell suspension was then passed through a 70µm cell strainer (Fisher 22-363-548) into an appropriate round bottomed polystyrene tube adaptable to the BD FACSAria III cell sorter (hereafter referred to as FACS collection tube) (Corning 352008). The cell suspension was placed on ice and in the dark until ready to be sorted. Resulting cell suspensions retained GFP expression (Figure 2-2C).

For limb disaggregation, media was discarded and replaced with 200µl of 0.035mg/ml liberase for one limb. The same procedures for disaggregation of the tail tissue were followed to bring limb buds to single cell suspension. Resulting suspensions yielded well disaggregated tissue (Figure 2-2F).

Trypsin

After spinning down tails in 1/9 MR as described previously, supernatant was discarded, and the pellet was resuspended in 100µL of 0.25% trypsin (ThermoFisher 25200056). Tails were incubated in trypsin for 20 minutes at room temperature. Tails were then gently mechanically dissociated with a p200 pipette until no visible tissue chunks were apparent (~2-5 minutes). Once tissue was fully dissociated, 100µL of newborn calf serum (NBCS) (ThermoFisher 26010066) was added to the cell suspension and mixed to stop enzymatic activity. The cell suspension was then spun down for 5 minutes at 1000g and supernatant was discarded. The pellet was washed with 180µl of 1xPBS and spun down for 1 minute at 1000g. The pellet was then resuspended in 250µl of 1xPBS and passed through a 70µm cell strainer into a flow cytometry tube. The cells were kept on ice and in the dark until sorting.

Papain

As with the liberase preparation, tails in 1/9 MR were spun down for 1 min at 500g to collect tails at bottom of Eppendorf tube and media was replaced with 100µl of papain solution (750 µL papain and 75 µL DNAse, following instructions in the Worthington Papain Dissociation System: catalog LK003160). The concentrations of papain and DNAse in solution are 20units/ml papain and 0.005% DNAse. Tails were incubated in papain solution for 10

minutes at room temperature followed by gentle agitation by pipetting up and down with a p200 for about 20 rounds. The tails were then left for an additional 10 minutes to incubate at room temperature. Tails were pipetted with a p200 gently until no more visible tissue chunks were apparent (~2-5 minutes). Following dissociation, 100µL of albumin-ovomucoid inhibitor solution was added to cell suspension to inactivate enzyme activity. Cell suspension was then spun down at 1000g for 3 minutes. Supernatant was discarded and the pellet was washed with 180µL of 1xPBS. The suspension was then spun down at 1000g for 1 minute, and the supernatant was replaced with 150µL of 1xPBS. The cell suspension was then pipetted through a 70µm cell strainer into a flow cytometry tube and kept on ice and in the dark until sorting.
For each enzyme preparation tested, cells were passed through the flow cytometer and the number of single cells were recorded and divided by the number of tails used for cell preparation (see **Figure 2-3**B-2). All three disaggregation systems resulted in a mean number of cells per tail recovered that were not significantly different from one another (**Figure 2-2**G). We did not find significant advantages for any one enzyme. We moved forward with Liberase because it had fewer components (no inactivating solution) and had less variability in cells per tail recovered.

2.3.4 *Viability analysis with Trypan blue*

To evaluate how many live cells we recovered from the tissue disaggregation process, we measured the cell viability from the single cell population. Trypan blue staining was used to distinguish dead cells from live cells; dead cells would stain with blue whereas live cells remain colorless. To do this analysis, 14µL of cell suspension was mixed with 7µL of 0.4% Trypan blue (Thermo 15250061). 10µL of solution was loaded on each side of a hemocytometer and cells were counted in a designated 25 squares corresponding to a cell count conversion of 1 cell

counted to 10,000 cells/ml solution. First all cells were counted, followed by only dead cells being counted. The percent viability was calculated by the following equation: (all cells - dead cells)/all cells * 100. We found that on average, disaggregation resulted in ~86% cell viability in suspensions collected by all three enzymatic methods (**Figure 2-2**H).

2.3.5 *Isolation of GFP+ single cells using FACS*

Setting gates for isolation of fluorescent cells from transgenic animals

The next step for purification of GFP^+ cells was to use the BD FACSAria III cell sorter to collect cells. One issue arising from sorting cells for GFP fluorescence from early stage *Xenopus* embryos is that each individual cell has a maternal yolk store for nutrition. This results in light scatter, which could skew the sorting toward false-positive cells. An example of the light scatter that occurs in wildtype tadpoles of this stage of tadpole can be seen in the GFP channel of **Figure 2-2**B. To mitigate the number of false-positives picked up as a result of the light scatter, we first used wild-type tadpole cells to set our gates (**Figure 2-3**A). A gating sample (wild-type cells) was prepared in addition to transgenic samples for collection in parallel. These cells were flowed through the BD FACS Aria III sorter at a flow rate of 3-6 µl/min, through a 70um nozzle. Once cells were flowing, cells were first gated by forward scatter (cell size) and side scatter (granularity) light to select against cellular debris (Figure 2-3B-1). Next, the cells from the "cell" population were gated by forward scatter to select for single cells apart from cells that had not been dissociated (Figure 2-3B-2). The cells selected by this gate were then visualized on a histogram where the x-axis represents GFP cells and the y-axis shows the count of them (**Figure 2-3**B-3). Two vertical gates were set, one selecting GFP^- cells and one selecting GFP^+ cells. Because no wild-type cells should have GFP reporter, all cells were selected for in the GFP^-

group. A second gate was then set to be the GFP^+ group. This step can also be visualized by a scatter plot, where the x-axis represents GFP and the y-axis represents side scatter light (**Figure 2-3**B-4). Here, we draw a gate encompassing all clustered cells representing the GFP^- cells and a second gate slightly offset to capture GFP^+ cells. Our goal was to use this gating scheme and apply it to transgenic samples, resulting in true GFP^+ cells being captured.

Sorting and collecting transgenic cells

To be sure the gates were representing GFP^+ signal and not excluding true positive cells from collection samples, a transgenic gating sample was also prepared. This sample was again flowed through the sorter with the same flow rate, and the gates from the wild-type gating sample were applied. Using the exact same gates as for wild-type cells, we saw an increase in the number of cells captured by the GFP^+ gate (**Figure 2-3**C/D). In the case of *eef1a*:GFP, this gating scheme recovered 88.1% of cells as GFP^+, with a clear division between the positive and negative cells (**Figure 2-3**C). This demonstrates that the FACS protocol was able to recover most cells, and therefore most cell types, counted by the cytometer. The strong split in signal between GFP^+ and GFP^- cells also suggests that a small proportion of cells in this transgenic line may not express detectable GFP at this stage. In addition to the sorting parameters, the collection action was now activated to collect GFP^+ and GFP^- cells. Cells were collected into 1.5ml Eppendorf tubes containing 100-200µl of chilled 1xPBS. Alternatively, cells destined for RNA extraction could be sorted into 1.5ml Eppendorf tubes containing 200µl of RNA-lysis buffer. Collected cells were kept on ice until nucleic acid extraction.

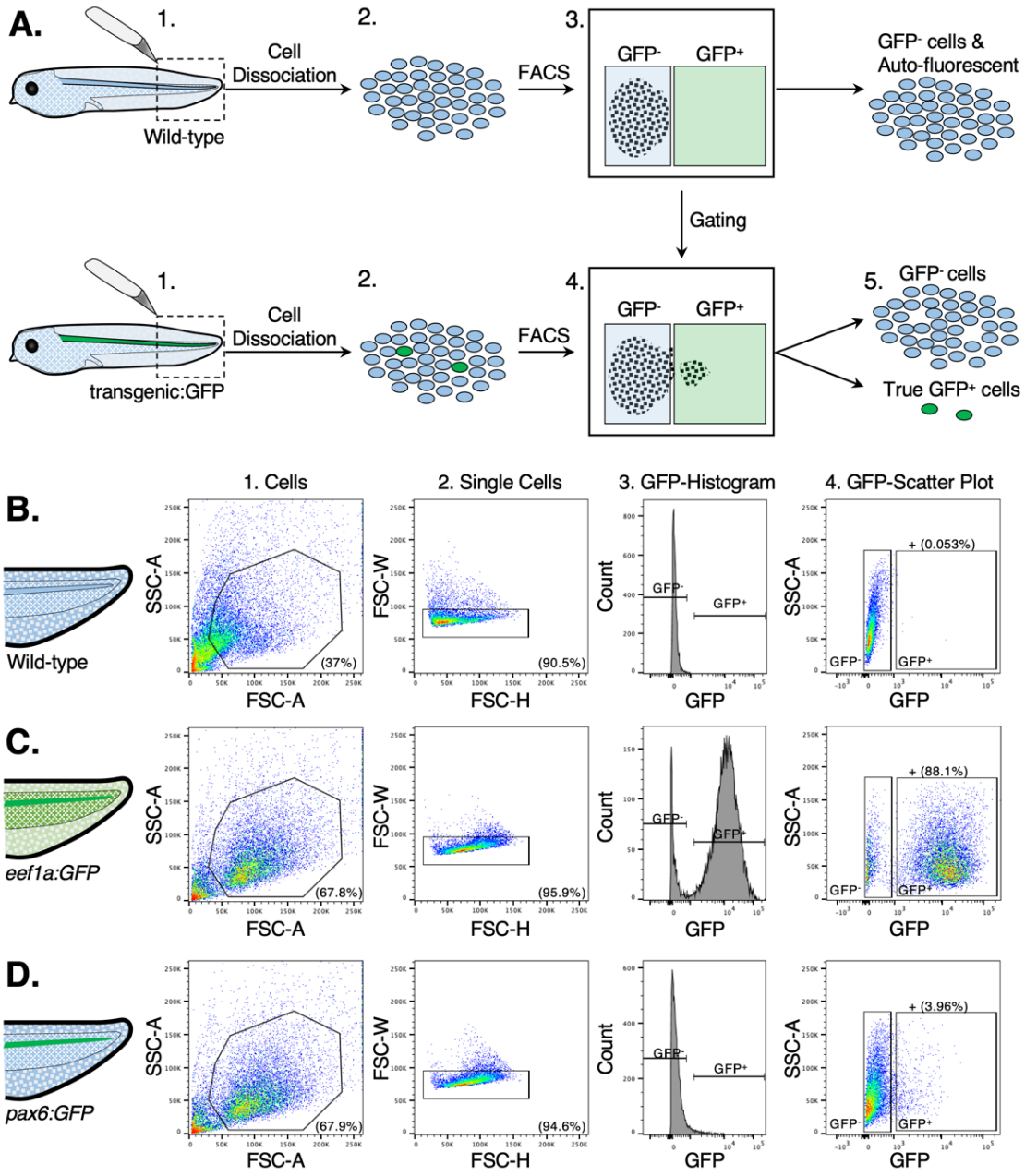

Figure 2-3 FACS gating settings for GFP+ cell isolation.
A) Gating plan for isolating true positive GFP cells from transgenic animals using wild-type tadpoles as controls. (B-D) The first column (1.Cells) refers to cells sorted by forward-scatter light (FSC-A) and side-scatter light (SSC-A) to select against cellular debris. Cells from this selected population were re-sorted by FSC-H and FSC-W to exclude multiples of cells (2.Single Cells). Single cells were then visualized by a histogram with GFP on the x-axis and counts on the y-axis (3.GFP-Hisotgram). The single cells w ere also visualized by a scatter plot with GFP on the x-axis and SSC-A on the y-axis (4. GFP-Scatter Plot). (B) Example of gating in wild-type cells. In the GFP gating, all cells are selected against to restrict auto-fluorescent cells from being picked up by the GFP gate. The gating from (B) was applied to transgenic cells from the (C) *eef1a:GFP* transgenic line and (D) *pax6:GFP* transgenic line. Cells in the GFP^+ gate in (C/D) represent true GFP positive reporter cells.

2.4 DOWNSTREAM LIBRARY PREPARATION AND ANALYSIS

The downstream use of isolated cells will vary among users. For the purpose of this paper, we extracted RNA for RT-qPCR and transposed DNA for ATAC-Seq, as described previously (Kakebeen et al., 2020). Specific alterations to the previously-described methods are noted.

2.5 STRATEGIES FOR VALIDATING CELL TYPE RECOVERY AND QUALITY

Sorted vs unsorted vs undissociated genomic library preparations

Our primary interest was in isolating specific cell types for analysis of gene expression and chromatin accessibility. A concern both when disaggregating tissue from its native state and when putting individual cells through the process of flow cytometry is that the cells may undergo damage as a consequence of either mechanical stress or of the time the procedure takes. While transcriptomic studies in leukocytes and mouse mammary tissue have suggested that the overall transcriptional profile is minimally changed by FACS (Beliakova-Bethell et al., 2014; Richardson et al., 2015), there is still relatively little information about how other tissues and cell types may be changed, particularly in *Xenopus*. Indeed, work conducted in mouse macrophages and cultured astrocytes suggests that there may be significant alterations to the metabolome and redox state of cells when subjected to FACS (Binek et al., 2019; Llufrio et al., 2018). Finally, the impact of FACS on the epigenome, and specifically on chromatin accessibility, has not been systematically interrogated. These motivations led us to optimize our disaggregation and FACS protocol for speed and gentleness, but to assess its performance we sought to test whether gene expression and chromatin accessibility were significantly disrupted in cells subjected to FACS when contrasted to comparable cells not subjected to FACS.

ATAC-seq library data analysis in sorted vs unsorted cells suggests that FACS does not disrupt chromatin accessibility

To test whether the physical force of flowing cells through a flow cytometer changes chromatin structure, we made ATAC-Seq libraries from cells that have been through a flow cytometer (sorted) and cells from the exact same tissue that have not (unsorted). For ATAC-Seq experiments, tadpole tails were amputated and the 0 hour post amputation tail tips were pooled (n=4 for ATAC-Seq) (**Figure 2-4**A). The pooled tails were then dissociated to single cell suspension. After washing, the cells were resuspended in 300µl of 1xPBS and split into two 150µl aliquots in separate FACS tubes; the two aliquots served as "sorted" and "unsorted" partner replicates. "Sorted" cells were flowed through the flow cytometer only utilizing gates 1:"Cells" and 2:"Single Cells" from Figure 2-3. All single cells were collected directly into chilled 1xPBS and placed on ice until all samples were collected. "Unsorted" samples were held in the FACS tubes on ice while its replicate pair was flowed through the flow cytometer (~7-10 minutes) (**Figure 2-4**A). When the sorted partner was finished running, the unsorted partner was transferred from the FACS tube to an Eppendorf tube containing chilled 1xPBS. We made ATAC-Seq libraries from these cells. When we ran libraries out on a bioanalyzer, we found that ATAC-Seq libraries made from either sorted or unsorted cells have a comparable single nucleosome peak at 180bp (**Figure 2-4**B). The presence of this peak is an important quality control step for efficient tagmentation and library amplification. We then went on to sequence these libraries.

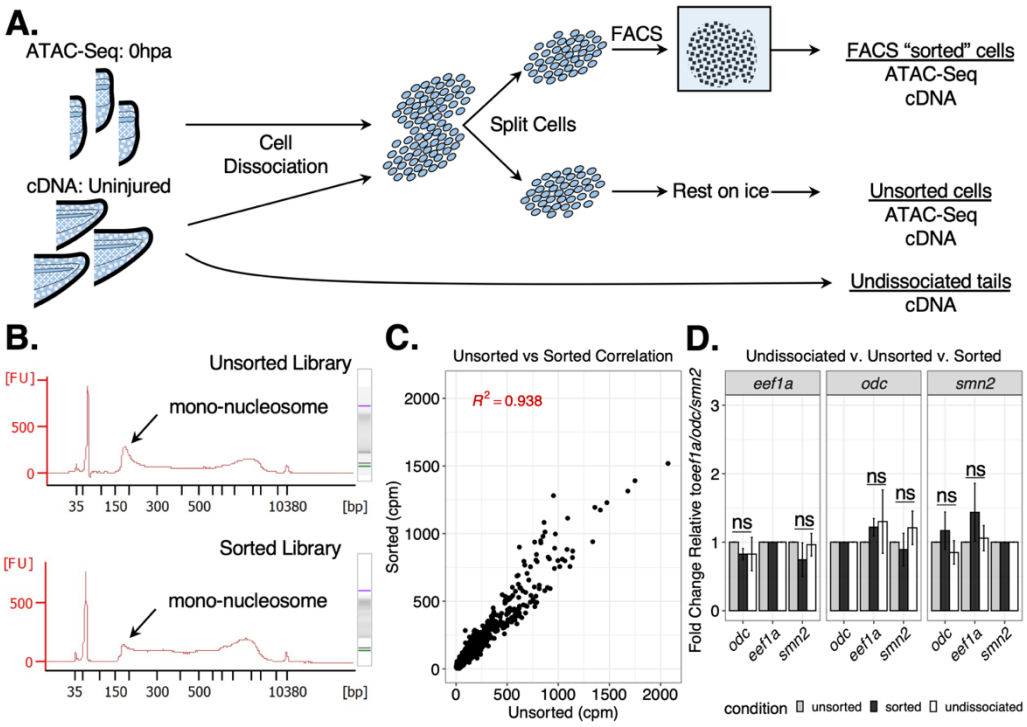

Figure 2-4 Sorting by FACS does not significantly change gene expression or chromatin quality. (A) Experimental design to test gene expression and chromatin architecture of cells that have been sorted and cells that have not been sorted and RNA from undissociated tails. (B) Bioanalyzer traces for the sorted and unsorted libraries. Arrow points to characteristic ~180bp. (C) Correlation plot of counts in counts per million (cpm) from ATAC-Seq samples that were sorted or unsorted. The samples correlated with an r^2 of 0.938. (D) RT-qPCR results from RNA extracted from cells that have been undissociated (n=3), unsorted(n=3), or sorted (n=3). We tested three house-keeping genes and quantified fold change expression between sorted or undissociated to the unsorted condition. Expression was independently normalized to *odc, eef1a,* or, *smn2*. T-tests between sorted and unsorted conditions were used to identify if FACS significantly altered gene expression. Error bars represent standard deviation. ns= not significant

Sequencing results were returned as *fastq* files and the pre-processing pipeline was performed as previously described (Kakebeen et al., 2020). Both types of libraries had comparable total read pairs sequenced, greater than 90% alignment to the genome, and a minimum of 25000 peaks called. These samples are within range of ENCODE's strictest QC parameters (https://www.encodeproject.org/atac-seq/). Next we ran a correlation analysis to test how well correlated the counts per peaks called were between the sorted and unsorted libraries. Correlation analysis between sorted and unsorted 0hpa libraries was performed by using the R

command “lm” to fit a linear model to the counts per million of peaks in each library type. Linear regression analysis called an R^2 of 0.98 showing that each sample has a similar number of counts per peak (**Figure 2-4**C). This suggests that sorted and unsorted libraries are not significantly divergent when evaluated by their peak calls.

Sorted vs unsorted vs undissociated gene expression

In addition to testing the differences in chromatin accessibility for sorted and unsorted cells, we added in one more technical consideration by also collecting RNA from whole-tail lysates without cell dissociation. For these experiments, 25 uninjured tails were pooled per sample. Sorted and unsorted samples were collected in the same manner as the ATAC-Seq samples, however cells were sorted or added to 200µl RNA-lysis buffer. In contrast, the “undissociated” tail tissue was spun down after amputation, supernatant was removed, and RNA-lysis buffer was added directly to tails. We elected to lyse the whole tails immediately after collection rather than incubating on ice, reasoning that cells in the context of whole tissue would be more vulnerable to long-range signals that might trigger apoptosis. Once cells were in lysis buffer, they were treated the same from RNA extraction to cDNA synthesis. To assure we controlled for cell loss in each condition, RNA concentration was normalized prior to cDNA synthesis.

To test whether gene expression is altered by the dissociation and FACS process, we extracted RNA from grouped cell populations that were FACS-sorted or not sorted and from cells collected from whole-undissociated tissue and made cDNA from each population. We collected sorted, unsorted, and undissociated cell populations in triplicate and used qRT-PCR to measure relative gene expression of commonly-used housekeeping genes, using previously

reported primers (Dhorne-Pollet et al., 2013). To test whether there were significant changes in gene expression between the different conditions, the fold change for each gene was calculated in the sorted and undissociated samples with respect to the unsorted samples. When evaluating the differences in gene expression we used a t-test between sorted and unsorted samples and undissociated and unsorted samples and found that there was not a significant change in gene expression for housekeeping genes in either of these conditions (**Figure 2-4**D). This analysis suggests that expression of these high-abundance housekeeping genes is not significantly changed when tissue is dissociated or sorted through the flow cytometer.

Validation of specific sorted cell identity using pax6:GFP

To test whether GFP-positive sorted cells from the *pax6:*GFP transgenic line had identity consistent with the transgenic reporter, we compared chromatin accessibility of sorted transgenic cells against cells from aggregate tissue. To assess whether ATAC-Seq could detect cell-type specific chromatin signatures we analyzed 0hpa *pax6:GFP* and 0hpa aggregate tissue samples from (Kakebeen et al., 2020) (GEO accession GSE146830) by identifying differentially accessible peaks enriched in *pax6:GFP*+ 0hpa samples relative to aggregate tissue 0hpa samples with edgeR (Robinson et al., 2010). Each peak was annotated to the nearest transcription start site. We then filtered the differentially accessible peaks to those within 15000 bases of a transcription start site (218 regions) and input the gene annotations to gProfileR2 (Raudvere et al., 2019) for gene ontology analysis. Gene ontology analysis revealed that the *pax6:GFP*+ libraries are enriched for regions annotated to GO terms such as “Nervous System Development” (p=0.0019), “Regulation of postsynaptic specialization assembly” (p=0.0145), and “Regulation of nervous system development” (p=0.0280) (**Figure 2-5**B). These GO terms confirm that the

pax6:GFP^{+} cells have a neural signature when compared to the aggregate tissue cells. If we look specifically at regions more accessible in the *pax6* libraries called in the GO term "Nervous system development", we see that known neural markers are more accessible (**Figure 2-5**C), thus indicating that purifying cells from a transgenic reporter by FACS presents a method to increase the resolution of bulk ATAC-Seq for a cell type of interest.

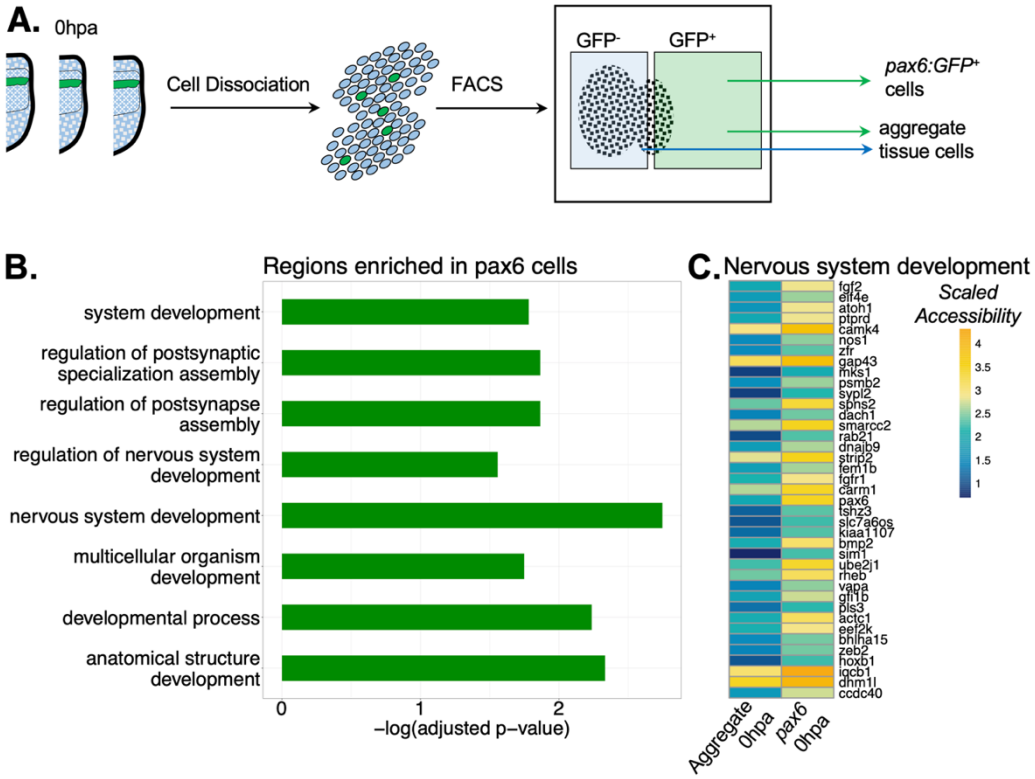

Figure 2-5 Sorting cells by FACS enriches specific cell types.
(A) Experimental design for collection of *pax6:*GFP^{+} cells, *pax6:*GFP^{-} cells, and wild-type, aggregate tissue cells (B) Reported -log(pvalue) for gene ontology terms returned for regions more accessible in *pax6:*GFP^{+} ATAC-Seq libraries (n=3) than aggregate tissue libraries (n=3). (C) Heatmap showing accessibility of regions of the chromatin annotated to the nearest transcription start site shown in aggregate tissue or sorted *pax6*, 0hpa ATAC-Seq libraries. Regions shown called the GO term "Nervous system development" in (B).

2.6 CONCLUSIONS

In this article we articulate a protocol to disaggregate *Xenopus tropicalis* tail and limb tissue to a single cell suspension and sort the cell suspension by transgenic fluorescent reporter expression. By gating against wild-type nonfluorescent cell populations, our protocol allows recovery of true positive cells, even in situations where a clear break-point in the fluorescent signal distribution is hard to identify. Downstream validation through Gene Ontology analysis of chromatin signatures from ATAC-Seq libraries confirms that this protocol can recover a rare target cell population, the neural progenitors, from small amounts of starting tissue.

In comparing multiple enzymatic approaches to disaggregate these tissues, we did not find a significant difference in performance between liberase, trypsin and papain. While we chose liberase for our downstream analysis, researchers using other starting tissues or with different target cells may wish to compare all three enzymes for performance in their specific application.

FACS offers the opportunity to isolate a population of interest and then sequence it in bulk, which is more cost effective than scRNA-Seq for situations where all members of a specific cell type are the target. For situations where single cell analysis is needed, our disaggregation protocol also offers a jumping-off point with high cell viability. We find that the FACS process does not drastically disrupt expression of well-known housekeeping genes, in agreement with existing literature that FACS has minimal effect on the transcriptome. Importantly, we find that FACS also causes minimal disruption to the epigenome as assayed by ATAC-Seq peak locations genome-wide. With the growing library of transgenic animals and expanding toolbox of sequencing techniques, we expect this protocol will be a useful resource for *Xenopus* researchers wanting to characterize specific cell types and may also offer a starting

point for researchers in emerging aquatic vertebrate systems seeking to develop their own protocols.

2.7 ACKNOWLEDGEMENTS

We thank the University of Washington Pathology Flow Cytometry Core for helping with the initial gating considerations and scheme. We thank the Reh lab for Papain reagents and guidance. We thank members of the Wills Lab for helpful discussions along the way. We thank Xenbase for curation of genomic and literature information and the National Xenopus Resource for frogs.

Chapter 3. Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors

Chapter 3 is adapted with minimal modification from:

Kakebeen AD, Chitsazan AD, Williams MC, Saunders LM, Wills AE. Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors. Elife. 2020 Apr 27;9:e52648. doi: 10.7554/eLife.52648. PMID: 32338593; PMCID: PMC7250574.

3.1 INTRODUCTION

In contrast to mammals, *Xenopus* tadpoles are able to undergo scarless healing and full regeneration of the limb, spinal cord, or tail after injury (Beck et al., 2009c; Kakebeen and Wills, 2019b; Lee-Liu et al., 2017; Tseng and Levin, 2008). While lifelong regenerative healing is a characteristic shared by many amphibians and fish, the regenerative capacity of *Xenopus* declines during metamorphosis, and is lost in the adult (Filoni and Bosco, 1981; Mitogawa et al., 2015; Suzuki et al., 2006). *Xenopus* therefore represents an especially useful model for understanding the cell-intrinsic and –extrinsic properties governing regeneration.

In *Xenopus* as in other regenerative animals, the event of a major injury triggers a rapid transcriptional remodeling of the injured tissue. It is now well-established that some aspects of this remodeling recapitulate developmental signaling events. In particular, developmental signaling pathways such as Wnt, FGF, BMP, TGF-ß, Notch and Shh are upregulated, and are required for full regeneration of the limb, tail, and spinal cord (Beck et al., 2003b; Ho and Whitman, 2008b; Slack et al., 2008; Taniguchi et al., 2014). Genome-wide transcriptomic studies have confirmed that numerous genes associated with embryonic development are re-expressed during regeneration (Chang et al., 2017b; Lee-Liu et al., 2014b; Love et al., 2011). However, these studies have been carried out on bulk regenerating tissue, making it difficult to identify what signals or factors are required to promote regeneration in specific cell types.

Recently, single-cell transcriptomic analysis (scRNA-Seq) of both the regenerating *Xenopus laevis* tail and the regenerating axolotl limb have begun to identify the transcriptional signatures associated with distinct cell types (Aztekin et al., 2019; Gerber et al., 2018; Pelzer et al., 2020). These studies also highlighted intriguing distinctions between the models. The regenerating axolotl limb shows a transcriptional convergence between all connective tissue cell types, associated with the formation of the undifferentiated limb blastema (Gerber et al., 2018), while in the *Xenopus* tail, cell type identities are kept clearly intact as regeneration progresses (Aztekin et al., 2019). While these studies substantially advance our atlas of the genes expressed by different cell types in regeneration, they do not identify the transcription factors that interpret and respond to injury cues, nor the transcriptional regulatory elements that trigger changes in expression profiles.

To address these gaps, we sought to articulate the gene regulatory network and transcriptional dynamics associated with regeneration by specifically targeting NPCs, a critical cell type for regeneration. NPCs represent the focus of some therapeutic efforts to restore human motor function following spinal cord injury in regenerative medicine (Khazaei et al., 2019; Levi et al., 2018; Shin et al., 2015) and much effort has been applied to defining the *in vivo* and *in vitro* programs that guide their cell fate decisions in the developing spinal cord as well as in culture. During development, neural stem cells give rise to distinct domains of lineage-restricted progenitor cells defined by their expression of specific transcription factors across the dorsal ventral axis (Alaynick et al., 2011; Lai et al., 2016). The decision to exit the cell cycle and undergo terminal neuronal differentiation is directed by extrinsic cues, particularly repression of Notch signaling, which is highly conserved across vertebrates (Lara-Ramirez et al., 2019). Following amputation, the spinal cord first begins to repair by closing

the open end to form the tube-like neural ampulla, a structure formed by ependymal cells. Progenitor cells then contribute to both growth and new differentiation as the newly regenerated spinal cord elongates (Beck et al., 2009c; Gaete et al., 2012; Stefanelli, 1951). Therefore, neural progenitors must balance their decisions between proliferating to repopulate lost tissue and differentiating to mature neurons. The global transcriptional dynamics that govern these decisions, including the transcription factors required and their target sites across the genome, remain poorly defined.

To identify gene regulatory dynamics in neural progenitors during regeneration, we analyzed the transcription and chromatin accessibility profiles of neural progenitors purified from *pax6:GFP* transgenic *Xenopus tropicalis* tails over a regenerative timecourse. Using both ATAC-Seq and scRNA-Seq, we show that by 24 hours post amputation (hpa), *pax6*+ neural progenitors place a high priority on differentiation to motor and interneuron subtypes, and only later at 72hpa do they accessibilize genes associated with proliferation and self-renewal. By intersecting transcriptional and chromatin accessibility data, we identified candidate transcriptional regulators associated with both early and later changes in this neural progenitor gene expression profile. Further analysis of two of these, Pbx3 and Meis1, revealed similar loss-of-function phenotypes, including reduced tail regeneration and disorganization of regenerated neural tissues. Taken together, we present an integrated transcriptional regulatory analysis of regeneration in *pax6*-expressing NPCs that highlights both parallels and distinctions relative to embryonic development, and present novel candidates regulating regeneration in this critical cell type.

3.2 RESULTS

3.2.1 *Transgenic pax6:GFP is expressed in NPCs during regeneration*

We selected the *Xtr.Tg(pax6:GFP;cryga:RFP;actc1:RFP)* transgenic line (hereafter abbreviated *pax6:GFP)* to track NPCs over the course of regeneration (Hartley et al., 2001; Hirsch et al., 2002). *Pax6* is a highly conserved paired box transcription factor essential for central nervous system development (Bel-Vialar et al., 2007; Osumi et al., 2008; Walther and Gruss, 1991). Because the expression pattern of this line has primarily been characterized earlier in development, we first wanted to confirm that GFP was found in a pattern consistent with *pax6* expression and NPC distribution (Ericson et al., 1997; Hartley et al., 2001). To do so, we imaged whole mount tadpoles to identify where the *pax6:GFP+* cells were localized anteroposteriorly in a stage 41 tadpole, finding that *pax6* is expressed in the forebrain, hindbrain, eye, and spinal cord (**Error! Reference source not found.**A). In lateral confocal images of whole-mount transgenic tails, we found that the GFP expression domain spanned the same dorsal ventral domain as the spinal cord (**Error! Reference source not found.**B). Transverse sections of a stage 41 tadpole also revealed a broad dorsoventral *pax6+* domain surrounding the central canal of the spinal cord (**Error! Reference source not found.**C). GFP expression in transgenic tadpoles included much of the same domain as the NSC marker Sox2 (**Error! Reference source not found.**D). Notably, both nuclear Sox2 and cytoplasmic GFP driven by *pax6* are localized medially around the central canal, while markers associated with differentiated neurons such as Dcx and neurofilament are largely found in the spinal cord periphery (**Error! Reference source not found.**E,F). Thus we concluded that we could use GFP expression to capture a large proportion of NPCs and thereby interrogate their population-level regulatory dynamics, although we can't conclusively rule out the possibility

that some NSCs are excluded from this domain, or that some differentiated neurons may be included in GFP-expressing cells.

To better understand when NPCs enter the regenerating spinal cord, we amputated *pax6:GFP* transgenic tadpoles and followed reentry of GFP+ cells into the regenerating spinal cord tissue. Transgenic tadpoles were collected at 0 hours post amputation (0hpa), 6hpa, 24hpa, and 72hpa and imaged for reporter GFP fluorescence. We found that *pax6:GFP+* cells can be clearly identified in the regenerating tail as early as 6hpa (**Error! Reference source not found.**G). By 24hpa, the GFP+ cells are found in the bulb of the neural ampulla (Beck et al., 2009c; Stefanelli, 1951). By 72hpa, GFP+ cells are present all along the anteroposterior axis of the regenerated spinal cord. From these experiments, we concluded that *pax6:GFP+* cells are appropriately localized as they repopulate the regenerated spinal cord.

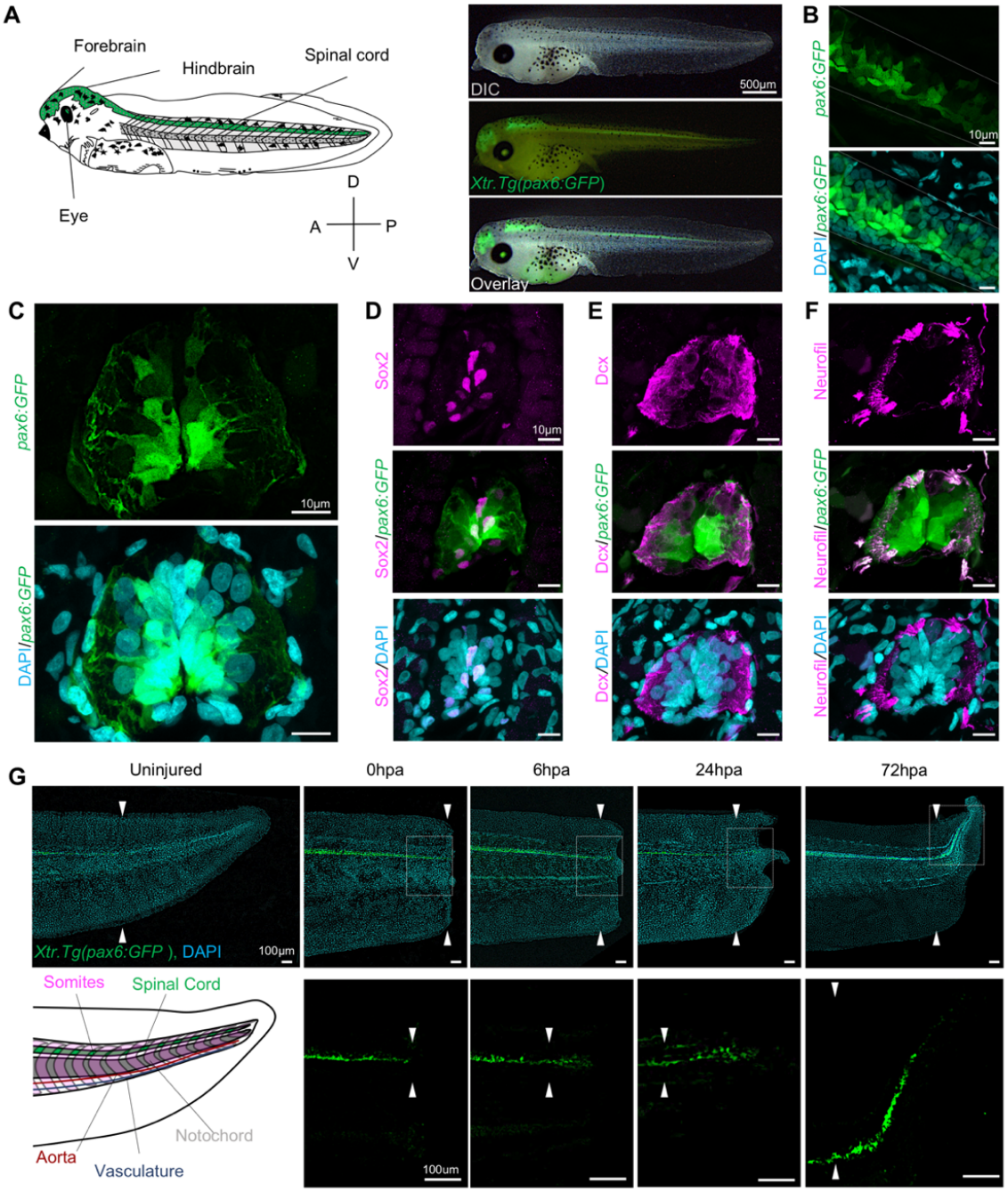

Figure 3-1 Expression dynamics of pax6:GFP in the regenerating spinal cord.
(A) Left, cartoon depiction of a NF stage 41 tadpoles with known *pax6*+ domains colored green. Right, Stage 41, transgenic tadpole expressing GFP under the control of *pax6* promoter tadpole. For all images in this figure, the green channel represents *pax6* reporter GFP fluorescence and the cyan channel represents DAPI staining. (B) Confocal image of a lateral view of a whole-mount transgenic stage 41 tadpole. (C-F) Transverse cryosections through the posterior spinal cord of a transgenic stage 41 tadpole. (D-F) Immunofluorescence images of the spinal cord sections stained with (D) anti-Sox2, (E) anti-Dcx, and (F) anti-neurofilament. (G) Regeneration timecourse of Xtr.Tg(*pax6*:*GFP*) tadpole over the first 72 hours following tail amputation. The white box in the top photos correspond with enlarged green channel below. White arrows indicate amputation plane.

3.2.2 *The chromatin accessibility profile of FACS-sorted NPCs is highly enriched for neural-specific regulatory regions*

We set out to identify the chromatin accessibility signature of NPCs over a regeneration timecourse using ATAC-Seq (Buenrostro et al., 2013). To perform this cell-type specific genomic analysis of regeneration in neural progenitors, we optimized a method to isolate *pax6:GFP+* fluorescent cells from transgenic tadpoles by flow cytometry (Figure 3-10) (see Methods for details). The posterior third of stage 41 tadpole tails were amputated and ATAC-Seq libraries were made from amputated tail tips (uninjured) or the newly regenerated tissue at 0hpa, 6hpa, 24hpa, or 72hpa (**Error! Reference source not found.**A). The timepoints were collected in two conditions. The first, "all-tissue" libraries refer to libraries made from total tail tissue, which includes all of the cell types in the tail. The second, "*pax6"* refers to libraries made from GFP+ sorted cells (Figure 3-2A). Each library was prepared from between 1200 and 4000 cells, either with no GFP gate applied and therefore containing all cell types ("all tissue") or with a GFP gate applied as in Figure 3-10C ("*pax6*"). Sample preparation details for each library are detailed in Supplementary File 1a. Quantitative RT-PCR of sorted GFP+ cells confirmed highly elevated expression of *gfp* and *pax6*, with low levels of neuronal-specific *tubb2b* or *cardiac actin*, confirming the identity of these cells with low levels of contamination (Figure 3-10D). Libraries for each timepoint and condition were made in triplicate, multiplexed and sequenced on the Illumina Next-Seq platform (see Methods for details). Read alignment and peak calling was performed using an inhouse pipeline detailed in Methods; library QC metrics are reported in Supplementary File 1b and Figure 3-11. Sequenced libraries were selected for analysis if they met minimum ENCODE standards for ATAC-Seq libraries (www.encodeproject.org/atac-seq).

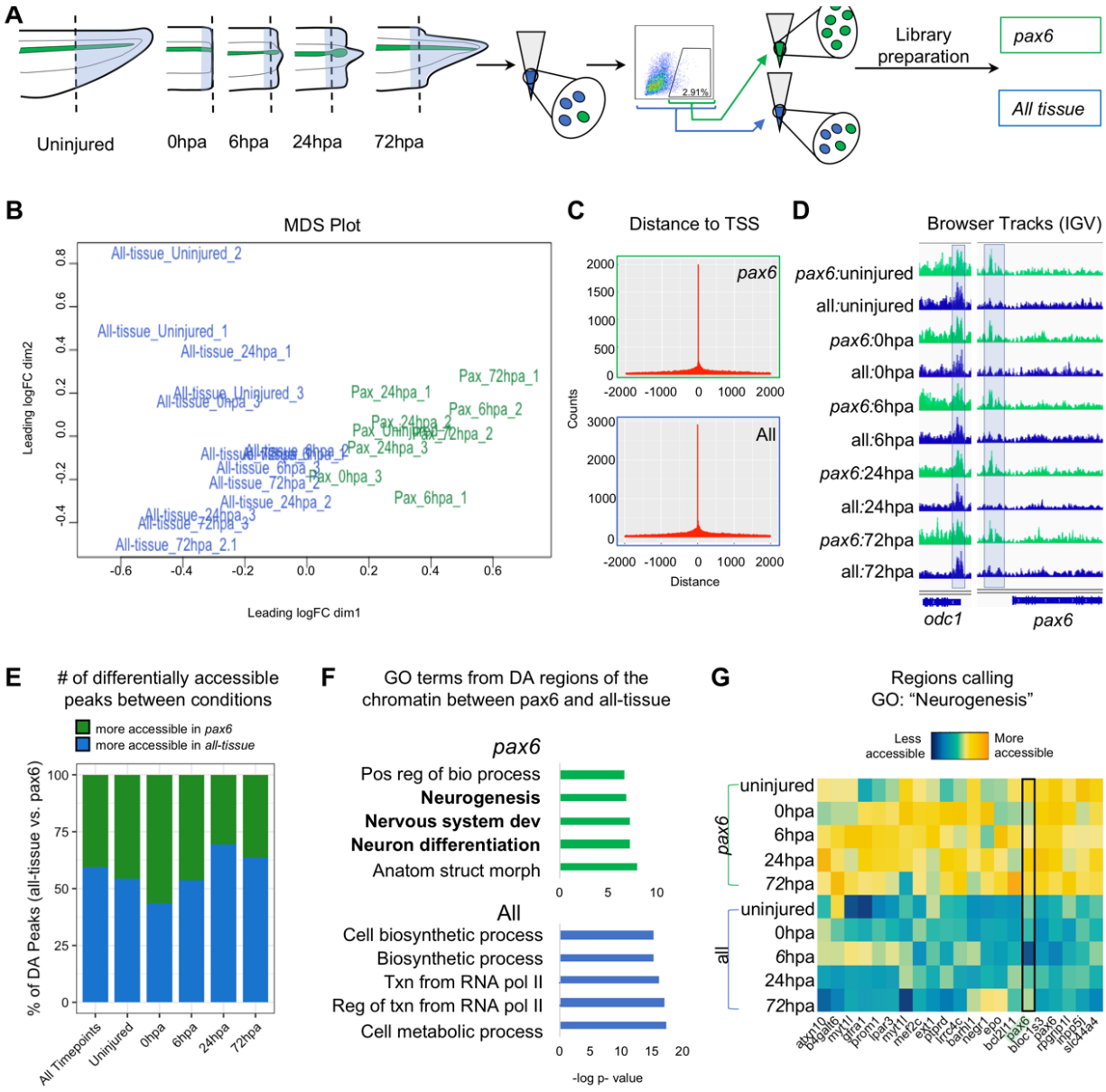

Figure 3-2 Pax6 ATAC-Seq libraries resolve accessible neural specific regions that were not identified in all-tissue libraries.

A) Experimental design for FACS isolation of reporter cells for sequencing library preparations. (B) MDS plot of sequenced *pax6* and all-tissue ATAC-Seq libraries. (C) Representative histograms showing distance of called peaks to transcription start sites (tss). (D) IGV browser tracks of a house keeping gene, *odc* and reporter line specific gene, *pax6*. Shaded bar in *odc* represents a peak retained in *pax6* and all-tissue libraries; shaded bar in *pax6* represents a peak called in *pax6* libraries, but not in all-tissue. (E) Bar plot showing percent of differentially accessible (DA) peaks that are more accessible in *pax6* or all-tissue libraries within an individual timepoint or across all timepoints. (F) Differentially accessible regions were annotated to the nearest TSS and used to call GO:BP terms. Top, GO terms from regions more accessible in *pax6* libraries than in all-tissue libraries. Bottom, GO terms from regions more accessible in all-tissue libraries than *pax6* libraries. (G) Heat map depicting accessibility of 21/254 randomly sampled peaks that call the GO term: Neurogenesis. *pax6* promoter region is boxed.

Having verified the quality of ATAC-Seq libraries made from sorted *pax6:GFP*+ cells and bulk tail tissue, we first confirmed that the overall tissue conditions had chromatin accessibility signatures that could be easily differentiated. Taking an unbiased approach, we used multi-dimensional scaling (MDS) to identify how similar or different the *pax6* and all-tissue libraries are overall. We found that the *pax6* libraries cluster away from the clustered all-tissue libraries (**Error! Reference source not found.**B). We then examined individual gene regulatory loci. The majority of peaks called were located within 500bp of a transcription start site for both all tissue and *pax6* libraries, confirming that the majority of ATAC-Seq signal is concentrated in nucleosome-depleted promoters, as expected (Buenrostro et al., 2013) (**Error! Reference source not found.**C). Both *pax6* and all-tissue libraries shared common peaks around a control housekeeping promoter (*odc*) (Dhorne-Pollet et al., 2013) without any apparent significant differences in accessibility (**Error! Reference source not found.**D, left). By contrast, upstream of the *pax6* TSS, all *pax6* library timepoints share a peak that is absent in all-tissue libraries (**Error! Reference source not found.**D, right). From this first proof-of-principle test, we concluded that ATAC-Seq libraries from sorted cells can be used to identify chromatin accessibility differences in NPCs relative to other tissues, and that ATAC-Seq of these sorted cells can be used to find tissue-specific accessible regions that are beneath the detection sensitivity of bulk analysis of all of the tissues together. We then progressed to genome-wide analysis of differences between *pax6*:*GFP*+ calls and bulk tissue.

Our next goal was to identify categories of gene regulatory regions that have increased accessibility in NPCs relative to bulk tail tissue. To this end, we interrogated the global NPC chromatin signature by compiling all *pax6* library peaks into one peak file and all of the all-tissue peaks into a second peak file and calling differentially accessible peaks between these

two compiled files. We identified 3604 regions that were more accessible in the *pax6* libraries than the all-tissue libraries, which we assigned to the nearest TSS, and used the associated gene names as input for gene ontology (GO) analysis using gProfileR2 (Raudvere et al., 2019) (**Error! Reference source not found.**E-F). Genes with increased accessibility in *pax6*:*GFP*+ cells showed statistically significant enrichment for neural-associated terms such as "Neurogenesis", "Nervous system development", and "Neuron differentiation," confirming the neural progenitor identity of our *pax6:GFP* cells (**Error! Reference source not found.**F, top, and Supplementary File 1c). We visually confirmed the increased accessibility of regions calling the term "Neurogenesis" with a heat map (**Error! Reference source not found.**G). The reverse analysis identified 5251 regions that are more accessible in all-tissue libraries than *pax6* libraries, which are associated with broad GO categories relating to ubiquitous cellular processes (i.e. "Cellular biosynthetic process", "Transcription from RNA Polymerase II", and Cellular metabolic process") (**Error! Reference source not found.**F, bottom, Supplementary File 1d). From this analysis, we concluded that the overall chromatin accessibility of NPCs is readily distinguished from bulk tissue and strongly reflects their distinct neural character.

3.2.3 *Analysis of differentially accessible peaks between pax6 libraries reveals an early prioritization of neuronal differentiation in regeneration*

Having established that the overall chromatin signature of NPCs could be used to identify neural-specific processes, we next identified the cellular processes being prioritized by NPCs at discrete regenerative stages. We therefore systematically identified differentially accessible regions of the chromatin that were uniquely accessible at each timepoint relative to

the flanking timepoints and performed GO analysis on the genes neighboring these regions (**Error! Reference source not found.**A-C). GO analysis returned ~30-100 terms, thus, to distill the list to main processes, we used ReviGO's semantics algorithm to reduce redundancy of the GO terms (Supek et al., 2011). Supplementary Files 1e-g contain the entire list of GO terms returned from gProfiler and Supplementary Files 1h-j contain the results from ReviGO.

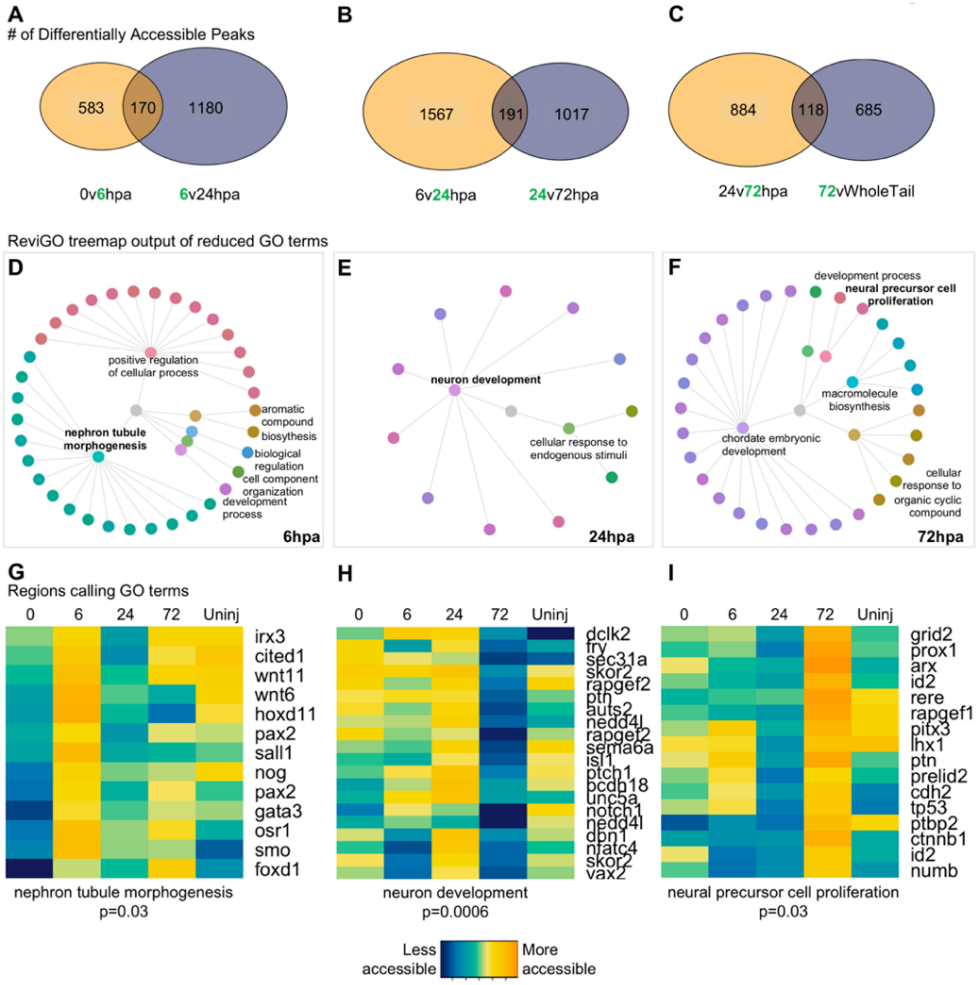

Figure 3-3 Differential accessibility analysis of pax6 libraries over regenerative time reveals chromatin accessibility prioritizes first tubule morphogenesis, followed by neural differentiation, and later proliferation.
From left to right, data represents the 6hpa timepoint, 24hpa timepoint, and 72hpa timepoint. (A-C) Venn diagrams depicting number of unique and shared peaks called in each of the timepoint contrasts within the *pax6* libraries. The intersect represents peaks that were both becoming accessible at a given timepoint, then losing accessibility. The green number in each contrast represents the timepoint that is more accessible. (D-F) All regions represented in (A-C) were used for GO:BP analysis with gProfileR2. ReviGO was then used to reduce redundancy in output GO terms and identify main families of terms. Tree graphs depict the hierarchy of reduced GO list. The central grey circle is the vertex of the graph

representing the set of all GO terms included in the data, the second level of circles are the families of GO terms called by ReviGO, and the third level of circles represents GO terms in each family. Terms provided are the family names. (G-I) Heatmap of accessibility of regions that were used to call GO terms: "Nephron Tubule Morphogenesis", "Neuron Differentiation", and "Neural Precursor Cell Proliferation".
We first examined genes associated with regulatory regions that are uniquely accessible at 6hpa. Genes preferentially accessible at 6hpa called six families of GO terms (**Error! Reference source not found.**D).

Notably, "Nephron Tubule Morphogenesis" was well represented, and included regulatory regions from *foxd1*, *gata3*, *smo, sall1, osr1, pax2, stat1, hoxd11, wnt6, nog, wnt11, irx3, cited1* (**Error! Reference source not found.**G). In addition to patterning the nephron, the majority of these genes are expressed in neural stem cells during dorsoventral neural patterning of the spinal cord and early differentiation (Alaynick et al., 2011; Delile et al., 2019). At this time in regeneration, we observe the spinal cord repairing itself by closing the neural tube and forming the neural ampulla. We have shown in **Error! Reference source not found.**D that pax6:*GFP*+ cells populate the regenerated neural ampulla, therefore, we hypothesize that these cells represent ependymoglial cells with NSC character that have been described previously (Beck et al., 2009c; Chernoff et al., 2018; McHedlishvili et al., 2007; Muñoz et al., 2015; O'Hara et al., 1992; Reimer et al., 2009).

We then identified regions that were most accessible at the 24hpa timepoint. This analysis recovered two main families of GO terms, the largest of which was "Neuron Development" (**Error! Reference source not found.**E). Strikingly, the majority of genes that call this GO term become more accessible from 6 to 24hpa, but lose that accessibility at 72hpa (**Error! Reference source not found.**H). Genes with preferentially accessible regulatory regions at 24hpa include factors known to function in multiple aspects of neuronal differentiation and growth. Examples of these include neuritogenesis and growth cone formation (*dclk2, rapgef2, auts2, nedd4l, unc5a*), axonogenesis and axon guidance (*dbn1, sema6a, fry, ptch1*), neuronal migration (*dclk2, rapgef2, auts2, pcdh18*), and transcriptional regulation of neuron

differentiation (*skor2, isl1, nfatc4, vax2*). The enrichment of these terms, and their specificity for functions carried out by differentiating neurons rather than proliferating progenitors, agrees well with previous published bulk RNA-Seq analysis (Chang et al., 2017) showing enrichment for similar neuronal morphogenesis terms, such as axonogenesis and dendritogenesis, in the first day post amputation.

Finally, we assessed the processes being prioritized at 72hpa. This analysis resulted in 5 main GO families being called (**Error! Reference source not found.**F), including "neuronal precursor cell proliferation". Notably, cell cycle and proliferation related terms were absent at the 6hpa and 24hpa timepoints. The regions calling this term were all more accessible at 72hpa than the 24hpa timepoint, however about half were not differentially accessible with respect to the uninjured tail (**Error! Reference source not found.**I), suggesting that by 72hpa, these regions are regaining similarity to their uninjured state. Altogether, we were able to resolve processes specific to neural progenitor cell function at each timepoint. These suggest that NPCs first prioritize ependymal tube morphogenesis at 6hpa, neural differentiation at 24hpa, and cell proliferation at 72hpa.

3.2.4 *Comparative analysis of uninjured and 24hpa tail single cell RNA-Seq data sets reveals expansion in differentiated neuronal clusters at 24hpa*

Our ATAC-Seq analysis suggested that neural progenitor cells may place a high transcriptional regulatory priority on neuronal differentiation early in regeneration, at 24hpa. To determine whether this was reflected in the transcriptional profile and cell-type composition of neural cells, we used single-cell RNA-Seq to interrogate the profile of neural cell types present before and 24 hours after amputation. We sequenced the transcriptomes of

individual NPCs from uninjured and regenerating tails, an analogous population to those collected for ATAC-seq. In total, we sequenced 2,617 and 1,090 cells from uninjured and 24hpa tissues, respectively. Using Seurat we aligned the two data sets with their integrated analyses and used UMAP dimensional reduction to find clusters (Becht et al., 2018; Butler et al., 2018; Stuart et al., 2019). From this analysis, we revealed 19 unique clusters (**Figure 3-12**A), which were enriched for but not fully restricted to the neural lineage. A total of 781 uninjured cells and 296 cells from 24hpa cells fell into neural clusters, which comprised NPC/NSCs (defined by expression of *sox2*), transitioning neurons (*neurog1; neurod1*), and differentiated neurons (*elavl4*) (**Figure 3-12**B). To gain the highest-possible resolution of cell types within these broad neural clusters, we subsetted neural cells, re-performed dimensionality reduction and called 8 unique sub-clusters, including spinal cord progenitors, floor plate progenitors, differentiating neurons, interneurons, motor neurons, vulnerable motor neurons, motor neurons (*leptin*+), and dopaminergic neurons (**Error! Reference source not found.**A). These clusters agreed well with the neural cell types recently identified for tail regeneration in *X. laevis,* and we applied the same naming conventions (Aztekin et al., 2019).

Several landmark insights emerged from this single-cell analysis of neural lineages in the tadpole tail. First, we were able to identify new candidate molecular markers for progenitors (*vwcx2l.2, iqca1, daw1*), floor plate progenitors (*bicc, vasn*), differentiating neurons *(nhlh1, tmem178.2),* motor neurons *(prdm14, prdm1, kcnb2),* leptin+ motor neurons *(abcb1, asns, lepr),* vulnerable motor neurons (*mmp17, slc18a3, colq-like),* interneurons *(cnrip1, slc17a7),* and dopaminergic neurons *(pkd2l1, pkd1l2, slc32a1)* (**Error! Reference source not found.**B). The "vulnerable motor neurons" term refers to vulnerability of specific neurons to degeneration in ALS (Kline et al., 2017). Representative genes for each

cell type were identified by using differential expression analysis between cell clusters. Second, we found that cell type identities could be clearly assigned to these clusters at 24 hours after injury (**Error! Reference source not found.**C). Unlike the connective tissues in the axolotl limb, which lose cell heterogeneity of gene expression after injury (Gerber et al., 2018), neural cells in the 24hpa *Xenopus* spinal cord maintain heterogeneous gene expression.

We next used our scRNA-Seq data to interrogate the hypothesis that neural lineages put an early priority on differentiation rather than proliferation. To address this question and to identify which neuronal types, if any, are being produced, we asked how the relative quantities of cell populations change following injury. To this end, we identified what percent of the total neural population is made up of each cluster under each condition. If neural lineages prioritize differentiation early in regeneration, we would expect to see a decline in proliferating progenitor cells and an increase in one or more differentiated neuronal types. Indeed, at 24hpa, we observe a decrease in the percent of NPCs and differentiating neurons with respect to the uninjured cells (**Error! Reference source not found.**D-G). In agreement with our GO analysis of 24hpa ATAC-Seq data, we see an expansion in differentiated neuronal clusters including interneurons, vulnerable motor neurons, and motor neurons (*leptin*+) at 24hpa. Overall, NSCs are reduced from 38% to 25%, differentiating neurons are reduced from 20% to 10%, and differentiated neurons increase from 42% to 65%. As previously described, we noted that motor neurons (*leptin*+) are virtually absent in the uninjured tail but are abundant at 24hpa (Aztekin et al., 2019). These transitions in cell type composition are consistent with the hypothesis that neuronal

differentiation, especially to specific motor neuron types, is prioritized early in regeneration, with a decline in the relative abundance of proliferative cells.

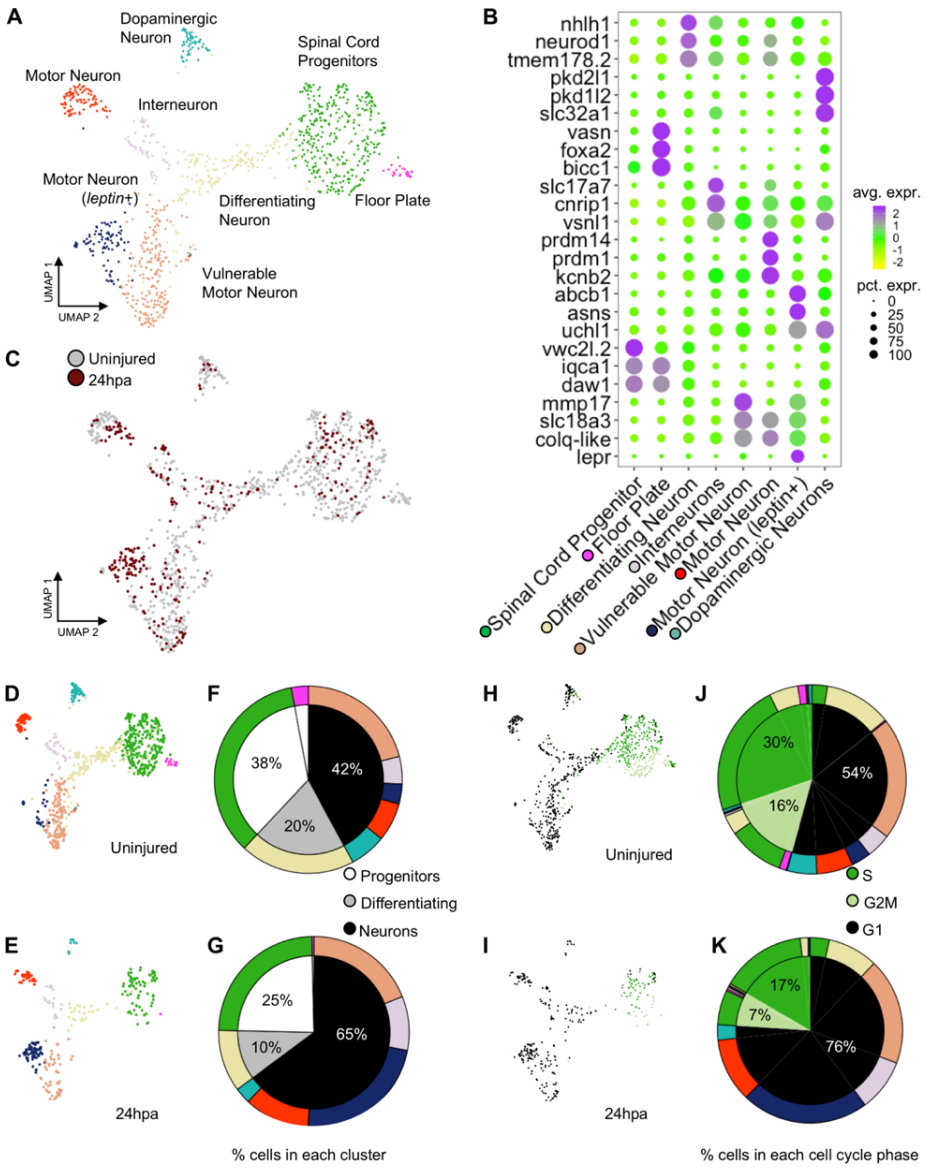

Figure 3-4 scRNA-Seq of uninjured and 24hpa tails reveals a transcriptional shift to differentiated neuronal types at 24hpa.
(A) UMAP projection of integrated uninjured and 24hpa neural lineage cells. 7 distinct clusters were identified. (B) Dot plot of genes identified as differentially expressed between cell clusters. The top 3 genes exclusively expressed in each of the 7 clusters are shown. Color of circle denotes average gene expression across one cluster and size of circle represents percent of cells in each cluster expressing each gene. (C) UMAP projection of neural cells colored by timepoint in the uninjured tails and 24hpa tails. UMAP projections of cell clusters split by timepoint to (D) uninjured and (E) 24hpa. (F-G) Sunburst diagrams showing (inside) the percentage of neural cells that are broadly stem cells, differentiating cells, or differentiated neurons and (outside) percentage of neural cells in each of the seven cell clusters called. Colors on outside of the sunburst chart correlate to cluster colors in (A/B). (H-I) UMAP projections of

neural clusters colored by predicted cell cycle phase in the (H) uninjured tail and (I) 24hpa tail. (J-K) Sunburst diagrams showing the percentage of all neural cells predicted to be in each cell cycle phase (inside) and the percentage of each cell cluster predicted to be in each cell cycle phase (outside). (J) represents the distribution in the uninjured tail and (K) represents the distribution in the 24hpa tail. Colors on outside of the sunburst chart correlate to cluster colors in (A/B).

Our hypothesis is that neuronal differentiation is prioritized early in regeneration while proliferation is decreased. To further test this hypothesis, we asked if we could see a shift in the transcriptomically defined cell cycle state accompanying the change in cell type representation. To this end we used the Seurat cell cycle phase predictor to predict the cell cycle phase of each cell in the neural lineage. In both the uninjured tail and at 24hpa, the NPCs are primarily predicted to be in S and G2M phases, the differentiating neurons are a mix of G2M and G1 phase, and the subtypes of differentiated neurons are primarily made up of G1 phase (**Error! Reference source not found.**H/I). The uninjured tail neural lineage has 30% cells predicted in S phase, 16% cells predicted in G2M phase, and 54% cells predicted in G1 phase (**Error! Reference source not found.**H/J). This distribution shifts markedly at the 24hpa timepoint, where the S phase falls to 17%, G2M falls to 7% and G1 increases to 76% (**Error! Reference source not found.**I/K). Thus, at the 24hpa timepoint, there is a relative increase in the proportion of cells predicted to be in G1 phase and decrease in cells in G2/M/S phases. This correlates with the relative increase in differentiated neurons and decrease in spinal cord progenitors and differentiating cells and supports the hypothesis that NPCs are undergoing differentiation.

We validated expression of several differentiated neuronal genes by *in situ* hybridization (ISH). We examined *l1cam*, *uchl1*, *nsg1* and *ass1* as candidate neuronal differentiation markers. All these genes are strongly expressed in one or more differentiated neuronal populations but are lowly expressed or absent in NPCs and non-neural cell types by scRNA-SEQ (**Error! Reference source not found.**A-C, G-I, M-O, S-U). When examined by ISH all four genes are most strongly

and distinctly expressed in the neural ampulla at 24hpa (**Error! Reference source not found.**E, K, Q, W). As a further test of the hypothesis that proliferation is not prioritized until 72hpa, we used phospho-histone H3 (PH3) staining to quantify mitotic cells in optical longitudinal sections of the spinal cord that overlap with the *pax6* domain, finding that proliferative cells are significantly more abundant at 72hpa relative to 24hpa (**Error! Reference source not found.**Y-AA). These findings further supported our model that differentiation is prioritized at 24hpa, and proliferation at 72hpa.

3.2.5 *Gene regulatory network prediction reveals Pbx3 and Meis1 as candidate regulators of neuronal regeneration*

To better understand the factors governing NPCs fate decisions, our next goal was to identify transcription factors that could be directing fate transitions. We therefore integrated our ATAC-Seq data, scRNA-Seq data, and used bulk RNA-Seq data to predict gene regulatory networks occurring over regenerative time. We used HOMER to identify transcription factor (TF) binding sites in differentially accessible peaks associated with called GO terms (**Error! Reference source not found.**A-C) (Heinz et al., 2010). We then used our previously published bulk-RNA-Seq data (Chang et al., 2017) to confirm each HOMER-identified transcription factor and its targets were expressed at the timepoint of interest. Next we restricted our analysis to differentially accessible regions of the chromatin where TF binding sites were accompanied by differential gene expression in the single-cell RNA-Seq. Differential gene expression was performed between the uninjured and 24hpa timepoints in scRNA-Seq following parameters detailed in Methods. These stringent criteria revealed two regulatory circuits spanning from 6hpa to 72hpa. One circuit predicts Meis1 as a regulatory

factor that binds motifs near *runx2* and *etv1* at 24hpa, and Etv1 as a regulatory factor that binds motifs near *pbx3* at 24hpa (**Error! Reference source not found.**A). Each of these four transcription factors also has numerous neuronal differentiation genes among its predicted targets. In our scRNA-Seq dataset, *etv1, meis1, pbx3, and runx2* are all expressed in the neural cells and are therefore reasonable candidates to regulate neural cell fate transitions. Of these, *pbx3* was the most restricted to the neural lineage (**Figure 3-12**E, G), and was found in several neuronal cell types. Pbx3 and Meis1 are homeodomain transcription factors thought to coregulate targets in B-cell leukemia and proximal-distal patterning of the regenerating axolotl limb (Li et al., 2016; Mercader et al., 2005). In *Xenopus*, Meis1 is well-studied for its roles in neural crest and posterior neural fate specification, working together with Pbx1 (Kelly et al., 2006; Maeda et al., 2001, 2002; Salzberg et al., 1999). Neither *pbx3* nor *meis1* have been described in neural regeneration, and so we decided to probe the function of these genes in neural development and regeneration in *Xenopus*.

We assessed target regions and differential accessibility for Pbx3 and Meis1 in the ATAC-Seq data, and gene expression in the scRNA-Seq data. Considered across all regenerative timepoints, Pbx3 and Meis1 have both independent and overlapping targets. A large fraction (26/61) of Pbx3 targets are shared with Meis1 (**Error! Reference source not found.**B). Meis1's target peaks are preferentially increased in accessibility at 24hpa, while Pbx3's are overall increased at 72hpa. Both transcription factors have numerous neuronal differentiation factors among their targets (**Error! Reference source not found.**C, D). When we examined expression of *meis1* and *pbx3* in our scRNA-Seq data, we found that transcripts for both factors were detected in multiple differentiated neuronal types, initially at low levels and in a small proportion of cells (**Error! Reference source not found.**E, and white portions

of left pie charts in F). Following injury, at 24hpa, expression of *meis1* increased to 30% of all neural cells (**Error! Reference source not found.**F, upper right), driven by increases in interneurons (IN), dopaminergic neuron (DN), and especially motor neuron (MN) clusters (Figure 6E, top). At 24hpa expression of *pbx3* increased to 11% of all neural cells (Figure 6F bottom right), driven by a strong increase in motor neuron expression (Figure 6E, bottom). Both factors are therefore expressed in cell types likely to be completing differentiation at 24hpa, most notably the motor neurons.

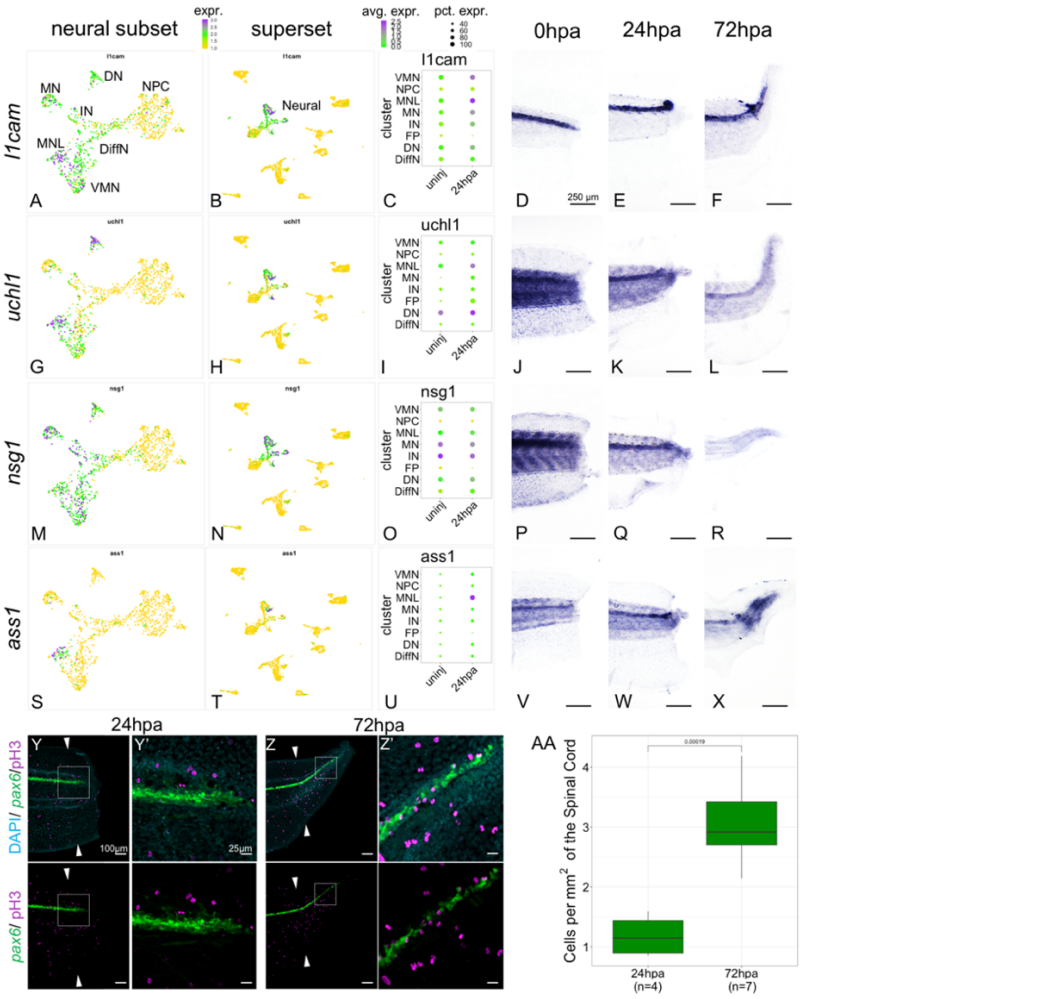

Figure 3-5 Markers of neuronal differentiation are increased at 24hpa and proliferation increases at 72hpa.
UMAP plots of expression in the neural restricted subset of scRNA-Seq for (A) *l1cam*, (G) *uchl1*, (M) *nsg1*, and (s) *ass1*. UMAP plots of expression in the superset scRNA-Seq dataset for (B) *l1cam*, (H) *uchl1*, (N) *nsg1*, and (T) *ass1*. Note that the topology of neural cell clusters is slightly different when

subset and reclustered to give high neural cluster resolution (A,G,M,S) than when neural cells are analyzed together with all other cell types (E,K, Q, W), but the same cells and expression data are included in both cases. Dot plots representing average expression of (C) *l1cam*, (I) *uchl1*, (O) *nsg1*, and (U) *ass1* per cluster and timepoint. *In situ* hybridization at 0hpa, 24hpa, and 72hpa for (D-F) *l1cam*, (J-L) *uchl1*, (P-R) *nsg1*, and (V-X) *ass1*. (Y-Z') Representative images of 24hpa (Y/Y') and 72hpa (Z/Z') regenerated tails from a *pax6:GFP* transgenic tadpole stained for DAPI (cyan) and mitotic cells with phospho-histone3 (magenta). The white arrows indicate regeneration plane. (Y') and (Z') are enlarged images of the boxed areas in (Y) and (Z). (AA) Boxplot representing the number of cells per regenerated spinal cord area (mm^2) at 24hpa and 72hpa. Statistics represent t-test performed between the two timepoints. Abbreviations: Neural Progenitor Cell (NPC), Differentiating Neuron (Diff), Interneuron (IN), Vulnerable Motor Neuron (VMN), Dopaminergic Neuron (DN), Motor Neuron Leptin+ (MNL), Motor Neuron (MN)

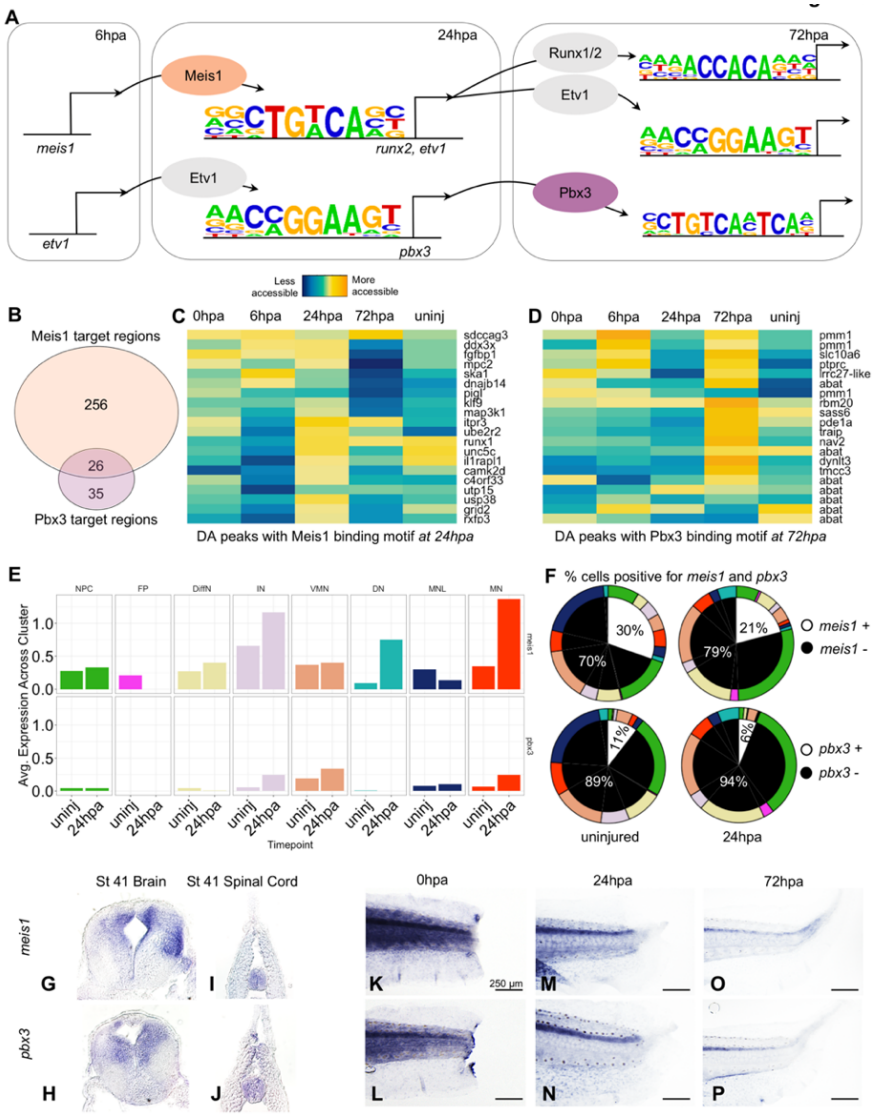

Figure 3-6 Gene regulatory network prediction identifies Meis1 and Pbx3 as key regulators of neural regeneration.

(A) Predicted gene regulatory networks across regenerative time were derived from integrative analysis between ATAC-Seq and RNA-Seq data (See analysis in Methods). Meis1 and Pbx3 emerged as candidate regulators of regeneration. (B) Venn diagram showing the number of Meis1 and Pbx3 target regions found in differentially accessible regions of the chromatin. The overlap represents regions where both Meis1 and Pbx3 binding motifs were found. (C, D) Heatmaps showing accessibility of 20 differentially accessible regions of the chromatin identified with binding sites for (C) Meis1 or (D) Pbx3. (E) Average gene expression of meis1 and pbx3 in each neural cell cluster in the scRNA-Seq data in the uninjured and 24hpa timepoints. (Abbreviations: Neural Progenitor Cell (NPC), Differentiating Neuron(DiffN), Interneuron (IN), Vulnerable Motor Neuron (VMN), Dopaminergic Neuron (DN), Motor Neuron Leptin+ (MNL), Motor Neuron (MN)). (F) Sunburst diagram showing the percentage of neural cells that are either positive or negative for meis1 or pbx3 gene expression (inside) and the percentage of each neural cluster positive or negative for the genes (outside). (G-P) In situ hybridizations for meis1 and pbx3 in transverse sections of the head (G, H) and tail (I, J) at stage 41. In situ hybridizations for meis1 and pbx3 at (K, L) 0hpa, (M, N) 24hpa, and (O, P) 72hpa.

Previously, spatial expression patterns for *meis1* have only been reported up to stage 29 (Maeda et al., 2002) and expression patterns for *pbx3* have not been reported in *Xenopus*. Therefore, we used *in situ* hybridization to characterize their expression in the tadpole. At stage 41, both *meis1* and *pbx3* are expressed in the uninjured brain and spinal cord (**Error! Reference source not found.**G-J). In the posterior tail, expression of both factors is weak and diffuse immediately following injury (**Error! Reference source not found.**K, L). However, by 24hpa, expression of both factors increases in the spinal cord (**Error! Reference source not found.**M,N). By 72hpa, expression persists more weakly in the spinal cord. These analyses therefore indicate that *pbx3* and *meis1* are expressed in the spinal cord, increase in neuronal subtypes at 24hpa, and contribute to increased promoter accessibility at their target genes at 24hpa (particularly for Meis1) and 72hpa (particularly for Pbx3), some of which are likely co-regulated by both factors.

3.2.6 *Meis1 and Pbx3 are necessary for successful spinal cord and tail regeneration*

We next asked whether Meis1 and Pbx3 are independently required for regeneration. To do so, we first injected morpholinos blocking translation of either *meis1* or *pbx3* into the dorsally fated blastomeres of 4-cell stage embryos to knockdown expression in the neural lineage (Figure 3-13B). Knockdown of either Meis1 or Pbx3 resulted in a similar phenotype in stage 41 tadpoles that included small or missing eyes, a shorter anteroposterior body axis and reduced pigment cell population (**Figure 3-13**C, D). We followed up on these phenotypes by repeating the injections with a second set of morpholinos against Pbx3 and Meis1. The second set of morpholinos phenocopy the original morpholino injections (**Figure 3-13**E, F). As a further confirmation, we designed and injected CRISPR guide RNAs targeting each of these genes together with Cas9 protein at the 4-cell stage, and were able to phenocopy the

morpholino effects with this strategy as well (**Figure 3-13**G-J). When we stained stage 41 morphants with anti-neurofilament antibody, we found that their spinal cord and intersomitic axons were mispatterned, suggesting Meis1 and Pbx3 are necessary for proper development of the nervous system (Figure 3-14B-F). To bypass the early embryonic effects of Meis1 and Pbx3 knockdown, we designed the same MO sequences as tissue-permeable vivo-MOs. We injected either Pbx3 or Meis1 vivo-MO into the embryonic tail vein at the late tailbud stage (NF 35), and reared these embryos to stage 41 (**Error! Reference source not found.**A). Neither vivo-MO caused gross morphological defects when delivered this way, but both vivo-MOs resulted in reduced axons in the spinal cord as assayed by neurofilament stains at stage 41 (**Error! Reference source not found.**C, D), relative to control tadpoles injected with tracer only (**Error! Reference source not found.**B).

To specifically assay the effect of Meis1 or Pbx3 knockdown on regeneration, we injected either Meis1 or Pbx3 vivo-MOs at stage 35 as described above, reared these embryos to stage 41 and amputated tails. At 72hpa, both Pbx3 and Meis1 knockdown tadpoles have reduced and disorganized neurofilament staining relative to tracer-injected controls (**Error! Reference source not found.**F,G, compare with E). We quantified the length of the regenerated tail in control and knockdown tadpoles at 72hpa, and found that both Meis1 and Pbx3 knockdown tadpoles have significantly shorter regenerated tails and spinal cords relative to tracer-injected control tadpoles (**Error! Reference source not found.**K). Spinal cord area (measured in confocal optical longitudinal sections) is also significantly reduced in both morphants at 72hpa (**Error! Reference source not found.**L), although the density of PH3 positive nuclei is not significantly changed relative to controls (**Error! Reference source not found.**H-J, M). Embryos injected with either Meis1 or Pbx3

MOs at the 4 cell stage also caused neurofilament defects during regeneration, as well as shortened regenerated tails and spinal cords (**Figure 3-14**G-AA), although, we cannot rule out that some aspects of these phenotypes may be attributable to embryonic growth and patterning defects in the starting tail tissue. We conclude that Meis1 and Pbx3 are not only required for proper neural development but also for proper regeneration of the tail and spinal cord.

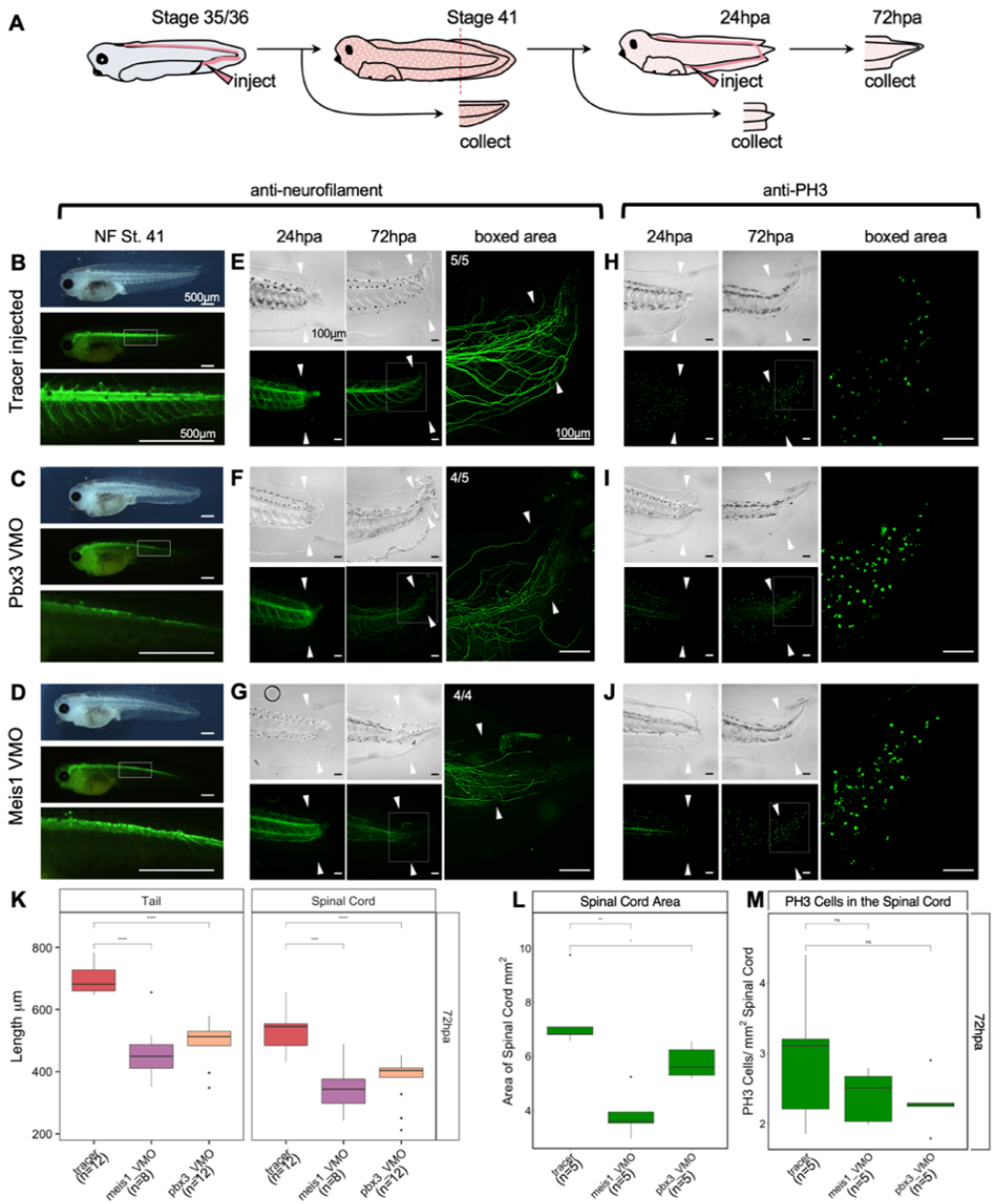

Figure 3-7 Pbx3 and Meis1 are independently required for successful regeneration of neural tissues and tails in response to injury.

(A) Injection scheme for administering vivo-morpholino. Stage 35 tadpoles were injected with a tracer or vivo-morpholinos (VMO) targeting *meis1* or *pbx3* and allowed to grow 24 hours to stage 41. Stage 41 tadpoles were amputated and 24hpa regenerates were collected. (B-D) Stage 41, whole-mount tadpoles shown in brightfield and immunostained against neurofilament. The box in the middle image corresponds to the enlarged image below. These images are shown for (B) tracer injected, (C) Meis1 VMO injected, and (D) Pbx3 VMO injected tadpoles. (E-J) 24hpa and 72hpa regenerates were collected and stained for neurofilament (E-G) or PH3 (H-J). Top photos are in DIC and bottom photos are immunostained as indicated. White arrows indicate amputation plane. These images were collected for (E, H) tracer, (F, I) Meis1 VMO, and (G, J) Pbx3 VMO. (K) Regenerated tail and spinal cord lengths were measured and reported in boxplots. (L) Boxplot representing regenerated spinal cord area. (M) Boxplot representing PH3 cells per regenerated spinal cord area. Statistics represent a two-tailed t-test to determine significance between conditions. (ns= not significant, *<0.05, **<0.005, ***<0.0005, ****<0.00005).

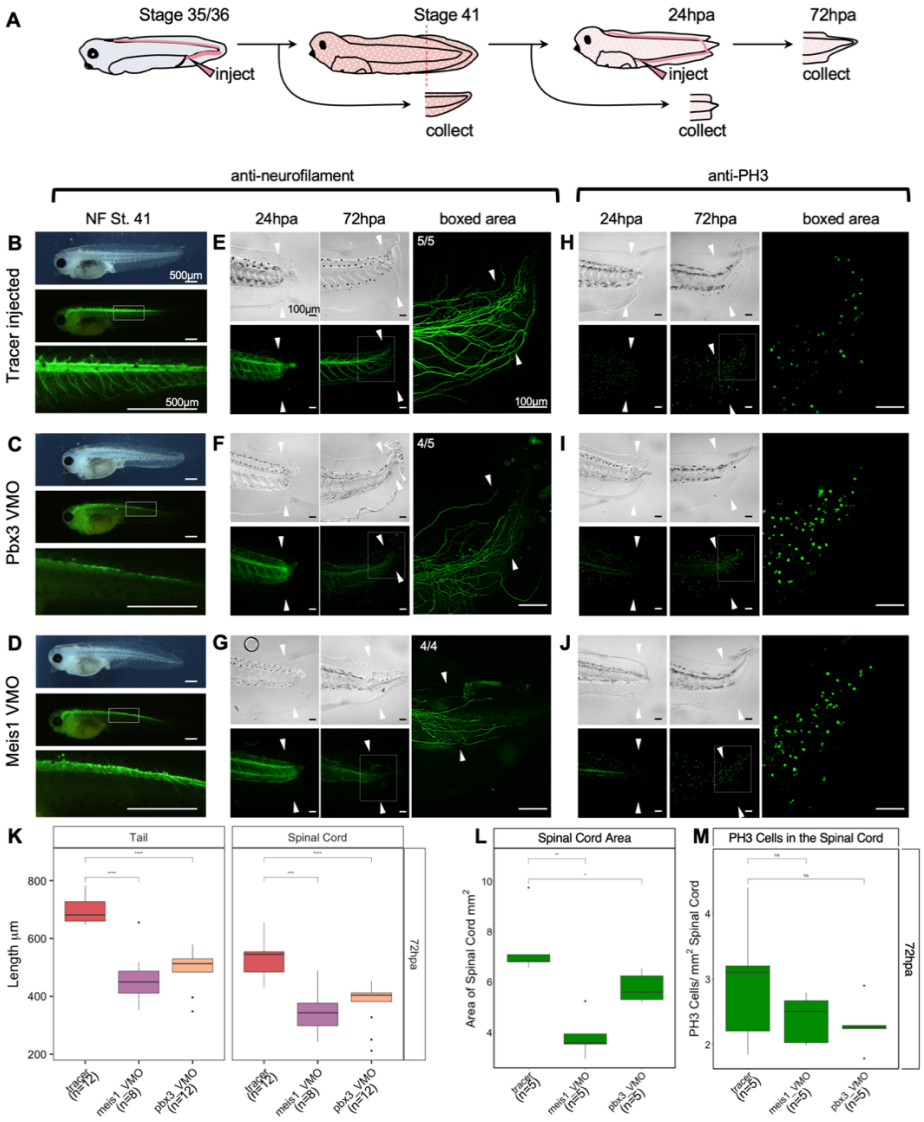

Figure 3-8 Pbx3 and Meis1 are independently required for successful regeneration of neural tissues and tails in response to injury.

(A) Injection scheme for administering vivo-morpholino. Stage 35 tadpoles were injected with a tracer or vivo-morpholinos (VMO) targeting *meis1* or *pbx3* and allowed to grow 24 hours to stage 41. Stage 41 tadpoles were amputated and 24hpa regenerates were collected. (B-D) Stage 41, whole-mount tadpoles shown in brightfield and immunostained against neurofilament. The box in the middle image corresponds to the enlarged image below. These images are shown for (B) tracer injected, (C) Meis1 VMO injected, and (D) Pbx3 VMO injected tadpoles. (E-J) 24hpa and 72hpa regenerates were collected and stained for neurofilament (E-G) or PH3 (H-J). Top photos are in DIC and bottom photos are immunostained as indicated. White arrows indicate amputation plane. These images were collected for (E, H) tracer, (F, I) Meis1 VMO, and (G, J) Pbx3 VMO. (K) Regenerated tail and spinal cord lengths were measured and reported in boxplots. (L) Boxplot representing regenerated spinal cord area. (M) Boxplot representing PH3 cells per regenerated spinal cord area. Statistics represent a two-tailed t-test to determine significance between conditions. (ns= not significant, *<0.05, **<0.005, ***<0.0005, ****<0.00005).

3.3 DISCUSSION

3.3.1 *Neural lineage-specific analysis of regeneration identifies new regulatory factors and target genes*

The full regeneration of a lost structure requires integrated gene regulatory decisions by multiple cell types. These have been documented through genome-wide transcriptomic analyses of the entire regenerating structure, such as the *Xenopus* tail (Chang et al., 2017; Love et al., 2011), spinal cord (Lee-Liu et al., 2014), or axolotl limb (Dwaraka et al., 2018). However, in bulk analyses it remains uncertain which cell types are up- or down-regulating particular genes, and increased expression in one cell type may be canceled out by decreases in another. These concerns are equally problematic in bulk-tissue analysis of gene regulatory dynamics (through methods such as ChIP-Seq or ATAC-Seq). In some cases, this problem can be addressed by explanting the tissue of interest (Chung et al., 2014), but for rare cell types or morphologically complex structures this is not a tenable solution. Here we present a major advance for vertebrate regeneration by using flow-cytometry and scRNA-Seq to coordinately study chromatin accessibility and gene expression in neural progenitors during regeneration, a critical cell type for spinal cord regeneration (Fei et al., 2014; Gaete et al., 2012; Muñoz et al., 2015). By isolating neural progenitors, we are able to identify a unique chromatin landscape for these cells, including regulatory peaks that cannot be detected in bulk tissue analysis.

It is important to note the potential limitations of our analysis as well. While our FACS approach utilizing the *pax6:GFP* transgenic line captures a broad domain of neural progenitor cells, we recognize that not every NPC may have been captured. For example, there may be some Sox2+ NPCs that are *pax6*-, and these cells may contribute to

regeneration in a distinct way that we did not capture. In particular, our confocal sections (**Error! Reference source not found.**) suggest that the *pax6:GFP*+ domain captures a large window of progenitors along the dorsoventral axis but does not capture the dorsal most Sox2+ NPCs. Future analyses using other NPC reporters, as well as domain-specific reporters for region-specific subtypes of both neural progenitors and differentiated neurons, would be necessary to fully round out the picture of what transcriptional priorities are shared among progenitor subtypes and which are region-specific. We may have also captured cells that were *pax6*+ and Sox2-. Specifically, we may have captured some cells that were recently *pax6*+ neural progenitors, and are now differentiating but retain GFP protein. This hypothesis would be well informed by birthdating neurons in the transgenic line to visualize if newborn neurons retain reporter GFP; however, for this set of experiments we acknowledge this potential caveat. Finally, while we find that Sox2 and *pax6*:*GFP* colocalize closely over much of their respective domains at stage 41, the very small size and delicacy of the regeneration bud at 24hpa precluded verification that these two markers continue to colocalize in the same way during the early stages of regeneration as the neural ampulla forms. Therefore while we expect that the spatial domains of these markers continue to overlap, and our scRNA-SEQ confirms that NPCs continue to express both of these markers, there may be subtleties to their spatial dynamics that reveal principles of NPC organization in regeneration that we did not capture.

Our study identified both novel candidate markers of specific neural lineages, and potential regulatory circuits directing regenerative decisions. Many of the genes we identified as differentially expressed and with differentially accessible promoters were not previously detected in bulk RNA-Seq or microarray analyses (Chang et al., 2017b; Love et

al., 2011). Of these, *pbx3* stands out as a gene with expression that is well-restricted to the neural lineage, becomes further restricted to differentiated neurons following injury, has a differentially accessible promoter itself, and has a target motif upstream of multiple differentially-expressed neuronal genes. Nevertheless, *pbx3* was not detected as differentially expressed in previous analyses of regeneration, likely because of the low overall expression levels of *pbx3*, and because expression changes are lineage-specific. Only by combining motif prediction from ATAC-Seq and the high-resolution expression data from scRNA-Seq did we bring this transcription factor to the top of our candidate list. Similarly, *meis1*, which is well-known as a regulator of neural and neural crest development (Erickson et al., 2010; Kelly et al., 2006; Machon et al., 2015; Mojsin and Stevanovic, 2009; Rataj-Baniowska et al., 2015; Stedman et al., 2009; Yamada et al., 2013), has not previously been detected in differential expression analyses during *Xenopus* regeneration. This may be because *meis1* is more broadly expressed, and bulk RNA-Seq was unable to capture the heterogeneity of expression changes *meis1* undergoes across multiple cell types.

3.3.2 *Evidence for a temporal uncoupling of differentiation and proliferation in the regenerating neural lineage*

We find that *pax6*+ neural progenitors place an early priority on neuronal differentiation, and specifically on formation of interneurons and motor neuron subtypes by 24hpa, shown as a model in Figure 8A-C. This is supported by the widespread accessibility enrichment of neuronal differentiation genes specifically at 24hpa, by the decline in relative

proportion of neural progenitors and transitioning neurons in favor of motor neurons and interneurons, and by a reduction in the number of cells in S/G2/M phases of the cell cycle.

Our finding that *pax6*+ progenitors prioritize differentiation before proliferation in regeneration represents an intriguing counterpoint to embryonic development, in which regional-specific proliferative neural progenitors can only later make the decision to undergo neurogenesis and exit the cell cycle (Hardwick et al., 2015). Transcriptional analysis of proliferating blastemal cells as well as recent pre-publication work in later-stage tadpoles of *Xenopus laevis* suggests the same model may also hold true in that species (Pelzer et al., 2020; Tsujioka et al., 2015).

The cell cycle exit and differentiation of some of the existing pool of neural progenitors and differentiating neurons at 24hpa results in a decline in these cell types and an increase in motor and interneurons. Not all progenitor cells exit the cell cycle however, and a residual population remains, from which we propose that additional progenitors are made by proliferation after 24hpa, in agreement with the emphasis placed on progenitor proliferation seen by ATAC-Seq at 72hpa and mitotic analysis by PH3 (**Error! Reference source not found.**D). We expect that the organized formation and outgrowth of neuronal structures remains important from 24 to 72hpa, as axons visibly elongate, and that *pbx3* and *meis1* are important regulators of these processes based on their target gene accessibility profiles (**Error! Reference source not found.**E/F). Our model agrees with numerous findings that there is little proliferation after tail amputation prior to 48hpa (Contreras et al., 2009; Gaete et al., 2012; Love et al., 2011b; Pelzer et al., 2020; Tsujioka et al., 2015) and with bulk RNA-Seq analysis showing a transcriptional upregulation of axonogenesis and

other neuronal differentiation genes (Chang et al., 2017). Our analysis suggests that early changes in neural cell identity and distribution, as well as the genes driving them, may represent an important early phase of regeneration that precede proliferation. In other regenerative vertebrates, specifically the anole lizards and geckos, spinal cord regeneration is incomplete; it includes formation of new axons along an ependymal cell tube, but does not include proliferation of progenitor cells to make a complete spinal cord (Duffy et al., 1990; Tokuyama et al., 2018; Wang et al., 2012). An interesting future direction would be to pursue whether program supporting axonogenesis may therefore have persisted in amniote lineages, where other features of complete spinal cord regeneration have been lost.

3.3.3 *Neural cell types preserve their identities in the regenerating Xenopus tail*

The identities of specific neural cell types are readily assigned to the same clusters both before and after tail transection in *Xenopus*. This agrees with recent single-cell transcriptomic analysis of the regenerating *Xenopus laevis* tail and spinal cord, in which the majority of cell types can be clearly assigned at 1, 2 and 3 days after amputation (Aztekin et al., 2019; Pelzer et al., 2020). The predominant model for vertebrate regeneration remains the axolotl limb, in which the formation of a morphologically undifferentiated mesenchymal blastema has been well characterized through morphological, histological, transcriptional, and single-cell analysis (Currie et al., 2016; Gerber et al., 2018; McCusker et al., 2015). In the axolotl limb, scRNA-Seq shows that different connective tissue cell types lose their unique transcriptional identities early in regeneration and converge on a similar signature, as the blastema forms (Gerber et al., 2018). While the early *Xenopus* tail regenerate has also been described as a blastema (Gargioli and Slack, 2004; Tsujioka et al., 2015), its parallels

to the axolotl blastema have been debated (Mochii et al., 2007), and our study confirms that diverse cell types retain well-discriminated transcriptional signatures as regeneration proceeds. This suggests that the molecular nature of the blastema may differ substantially between axolotls and *Xenopus*. In other regenerative contexts such as bone regeneration of the zebrafish fin, the formation of a blastema at the tip of each bony ray has been well documented (Knopf et al., 2011; Münch et al., 2013; Sousa et al., 2011). Global transcriptional responses to injury can therefore vary among regenerative vertebrate structures, which highlights the importance of multiple models for appendage regeneration.

3.3.4 *Pbx3 and Meis1 are required for axon organization and for tail regeneration*

We identify two transcription factors, Pbx3 and Meis1, as transcriptional regulators acting upstream of neuronal differentiation genes during regeneration (**Error! Reference source not found.**E). We show that these factors are both expressed in the regenerating spinal cord, with increased expression in motor neuron subtypes during regeneration. Loss of function of either factor results in disorganized neurofilament outgrowth with a failure of successful axial tissue regeneration. These factors are predicted to bind upstream of several dozen differentially accessible neuronal differentiation factors and are thus well-placed to drive the prioritization of neuronal differentiation we observe. Loss of either factor does not result in a failure to form neurofilaments, but does result in grossly disrupted organization of neurofilaments and reduced regeneration of the tail and spinal cord. Our data therefore suggests these factors may be critical for the proper re-formation and organization of differentiated neuronal structures. We note that even though the differentially accessible targets of Pbx3 are preferentially found at 72hpa, when proliferation is increasing, many

targets of Pbx3 are shared with Meis1, and Pbx3 itself increases in expression in neuronal types by 24hpa. We therefore propose that Pbx3 and Meis1 contribute to the ongoing differentiation of neuronal structures from 24hpa to 72hpa (**Error! Reference source not found.**F). Intriguingly, loss of either Pbx3 or Meis1 results in reduced tail regeneration overall. This may be due to interactions between regenerating nerves or neuronal structures and other axial structures, as regeneration of appendages has been shown to be nerve dependent in *Xenopus* and other systems (Farkas and Monaghan, 2017).

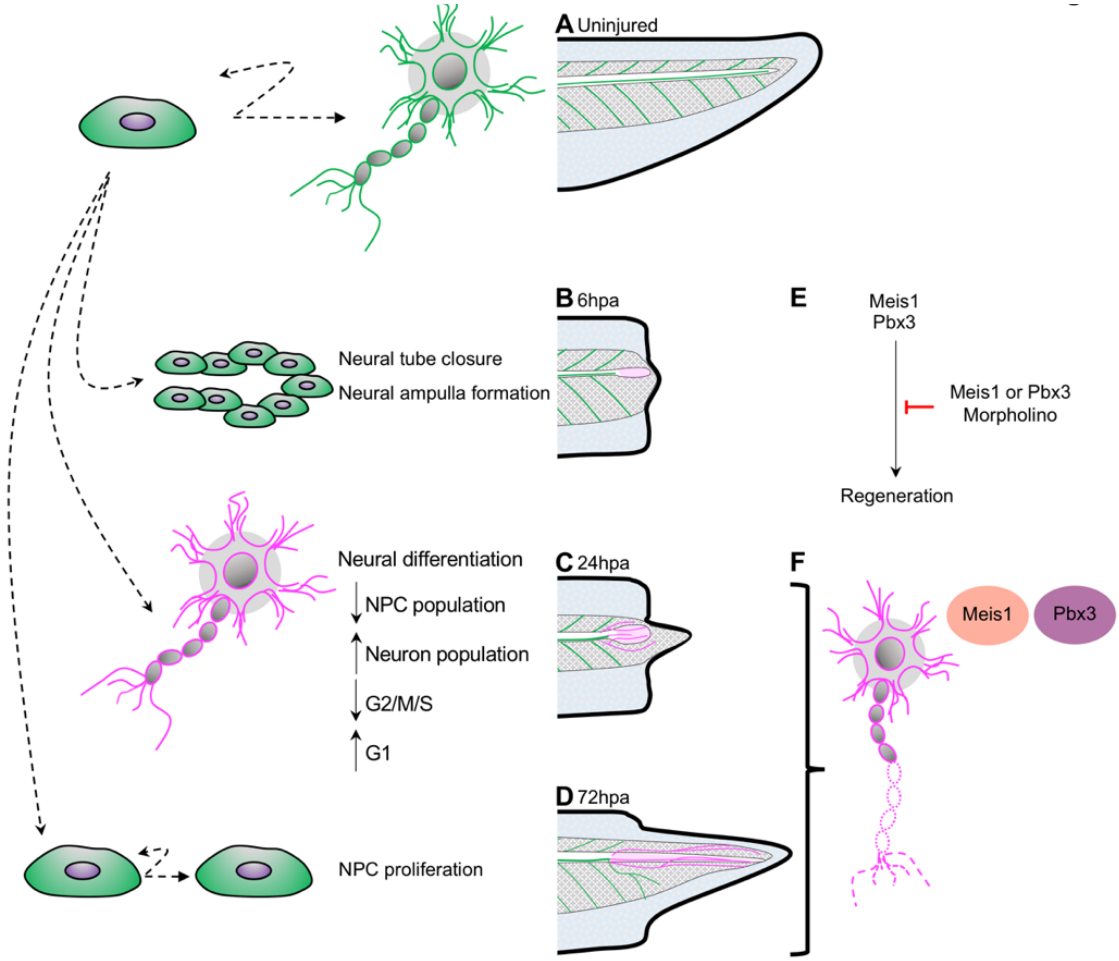

Figure 3-9 Model for neural progenitor regeneration.
(A) In normal development, neural progenitor cells make cell fate decisions to proliferate or differentiate. From this study we predict from our ATAC-Seq and scRNA-Seq analysis that neural progenitors first prioritize (B) migration and tube morphogenesis at 6hpa, followed by (C) neural differentiation at 24hpa, and then (D) turn on proliferation programs at 72hpa. (E) From integrative analysis of these data sets, we identified *meis1* and *pbx3* as candidate regulators of NPC regeneration at 24hpa and 72hpa. Loss-of-function experiments suggest Meis1 and Pbx3 are necessary for successful regeneration as well as proper

axonal patterning. Magenta shaded spinal cord represents NPCs in regenerated spinal cord. Magenta axons represent axons in the regenerated spinal cord and tail tissue.

Pbx3 and Meis1 may work independently or as part of a complex. Pbx3 and Meis1 are known to heterodimerize and coregulate targets upstream of *hoxa9* during leukemia progression (Li et al., 2016; Machon et al., 2015). Meis1 also partners with the related protein Pbx1 in axolotl limb regeneration and neural development (French et al., 2007; Mercader et al., 2005; Mojsin and Stevanovic, 2009). Although their targets are incompletely overlapping in our regeneration data, they are co-expressed in several cell types, and so we do not rule out the possibility that they may also act as a complex in the tadpole tail. Neither *pbx3* nor *meis1* have been studied in neural regeneration. These two factors therefore represent exciting new candidates for regulating neuronal differentiation in regeneration.

3.4 Conclusion

Taken as a whole, our study presents a coordinated analysis of chromatin accessibility and single cell transcriptomics in the regenerating neural lineage. We show that neural progenitors temporally uncouple their regulation of differentiation and proliferation, placing an earlier priority on differentiation of motor neuron subtypes. This new model represents an intriguing contrast to embryonic development. We computationally identify multiple candidate regulators of neural progenitor fate decisions in a regenerative context, and specifically highlight *pbx3* and *meis1* as critical regulators of neural development and regeneration.

3.5 METHODS

Key resources can be found in the key resource table,

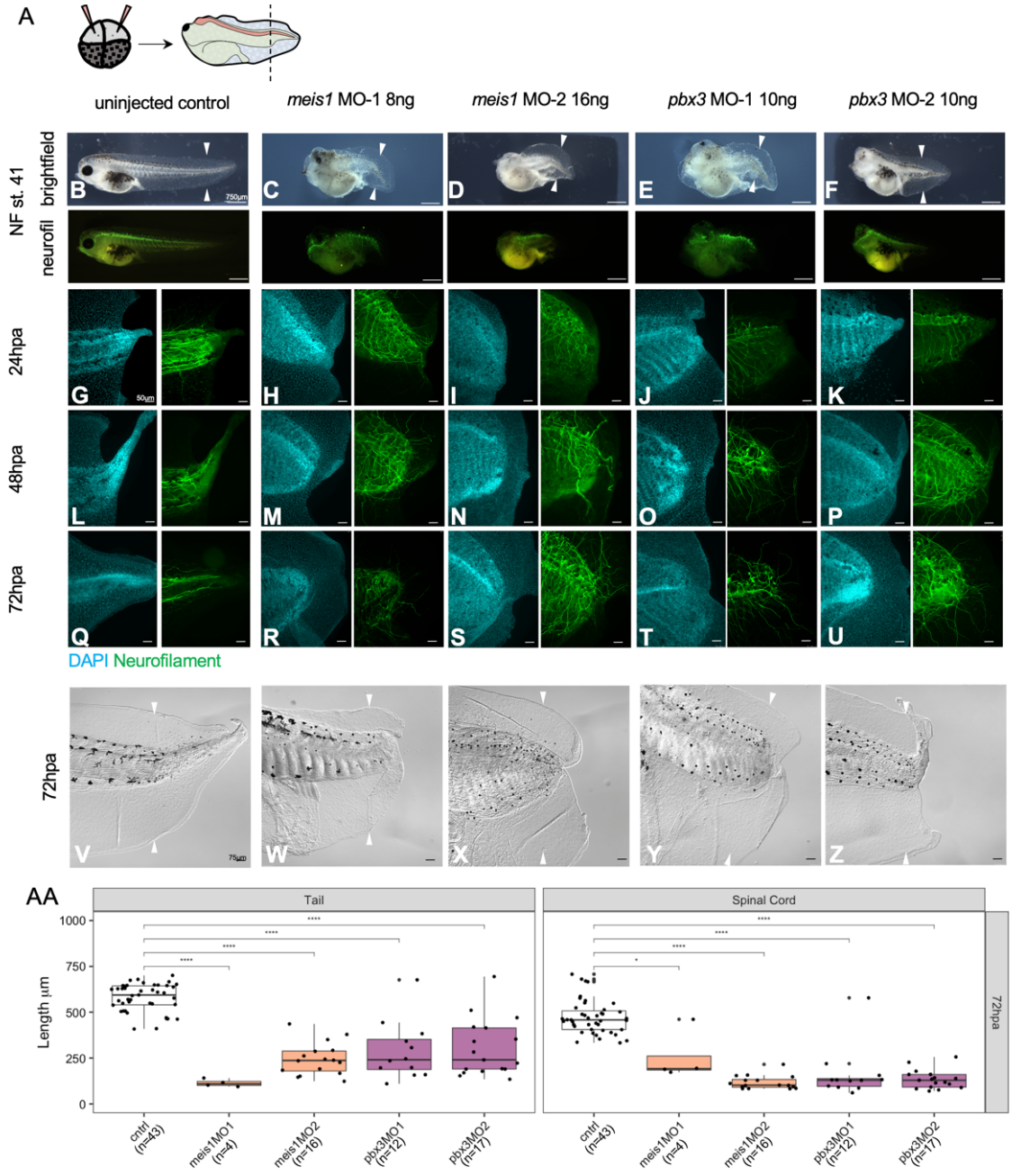

Figure **3-14 Meis1 and Pbx3 are independently necessary for successful regeneration of the axial tissue and neuronal patterning.**

(A) Experimental set up for morpholino injections and amputations. Two sets of morpholinos were injected per gene to dorsally fated blastomeres at the 4-cell stage. Embryos were reared to stage 41 for amputation. Representative images of uninjured stage 41 (B) control, (C, D) Meis1 morphants 1 & 2, and (E, F) Pbx3 morphants 1 & 2. Images are shown in bright field and immunostained for neurofilament. Stage 41 tadpoles were amputated and collected at (G-K) 24hpa, (L-P) 48hpa, and (Q-U) 72hpa. The cyan channel represents tails stained with DAPI and the green channel represents tails immunostained for neurofilament. Scale bar: 75um. (V-Z) Brightfield images of the 72hpa tails. Scale bar: 50um. (AA) Boxplots representing length of regenerated tail or spinal cord for each tadpole. Statistics represent a two-tailed t-test to determine significance between conditions. (*<0.05, **<0.005, ***<0.0005, ****<0.00005).

Table 3-1. All data generated and analyzed in this paper can be found with the GEO accession GSE146837 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146837). Source files and analysis code can be found at https://gitlab.com/akakebee/kakebeen-et-al-2020.

3.5.1 *X. tropicalis husbandry and use*

Use of *Xenopus tropicalis* was carried out under the approval and oversight of the IACUC committee at UW, an AALAC-accredited institution. Ovulation of adult *X. tropicalis* and generation of embryos by natural matings were performed according to published methods (Khokha et al., 2002; Sive et al., 2010). Fertilized eggs were de-jellied in 3% cysteine in 1/9x modified frog ringer's solution (MR) for 10-15 minutes. Embryos were reared as described (Khokha et al., 2002). Staging was assessed by the Nieuwkoop and Faber staging series (Nieuwkoop and Faber, 1994). In this study we used both wild-type frogs and frogs from the triple transgenic line Xtr.Tg(pax6:*GFP*;cryga:*RFP*;actc1:*RFP*)Papal (Hartley et al., 2001; Hirsch et al., 2002) reared and purchased from the National *Xenopus* Resource (https://www.mbl.edu/xenopus/). *Pax6* transgenic matings were performed by crossing a heterozygous transgenic frog to a wild-type frog. These matings yielded clutches with 50/50 wt/*pax6*:*GFP*+ populations.

3.5.2 *Sectioning and imaging*

For sectioning, tadpoles were fixed for 30 minutes in 10xMEMFA, 3.7% formaldehyde at room temperature. Tadpoles were then washed briefly in PBT and incubated in 4% sucrose in 1xPBS at 4C overnight. Tadpoles were embedded in Tissue Tek O.C.T. (Sakura 4583) in 10mmx10mmx5mm molds (Sakura 4565) and stored in the -80C until sectioning. Tadpoles were sectioned transversely at 14µm with a cryostat (Leica). Sections that were not stained were baked onto slides on a heatblock and imaged within 24hours to capture endogenous *pax6* reporter fluorescence. Images were acquired with a Lecia SP8 and UPDATE camera.

3.5.3 *X. tropicalis amputation assay*

X. tropicalis were sorted for transgenic reporter fluorescence for regeneration assays. Tadpoles were anesthetized with 0.05% ms-222 in 1/9x MR and tested for response to touching prior to amputation surgery. Once fully anesthetized, a sterilized scalpel was used to amputate the posterior third of the tail. Amputated tadpoles were removed from anesthetic media within 10 minutes of amputation into new 1/9x MR. Tadpoles were left to regenerate at a density of no more than 3 tadpoles to 1 mL of media.

3.5.4 *Cell dissociation and fluorescent activated cell sorting (FACS)*

Cell dissociation: At each timepoint, regenerating tadpoles were anesthetized until non-responsive to touch. A scalpel was used to amputate to the regenerated tail tissue (most posterior point to ~500µm anterior of that point). Regenerated tail tips were collected and

spun down at 1000xg. Media was taken off and ~60 tails were resuspended in 200uL of 0.035mg/ml liberase in PBS (Roche 05401119001). Tails were left to incubate for 20 minutes at room temperature. A p-200 was then used to pipet the tails up and down gently and break up tail tissue. Once no visible chunks were apparent, cells were spun down for 3 minutes at 1000xg. Supernatant was disposed and cells were washed with 180uL of 1xPBS. Cells were then spun once more for 1 minute at 1000xg and supernatant disposed. Cells were resuspended in 200-300uL of 1xPBS and filtered through a 70µm cell filter (Fisher 22-363-548) into a FACS tube. Cells were kept on ice until sorting.

FACS: Cell sorting and collecting were performed through on the BD FACS Aria III Cell Sorter (BD Biosciences). Wild-type samples were used as negative controls to set sorting gates to exclude any autofluorescence from our true GFP signal. Cells were sorted based on selection from 3 criteria. FSC x SSC were used to select for cells based on size and granularity. From this population, cells were visualized with the FSC-W x FSC-H dimensions to select for singlets and exclude doublets. The doublets were then visualized in a histogram where GFP was the x-axis. For GFP gate setting, selection gates were drawn for all area without cells in the wild-type tails. Here we assume that any GFP-signal is autofluorescence and need be selected against. Positive control transgenic tadpoles were used to test for GFP signal. Cells were passed through the sorter and collected through a 70µm nozzle. Cells were collected into 50uL of 1xPBS. For *pax6* samples, GFP+ and GFP- cells were collected into separate tubes and for all-tissue samples, GFP+ and GFP- cells were collected into the same tube.

3.5.5 *Quantitative RT-PCR*

For RT-PCR, either GFP+ cells or GFP- cells were collected from the flow cytometer directly into RNA lysis buffer. Between 2000-4000 cells were collected for each condition from uninjured tail tips. Cells were then agitated at 42C for 45 min. RNA/DNA were extracted and purified from the lysed cells by phenol:chloroform extraction with subsequent ethanol precipitation. Samples were then DNAse treated to remove DNA. To make cDNA from the RNA, we used the Superscript III reverse transcriptase kit (ThermoFisher 18080044). Amplifications were carried with iQ SYBR Green Supermix (BioRad 1708882) for 40 cycles of: 95C:10sec, 57C:30sec after initial 2minutest at 95C. Primers used for *gfp, pax6, actc1, tubb2* are provided Supplementary File 2.

3.5.6 *ATAC-Seq library preparation*

Within 30 minutes of cell sorting, cells were prepared for transposition. (See Supplementary File 1 for details of cells sorted for each sample). Cells were spun down for 3 minutes at 1000xg and supernatant was discarded. Cells were resuspended in 15uL of TN5 reaction mix using the TN5 enzyme, buffer, and water from the Nextera DNA Sample prep kit (Illumina 15028212). Cells were transposed for 1 hour at 37C. Post transposition, cells were removed from 37C and DNA was purified using the Qiagen MinElute kit (Qiagen 28206). DNA was eluted in 10uL of buffer EB. Purified DNA was then amplified using the 2xNEB PCR Master Mix (NEB M0541L), a universal forward primer, and an indexed reverse primer from the i7 Illumina series. Initial amplification as follows: 72C:5min, 98C:30sec, then 5 cycles of 98C:10sec, 63C:30sec, 72C:1min. 5uL of the initial PCR product was then used for a 15µl side qPCR reaction to determine how many more cycles to

run to stop amplification prior to saturation. The side reaction was carried out as follows: 98C:30sec, then 30 cycles of 98C:10sec, 63C:30sec, 72C:1min. The number of cycles was determined by taking the range of the starting and ending CT values and dividing that number by 4, then identifying how many cycles the sample had gone through at the derived CT value. The initial PCR reactions were then run for the calculated number of extra cycles and removed from the thermo cycler. Amplified samples were purified with the Qiagen MinElute kit and eluted in 20uL of buffer EB. 5uL of purified PCR product were run out on a 5% acrylamide gel in 1xTBE at 100V for ~45 minutes. Gels were stained with Ethidium Bromide and imaged on the gel imager (BioRad). Samples with sufficient periodic bands, relating to length of DNA to wrap a mono-, di-, tri- nucleosome, were then run out on a bioanalyzer. Library concentrations were taken using the Qubit high sensitivity dsDNA assay. Libraries were pooled by normalizing library concentrations to the lowest concentration and adding equal volumes of libraries for a minimum of 5ng of each library.

Libraries were sequenced on the Illumina Next-Seq platform across three NextSeq 500 High Output Kit v2.5 (150 cycles, Illumina CAT. 20024907). Libraries were sequenced with, paired end, 75 bp reads at the Sound Genomics facility.

3.5.7 *ATAC-Seq analysis pipeline*

Trimming adaptors and alignment: Adapters were trimmed from reads and low-quality sequences (Phred <33) were removed using Trim Galore!. Reads were aligned to xtropicalis9.0 using Bowtie2 (option:–very-sensitive) (Langmead and Salzberg, 2012; Langmead et al., 2019). Duplicate reads were marked using Picard ‘MarkDuplicates’

(http://broadinstitute.github.io/picard/). Duplicate reads were removed using SAMtools. Bigwigs were generated using a custom script (https:// rpubs.com/achitsaz/98857) and visualized in IGV.

Peak calling: Peaks were called using MACS2 (options:–nomodel–shif -100–extsize 200) (Zhang et al., 2008). Consensus peak set used for differential accessibility were called on a merged bam of all ATAC-seq files used in this experiment.

Peak annotation: Annotation and genomic feature enrichment analysis. Annotation of ATAC-seq peaks and genomic annotation enrichment analysis were performed using GenomicRanges (Lawrence et al., 2013), matching each peak to the nearest TSS using Xtropicalisv9.0.Named.primaryTrs.gff3 from Xenbase.

Differential accessibility analysis: Differential accessibility analysis in ATAC-seq peaks. Differential accessibility across ATAC-seq sample groups was determined as detailed in the edgeR users guide (Robinson et al., 2010). To define the regions for differential accessibility analysis, peaks were called on a merged bam of all ATAC-seq samples in this experiment. A matrix of counts for all samples was then generated using the ‘featureCounts’ in the Rsubread package (options: isPairedEnd=TRUE, maxFragLength=2000). The counts matrix was filtered to select rows having at least 1 count per million in n−1 samples to minimize the influence of variability at the threshold for sensitivity on the analysis.

3.5.8 *Gene Ontology Analysis*

GO:BP analysis (gprofiler2): Using differential accessibility flags, we identified all regions more accessible in the indicated contrast. The TSS annotation of these regions was used as input for gene ontology analysis. GO analysis was carried out using the 'gprofiler2' R package (Raudvere et al., 2019). The command 'gost' was used to call GO terms (options: query= DA peak list, organism="hsapiens", sources="GO:BP", evcodes= TRUE).

Reduce Redudancy (ReviGO): To reduce redundancy of GO terms, we used the online resource ReviGO (http://revigo.irb.hr/) (Supek et al., 2011). We input GO IDs and p-values into ReviGO and used the following parameters to run the analysis: allowed similarity=tiny(0.4), p-values, database with GO term sizes=Homo sapiens, Semantic similarity measure=SimRel. We then saved the output .CSV file for further plotting and analysis in R.

3.5.9 *Single cell RNA-Seq library preparation*

GFP+ cells from ~500 tails were collected for single cell analysis using the above methods but scaled for 500 tails. After sorting, cells were immediately washed with 1xPBS and diluted to a concentration of 1000 cells/µl based on the 10x Genomics guidelines. We aimed for capture of 10,000 cells using the 10x Genomics Platform. Single-cell mRNA libraries were prepared using the single-cell 3' solution V2 kit (10x Genomics). Quality control and quantification assays were performed using a Qubit fluorometer (Thermo Fisher) and a D1000 Screentape Assay (Agilent). Libraries were sequences on an Illumina NextSeq 500 using 75-cycle, high output kits (reads 1: 26 cycles, i7 Index: eight cycles, read 2: 57

cycles). Each sample was sequenced to average read depth of 16 million total reads. This resulted in an average read depth of ~15,000 reads/cell after read depth normalization.

3.5.10 *Single cell RNA-Seq data processing*

Cellular barcodes and μMis were determined using Cell Ranger 2.0.2 (10X Genomics) and cells were filtered to include only high-quality cells. Cell ranger defaults for selecting cell-associated barcodes versus barcodes associated with empty partitions were used.

UMAP visualization and clustering

Clustering and integration of all cells in uninjured tail and 24hpa: We used Seurat V3.0 for analysis of scRNA-Seq data (Butler et al., 2018, 2018). To reduce dimensions and cluster cells for the uninjured and 24hpa data sets, we followed Seurat's "Integrating stimulated vs. control PBMC datasets to learn cell-type specific responses" vignette (https://satijalab.org/seurat/v3.1/immune_alignment.html). We created Seurat objects from CellRanger processed data with the command 'CreateSeuratObject' (options: min.cells=5). We then removed cells that had less than 500 RNA features by subsetting data (options: subset=nFeature_RNA > 500). Cells were normalized ('NormalizeData') and variable features were found with 'FindVariableFeatures' (options: selection.method = "vst"). Integration anchors for the two data sets were identified with 'FindIntegrationAnchors' (options: dims=1:20) and the integration of the data sets was performed with 'IntegrateData' (options: dims=1:20). The integrated data set was then scaled ('ScaleData') and principle components were found with the command 'RunPCA' (options: npcs=60). We then reduced dimensions using Uniform Manifold Approximation and Projection (UMAP). We did so by

using ‘RunUMAP’ (options: npcs = 40). Next we clustered cells by first finding nearest neighbors with ‘FindNeighbors’ (options: reduction=”pca”, dims=40) and then calling clusters with ‘FindClusters’ (options: resolution=0.7).

Subsetting of neural cells in uninjured tail and 24hpa: We used neural marker genes (*sox2, neurog1, neurod1,* and *elavl4*) to determine which cell clusters were neural cells. To subset the data set, we used ‘SubsetData’ (options: ident.use=c(1,4,5,13,15). The subset data was scaled, and principle components were called. ‘RunUMAP’ and “FindNeighbors’ was run (options: reduction=”pca”, dims=1:20). ‘FindClusters’ was then run to establish neural clusters (options: resolution=0.4).

Marker identification and differential expression analysis

Cluster markers: To identify genes that define each cluster, we used Seurat’s ‘FindMarkers’ command (options: ident.1=Cluster of interest, min.pct=0.3, only.pos=TRUE, assay= “RNA”). This command was run for all 7 subset neural clusters.

Differential gene expression across timepoints: To identify genes that were differentially expressed between uninjured and 24hpa tail cells, the Seurat function ‘FindMarkers’ was used. (options: ident.1=Cluster of interest_uninjured, ident.2=Cluster of interest_24hpa, min.pct=0.3, assay= “RNA”).

Cell cycle prediction analysis

Cell cycle prediction: To predict the cell cycle of cells sequenced in our scRNA-Seq, we followed the methods outlined in the Seurat Vignette “Cell-Cycle Scoring and Regression”

(https://satijalab.org/seurat/v3.1/cell_cycle_vignette.html). Genes defining "S" phase and "G2M" phase were called from a pre-loaded list of cell cycle markers from (Kowalczyk et al., 2015). These genes were then used in the Seurat command "CellCycleScoring" to predict cell cycle phases (options: s.features=s.genes (from Seurat), g2m.features=g2m.genes(from Seurat), set.ident =TRUE).

Gene regulatory predictions

Motif Finding: We used the 'findMotifsGenome.pl' command to identify all known binding motifs from the Homer motif library in differentially accessible peaks (Heinz et al., 2010). The motifs were then appended to the ATAC-Seq differential accessibility table by matching peak names. This table gave us a called peak, the TSS annotation of that peak, motifs called in the peak, and transcription factor binders of the called motif.

Bulk RNA-Seq integration: TSS annotations were used to vet table for genes that are expressed at the given timepoint. To do so, we subsetted the table for genes that were expressed in bulk RNA-Seq data from (Chang et al., 2017).

GRN prediction across regenerative time: Using a homemade script, we identified DA regions at 6hpa that had binding motifs in 24hpa DA peaks. We then identified DA regions at 24hpa that had binding motifs in 72hpa DA peaks. To connect the timepoints, we identified targets of TFs with motifs at 24hpa, then asked whether those targets had binding motifs at 72hpa. To narrow down potential GRNs to follow up on, we asked whether the genes were expressed in neural cells in the scRNA-Seq data in the uninjured at 24hpa cells and whether they were differentially expressed.

3.5.11 *Whole-Mount In situ hybridization*

Embryos and larval tadpoles were fixed in 1xMEMFA, 3.7% Formaldehyde for 2-6 hr at room temperature or overnight at 4C. *X. tropicalis* multibasket in situ hybridization protocols were followed as described in (Khokha et al., 2002). For plasmids used for in situ hybridization probes, see reagents table). Whole mount in situ tadpoles were imaged on a Leica M205-FA with a D5C550 camera.

3.5.12 *Morpholino microinjections*

Morpholinos were ordered from Gene Tools. BLAST searches for both morpholinos for Meis1 and Pbx3 using the *X. tropicalis* v9.1 genome identified just one target in the correct gene for each of the 4 MOs, and no off target binding elsewhere in the genome.

Dorsal blastomere injections: Wild-type *X. tropicalis* embryos were staged to NF stage 3 (4-cell). Using a microinjector, the two dorsally-fated blastomeres were both injected with 2nL of morpholino mix. Morpholino mix contained a morpholino against *meis1 (*MO1:8ng or MO2:16ng*)* or against *pbx3* (MO1:10ng or MO2:10ng*)* and a labeled dextran tracer (used at 2mg/mL , ThermoFisher D1817). Embryos were screened at stage 18 for bilateral red fluorescence in the nervous system tissues. Neurulas were raised to stage 41 and used for regeneration assays. Morphant regenerates were collected at 24hpa, 48hpa, and 72hpa for measurement analysis.

Tail vein injections: Wild-type *X. tropicalis* embryos were staged to NF embryonic stage 35-36 (pre-hatching stage). Tadpoles were anaesthetized with MS-222 and moved from culture

dish to an agarose coated dish in one drop. Surrounding media was removed. Using a microinjector, a pulled needle containing VIVO-morpholino and labeled dextran tracer was inserted into the ventral tail vein in the posterior tail and 3x2nL injections were delivered (20ng). Embryos were returned to fresh media and screened for tracer fluorescence in the blood stream. Injected tadpoles were grown 24 hours to stage 41 and amputated. At 24hpa, tadpoles to be collected at 72hpa were re-injected into the tail vein as before.

3.5.13 *CRISPR-mediated mutation*

sgRNA Design, Preparation, and Injection: sgRNAs were designed targeting meis1 and pbx3 using CRISPRSCAN (https://www.crisprscan.org/?page=gene) (options: organism= "Frog-Xenopus Tropicalis", Gene= "pbx3" or "meis1", Cas9-NGG, In Vitro T7 promoter). PCR was performed as described in Bhattacharya et al. (2015). SgRNA was transcribed using T7 mMachine kit (Ambion). Guides were injected into 2 blastomeres of 4-cell embryos with 1.5 ng Cas9 (Bhattacharya et al., 2015; Nakayama et al., 2013). All guides were injected at a dose of 400 pg/embryo.

3.5.14 *Immunohistochemistry*

Whole-mount*: X. tropicalis* tadpoles were fixed for 20 minutes in 10xMEMFA, 3.7% formaldehyde at room temperature. To preserve reporter fluorescence, tadpoles were kept dark upon fixing. Tadpoles were permeabilized by washing 3X 20 minutes in PBS + 0.01% Triton x-100 (PBT). Tadpoles were blocked for 1 hour at room temperature in 10% CASblock (Invitrogen #00-8120) in PBT. Then tadpoles were incubated in primary antibody (1:50

neurofilament associated antigen, DSHB 3A10; 1:1000 Anti-Histone H3 (tri methly K9, Phospho S10), Abcam 14955) diluted in 100% CAS-block overnight at 4°C. Tadpoles were then washed 3X10 minutes at room temperature in PBT and reblocked for 30 minutes in 10% CAS-block in PBT. Secondary antibody (goat anti-mouse 488, ThermoFisher A11001 or goat anti-mouse 555, ThermoFisher A21422) were diluted 1:500 in 100% CAS-block and incubated for 2 hours at room temperature.

Sections: Sections were baked onto slides on a heat block for 1 hour. Once slide had cooled, a hydrophobic barrier was drawn around the sections with a PAP pen (cat) and left to dry. Once dry, sections were washed with PBT 3x10minutes. Sections were then blocked for 1 hour in 10% CASBlock followed by incubation in primary antibody overnight at 4C (1:50 neurofilament associated antigen, DSHB 3A10; 1:100 anti-DCX, Cell Signaling Technology 4604S; 1:200 anti-Sox2, Cell Signaling Technology 2748S). The following day, primary antibody was removed and replaced with a brief wash in PBT. Sections were blocked for 30 minutes followed by 2 hours incubation in secondary antibody at room temperature (1:500 goat anti-mouse 555, ThermoFisher A21422; 1:500 goat anti-rabbit 594, ThermoFisher Cat. A-21422).

For both whole-mount and sections, tadpoles were then washed 3X 20 minutes in PBT. Tadpoles were then incubated in 1:2000 DAPI (Sigma D9542) for 10 minutes before being washed with 1xPBS for 10 more minutes. Isolated tails were mounted on slides in ProLong Gold (ThermoFisher P36930). Images were acquired with a Leica SP8 confocal microscope.

3.5.15 *Regenerated tissue measurement and analysis*

Measurement: To measure regenerated length of spinal cord and tail, regenerated tails were mounted on slides and were imaged on the Images were acquired with a Lecia DM 5500 B microscope with the 10x or 20x objective. Using the Leica imaging software (LasX), measurements were taken from the amputation plane to the most posterior point of the axial tissues for tail length and amputation plane to most posterior point of the spinal cord for spinal cord length. We used morphological features such as disruption to somite chevron morphology to identify the amputation plane.

Analysis: We used the R package ggplot2 to plot boxplots of the regenerate measurement data. A single regenerated tadpole was treated as one measurement. A two-tailed t-test was used to test how different the uninjected controls were from the morphants. Statistical analysis was performed using the R package ggpubr.

3.5.16 *PH3 cell quantification*

24hpa and 72hpa regenerated tails stained with anti-PH3 and DAPI were imaged on a Leica DM 5500 B Microscope with a 20x objective. Z-Stacks were acquired encompassing the entire spinal cord in focus, while excluding the most outer structures (i.e. epidermis) and max projected. In FIJI, the regenerated spinal cord in was identified and outlined in the DAPI channel to create a region of interest. The region of interest was transferred to the PH3 channel, and the area of the region was taken. PH3+ cells were counted manually within the region of interest. Quantifications were reported as PH3+ cells

per spinal cord area. Significance was tested by using a two tailed t-test between conditions in question. This test was performed with the R package ggpubr.

3.6 Data Availability

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE146837 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146837).

3.7 Acknowledgements

We thank members of the Wills lab for critical comments during the preparation of this manuscript and support in frog husbandry. We thank Rosalind Bump for the 72hpa *ass1 in situ* hybridization image. We thank members of MF4 Supergroup and Regeneration Club for productive discussion and helpful suggestions throughout the course of this project. We thank the UW Pathology Flow Cytometry Core Facility for help with FACS. We thank the Reh lab for access and training on their cryostat. We thank Xenbase for curation of genomic and literature information and the National *Xenopus* Resource for frogs. ADK was supported by the Cellular and Molecular Biology Training Grant PHS NRSA T32GM007270 from NIGMS. This work was supported by NINDS R01NS099124 to AEW.

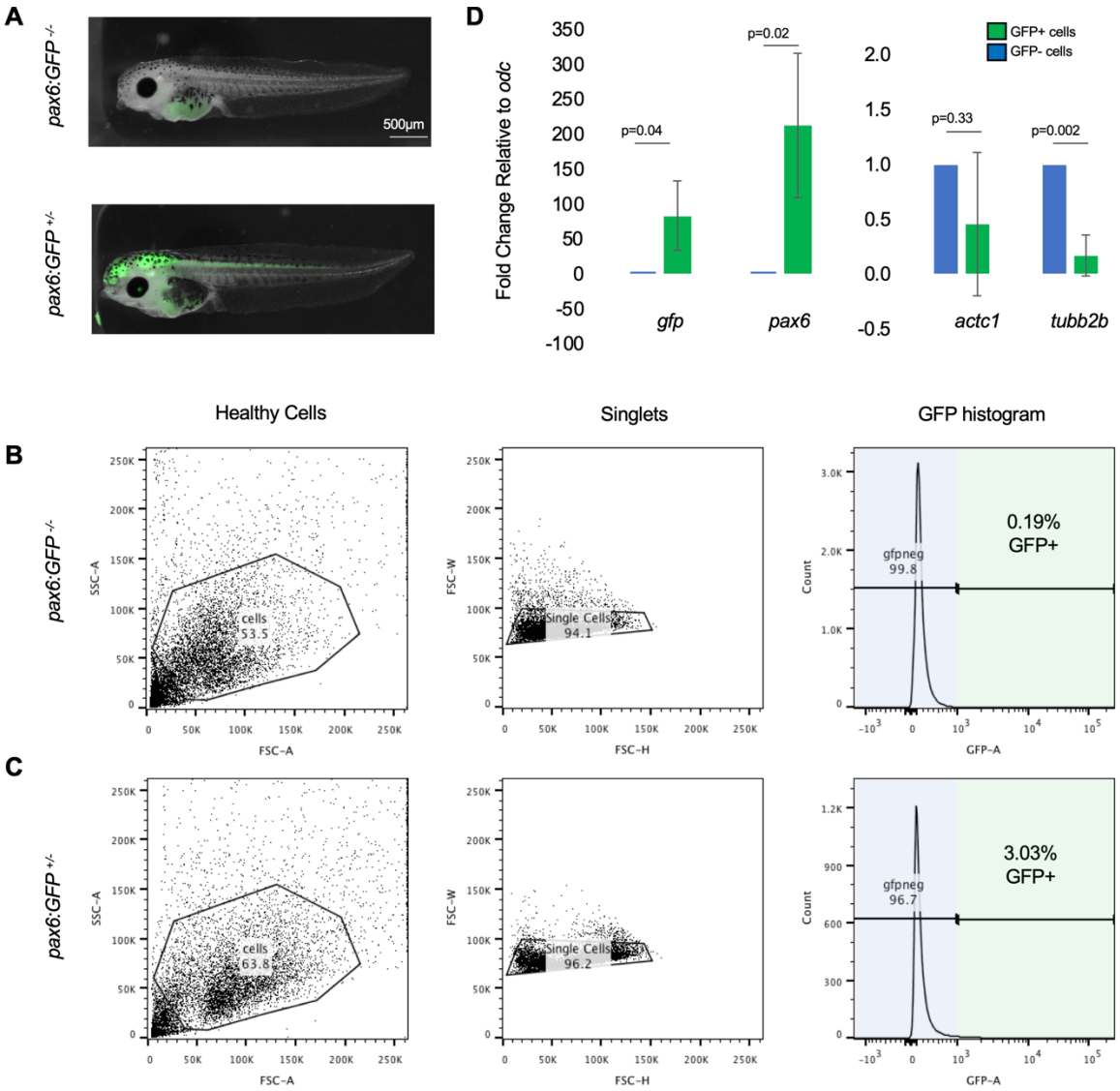

Figure 3-10 FACS gating was set using wild-type tadpoles to avoid false-positive GFP+ cells. (A) Representative images of stage 41 (top) wild-type tadpoles and (bottom) Xtr.Tg(*pax6*:*GFP*) tadpoles. (B) Representative gating for wild-type tadpoles. Gates were drawn to exclude all cells in the wild-type sample from being part of GFP signal. This gate setting was used to exclude autofluorescence from yolk in tadpole cells. The same gates from (B) were applied in (C) to collect *pax6*:*GFP*+ tadpole cells. Cells that fall within the gate represent true GFP+ signal. (D) qRT-PCR on GFP+ and GFP- cells from uninjured tails. *GFP* was used as a positive control, *pax6* was used to assay for NPC marker, *actc1* was used as a non-neural marker, and *tubb2b* was used as a differentiated neural marker.

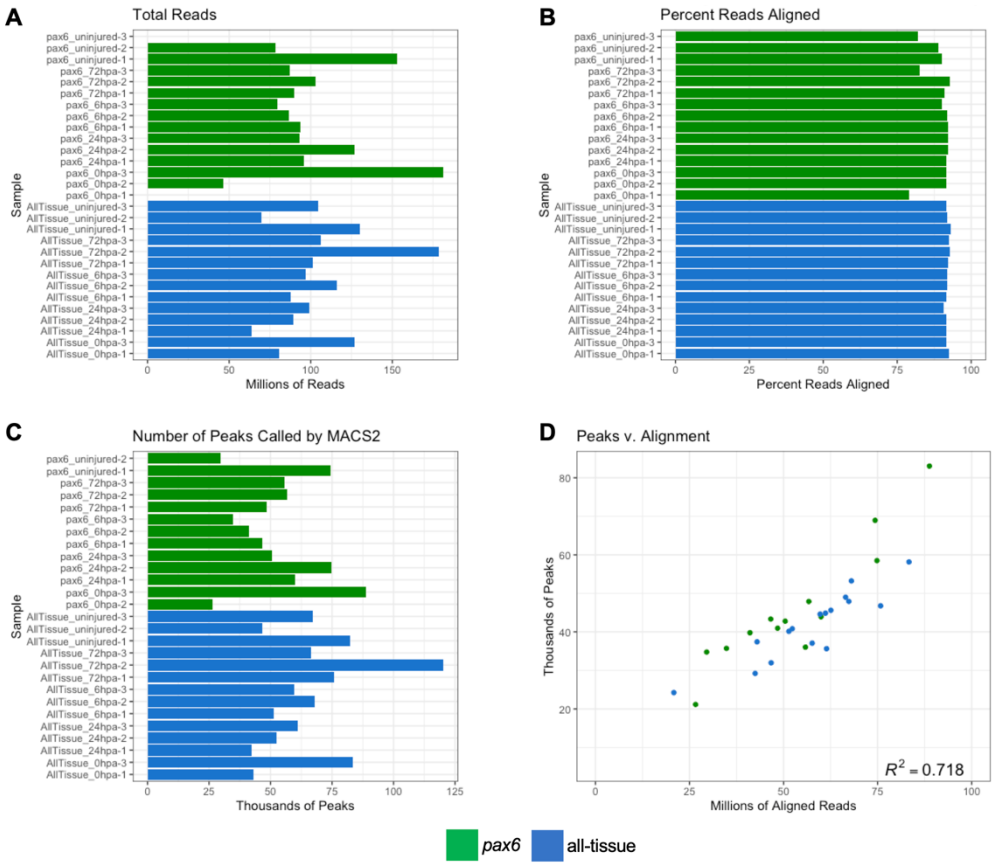

Figure 3-11 Low input ATAC-Seq libraries meet sufficient quality for downstream analysis.
(A) Barplot of total number of reads sequenced for each library. (B) Barplot of the percent reads aligned to the genome by Bowtie2 (C) Barplot of the number of peaks called from aligned reads by MACS2. (D) Correlation plot of number of aligned reads to number of called peaks. Reads and peaks aligns with correlation coefficient of r^2=0.785.

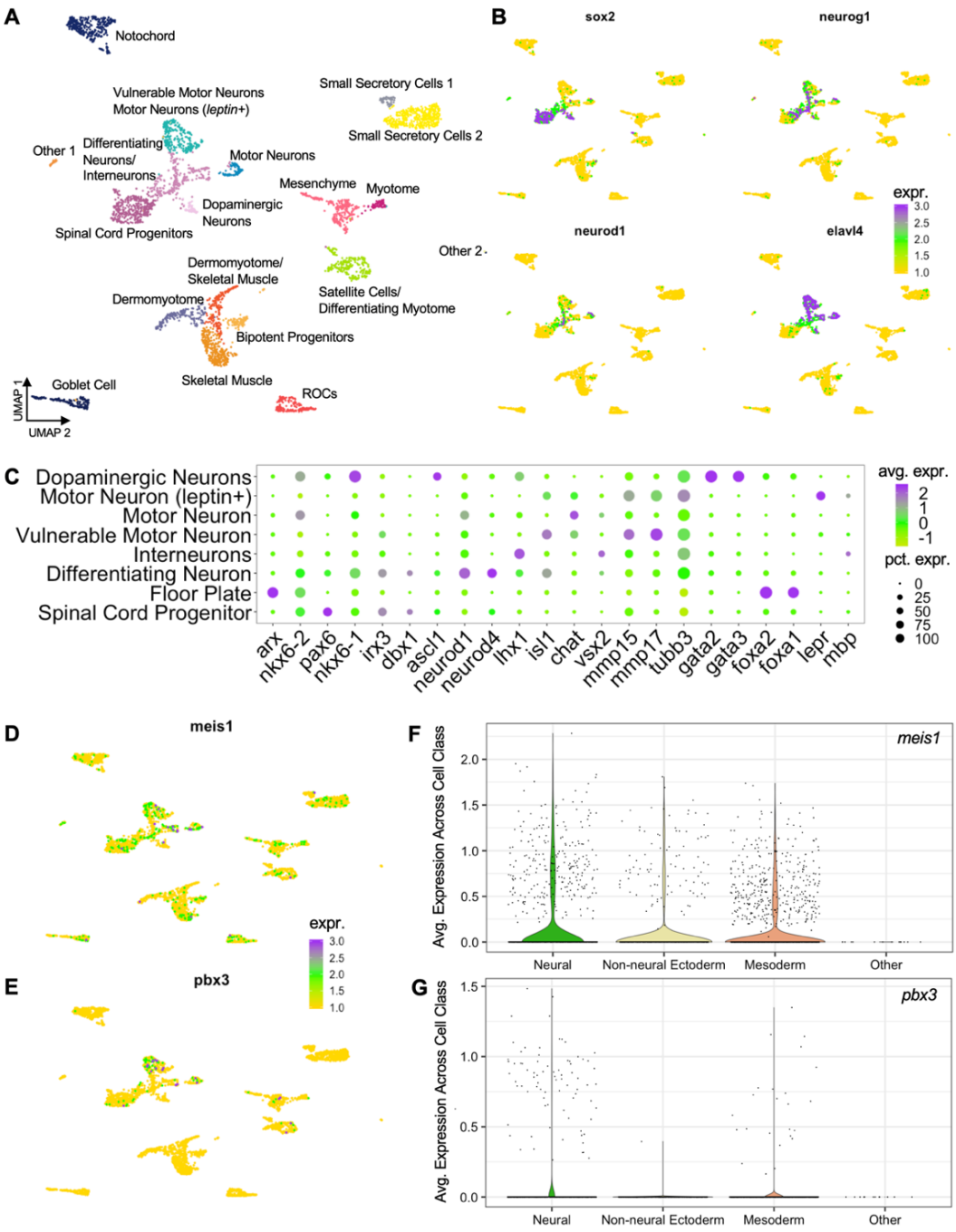

Figure 3-12 scRNA-Seq analysis of all cells sequenced in uninjured and 24hpa tails.
(A) UMAP dimensionality reduction and cluster assignment of integrated cells from uninjured and 24hpa tails. These cells are a superset of the 7 clusters shown in Figure 4. (B) UMAP plots of cell colored by expression of neural lineage markers. *Sox2* was used to identify neural stem cells, *neurod1* and *neurog1* were used to identify differentiating neurons, and *elavl4* was used to identify differentiated neurons. 6 cell clusters were identified by these markers as neural lineage clusters. (C) Dotplot showing expression of genes used to determine neural identities as defined by Aztekin et. al. Size of circle represents percentage of cells in cluster expressing gene and color denotes average expression of a gene across a cluster. UMAP plots of cells colored by expression for (D) *meis1* and (E) *pbx3* across all cells. Violin plots showing distribution of gene expression for (F) *meis1* and (G) *pbx3*. In the violin plots, cell types from (A) were grouped into broader cell lineages.

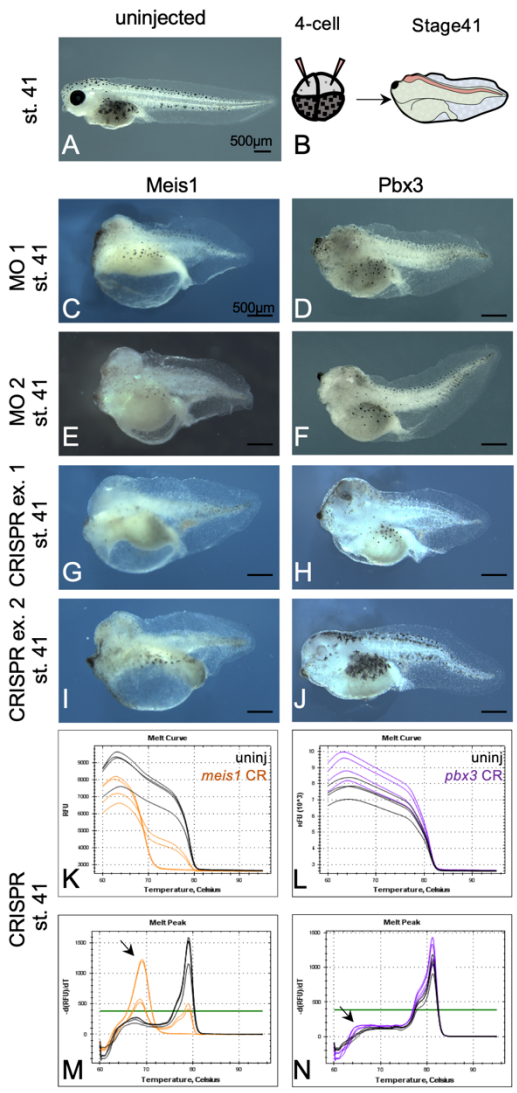

Figure 3-13 Embryonic Meis1 and Pbx3 morphants are characterized by small heads and missing or small eyes.

(A) Uninjected, stage 41 tadpole. (B) Injection schematic showing injections were made in the 2 dorsally fated blastomeres at the 4-cell stage. (C-F) Stage 41, morphant tadpoles that were independently injected with two sets of translation blocking morpholinos against (C, E) Meis1 and (D, F) Pbx3. Two different MO sequences were used for each gene. (G-J) Stage 41, mutant tadpoles that were injected with Cas9 and gRNA against (G,I) *meis1* and (H,J) *pbx3*. Two representative animals are shown for each gRNA set; only one gRNA set was used per gene. (K-N) CRISPR mutant tadpoles were analyzed for mutations by high resolution melt (HRM) analysis. Melt curve and peak chart for (K,M) *meis1* and (L,N) *pbx3* are shown. Black lines depict wild-type siblings from the same clutch whereas orange lines represent *meis1* mutants and purple lines depict *pbx3* mutants. Black arrowheads indicate the position of curve shift for mutants relative to wild-type siblings.

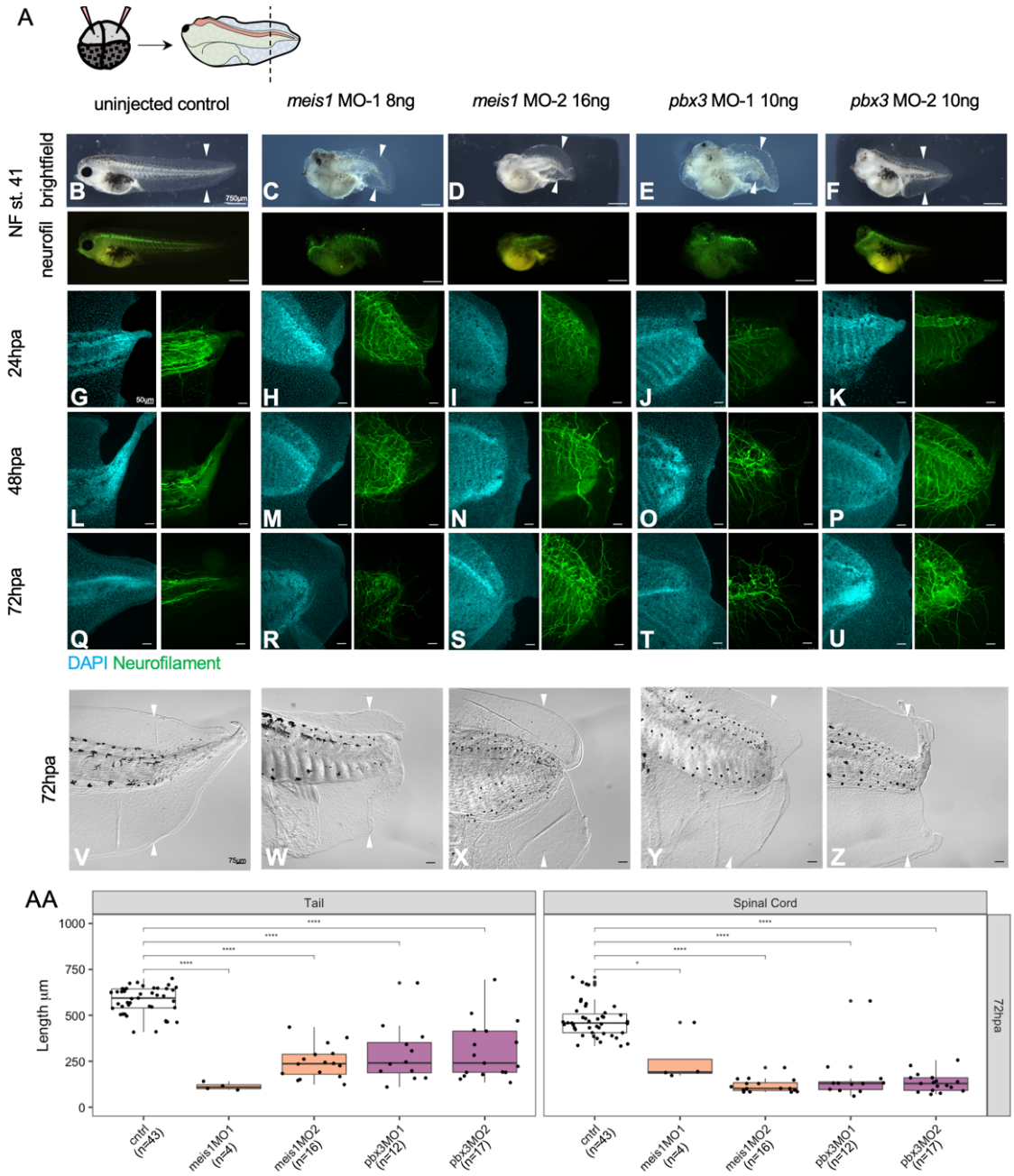

Figure 3-14 Meis1 and Pbx3 are independently necessary for successful regeneration of the axial tissue and neuronal patterning.

(A) Experimental set up for morpholino injections and amputations. Two sets of morpholinos were injected per gene to dorsally fated blastomeres at the 4-cell stage. Embryos were reared to stage 41 for amputation. Representative images of uninjured stage 41 (B) control, (C, D) Meis1 morphants 1 & 2, and (E, F) Pbx3 morphants 1 & 2. Images are shown in bright field and immunostained for neurofilament. Stage 41 tadpoles were amputated and collected at (G-K) 24hpa, (L-P) 48hpa, and (Q-U) 72hpa. The cyan channel represents tails stained with DAPI and the green channel represents tails immunostained for neurofilament. Scale bar: 75um. (V-Z) Brightfield images of the 72hpa tails. Scale bar: 50um. (AA) Boxplots representing length of regenerated tail or spinal cord for each tadpole. Statistics represent a two-tailed t-test to determine significance between conditions. (*<0.05, **<0.005, ***<0.0005, ****<0.00005).

Table 3-1 Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Xenopus tropicalis, female/male)*	Wild-type, Nigerian	NASCO	NASCO LM00822	
Strain, strain background (*Xenopus tropicalis, female/male)*	Strain, strain background (*Xenopus tropicalis, female/male)*	National *Xenopus* Resource, PMID: 12454928	RRID:NXR_1.0021	
Antibody	Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	ThermoFisher Cat. A-11001	RRID:AB_2534069	1:500 2hours at RT
Antibody	Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	ThermoFisher Cat. A-21422	RRID:AB_2535844	1:500 2hours at RT
Antibody	Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary antibody, Alexa Fluor 594	ThermoFisher Cat. A-11012	RRID:AB_2534079	1:500 2hours at RT
Antibody	Anti-Neurofilament associated, mouse monoclonal	Developmental Studies Hybridoma Bank, Cat. 3A10	RRID:AB_531874	1:50 overnight at 4C
Antibody	Anti-Doublecortin, rabit polyclonal	Cell Signaling Technology, 4604S	RRID:AB_561007	1:200 overnight at 4C
Antibody	Anti-Sox2, rabit polyclonal	Cell Signaling Technology, 2748S	RRID:AB_823640	1:100 overnight at 4C
Antibody	Anti-Histone H3 (tri methly K9, Phospho S10), mouse monoclonal	Abcam 14955	RRID:AB_443110	1:1000 overnight at 4C

Commerical assay or kit	DAPI	Sigma D9542	Sigma D9542	1:2000 10min at RT
Commerical assay or kit	Dextran (fluoro-tracer)	ThermoFisher D1817	ThermoFisher D1817	injected at 2mg/mL
Software, algorithm	Cellranger (v2.0.2)	10x Genomics	RRID:SCR_017344	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
Software, algorithm	Seurat (V3.0)	PMID: 31178118	RRID:SCR_016341	https://satijalab.org/seurat/
Software, algorithm	Trim Galore!	Felix Krueger	RRID:SCR_011847	https://github.com/FelixKrueger/TrimGalore
Software, algorithm	Bowtie2 (V2)	PMID: 22388286 ; PMID: 30020410	RRID:SCR_005476	http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml
Software, algorithm	MACS2 (V2)	PMID: 18798982	RRID:SCR_013291	https://github.com/taoliu/MACS
Software, algorithm	GenomicRanges	PMID: 23950696		https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
Software, algorithm	edgeR	PMID: 19910308; PMID: 22287627	RRID:SCR_012802	https://bioconductor.org/packages/release/bioc/html/edgeR.html
Software, algorithm	gprofiler2 (V2)	PMID: 31066453	RRID:SCR_018190	https://cran.r-project.org/web/packages/gprofiler2/index.html
Software, algorithm	ReviGO	PMID: 21789182	RRID:SCR_005825	http://revigo.irb.hr/
Software, algorithm	HOMER (V4.10)	PMID: 20513432	RRID:SCR_010881	http://homer.ucsd.edu/homer/

Sequence-based reagent	PMID: 24097267: Ad2.1; Nextera barcode 701	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.2; Nextera barcode 702	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.3; Nextera barcode 703	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.4; Nextera barcode 704	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.5; Nextera barcode 705	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.6; Nextera barcode 706	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.7; Nextera barcode 707	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.8; Nextera barcode 708	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.9; Nextera barcode 709	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT

Sequence-based reagent	PMID: 24097267: Ad2.10; Nextera barcode 710	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.11; Nextera barcode 711	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad2.12; Nextera barcode 712	PMID: 24097267	library primers	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGGAGATGT
Sequence-based reagent	PMID: 24097267: Ad1_noMX	PMID: 24097267	library primers	AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG
Sequence-based reagent	gfp_1F	this paper	RT-qpcr primers	ATGGCCAACACTTGTCACTA
Sequence-based reagent	gfp_1R	this paper	RT-qpcr primers	GGCATGGCACTCTTGAAAAA
Sequence-based reagent	pax6_1F	this paper	RT-qpcr primers	AGACGGACAGACTGGCAAAC
Sequence-based reagent	pax6_1R	this paper	RT-qpcr primers	TTCCATGTGCAAAAAGTCCA
Sequence-based reagent	actc_1F	this paper	RT-qpcr primers	GAGACCCTCTTCCAGCCTTC
Sequence-based reagent	actc_1R	this paper	RT-qpcr primers	GGTCCTTGCGGATATCAATG
Sequence-based reagent	tubb2b_1F	this paper	RT-qpcr primers	GCTGAAAGCAACATGAACGA
Sequence-based reagent	tubb2b_1R	this paper	RT-qpcr primers	CATCTTCCCCTTCCTCTTCC
Sequence-based reagent	odc1_1F	PMID: 23559567	RT-qpcr primers, PMID: 23559567: odc1	AGGCCACACTGGCAACTCA
Sequence-based reagent	odc1_1R	PMID: 23559567	RT-qpcr primers, PMID: 23559567: odc1	TGCGCTCAGTTCTGGTACTTCA

Sequence-based reagent	pbx3_ish_F	this paper	In situ hybridization primers	TGCAGGCAAC TACAGTCCAG
Sequence-based reagent	pbx3_ish_R	this paper	In situ hybridization primers	TAATACGACT CACTATAGGG AGTAGTGATG TGCGGGTTGG
Sequence-based reagent	meis1_ish_F	this paper	In situ hybridization primers	ACGGGACTTC TGTGCATACC
Sequence-based reagent	meis1_ish_R	this paper	In situ hybridization primers	TAATACGACT CACTATAGGG GTGGCTGTTTT CTGCAATGA
Sequence-based reagent	l1cam_ish_F	this paper	In situ hybridization primers	CTCACTAGAG TGGACGGGGA
Sequence-based reagent	l1cam_ish_R	this paper	In situ hybridization primers	TAATACGACT CACTATAGGG AGTGCCTGAC ACAAGGACAG
Sequence-based reagent	uchl1_ish_F	this paper	In situ hybridization primers	AAATGTCCCT GGGGTGTTCC
Sequence-based reagent	uchl1_ish_R	this paper	In situ hybridization primers	TAATACGACT CACTATAGGG CACGGAGTTG TGAGCAGACT
Sequence-based reagent	nsg1_ish_F	this paper	In situ hybridization primers	GGGCATGGCT GAGATGACAA
Sequence-based reagent	nsg1_ish_R	this paper	In situ hybridization primers	TAATACGACT CACTATAGGG CGTAGGGCAA GGGGTAAGTG
Sequence-based reagent	ass1_ish_F	this paper	In situ hybridization primers	CCGGAATGTG AGTTTGTGCG
Sequence-based reagent	ass1_ish_R	this paper	In situ hybridization primers	TAATACGACT CACTATAGGG ACTGCGCAAA GGGGTCTTTA
Sequence-based reagent	pbx3 MO1	this paper	morpholino	CAAATCGTCG TACCTTTGAG CCATC
Sequence-based reagent	pbx3 MO2	this paper	morpholino	TGACAACACA GACTGCAAGG AGACC

Sequence-based reagent	pbx3 MO2 VIVO	this paper	vivo-morpholino	TGACAACACA GACTGCAAGG AGACC
Sequence-based reagent	meis1 MO1	this paper	morpholino	CATACTCTTCT CCGCAGCCTA ATCA
Sequence-based reagent	meis1 MO2	this paper	morpholino	AACTCCTTCCT ACTTCAACTT CAGA
Sequence-based reagent	meis1 MO2 VIVO	this paper	vivo-morpholino	AACTCCTTCCT ACTTCAACTT CAGA
Sequence-based reagent	pbx3 gRNA	this paper	sgRNA primers	CTAGCTAATA CGACTCACTA TAGGTGCCCG GCCCTTCTGT GGGTTTTAGA GCTAGAA
Sequence-based reagent	meis1 gRNA	this paper	sgRNA primers	CTAGCTAATA CGACTCACTA TAGGCATGGG GATCGCCATA CAGTTTTAGA GCTAGAA
Sequence-based reagent	pbx3_hrm_1F	this paper	pcr primers	TCTCTTTCGAG TTAACGCATG A
Sequence-based reagent	pbx3_hrm_1R	this paper	pcr primers	GCACCCTCAG CTTGTAGAGA AT
Sequence-based reagent	meis1_hrm_1F	this paper	pcr primers	TTACACAGTA CGACGATTTG CC
Sequence-based reagent	meis1_hrm_1R	this paper	pcr primers	TAAGGGAAGA AGTGACAGAG GG
Software, algorithm	Data analysis and figure production	this paper	source code and data analysis	https://gitlab.com/akakebee/kakebeen-et-al-2019

Chapter 4. CONCLUDING REMARKS AND FUTURE DIRECTIONS

In this thesis, I utilized cell type-specific genomics and transcriptomics to investigate the prioritization of cell fate decisions in neural progenitor cells in regeneration, and to identify novel transcription factors involved in spinal cord regeneration. I took a multi-modal approach using genomics and transcriptomics to address the guiding hypothesis that neural progenitors have temporally restricted gene regulation in regeneration. My work has informed our basic biological understanding of regeneration in two major ways: (1) technical innovation to study gene regulation in specific cell types and (2) novel insight into how regeneration uniquely repurposes developmental programs. In this final chapter, I reflect on how my work weaves into the broader framework of regeneration research and the new questions which my conclusions elicit.

4.1 CELL TYPE-SPECIFIC DISCOVERY APPROACH

Genomics and transcriptomics are becoming more common in regeneration research due to their ability to generate many hypotheses from one set of experiments. Advancements in *Xenopus* genomes and adaptation of next generation sequencing techniques has made discovery-based approaches tractable for genome-wide profiling of regenerating tissue. Prior to my study, bulk genomics and transcriptomics approaches were conducted in whole tails or whole spinal cords (Chang et al., 2017; Lee-Liu et al., 2014; Love et al., 2011). Bulk transcriptomics boasts the ability to detect master-regulators of regeneration which may be present across various cell types and organize a regenerative program. However, these approaches lack the ability to resolve less pronounced gene expression changes that may be specific to one cell type. Since publication of those bulk studies, a scRNA-Seq study presented a gene-expression atlas for the regenerating

Xenopus laevis whole-tail at single cell level (Aztekin et al., 2019). While the atlas strategy was used to uncover novel cell-types that may be important for regeneration, it did not predict specific transcription factors or regulatory elements that interpret and respond to injury cues.

In contrast, my work used scRNA-Seq and cell-type specific ATAC-Seq to provide new insight into a cell type known to be critical to regeneration: the neural progenitor cell. ATAC-Seq analysis of neural progenitor cells allowed me to look genome-wide at shifts in chromatin accessibility and identify the processes the accessible chromatin promotes. Additionally, I was able to identify transcription factor binding sites with temporal-specific accessibility and their target genes. By intersecting this chromatin accessibility data with gene expression changes in the scRNA-Seq, I added a valuable dimension to both predictions of the cellular processes temporally prioritized by neural progenitor cells and predictions of temporal sequences of gene-regulatory cascades. Although the approaches I used yielded novel hypotheses, I am excited by emerging technical innovations that will further resolve the picture of regeneration.

Proteomics are becoming more common in *Xenopus* research, particularly in uncovering global protein dynamics across embryological development (Peshkin et al., 2015; Peshkin et al., preprint 2019; Sonnett et al., 2018; Sun et al., 2014). Previously, quantitative proteomics has been performed in the regenerating and non-regenerating *X. laevis* spinal cord, revealing unique sets of protein dynamics between the two contexts (Lee-Liu et al., 2018). (Peshkin et al., 2015; Peshkin et al., preprint 2019). Both this study as well as the developmental studies reveal that gene expression does not always accurately predict protein expression, suggesting that a direct read out for protein expression is necessary to more accurately predict regulatory networks. These studies also share a lack of cell-type specific protein information. Similar to how scRNA-Seq and cell-type specific bulk approaches increase resolution to changes in rare cell types, cell-types specific

proteomics will be able to better resolve dynamic protein changes that may not be picked up otherwise (Wilson and Nairn 2018). Having quantitative proteomic data in addition to genomic and transcriptomic data can be used to more accurately predict networks underlying regeneration and identify high fidelity targets.

One other technique I believe to be crucial in the upcoming years is the adaptation of multiplexed-FISH. With the ever-growing list of scRNA-Seq atlases, we are given a plethora of transcriptionally defined cell-types without high throughput ways to spatially identify expression of marker genes and thus identify new cell types *in situ.* While fluorescent *in situ* hybridization provides greater spatial resolution than colorimetric *in situ* hybridization*,* visualizing individual markers presents a bottleneck to identifying new cells types which require several markers to distinguish. Emerging multiplexed techniques, such as SABER-FISH (Kishi et al., 2019), provide a way to visualize several different transcripts at a time in the same tissue which would help add spatial information to transcriptomics data and provide an avenue to identify where transcriptionally defined cell types are *in situ*.

4.2 REGENERATION UNIQUELY UTILIZES DEVELOPMENTAL PROGRAMS

My gene ontology analysis of differentially accessible regions of the chromatin led to the hypothesis that neural progenitor differentiation was prioritized prior to proliferation. The addition of quantification of the relative number of transcriptionally defined neural progenitor cells and neurons in regeneration as well as predicted cell-cycle phases added valuable complementary data to this hypothesis. Hypothesizing that neural differentiation is being prioritized prior to proliferation is interesting due to the contrary nature of the temporal timing of differentiation and proliferation in regeneration with respect to development. If we followed the hypothesis that

regeneration is the recapitulation of development, we would have expected the regenerating tail would undergo a necessary stage of proliferation to expand the progenitor pool prior to committing some of those cells to differentiation to then give the spinal cord its function. However, our data predicts the opposite ordering of cell fate decisions. Encouragingly, over the course of this work Pelzer et al., and Aztekin et al., produced transcriptional profiling studies that support our hypothesis. Analysis of Pelzer's scRNA-Seq and bulk RNA-Seq of the regenerating spinal cord further demonstrate an early transcriptional prioritization of neural differentiation prior to proliferation in the regenerating NF stage 50 *X. tropicalis* spinal cord (Pelzer et al., preprint 2020). Aztekin's scRNA-Seq atlas of the regenerating *X. laevis* tail shows that over the first day of regeneration the relative population of several neuronal clusters increases (Aztekin et al., 2019). Together, our studies support a novel temporal ordering of neural progenitor cell fate decisions in regeneration.

While the first next steps would be to characterize phases of differentiation and proliferation in situ, the main line of interrogations that this hypothesis necessitates, is testing whether the temporal ordering of neural progenitor differentiation and proliferation is critical to completion of successful regeneration. Namely, questioning if prematurely promoting neural progenitor cell proliferation and restricting differentiation disrupt how the spinal cord regenerates. Previous studies in regeneration of the tail and limb in *Xenopus* and other systems suggest regeneration is nerve dependent (Farkas and Monaghan, 2017). Thus, we might predict that establishing early axonal connections from trunk to regeneration bud is necessary for regeneration to continue and restriction of differentiation and axonogenesis would inhibit regeneration. Our dataset provides a rich reference to examine gene expression changes when modulating neural

progenitor cell proliferation and differentiation. Thus, we can continue to explore the functional and regulatory effects of how the temporal ordering of cell fate decisions confers regeneration.

Speculatively, this hypothesis also serves an interesting purpose to promote the study of evolution and regeneration. The formation of the ependymal tube-neural ampulla followed by differentiation and axonogenesis (as described in this thesis) may represent basic components of regenerative capacity with respect to spinal cord regeneration. In the anole lizard and gecko, spinal cord regeneration results in a functional, however incomplete spinal cord composed of a simple ependymal tube with axonal tracts lining it (Duffy et al., 1990; Tokuyama et al., 2018; Wang et al., 2012). Thus, our model coupled with what is known in the lizard suggests that regeneration may be an ancestral trait which has been adapted for increasing neural complexity rather than an adaptive trait (Tanaka and Ferretti, 2009). An interesting follow up to begin testing this hypothesis would be to do a comparative study of the temporal preference of neural progenitor cells for differentiation and proliferation in excellent regenerators such as axolotls and newts to ask whether this same module is conserved.

4.3 IDENTIFICATION OF NOVEL TRANSCRIPTION FACTORS IN REGENERATION

In addition to uncovering process prioritization, I also used the multi-modal approach at the level of individual transcription factors to address one of the goals of this project: identify candidate transcription factor regulators of neural progenitor cell fate decisions in regeneration. Identification of differentially accessible transcription factor binding motifs and their targets were identified by ATAC-Seq and used in parallel with differential expression analysis to predict gene regulatory networks in neural progenitor cells. The two transcription factors we were most interested in studying were *pbx3* and *meis1*. *Meis1* has been previously implicated in neural

development in *Xenopus* (Maeda et al., 2002) and *pbx3* in zebrafish and mice (French et al., 2007; Rhee et al., 2004). While *meis1* has been studied in cardiomyocyte regeneration and limb regeneration (Aksoz et al., 2018; Mercader et al., 2005; Paul et al., 2020), neither of these genes have been investigated in neural regeneration. We first observed that morpholino knockdowns of either transcription factor resulted in changes in the neural plate and the domain of *pax6* expression as well as lack of developing eyes in neurulation to tailbud stage. In regeneration we found that loss of either transcription factor inhibited successful regeneration. More interestingly, *pbx3* and *meis1* morphants exhibited a loss of intersomitic and spinal cord axons after delivery of morpholino; amputated axons grew back, but in a wildly irregular way. Although we detected a change in axonal morphology, we did not observe a significant change in the rate of proliferation in the regenerate suggesting that these transcription factors are involved in the neuron development side of neural progenitor cell fate decisions.

Results from the preliminary investigation of *pbx3* and *meis1* suggest several intriguing follow up directions to their function in regeneration. Evidence from knockdown experiments suggest the transcription factors have necessary roles in both neural development as well as regeneration, however it remains unclear whether they perform unique roles in each context Elucidating roles in development and regeneration will provide an interesting platform to test context-specific roles at the level of the transcription factor. Interestingly, Meis and Pbx family members are known to work in co-complexes to regulate transcription of their target genes in development and other contexts (Chang et al., 1997; Maeda et al., 2002). Specifically, Pbx3 and Meis1 have been shown to work together as co-factors to bind DNA with Hoxa9 in leukemia (Garcia-Cuellar et al., 2015; Li et al., 2016). In the tadpole, Hoxa9 is known to pattern the posterior tail including the spinal cord (Lombardo and Slack 2001). Given its role in development and

characterizations of family member *hoxa13* in regeneration (Christen et al., 2003), we suspect this factor may be involved in patterning the regenerating spinal cord as well. The known interactions of these three factors suggest tempting questions in regeneration: (1) Does knockdown of *meis1* and *pbx3* have synergistic inhibitory effects in regeneration? (2) Do *meis1* and *pbx3* both need to be present to function in regeneration? (3) Is *hoxa9* a necessary factor for regeneration? (4) Is the presence of *meis1* and *pbx3* necessary for *hoxa9* to function?

The targeted characterization of how these factors work together in a complex in regeneration could be followed up by dissection of regulatory networks they are involved in. The goal of my work was to predict gene regulatory networks in regeneration by computational analysis, however the predicted networks are left to be tested in the tadpole. Further analysis of the datasets presented in this thesis can serve to scaffold a more open-ended investigation of upstream regulators and downstream targets of *meis1* and *pbx3* in regeneration. In depth interrogation of the transcription factor networks involving *meis1* and *pbx3* will provide novel molecular detail to how axonogenesis and axonal patterning are regulated in tail regeneration.

4.4 FUTURE USE OF GENERATED DATA-SETS

One of the most common questions I received when presenting my thesis work can be summed up by “Have you analyzed your data in X way”. This question demonstrates the power of omics approaches in that a different perspective taken to the same data set can yield new and interesting hypotheses. Besides the suggested future directions based on conclusions I have already made, I hope these data sets can be mined to identify other novel transcriptional factors or regulatory patterns observed during neural progenitor cells in regeneration. With growing computational toolkits and folks with bioinformatic experience, these data sets can be integrated

with previously published and soon-to-be published data sets for fresh perspective on how the tadpole regenerates its spinal cord.

4.5 MILE HIGH VIEW

The main goal of this work is to understand the biological mechanisms underlying naturally-occurring regeneration; however, the driving motivation to study this topic is the inability of humans to regenerate the spinal cord. Each year, the U.S. sees 17,730 new cases of spinal cord injury resulting in varying degrees of paralysis, surmounting medical debt, and immense changes to quality of life (National Spinal Cord Injury Statistical Center, Facts and Figures at a Glance. Birmingham, AL: University of Alabama at Birmingham, 2019). While I do not expect work from *Xenopus* to jump straight to non-human apes, *Xenopus* can play a unique role in informing how our clinicians develop therapies. The stage-specific regenerative competence of *Xenopus* (discussed in section 1.1) presents a unique research model to understand how regeneration works and test different hypotheses for rescue of regeneration in the same system. The Larraín group at Pontificia Universidad Católica de Chile has pioneered several uses of this property. Comparative RNA-Seq and proteomics between the regenerative and non-regenerative spinal cord have revealed differentially regulated genes and proteins which may be strong targets for stimulating regeneration (Lee-Liu et al., 2014; 2018). Most excitingly, the Larraín group created a system to test if putative factors restore spinal cord regeneration in non-regenerative frogs (Méndez-Olivos et al., 2017). In this study, the authors added cells from dissociated spinal cords from regenerative tadpoles to a thrombin clot in an adult transected spinal cord and found that *sox2*$^{+}$ cells proliferated and could even differentiate; however, this did not confer recovery. Although adding the neural stem cells and growth factors to the injury did not promote successful regeneration, both

approaches serve to reduce the search space for factors that may be critical in restoring regeneration.

The main conclusions I have presented in this thesis are interesting mechanisms of how regeneration occurs, however we did not test if they can promote regeneration in a non-regenerative context. The hypothesis that neural differentiation precedes proliferation suggests that to promote regeneration: (1) a neural progenitor population must exist or be supplied, (2) factors promoting neural differentiation should be supplied early while restricting proliferation and (3) a phase of controlled proliferation must be stimulated when axons have breached the transection area. Using the Larraín lab method, an interesting direction would be to supply neural progenitor cells and attempt to add and tune factors which push neural progenitors toward differentiation and proliferation. Although complex, this approach may provide novel insight into how we can control neural progenitor cell fates in an injury context. Our second hypothesis, that *meis1* and *pbx3* are critical for axonogenesis and neural patterning in regeneration, suggests that they may also be important in promoting regeneration in a non-regenerative context. This could also be tested with the Larraín lab method by introducing neural progenitor cells from regenerative tadpoles as well as either transcription factor and assaying for axonal patterning. I suspect that testing the role of individual factors and if they behave similarly in a non-regenerative context will become common-place soon enough.

Beyond my own results, continuing to develop the study of spinal cord regeneration with a multidisciplinary approach will be crucial to understanding how regeneration works and how we can promote it. For example, basic biologists may provide insight into factors to add and when to add them, but biomedical engineers are developing substrates to deliver these factors. Multidisciplinary approaches have been employed in mice to develop therapies to restore

regeneration (Ashammakhi et al., 2019), however these lack the benefit of reference for what regenerative processes should look like. Thus, while I believe our basic biological understanding of regeneration will provide a blueprint, I think *Xenopus* researchers could take a step farther and start new collaborations across fields to make the work more impactful, and hopefully lead to clinical gains.

Chapter 5. REFERENCES

Aksoz, M., Turan, R.D., Albayrak, E., and Kocabas, F. (2018). Emerging Roles of Meis1 in Cardiac Regeneration, Stem Cells and Cancer. Curr Drug Targets *19*, 181–190.

Alaynick, W.A., Jessell, T.M., and Pfaff, S.L. (2011). SnapShot: spinal cord development. Cell *146*, 178-178.e1.

Alunni, A., and Bally-Cuif, L. (2016). A comparative view of regenerative neurogenesis in vertebrates. Development *143*, 741–753.

Ashammakhi N, Kim HJ, Ehsanipour A, Bierman RD, Kaarela O, Xue C, Khademhosseini A, Seidlits SK (2019). Regenerative Therapies for Spinal Cord Injury. Tissue Eng Part B Rev. 25(6):471-491.

Aztekin, C., Hiscock, T.W., Marioni, J.C., Gurdon, J.B., Simons, B.D., and Jullien, J. (2019). Identification of a regeneration-organizing cell in the Xenopus tail. Science *364*, 653–658.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol.

Beck, C.W., Christen, B., and Slack, J.M.W. (2003). Molecular Pathways Needed for Regeneration of Spinal Cord and Muscle in a Vertebrate. Developmental Cell *5*, 429–439.

Beck, C.W., Izpisúa Belmonte, J.C., and Christen, B. (2009). Beyond early development: Xenopus as an emerging model for the study of regenerative mechanisms. Developmental Dynamics : An Official Publication of the American Association of Anatomists *238*, 1226–1248.

Beliakova-Bethell, N., Massanella, M., White, C., Lada, S., Du, P., Vaida, F., Blanco, J., Spina, C.A., and Woelk, C.H. (2014). The effect of cell subset isolation method on gene expression in leukocytes. Cytometry A *85*, 94–104.

Bel-Vialar, S., Medevielle, F., and Pituello, F. (2007). The on/off of Pax6 controls the tempo of neuronal differentiation in the developing spinal cord. Dev. Biol. *305*, 659–673.

Bhattacharya, D., Marfo, C.A., Li, D., Lane, M., and Khokha, M.K. (2015). CRISPR/Cas9: An inexpensive, efficient loss of function tool to screen human disease genes in Xenopus. Dev. Biol. *408*, 196–204.

Binek, A., Rojo, D., Godzien, J., Rupérez, F.J., Nuñez, V., Jorge, I., Ricote, M., Vázquez, J., and Barbas, C. (2019). Flow Cytometry Has a Significant Impact on the Cellular Metabolome. J. Proteome Res. *18*, 169–181.

Bonner, W.A., Hulett, H.R., Sweet, R.G., and Herzenberg, L.A. (1972). Fluorescence activated cell sorting. Rev Sci Instrum *43*, 404–409.

Briggs, J.A., Weinreb, C., Wagner, D.E., Megason, S., Peshkin, L., Kirschner, M.W., and Klein, A.M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. Science *360*.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Cannata, S.M., Bagni, C., Bernardini, S., Christen, B., and Filoni, S. (2001). Nerve-independence of limb regeneration in larval Xenopus laevis is correlated to the level of fgf-2 mRNA expression in limb tissues. Developmental Biology *231*, 436–446.

Cardozo MJ, Mysiak KS, Becker T, Becker CG (2017). Reduce, reuse, recycle - Developmental signals in spinal cord regeneration. Dev Biol. 432(1):53-62. PMID: 28502615.

Cerda, G.A., Hargrave, M., and Lewis, K.E. (2009). RNA profiling of FAC-sorted neurons from the developing zebrafish spinal cord. Dev. Dyn. *238*, 150–161.

Chang, C.P., Jacobs, Y., Nakamura, T., Jenkins, N.A., Copeland, N.G., and Cleary, M.L. (1997). Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. Mol. Cell. Biol. *17*, 5679–5687.

Chang, J., Baker, J., and Wills, A. (2017). Transcriptional dynamics of tail regeneration in *Xenopus tropicalis*. Genesis *55*, e23015.

Chen, Y., Love, N.R., Amaya, E., Gurtner, G.C., Werner, S., Barrandon, Y., Longaker, M.T., Brockes, J.P., Kumar, A., Alvarado, A.S., et al. (2014). Tadpole tail regeneration in Xenopus. Biochemical Society Transactions *42*, 617–623.

Chernoff, E.A.G., Sato, K., Salfity, H.V.N., Sarria, D.A., and Belecky-Adams, T. (2018). Musashi and Plasticity of Xenopus and Axolotl Spinal Cord Ependymal Cells. Front. Cell. Neurosci. *12*.

Christen, B., Beck, C.W., Lombardo, A., and Slack, J.M.W. (2003). Regeneration-specific expression pattern of three posterior Hox genes. Developmental Dynamics : An Official Publication of the American Association of Anatomists *226*, 349–355.

Chung, M.-I., Kwon, T., Tu, F., Brooks, E.R., Gupta, R., Meyer, M., Baker, J.C., Marcotte, E.M., and Wallingford, J.B. (2014). Coordinated genomic control of ciliogenesis and cell movement by RFX2. Elife *3*, e01439.

Contreras, E.G., Gaete, M., Sánchez, N., Carrasco, H., and Larraín, J. (2009). Early requirement of Hyaluronan for tail regeneration in Xenopus tadpoles. Development *136*, 2987–2996.

Currie, J.D., Kawaguchi, A., Traspas, R.M., Schuez, M., Chara, O., and Tanaka, E.M. (2016). Live Imaging of Axolotl Digit Regeneration Reveals Spatiotemporal Choreography of Diverse Connective Tissue Progenitor Pools. Dev. Cell *39*, 411–423.

Delile, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J., and Sagner, A. (2019). Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. Development *146*.

DENT, J.N. (1962). Limb regeneration in larvae and metamorphosing individuals of the South African clawed toad. Journal of Morphology *110*, 61–77.

Dhorne-Pollet, S., Thélie, A., and Pollet, N. (2013). Validation of novel reference genes for RT-qPCR studies of gene expression in Xenopus tropicalis during embryonic and post-embryonic development. Dev. Dyn. *242*, 709–717.

Duffy, M.T., Simpson, S.B., Liebich, D.R., and Davis, B.M. (1990). Origin of spinal cord axons in the lizard regenerated tail: supernormal projections from local spinal neurons. J. Comp. Neurol. *293*, 208–222.

Duncan, A.R., and Khokha, M.K. (2016). Xenopus as a model organism for birth defects-Congenital heart disease and heterotaxy. Semin. Cell Dev. Biol. *51*, 73–79.

Dwaraka, V.B., Smith, J.J., Woodcock, M.R., and Voss, S.R. (2018). Comparative transcriptomics of limb regeneration: Identification of conserved expression changes among three species of Ambystoma. Genomics.

Erickson, T., French, C.R., and Waskiewicz, A.J. (2010). Meis1 specifies positional information in the retina and tectum to organize the zebrafish visual system. Neural Dev *5*, 22.

Ericson, J., Rashbass, P., Schedl, A., Brenner-Morton, S., Kawakami, A., van Heyningen, V., Jessell, T.M., and Briscoe, J. (1997). Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. Cell *90*, 169–180.

Farkas, J.E., and Monaghan, J.R. (2017). A brief history of the study of nerve dependent regeneration. Neurogenesis (Austin) *4*, e1302216.

Fei, J.-F., Schuez, M., Tazaki, A., Taniguchi, Y., Roensch, K., and Tanaka, E.M. (2014). CRISPR-mediated genomic deletion of Sox2 in the axolotl shows a requirement in spinal cord neural stem cell amplification during tail regeneration. Stem Cell Reports *3*, 444–459.

Filoni, S., and Bosco, L. Comparative analysis of the regenerative capacity of caudal spinal cord in larvae of serveral Anuran amphibian species. Acta Embryologiae et Morphologiae Experimentalis (“Halocynthia” Association") *2*, 199–226.

Filoni, S., and Bosco, L. (1981). Comparative analysis of the regenerative capacity of caudal spinal cord in larvae of serveral Anuran amphibian species. Acta Embryol Morphol Exp *2*, 199–226.

French, C.R., Erickson, T., Callander, D., Berry, K.M., Koss, R., Hagey, D.W., Stout, J., Wuennenberg-Stapleton, K., Ngai, J., Moens, C.B., et al. (2007). Pbx homeodomain proteins pattern both the zebrafish retina and tectum. BMC Dev. Biol. *7*, 85.

Freshney, R.I. (2015). Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications (John Wiley & Sons).

Gaete, M., Muñoz, R., Sánchez, N., Tampe, R., Moreno, M., Contreras, E.G., Lee-Liu, D., and Larraín, J. (2012). Spinal cord regeneration in Xenopus tadpoles proceeds through activation of Sox2-positive cells. Neural Dev *7*, 13.

Gallardo, V.E., and Behra, M. (2013). Fluorescent Activated Cell Sorting (FACS) Combined with Gene Expression Microarrays for Transcription Enrichment Profiling of Zebrafish Lateral Line Cells. Methods *62*, 226–231.

Garcia-Cuellar, M.-P., Steger, J., Füller, E., Hetzner, K., and Slany, R.K. (2015). Pbx3 and Meis1 cooperate through multiple mechanisms to support Hox-induced murine leukemia. Haematologica *100*, 905–913.

Gargioli, C., and Slack, J.M.W. (2004). Cell lineage tracing during Xenopus tail regeneration. Development *131*, 2669–2679.

Gemberling, M., Bailey, T.J., Hyde, D.R., and Poss, K.D. (2013). The zebrafish as a model for complex tissue regeneration. Trends Genet *29*.

Gerber, T., Murawala, P., Knapp, D., Masselink, W., Schuez, M., Hermann, S., Gac-Santel, M., Nowoshilow, S., Kageyama, J., Khattak, S., et al. (2018). Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. Science *362*.

Gilbert, D.M. (1986). Temporal order of replication of Xenopus laevis 5S ribosomal RNA genes in somatic cells. Proc Natl Acad Sci U S A *83*, 2924–2928.

Godsave, S.F., and Slack, J.M. (1989). Clonal analysis of mesoderm induction in Xenopus laevis. Dev. Biol. *134*, 486–490.

Grainger, R.M. (2012). Xenopus tropicalis as a Model Organism for Genetics and Genomics: Past, Present and Future. Methods Mol Biol *917*, 3–15.

Graham V, Khudyakov J, Ellis P, Pevny L (2003). SOX2 functions to maintain neural progenitor identity. Neuron. 39(5):749-65.

Hadjantonakis, A.K., and Nagy, A. (2000). FACS for the isolation of individual cells from transgenic mice harboring a fluorescent protein reporter. Genesis *27*, 95–98.

Hardwick, L.J.A., Ali, F.R., Azzarelli, R., and Philpott, A. (2015). Cell cycle regulation of proliferation versus differentiation in the central nervous system. Cell Tissue Res *359*, 187–200.

Harland, R.M., and Grainger, R.M. (2011). Xenopus Research: Metamorphosed by Genetics and Genomics. Trends Genet *27*, 507–515.

Hartley, K.O., Hardcastle, Z., Friday, R.V., Amaya, E., and Papalopulu, N. (2001). Transgenic Xenopus embryos reveal that anterior neural development requires continued suppression of BMP signaling after gastrulation. Dev. Biol. *238*, 168–184.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hirsch, N., Zimmerman, L.B., Gray, J., Chae, J., Curran, K.L., Fisher, M., Ogino, H., and Grainger, R.M. (2002). Xenopus tropicalis transgenic lines and their use in the study of embryonic induction. Dev. Dyn. *225*, 522–535.

Ho, D.M., and Whitman, M. (2008). TGF-beta signaling is required for multiple processes during Xenopus tail regeneration. Developmental Biology *315*, 203–216.

Horb, M., Wlizla, M., Abu-Daya, A., McNamara, S., Gajdasik, D., Igawa, T., Suzuki, A., Ogino, H., Noble, A., Centre de Ressource Biologique Xenope team in France, et al. (2019). Xenopus Resources: Transgenic, Inbred and Mutant Animals, Training Opportunities, and Web-Based Support. Front Physiol *10*, 387.

Johnson, A.D., and Krieg, P.A. (1994). pXeX, a vector for efficient expression of cloned sequences in Xenopus embryos. Gene *147*, 223–226.

Kakebeen, A., and Wills, A. (2019a). Advancing genetic and genomic technologies deepen the pool for discovery in Xenopus tropicalis. Dev. Dyn. *248*, 620–625.

Kakebeen, A.D., and Wills, A.E. (2019b). More Than Just a Bandage: Closing the Gap Between Injury and Appendage Regeneration. Front Physiol *10*, 81.

Kakebeen, A.D., Chitsazan, A.D., Williams, M.C., Saunders, L.M., and Wills, A.E. (2020). Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors. Elife *9*.

Kay, B., and Peng, H. (1991). Appendix A: solutions and protocols. In Xenopus Laevis: Practical Uses in Cell and Molecular Biology, pp. 661–662.

Kelly, L.E., Carrel, T.L., Herman, G.E., and El-Hodiri, H.M. (2006). Pbx1 and Meis1 regulate activity of the Xenopus laevis Zic3 promoter through a highly conserved region. Biochem. Biophys. Res. Commun. *344*, 1031–1037.

Khazaei, M., Ahuja, C.S., Rodgers, C.E., Chan, P., and Fehlings, M.G. (2019). Generation of Definitive Neural Progenitor Cells from Human Pluripotent Stem Cells for Transplantation into Spinal Cord Injury. Methods Mol. Biol. *1919*, 25–41.

Khokha, M.K., Chung, C., Bustamante, E.L., Gaw, L.W.K., Trott, K.A., Yeh, J., Lim, N., Lin, J.C.Y., Taverner, N., Amaya, E., et al. (2002). Techniques and probes for the study of Xenopus tropicalis development. Dev. Dyn. *225*, 499–510.

Kishi, J.Y., Lapan, S.W., Beliveau, B.J., West, E.R., Zhu, A., Sasaki, H.M., Saka, S.K., Wang, Y., Cepko, C.L., and Yin, P. (2019). SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. Nat Methods *16*, 533–544.

Kline, R.A., Kaifer, K.A., Osman, E.Y., Carella, F., Tiberi, A., Ross, J., Pennetta, G., Lorson, C.L., and Murray, L.M. (2017). Comparison of independent screens on differentially vulnerable motor neurons reveals alpha-synuclein as a common modifier in motor neuron diseases. PLoS Genet *13*.

Knopf, F., Hammond, C., Chekuru, A., Kurth, T., Hans, S., Weber, C.W., Mahatma, G., Fisher, S., Brand, M., Schulte-Merker, S., et al. (2011). Bone regenerates via dedifferentiation of osteoblasts in the zebrafish fin. Dev. Cell *20*, 713–724.

Koniaris, L.G., McKillop, I.H., Schwartz, S.I., and Zimmers, T.A. (2003). Liver regeneration. J Am Coll Surg *197*, 634–659.

Kowalczyk, M.S., Tirosh, I., Heckl, D., Rao, T.N., Dixit, A., Haas, B.J., Schneider, R.K., Wagers, A.J., Ebert, B.L., and Regev, A. (2015). Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res *25*, 1860–1872.

Kragl M, Knapp D, Nacu E, Khattak S, Maden M, Epperlein HH, Tanaka EM (2009). Cells keep a memory of their tissue origin during axolotl limb regeneration. Nature. 460(7251):60-5.

Lai, H.C., Seal, R.P., and Johnson, J.E. (2016). Making sense out of spinal cord somatosensory development. Development *143*, 3434–3448.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357–359.

Langmead, B., Wilks, C., Antonescu, V., and Charles, R. (2019). Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics *35*, 421–432.

Lara-Ramirez, R., Pérez-González, C., Anselmi, C., Patthey, C., and Shimeld, S.M. (2019). A Notch-regulated proliferative stem cell zone in the developing spinal cord is an ancestral vertebrate trait. Development *146*.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

Lee-Liu, D., Moreno, M., Almonacid, L.I., Tapia, V.S., Muñoz, R., von Marées, J., Gaete, M., Melo, F., and Larraín, J. (2014). Genome-wide expression profile of the response to spinal cord

injury in Xenopus laevis reveals extensive differences between regenerative and non-regenerative stages. Neural Development *9*, 12.

Lee-Liu, D., Sun, L., Dovichi, N.J., and Larraín, J. (2018). Quantitative Proteomics After Spinal Cord Injury (SCI) in a Regenerative and a Nonregenerative Stage in the Frog Xenopus laevis. Mol Cell Proteomics *17*, 592–606.

Levi, A.D., Okonkwo, D.O., Park, P., Jenkins, A.L., Kurpad, S.N., Parr, A.M., Ganju, A., Aarabi, B., Kim, D., Casha, S., et al. (2018). Emerging Safety of Intramedullary Transplantation of Human Neural Stem Cells in Chronic Cervical and Thoracic Spinal Cord Injury. Neurosurgery *82*, 562–575.

Li, Z., Chen, P., Su, R., Hu, C., Li, Y., Elkahloun, A.G., Zuo, Z., Gurbuxani, S., Arnovitz, S., Weng, H., et al. (2016). PBX3 and MEIS1 Cooperate in Hematopoietic Cells to Drive Acute Myeloid Leukemias Characterized by a Core Transcriptome of the MLL-Rearranged Disease. Cancer Res. *76*, 619–629.

Lin, G., and Slack, J.M.W. (2008). Requirement for Wnt and FGF signaling in Xenopus tadpole tail regeneration. Developmental Biology *316*, 323–335.

Lin, G., Chen, Y., and Slack, J.M.W. (2007). Regeneration of neural crest derivatives in the Xenopus tadpole tail. BMC Dev Biol *7*, 56.

Llufrio, E.M., Wang, L., Naser, F.J., and Patti, G.J. (2018). Sorting cells alters their redox state and cellular metabolome. Redox Biol *16*, 381–387.

Lombardo A, Slack JM (2001). Abdominal B-type Hox gene expression in Xenopus laevis. Mech Dev. 106(1-2):191-5.

Love, N.R., Chen, Y., Bonev, B., Gilchrist, M.J., Fairclough, L., Lea, R., Mohun, T.J., Paredes, R., Zeef, L.A.H., and Amaya, E. (2011). Genome-wide analysis of gene expression during Xenopus tropicalis tadpole tail regeneration. BMC Developmental Biology *11*, 70.

Machon, O., Masek, J., Machonova, O., Krauss, S., and Kozmik, Z. (2015). Meis2 is essential for cranial and cardiac neural crest development. BMC Dev. Biol. *15*, 40.

Maeda, R., Mood, K., Jones, T.L., Aruga, J., Buchberg, A.M., and Daar, I.O. (2001). Xmeis1, a protooncogene involved in specifying neural crest cell fate in Xenopus embryos. Oncogene *20*, 1329–1342.

Maeda, R., Ishimura, A., Mood, K., Park, E.K., Buchberg, A.M., and Daar, I.O. (2002). Xpbx1b and Xmeis1b play a collaborative role in hindbrain and neural crest gene expression in Xenopus embryos. Proc. Natl. Acad. Sci. U.S.A. *99*, 5448–5453.

Manoli, M., and Driever, W. (2012). Fluorescence-Activated Cell Sorting (FACS) of Fluorescently Tagged Cells from Zebrafish Larvae for RNA Isolation. Cold Spring Harb Protoc *2012*, pdb.prot069633.

Martínez-Cerdeño V, Noctor SC (2018). Neural Progenitor Cell Terminology. Front Neuroanat. 2018;12:104.

McCusker, C., Bryant, S.V., and Gardiner, D.M. (2015). The axolotl limb blastema: cellular and molecular mechanisms driving blastema formation and limb regeneration in tetrapods. Regeneration (Oxf) *2*, 54–71.

McCusker CD, Diaz-Castillo C, Sosnik J, Q Phan A, Gardiner DM (2016). Cartilage and bone cells do not participate in skeletal regeneration in Ambystoma mexicanum limbs. Dev Biol. 416(1):26-33

McHedlishvili, L., Epperlein, H.H., Telzerow, A., and Tanaka, E.M. (2007). A clonal analysis of neural progenitors during axolotl spinal cord regeneration reveals evidence for both spatially restricted and multipotent progenitors. Development *134*, 2083–2093.

Méndez-Olivos, E. E., Muñoz, R., & Larraín, J. (2017). Spinal Cord Cells from Pre-metamorphic Stages Differentiate into Neurons and Promote Axon Growth and Regeneration after Transplantation into the Injured Spinal Cord of Non-regenerative Xenopus laevis Froglets. Frontiers in cellular neuroscience, 11, 398.

Mercader, N., Tanaka, E.M., and Torres, M. (2005). Proximodistal identity during vertebrate limb regeneration is regulated by Meis homeodomain proteins. Development *132*, 4131–4142.

Mitogawa, K., Makanae, A., Satoh, A., and Satoh, A. (2015). Comparative Analysis of Cartilage Marker Gene Expression Patterns during Axolotl and Xenopus Limb Regeneration. PLoS One *10*.

Mochii, M., Taniguchi, Y., and Shikata, I. (2007). Tail regeneration in the Xenopus tadpole. Dev. Growth Differ. *49*, 155–161.

Mojsin, M., and Stevanovic, M. (2009). PBX1 and MEIS1 up-regulate SOX3 gene expression by direct interaction with a consensus binding site within the basal promoter region. Biochem. J. *425*, 107–116.

Münch, J., González-Rajal, A., and de la Pompa, J.L. (2013). Notch regulates blastema proliferation and prevents differentiation during adult zebrafish fin regeneration. Development *140*, 1402–1411.

Muñoz, R., Edwards-Faret, G., Moreno, M., Zuñiga, N., Cline, H., and Larraín, J. (2015). Regeneration of Xenopus laevis spinal cord requires Sox2/3 expressing cells. Dev Biol *408*, 229–243.

Nakayama, T., Fish, M.B., Fisher, M., Oomen-Hajagos, J., Thomsen, G.H., and Grainger, R.M. (2013). Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis. Genesis *51*, 835–843.

Nenni, M.J., Fisher, M.E., James-Zorn, C., Pells, T.J., Ponferrada, V., Chu, S., Fortriede, J.D., Burns, K.A., Wang, Y., Lotay, V.S., et al. (2019). Xenbase: Facilitating the Use of Xenopus to Model Human Disease. Front Physiol *10*, 154.

Nieuwkoop, P.D., and Faber, J. (1956). Normal table of Xenopus laevis (daudin): A systematical and chronologica survey of the development from the fertilized egg till the end of metamorphosis (North Holland, Amsterdam).

Nieuwkoop, P.D., and Faber, J. (1994). Normal table of Xenopus laevis (daudin) (New York: Garland Publishing Inc).

Noelanders, R., and Vleminckx, K. (2018). Cell Cycle Analysis of the Embryonic Brain of Fluorescent Reporter Xenopus tropicalis by Flow Cytometry. In Xenopus: Methods and Protocols, K. Vleminckx, ed. (New York, NY: Springer), pp. 243–250.

Ny, A., Vandevelde, W., Hohensinner, P., Beerens, M., Geudens, I., Diez-Juan, A., Brepoels, K., Plaisance, S., Krieg, P.A., Langenberg, T., et al. (2013). A transgenic Xenopus laevis reporter model to study lymphangiogenesis. Biol Open *2*, 882–890.

Ogai, K., Nakatani, K., Hisano, S., Sugitani, K., Koriyama, Y., and Kato, S. (2014). Function of Sox2 in ependymal cells of lesioned spinal cords in adult zebrafish. Neurosci Res *88*, 84–87.

O'Hara, C.M., Egar, M.W., and Chernoff, E.A. (1992). Reorganization of the ependyma during axolotl spinal cord regeneration: changes in intermediate filament and fibronectin expression. Dev. Dyn. *193*, 103–115.

Osumi, N., Shinohara, H., Numayama-Tsuruta, K., and Maekawa, M. (2008). Concise review: Pax6 transcription factor contributes to both embryonic and adult neurogenesis as a multifunctional regulator. Stem Cells *26*, 1663–1672.

Paul, S., Zhang, X., and He, J.-Q. (2020). Homeobox gene Meis1 modulates cardiovascular regeneration. Semin Cell Dev Biol *100*, 52–61.

Pelzer, D., Phipps, L.S., Thuret, R., Baker, S.M., and Dorey, K. (2020). Foxm1 regulates neuronal progenitor fate during spinal cord regeneration. BioRxiv 2020.02.26.962977.

Peshkin L, Wühr M, Pearl E, Haas W, Freeman RM Jr, Gerhart JC, Klein AM, Horb M, Gygi SP, Kirschner MW (2015). On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development. Dev Cell. 35(3):383-94.

Peshkin, L., Lukyanov, A., Kalocsay, M., Gage, R.M., Wang, D., Pells, T.J., Karimi, K., Vize, P.D., Wühr, M., and Kirschner, M.W. (2019). The protein repertoire in early vertebrate embryogenesis (Developmental Biology).

Phipps, L.S., Marshall, L., Dorey, K., and Amaya, E. (2020). Model systems for regeneration: Xenopus. Development *147*.

Power, C., and Rasko, J.E.J. (2008). Whither prometheus' liver? Greek myth and the science of regeneration. Ann Intern Med *149*, 421–426.

Rataj-Baniowska, M., Niewiadomska-Cimicka, A., Paschaki, M., Szyszka-Niagolov, M., Carramolino, L., Torres, M., Dollé, P., and Krężel, W. (2015). Retinoic Acid Receptor β Controls Development of Striatonigral Projection Neurons through FGF-Dependent and Meis1-Dependent Mechanisms. J. Neurosci. *35*, 14467–14475.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198.

Reichard, A., and Asosingh, K. (2019). Best Practices for Preparing a Single Cell Suspension from Solid Tissues for Flow Cytometry. Cytometry Part A *95*, 219–226.

Reimer, M.M., Kuscha, V., Wyatt, C., Sörensen, I., Frank, R.E., Knüwer, M., Becker, T., and Becker, C.G. (2009). Sonic hedgehog is a polarized signal for motor neuron regeneration in adult zebrafish. J. Neurosci. *29*, 15073–15082.

Rhee, J.W., Arata, A., Selleri, L., Jacobs, Y., Arata, S., Onimaru, H., and Cleary, M.L. (2004). Pbx3 Deficiency Results in Central Hypoventilation. Am J Pathol *165*, 1343–1350.

Richardson, G.M., Lannigan, J., and Macara, I.G. (2015). Does FACS perturb gene expression? Cytometry A *87*, 166–175.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Salzberg, A., Elias, S., Nachaliel, N., Bonstein, L., Henig, C., and Frank, D. (1999). A Meis family protein caudalizes neural cell fates in Xenopus. Mech. Dev. *80*, 3–13.

Sargent, T.D., Jamrich, M., and Dawid, I.B. (1986). Cell interactions and the control of gene activity during early development of Xenopus laevis. Developmental Biology *114*, 238–246.

Sato, K., Uehara, A., Kinoshita, S., Nomura, I., Yagi, M., Tanizaki, Y., Matsuda-shoji, Y., Matsubayashi, A., Endo, N., Nagai, Y., et al. (2018). Flow cytometric analysis of Xenopus laevis and X. tropicalis blood cells using acridine orange. Scientific Reports *8*, 16245.

Selman, G.G., and Pawsey, G.J. (1965). The utilization of yolk platelets by tissues of Xenopus embryos studied by a safranin staining method. J Embryol Exp Morphol *14*, 191–212.

Shin, J.C., Kim, K.N., Yoo, J., Kim, I.-S., Yun, S., Lee, H., Jung, K., Hwang, K., Kim, M., Lee, I.-S., et al. (2015). Clinical Trial of Human Fetal Brain-Derived Neural Stem/Progenitor Cell Transplantation in Patients with Traumatic Cervical Spinal Cord Injury. Neural Plast. *2015*, 630932.

Sive, H.L., Grainger, R.M., and Harland, R.M. (2010). Defolliculation of Xenopus oocytes. Cold Spring Harb Protoc *2010*, pdb.prot5535.

Slack, J.M.W., Lin, G., and Chen, Y. (2008). The Xenopus tadpole: a new model for regeneration research. Cell. Mol. Life Sci. *65*, 54–63.

Sonnett M, Gupta M, Nguyen T, Wühr M (2018). Quantitative Proteomics for Xenopus Embryos II, Data Analysis. Methods Mol Biol. 1865:195-215.

Sousa, S., Afonso, N., Bensimon-Brito, A., Fonseca, M., Simões, M., Leon, J., Roehl, H., Cancela, M.L., and Jacinto, A. (2011). Differentiated skeletal cells contribute to blastema formation during zebrafish fin regeneration. Development *138*, 3897–3905.

Stedman, A., Lecaudey, V., Havis, E., Anselme, I., Wassef, M., Gilardi-Hebenstreit, P., and Schneider-Maunoury, S. (2009). A functional interaction between Irx and Meis patterns the anterior hindbrain and activates krox20 expression in rhombomere 3. Dev. Biol. *327*, 566–577.

Stefanelli, A. (1951). I fenomeni rigenerativi e degenerativi del midollo spinale caudale degli Anfibi e dei Rettili. Bolletino di zoologia *18*, 279–290.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Sun L, Bertke MM, Champion MM, Zhu G, Huber PW, Dovichi NJ (2014). Quantitative proteomics of Xenopus laevis embryos: expression kinetics of nearly 4000 proteins during early development. Sci Rep. 4:4365.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE *6*, e21800.

Suzuki, M., Yakushiji, N., Nakada, Y., Satoh, A., Ide, H., and Tamura, K. (2006). Limb Regeneration in *Xenopus laevis* Froglet. The Scientific World JOURNAL *6*, 26–37.

Tanaka, E.M. (2016). The Molecular and Cellular Choreography of Appendage Regeneration. Cell *165*, 1598–1608.

Tanaka, E.M., and Ferretti, P. (2009). Considering the evolution of regeneration in the central nervous system. Nat. Rev. Neurosci. *10*, 713–723.

Taniguchi, Y., Watanabe, K., and Mochii, M. (2014). Notochord-derived hedgehog is essential for tail regeneration in Xenopus tadpole. BMC Developmental Biology *14*, 27.

Thuret, R., Auger, H., and Papalopulu, N. (2015). Analysis of neural progenitors from embryogenesis to juvenile adult in Xenopus laevis reveals biphasic neurogenesis and continuous lengthening of the cell cycle. Biol Open *4*, 1772–1781.

Tokuyama, M.A., Xu, C., Fisher, R.E., Wilson-Rawls, J., Kusumi, K., and Newbern, J.M. (2018). Developmental and adult-specific processes contribute to de novo neuromuscular regeneration in the lizard tail. Dev Biol *433*, 287–296.

Tseng, A.-S., and Levin, M. (2008). Tail regeneration in Xenopus laevis as a model for understanding tissue repair. J. Dent. Res. *87*, 806–816.

Tsujioka, H., Kunieda, T., Katou, Y., Shirahige, K., and Kubo, T. (2015). Unique Gene Expression Profile of the Proliferating Xenopus Tadpole Tail Blastema Cells Deciphered by RNA-Sequencing Analysis. PLOS ONE *10*, e0111655.

Walther, C., and Gruss, P. (1991). Pax-6, a murine paired box gene, is expressed in the developing CNS. Development *113*, 1435–1449.

Wang, Y., Dong, Y., Song, H., Liu, Y., Liu, M., Yuan, Y., Ding, F., Gu, X., and Wang, Y. (2012). Involvement of gecko SNAP25b in spinal cord regeneration by promoting outgrowth and elongation of neurites. Int. J. Biochem. Cell Biol. *44*, 2288–2298.

Wilson, R. S., & Nairn, A. C. (2018). Cell-Type-Specific Proteomics: A Neuroscience Perspective. *Proteomes*, *6*(4), 51.

Yamada, T., Urano-Tashiro, Y., Tanaka, S., Akiyama, H., and Tashiro, F. (2013). Involvement of crosstalk between Oct4 and Meis1a in neural cell fate decision. PLoS ONE *8*, e56997.

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

VITA

Anneke D. Kakebeen grew up in the Los Angeles suburb of La Crescenta, California. She graduated from the University of California at San Diego in La Jolla, CA with a B.S in Biochemistry and a minor in Philosophy while running cross country and track. While at UCSD, she first interned as a Quality Control Assistant at Pharmatek Laboratories. She then shifted to academic science in the lab of Elizabeth Komives and studied the protein complex between Creatine Kinases Brain type and Ankyrin repeat Socs Box containing protein 9 by way of protein crystallography. Anneke proceeded straight from UCSD to begin her graduate studies at the University of Washington. While in the lab of Andrea Wills, Anneke discovered her love for developmental biology and became a certified embryologist by attending the Embryology Course at the Marine Biological Labs (Woods Hole, MA) in 2018. When not in lab, Anneke spends her time running, in the mountains climbing, cooking and eating any and all food.

Work published during this thesis:

Kakebeen AD, Chitsazan AD, Wills AE. Tissue disaggregation and isolation of specific cell types from transgenic Xenopus appendages for transcriptional analysis by FACS. Dev Dyn. 2020 Nov 2. doi: 10.1002/dvdy.268. Epub ahead of print. PMID: 33137227.

Kakebeen AD, Chitsazan AD, Williams MC*, Saunders LM, Wills AE. Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors. Elife. 2020;9:e52648. Published 2020 Apr 27. doi:10.7554/eLife.52648

Kakebeen A, Wills A. Advancing genetic and genomic technologies deepen the pool for discovery in Xenopus tropicalis. Dev Dyn. 2019;248(8):620-625. doi:10.1002/dvdy.80

Kakebeen AD, Wills AE. More Than Just a Bandage: Closing the Gap Between Injury and Appendage Regeneration. Front Physiol. 2019;10:81. Published 2019 Feb 8. doi:10.3389/fphys.2019.00081

Kakebeen AD, Huebner R, Shindo A, Kwon K, Kwon T, Wills AE, Wallingford JB**.** A temporally resolved transcriptome for developing "Keller" explants of the Xenopus laevis dorsal marginal zone. **Accepted 2020 December**

Williams MC, Patel JH, **Kakebeen AD**, Wills AE. Nutrient availability and sensing contribute to a graded refractory period for regeneration in *Xenopus tropicalis*. **in review**