UNIVERSITY OF CALIFORNIA SAN DIEGO

Using Differential Expression Analysis to Explore Temporal and Spatial Development in the Tadpole Brain

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Biology

by

Aaron Ta

Committee in charge:

Hollis Cline, Chair
Nicholas Spitzer, Co-Chair
Barry Grant

2020

The Thesis of Aaron Ta is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____
Co-chair

_____
Chair

University of California San Diego

2020

DEDICATION

This thesis is dedicated to Michelle, Tom, and Felicity Ta, for their unconditional love and support.

TABLE OF CONTENTS

List of Figures

List of Abbreviations

FB: Forebrain

MB: Midbrain

HB: Hindbrain

SC: Spinal Cord

NPC: Neural Progenitor Cell

IN: Immature Neuron

ACKNOWLEDGEMENTS

I would like to thank Dr. Hollis Cline for her thoughtful insight and mentorship as my advisor. In addition, I would like to acknowledge Dr. Nicholas Spitzer and Dr. Barry Grant for serving as members of my committee. I would also like to acknowledge Dr. Caroline McKeown for her continued support and guidance throughout this entire project, and Dr. Lin-Chien Huang for helping me through the first steps of the data processing and analysis. Finally, I would like to thank the current and prior members of the Cline Lab for their valuable input and suggestions: Dr. Masateru Hiramoto, Dr. Regina Faulkner, Dr. Pranav Sharma, Dr. Lucio Shiapparelli, Dr. Haiyan He, Ana Verduzco, Ana Araujo, Yi Xie, Vi Dang, Kailey Archuleta, and Erika Barth. This thesis would not have been possible without their contributions.

The introduction, results, and discussion sections are currently being prepared for submission for publication of the material. The thesis author was the primary investigator and author of this material.

The methods section is coauthored with Huang, Lin-Chien. The thesis author was the primary author of this material.

ABSTRACT OF THE THESIS

Using Differential Expression Analysis to Explore
Temporal and Spatial Development in the Tadpole Brain

by

Aaron Ta

Master of Science in Biology

University of California San Diego, 2020

Professor Hollis Cline, Chair
Professor Nicholas Spitzer, Co-Chair

Amphibian metamorphosis is a transitional period that involves significant changes in the cell type populations and biological processes occurring in the brain. However, gene expression dynamics during this process are not as well-studied as they are in earlier development. To rectify this, we sought to quantify gene expression in the X. laevis brain over this time period in two ways: firstly over stages of development in the midbrain, and secondly across regions of the brain at a single developmental stage. We found that genes pertaining to positive regulation of neuronal proliferation as well as known progenitor cell markers were upregulated prior to metamorphic climax in the midbrain; concurrently, expression of cell cycle timing regulators decreased in expression across this period, supporting the notion that cell cycle lengthening contributes to a decrease in proliferation by the end of metamorphosis. We also found that at the start of metamorphosis, neural progenitor populations appeared to be similar across the fore-,

mid-, and hindbrain. Genes pertaining to negative regulation of differentiation were upregulated in the spinal cord compared to the rest of the brain, however, suggesting that a different program of neurogenesis regulation may be occurring there. Finally, we found that regulation of biological processes like cell fate commitment and synaptic signaling follow similar trajectories in the brain across early tadpole metamorphosis and mid- to late-embryonic mouse development. By comparing expression across both temporal and spatial conditions, we were able to illuminate cell type and biological pathway dynamics in the brain during metamorphosis.

The introduction, methods, results, and discussion sections are currently being prepared for submission for publication of the material. The thesis author was the primary investigator and author of this material.

Introduction

Xenopus laevis, the African clawed toad, is a ubiquitous model for neural studies (Blum & Ott, 2018; Pratt & Khakhalin, 2013); it possesses a nervous system broadly analogous to that of *Homo sapiens* and is both cheaper and quicker to raise than common mammalian models like *Mus musculus* (Blum & Ott, 2018), making it a resource-efficient model organism. As a popular model in neurobiological research, gene expression in the entire *X. laevis* body at successive stages of early development has been well-documented in literature (Pratt & Khakhalin, 2013), and is compiled in the online database Xenbase (Bowes et al., 2011; Segerdell et al., 2013). Prior experiments have utilized this data in differential expression analyses in order to identify genes that are specific to certain developmental time points or individual regions of the body (Briggs et al., 2018). In the *X. laevis* brain specifically, RNA expression has been measured in the first three days of post-fertilization development, to NF stage 44 (Nieuwkoop & Faber, 1994; Session et al., 2018). However, expression in the brain following this developmental period has not yet been thoroughly quantified, nor has expression in individual regions of the brain (as opposed to a whole-brain approach) been examined.

Metamorphosis is a significant transitional period in amphibian development, marked by drastic changes in physiology and gene expression across the body, including the central nervous system (Yaoita, 2018). Xenopus metamorphosis can be subdivided into three stages. The first, premetamorphosis, occurs from NF stages 46 to 53/54 and is marked by low levels of mitosis in the brain (Miyata, 2012; Thuret, 2015). It precedes prometamorphosis (NF stages 53/54 – 57/58), a period in which neurogenesis dramatically increases (Wen, 2019). The third stage is the metamorphic climax (NF stages 57/58 – 66), in which proliferation and neurogenesis fall below even premetamorphic levels (Thuret, 2015). The metamorphic climax is also significant in that it

is thought to parallel many aspects of perinatal mammalian development, including a reorganization of the nervous system accompanying the transition from aquatic to terrestrial life (Yaoita, 2018; Holzer, 2013). Given these observations, metamorphosis appears to be an important period in *Xenopus* brain development that remains understudied.

We thus sought to accomplish two goals with this research project. Firstly, we wished to expand the existing dataset of *X. laevis* gene expression over brain development and region, quantifying expression at developmental stages following 3 days post-fertilization as well as in the fore-, mid-, hindbrain, and spinal cord regions. Secondly, we wished to utilize this data to gain insight on what cell types and biological processes are altered over the course of metamorphosis. As changes in neuronal proliferation and differentiation rates have already been shown to occur during this period, we were particularly interested in exploring how these changes might be regulated at the level of gene expression. We utilized neural progenitor cell- (NPC) and immature neuron-associated (IN) genes in addition to known cell type markers in order to investigate relative changes in their population over development. Additionally, given the similar programs of expression between the tadpole brain at metamorphic climax and the perinatal mouse brain, we examined if similar patterns of expression also exist in the prior developmental stages of the prometamorphic tadpole and late-prenatal mouse. Utilizing both temporal comparisons across development as well as spatial comparisons across brain regions, then, we sought to illuminate the chronology and locality in which these cell types and biological processes are regulated in the developing *Xenopus* brain.

The introduction section is currently being prepared for submission for publication of the material. The thesis author was the primary investigator and author of this material.

Methods

Sample Collection and Processing

Three tissue samples were collected from the *X. laevis* midbrain at each of the NF stages 44, 46, 55, and 61. Samples were also taken from the forebrain, hindbrain, and spinal cord at stage 46. A second set of samples had been previously collected from animals enriched in either immature neurons (INs) or neural progenitor cells (NPCs) as detailed in Sharma and Cline, 2010. In brief, tadpoles were reared in darkness for 2 consecutive days to obtain heightened proliferation and NPC enrichment, or alternatively reared with enhanced visual stimulation to obtain increased differentiation and IN enrichment.

We aligned the samples against the *X. laevis* v9.2 genome assembly on Xenbase (Karimi et al., 2018) using two read mapping programs, STAR v2.5.2a (Dobin et al., 2012) and HISAT2 v2.0.4 (Langmead & Salzberg, 2015). As both aligners yielded comparable results, we chose to proceed with STAR. We then counted and assigned the alignments to genes using HTSeq v0.11.1 (Anders et al, 2015). Finally, we normalized the counts using the DESeq2 v1.22.2 package (Young et al., 2010) for further analysis.

For the comparison to mouse brain development, we utilized RNA-seq data collected from the cerebrum and cerebellum at timepoints from embryonic day 13.5 to 18.5 (PRJEB26869) (Heidelberg, 2019; Leihonen, 2011). We aligned the samples to the Ensembl GRCm38.p6 assembly (Howe et al., 2020) and analyzed the data using the same pipeline as we did for *X. laevis*.

Data Analysis and Visualization

We also used DESeq2 to perform differential analysis. Unless otherwise specified, we considered genes with both a log2 fold change magnitude of > 2 and a FDR < 0.05 to be differentially expressed. We used DEGreport v1.18.1 (Pantano, 2019) with default settings for cluster analysis of differential expression analysis results. We used GOseq v1.34.1(Young et al., 2010) to identify GO-term enrichment in sets of differentially-expressed genes. To determine enrichment of gene lists within one another, we used a two-sided Fisher's exact test. We used the R package pheatmap v1.0.12 (Kolde, 2013) to generate the heatmaps.

Interspecies Gene Conversion

James A. Briggs and colleagues have previously mapped *Xenopus tropicalis* gene symbols to the *H. sapiens* genome. To identify homologous genes between *X. laevis* and *H. sapiens*, we first aligned *X. laevis* gene symbols using BLAST to the closely-related *X. tropicalis* genome. This table was then used to map the *X. laevis* gene symbols to *H. sapiens*, with *X. tropicalis* as an intermediary. Of the 91611 *X. laevis* transcripts, 90.77% (83159) were mapped to the *X. tropicalis* transcriptome with high confidence. 70.23% (64399) of the *X. laevis* transcripts were ultimately mapped to *H. sapiens* gene symbols. We used biomaRt v2.38.0 (Kinsella, 2011) to identify homologous genes between *M. musculus* and *H. sapiens*, as a bridge between the mouse and tadpole gene lists.

The materials methods section is currently being prepared for submission for publication of the material. The thesis author was the primary investigator and author of this material.

The methods section is coauthored with Huang, Lin-Chien. The thesis author was the primary author of this material.

Results

Identification of Neural Progenitor Cell- and Immature Neuron-Associated Genes

We have previously isolated and sequenced RNA from stage 46 midbrain samples enriched in neural progenitor cells (NPCs) or immature neurons (INs) (Huang, 2016). We performed differential expression analysis between these groups to identify transcripts that were upregulated in one compared to the other, considering them NPC-associated or IN-associated. In total, we identified 547 NPC-associated and 865 IN-associated genes. We found that the NPC-associated genes included genes already known to be involved in neuronal differentiation such as *wnt1* and *neurod1*; similarly, the IN-associated genes included genes involved in proliferation like *bmp4* and *elk1* (Figure 2). This helped validate our classification of these genes as cell type-associated. As the genes in each set are differentially expressed when comparing NPC- and IN-enriched conditions, we sought to use these associated genes to detect enrichment of one cell type or the other across brain development and region.

Changes in Gene Expression in the Developing Tadpole Midbrain

We compared samples taken from the midbrain at NF stages 44, 46, 55, and 61 (Figure 1.1) through differential expression analysis. Principal component analysis showed that the samples clustered by developmental stage rather than batch (Figure 3.1). In total, we found 3,358 genes that were differentially expressed between at least two stages. The analysis also revealed that samples taken from stage 44 and 46 were similar in expression to one another. Only 4 genes were differentially expressed between these two stages, about tenfold less than any other pairwise analysis. Thus, we considered the two stages jointly in comparisons with stages 55 and 61.

We found that genes that were differentially expressed between two or more stages were enriched for GO terms pertaining to cell division, development, and cell cycle transitions (Figure 3.2). In particular, we found that positive regulation of cell proliferation as well as negative regulation of cell differentiation were highly enriched. Cluster analysis of genes involved in positive regulation of proliferation revealed three distinct groups that exhibited high expression specifically at stages 44 and 46, stage 55, or stage 61 (Figure 3.3). The group with elevated expression at the early stages were highly enriched in cell cycle regulation genes, including a number of cyclins and cyclin-dependent kinases. The groups with elevated expression at stages 55 or 61 were enriched in genes involved in the PI3K-AKT signaling pathway. In addition to broad processes such as proliferation being differentially expressed over midbrain development, more specific pathways like PI3K-AKT also showed temporal enrichment at particular stages of development.

We next investigated how relative populations of NPCs and INs may change in the tadpole midbrain over development. We used the cell type-associated gene lists obtained previously as indicators of a transition from a timepoint enriched in one cell type to a timepoint enriched in the other. Testing for enrichment, we found that IN-associated genes were significantly depleted both over genes that were downregulated in stages 44 and 46 as well as in genes that were upregulated at stage 61 ($p < 0.0001$).

Finally, we examined how other known cell type markers are expressed in the midbrain over time. Using a list of neural cell type markers (Abcam, 2020), we tested for enrichment over the developmental stages (Figure 3.4). We found that glial markers were enriched in stages 55 ($p = 0.003$) and 61 ($p = 0.008$) compared to the earlier two stages, while progenitor cell markers were enriched in stages 44, 46, and 55 compared to stage 61 ($p = 0.0006$). Specifically, radial

glia ($p = 0.001$) and immature neuron markers ($p = 0.008$) were enriched in these three earlier stages compared to the latest stage.

Differences in Regional Gene Expression at NF Stage 46

We also performed differential expression analysis between samples taken at NF stage 46 from the fore-, mid-, hindbrain, and spinal cord. Principal component analysis showed clustering by brain region as expected (Figure 4.1). We identified genes that were differentially expressed at one brain region against all other brain regions. From 1,267 genes that were differentially expressed between at least two brain regions, we found that 138 genes were specifically differentially expressed in the forebrain; 301 in the midbrain; 45 in the hindbrain; 279 in the spinal cord. Anterior-posterior planning genes were present in the expected regions: *hox* genes in the forebrain and midbrain, and *otx* genes in the hindbrain and spinal cord (Schilling & Knight, 2001) (Figure 4.3). Testing for GO term enrichment, in addition to terms pertaining to regional development and pattern specification, many genes related to neurogenesis and fate commitment were also differentially expressed between brain regions (Figure 4.2).

We found that NPC-associated genes were depleted in the set of genes upregulated in the midbrain compared to both the forebrain ($p = 0.0282$) and the hindbrain ($p = 0.0058$), but not the spinal cord. Using the list of cell markers, we found that progenitor markers ($p = 0.002$) and neuronal markers ($p < 0.0001$) were enriched amongst genes that were differentially expressed across regions. Specifically, progenitor/immature neuron markers were downregulated in the spinal cord compared to the other brain regions ($p < 0.0001$). Dopaminergic neuron markers were enriched in the midbrain and hindbrain, while cholinergic neuron markers were enriched in the hindbrain and spinal cord ($p < 0.0001$) (Figure 4.4).

Comparing Temporal Expression Between X. laevis and M. musculus

Previous studies have shown that similar genes are upregulated in the postnatal rodent brain as in the *Xenopus* brain during metamorphosis (Yaoita, 2018). Similar patterns of thyroid hormone-associated gene expression have been observed in both species during the first three postnatal weeks in mice and the metamorphic climax (NF stages 57/58 – stage 66) in *Xenopus* (Miyata, 2012; Holzer, 2013). Given this, we sought to examine if this interspecies parallel extends earlier into the developmental process.

Separately in the mouse cerebrum and cerebellum, we identified genes that were upregulated at each of E13.5/E14.5, E15.5/E16.5, and E17.5/E18.5 compared to the other timepoints. From 3525 total differentially expressed genes in the cerebrum and 2339 in the cerebellum, 1092 displayed the same pattern of expression (upregulated at the same timepoints) in both regions; by requiring the genes to have similar expression in both regions, we aimed to exclude region-specific genes and obtain a more general picture of temporal expression across the brain. We found that genes upregulated at each timepoint displayed enrichment in distinct GO biological pathways. Fate commitment and differentiation were enriched at E13.5/E14.5; lipoproteins including Apoa1 and Apob were enriched at E15.5/E16.5; synaptic signaling was enriched at E17.5/E18.5 (Figure 5.3).

Examining these genes over *Xenopus* midbrain development, many of them displayed a similar pattern of temporal expression (Figure 5.4). Genes upregulated at E13.5/E14.5 in the mouse model significantly tended to also be upregulated at stages 44 and 46 in the tadpole midbrain ($p = 0.0001$), while those upregulated at E17.5/E18.5 were significantly upregulated at stage 61 ($p < 0.0001$). The intermediate E15.5/E16.5 genes were somewhat biased towards

stages 44 and 46 ($p = 0.04$), but also had a fair number of genes highly expressed at the later timepoints as well.

Comparing Regional Expression Between X. laevis and M. musculus

We also compared samples between the *M. musculus* cerebrum and cerebellum at the six developmental stages between E13.5 and E18.5. We found 2,290 genes that were differentially expressed between at least two stages in the cerebellum and 3,585 genes in the cerebrum, both enriched in GO terms pertaining to nervous system development and regionalization. PCA analysis showed expected clustering by developmental stage in both regions (Figure 5.1). In the tadpole forebrain, we found 107 genes that were differentially expressed between stages 46 and 49; in the hindbrain, there were 192 differentially expressed genes between these stages.

Comparing the mouse cerebrum and cerebellum, differentially expressed genes were highly conserved at each timepoint. 183 genes were differentially expressed between the two regions at all timepoints, composing a robust set of genes that varied in expression regionally but not temporally ($p < 0.0001$). Examining homologous genes in *X. laevis*, we found that about a third of these genes were also differentially expressed between the hindbrain and forebrain at stage 46 ($p < 0.0001$). These genes consisted primarily of homeobox and transcriptional regulation genes, and as such were significantly enriched in GO terms pertaining to regionalization and pattern specification (Figure 5.2). As there was not a significant difference in the proportion of genes shared between any particular timepoint in the mouse model and the tadpole model at stage 46 (Figure 5.3), these homeobox genes seem to possess a region-specific pattern of expression that is conserved both across species as well as development. Excluding these invariable genes, we identified three additional homeobox genes that were differentially

expressed between the tadpole stage 46 forebrain and hindbrain as well as between the cerebrum and cerebellum at a single developmental stage in mouse: Hoxb5 at E17.5, Hoxe4 at E18.5, and Hoxd4 at E18.5.

The results section is currently being prepared for submission for publication of the material. The thesis author was the primary investigator and author of this material.

Discussion

Gene Expression Over X. Laevis Midbrain Development

We observed a sharp decrease in radial glial and immature neuron markers at stage 61 in the tadpole midbrain compared to the three earlier stages. Prior studies have shown that both neuronal birth and neural progenitor proliferation decrease dramatically between stages 56 and 66 in the posterior hindbrain (Thuret, 2015) following prometamorphosis, as well as by stage 66 in the optic tectum (D'Amico, 2011). As such, this observation seems in line with these previous findings. However, we did not clearly observe a similarly-reported secondary wave of neurogenesis at the onset of prometamorphosis, between stages 54 and 56. Though Stage 55 samples displayed the highest expression of many radial glia and immature neuron markers, this difference was not significant compared to stages 44 and 46. As proliferation activity decreases from its observed peak at stage 54 in the hindbrain to premetamorphosis levels by stage 56 (Thuret, 2015), it is possible that this decrease is more abrupt than previously suspected. Proliferation in the midbrain may also express an early-shifted pattern of activity compared to the hindbrain, peaking and returning to pre-metamorphic levels within the stage 46 to stage 54 window. Collection of additional data at timepoints within this window across the different brain regions would allow investigation of both possibilities.

Using our cell type-associated gene lists, we observed significant depletion of IN-associated genes in the set of genes downregulated at stages 44 and 46, as well as in the set upregulated at stage 61. It is likely that changes in expression pertaining to cell-type enrichment are somewhat obscured by other developmental changes; we observed no enrichment of the NPC-associated genes over development. However, this depletion does fall in line with our previous observations using marker genes, in that it suggests an increase in the relative IN

population does not occur between these timepoints. Future studies comparing less dynamic conditions than developmental timepoints may be able to leverage these progenitor- and immature neuron-associated genes with greater resolution.

Examining expression of proliferation-promoting genes across midbrain development, we found that cyclins and cyclin-dependent kinases decreased in expression over time. For instance, we found cyclin D1 and cyclin-dependent kinase 2 to be significantly downregulated post-prometamorphosis. Such genes are critical for the timing of progression through the cell cycle (Vernon, 2003). In mice, their inhibition has been shown to lead to a lengthening of the G1 phase and a decrease in proliferative division (Calegari & Huttner, 2003). This again agrees with the expected decline in proliferation after prometamorphosis. Previous studies have also found that the average cell cycle length in the *Xenopus* brain increases during this developmental period alongside proliferation (Thuret, 2015). This is in contrast to the primate brain, in which cell cycle length begins decreasing partway through cortical neurogenesis (Kornack & Rakic, 1998). Taken together, the downregulation of cyclin D1 and other G1 phase regulators in tandem with a decrease in neural progenitor markers seems to suggest that cell cycle length and neuronal proliferation in *Xenopus* may share a similar mechanistic relationship as in mice.

We also observed that many genes involved in the PI3K-AKT signaling pathway, a positive regulator of proliferation (Roberts, 2002), were upregulated in later stages at either stage 55 or 61. This included genes such as *rictor* and *fgf1*. A partial explanation for the discrepancy between the expected decrease in proliferation and the observed mRNA expression of proliferation-upregulating genes could be a concurrent change in post-transcriptional regulation. Studies in the closely-related anuran *Microhyla fissipes* found that during metamorphosis, miRNAs targeting the PI3K-AKT pathway were the most significantly enriched group,

indicating that miRNA regulation plays a significant role in metamorphic development (Liu, 2018). miRNA expression during *Xenopus* metamorphosis has previously been profiled (Hikosaka, 2007), but target genes for these miRNAs have yet to be identified. Characterization of these differentially-expressed miRNAs could reveal a similar increase in PI3K-AKT pathway targeting in *Xenopus* and underline the importance of post-transcriptional regulation in developmental processes.

Gene Expression Across X. Laevis Brain Regions

As expected, many genes that were differentially-expressed between brain regions at stage 46 pertained to pattern specification. In addition, many genes related to neurogenesis were also differentially expressed between brain areas at the same timepoint. We particularly noted that negative regulators of differentiation were highly expressed in the spinal cord compared to other regions. Furthermore, the spinal cord showed significantly reduced expression of intermediate progenitor markers like *eomes* at this timepoint. While *Xenopus* tadpoles are usually able to regenerate spinal cord injuries prior to adulthood, it is known that they experience a refractory period between stages 45 and 47 in which this ability is briefly lost (Beck, 2003). Previous studies have found that following tail amputation in non-refractory period tadpoles, proliferation-regulating pathways such as Wnt and BMP become active in the regenerating area; significantly lower activity of such pathways is observed in refractory period individuals, indicating that inhibition of cell proliferation likely plays a role in this phenomenon (Kabakeen & Wills, 2019; Lin & Slack, 2008). Our observations suggest that differentiation-inhibiting activity may also be already present in the uninjured spinal cord during this period independent of injury, which may contribute to the transient loss of regenerative ability during this period.

Further investigation of gene expression in the pre- and post-refractory period spinal cord would help illuminate if this inhibition is temporally specific to this timeframe in addition to being regionally-specific.

The depletion of NPC-associated genes in genes differentially expressed in the stage 46 midbrain compared to the forebrain and hindbrain may suggest that the relative population of NPCs at this timepoint is similar across all three regions. It has previously been shown that from stage 52/53 to 54/55 at the tail end of premetamorphosis, NPC populations are similar in the *X. laevis* telencephalon, diencephalon, tectum, and cerebellum (Denver, 2009). Additionally, we have previously discussed that neural progenitor proliferation activity does not appear to change significantly in the midbrain from stages 44-55 (Figure 3.4), and that a period of quiescence precedes a secondary wave of neurogenesis and decreased proliferation in the *Xenopus* hindbrain during prometamorphosis (Thuret, 2015). As stage 46 constitutes the start of premetamorphosis, it is plausible that neural progenitor populations remain similar both across all three brain regions as well as developmental timepoints during this period. As previously mentioned, there remains the possibility that a sharp spike in proliferation may occur at NF stage 54 or earlier, which we would not be able to detect with our current data. Again, additional timepoints are necessary to investigate this further.

Comparing Gene Expression Between X. laevis and M. musculus

We found that the dominant biological processes over brain development during E13.5 – E18.5 in mouse and NF stages 44 – 61 in *Xenopus* were generally similar. In both timeframes, expression of neuron cell fate commitment and differentiation-related genes decreased consistently with each consecutive stage. Meanwhile, synaptic signaling was highest at stage 61

in tadpoles and at the last timepoints in mice; this pattern has previously been observed comparing ISH data in embryonic mice (Thompson, 2014). This is not to say that the two periods parallel one another in development, but rather that they appear to possess similar patterns of overall expression regarding their major enriched pathways. In fact, enrichment of E15.5/E16.5 genes in the tadpole stage 44/46 midbrain rather than the intermediate stage 55 suggests that this span of mouse brain development may more closely resemble an earlier window of tadpole development.

Regarding differences in regional expression between the species, we observed a high proportion of genes that were differentially expressed in both tadpole and mouse. Considering the robust conservation of homeobox genes across species (Santini, 2003), this was to be expected. The only homeobox genes that were not consistently differentially expressed between the mouse cerebrum and cerebellum at all time points were hoxb5, hoxc4, and hoxd4. All are associated with hindbrain expression in the brain (Sjöstedt, 2020), and hoxc4 and hoxd4 have additionally been shown to continue increasing in hindbrain expression during the first postnatal stages in mice (Sunkin, 2012). Because these genes were only differentially expressed between the cerebrum and cerebellum at E17.5 and E18.5 in addition to the tadpole midbrain at stage 46, it is possible that regional development of the premetamorphic tadpole brain at this time is broadly similar to that of the late prenatal mouse. This would support the previous notion that the analyzed window of mouse development is in fact more similar to an earlier period of tadpole development. Further studies experimentally validating this observed differential expression and tracing its downstream effects will help illuminate the extent of this comparability between models.

The discussion section is currently being prepared for submission for publication of the material. The thesis author was the primary investigator and author of this material.

Figures and Tables

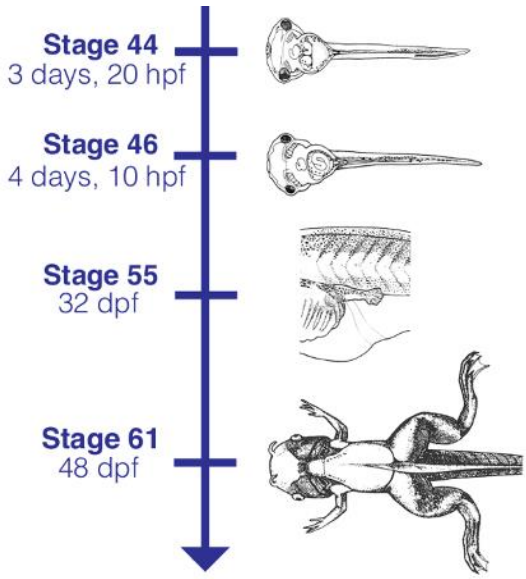


Figure 1.1: Development of *X. laevis* at each NF stage sequenced. Adapted from the *Xenopus Anatomy Ontology Database* by E. Segerdell et al. Retrieved June 4, 2019, from https://www.xenbase.org/anatomy/alldev.do. Copyright Xenbase 2020.

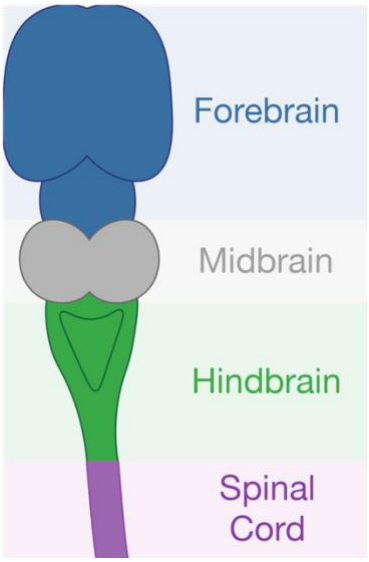


Figure 1.2: Schematic depicting the four regions of the NF stage 46 *X. laevis* brain sequenced.

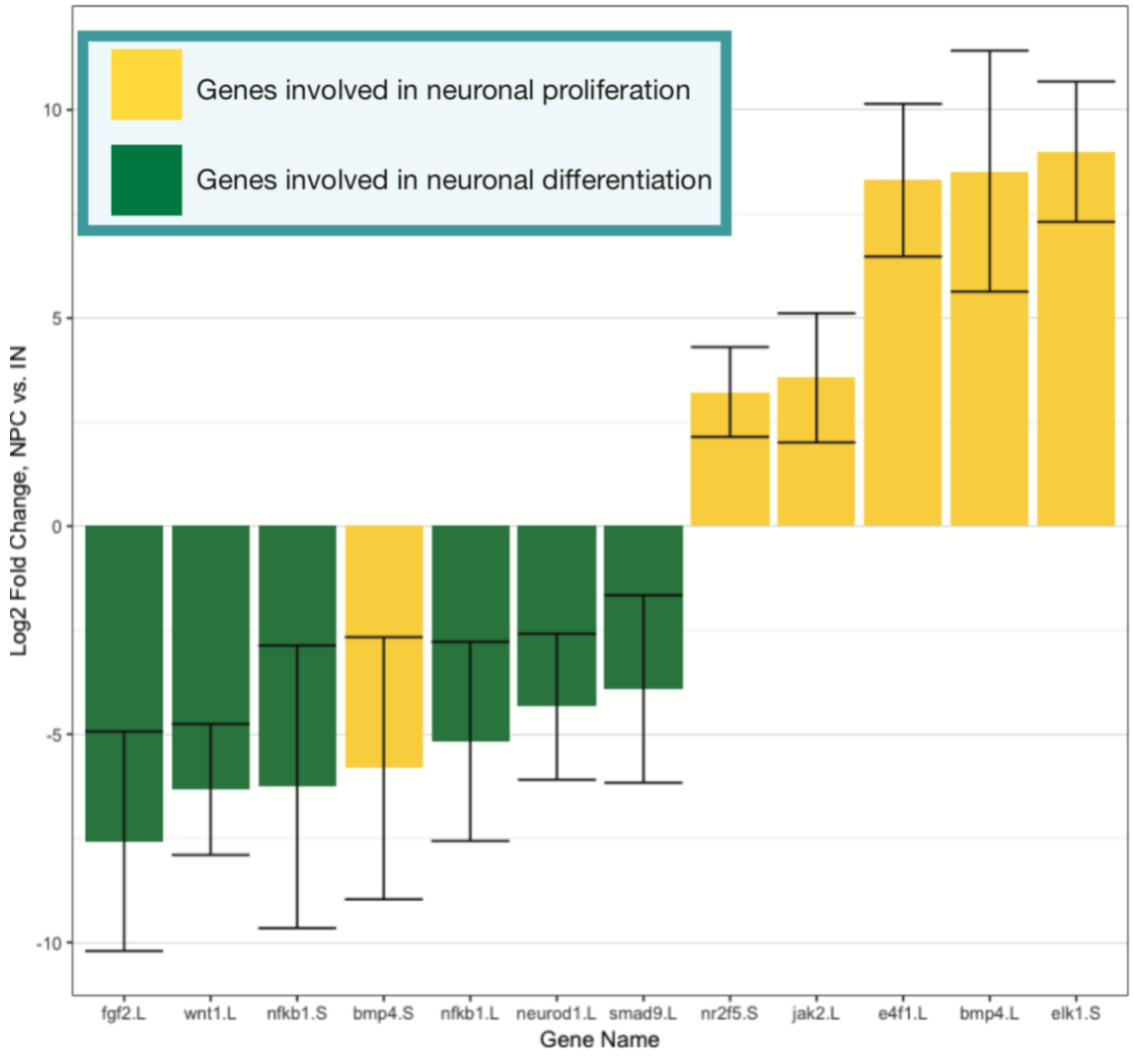


Figure 2: Expression of genes identified in existing literature as associated with either neuronal proliferation or neuronal differentiation. Yellow corresponds to proliferation, and green to differentiation. A negative log2 fold change indicates upregulation in NPCs compared to INs, while a positive log2 fold change indicates upregulation in INs compared to NPCs. Error bars represent standard error of the log2 fold change.

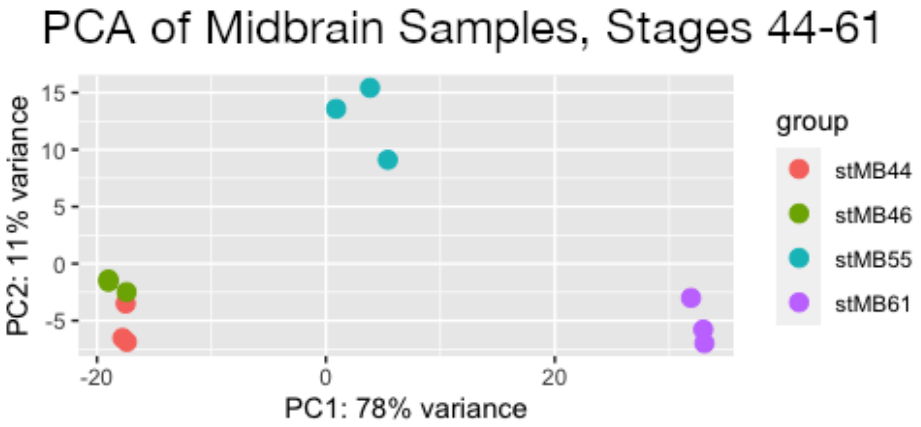


Figure 3.1: Principal component analysis of samples collected from the *X. laevis* midbrain at developmental stages NF 44-61.

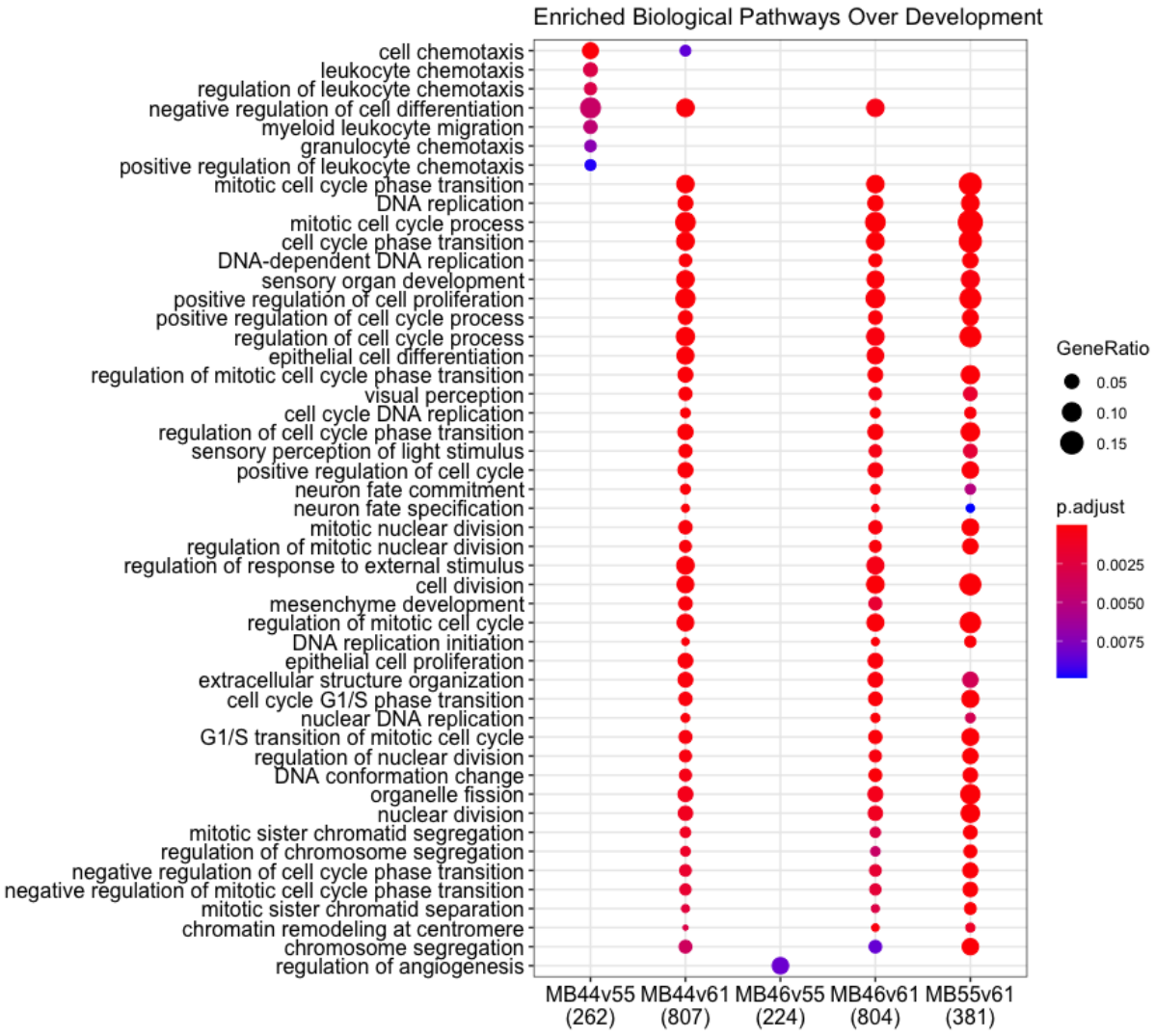


Figure 3.2: GO biological pathways enriched in the set of genes differentially expressed between two stages of tadpole midbrain development. Dot size corresponds to the proportion of differentially expressed genes associated with the biological pathway, and dot color to the adjusted p-value. The numbers along the x-axis correspond to the total number of genes differentially expressed in this comparison.

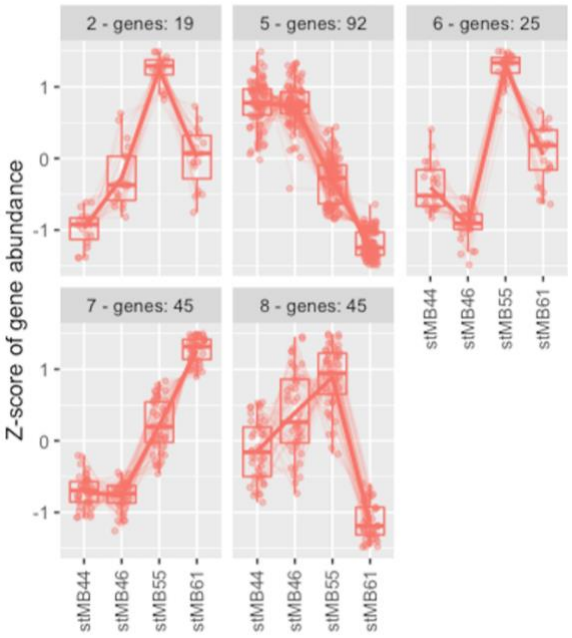


Figure 3.3: Clusters of genes differentially expressed during tadpole midbrain development that are associated with positive regulation of cell proliferation. The five clusters shown (2, 5, 6, 7, and 8) contained at least 10 genes and showed a distinct pattern of expression. Cluster 5 peaked at stages 44 and 46 and was enriched in cell cycle-related genes ($p < 0.001$); Clusters 2, 6, and 8 peaked in expression at stage 55 ($p = 0.019$), and cluster 7 peaked at stage 61 ($p < 0.001$); these were enriched in PI3K-AKT pathway associated genes.

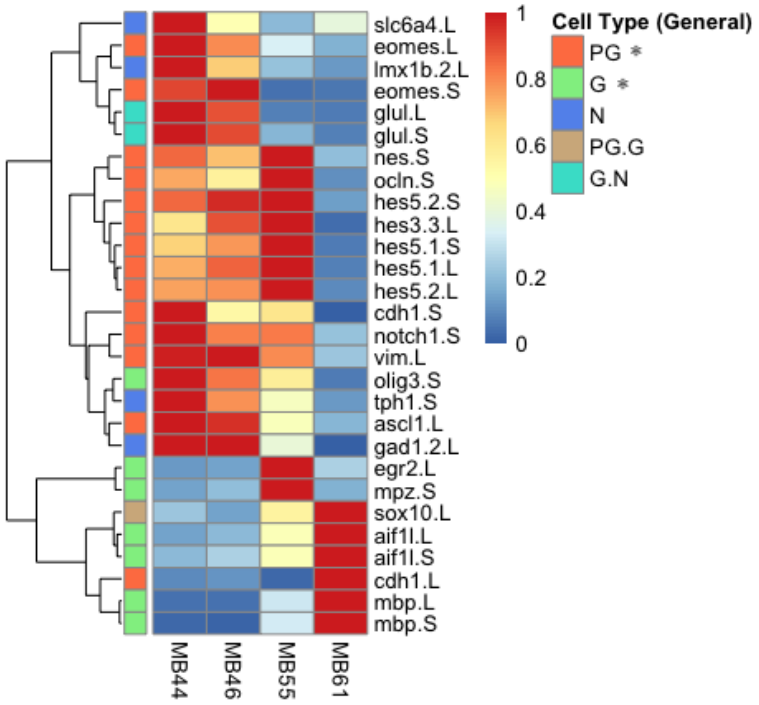


Figure 3.4: Cell type marker expression over tadpole midbrain development. PG = progenitor cell, G = glial cell, N = neuron, PG.G = progenitor cell/neuron marker, G.N = glial/neuron marker. Asterisks indicate significantly enriched cell types ($p < 0.001$ for progenitor cell markers, $p = 0.005$ for glial cell markers). Expression is normalized per-gene to the highest expression for gene.

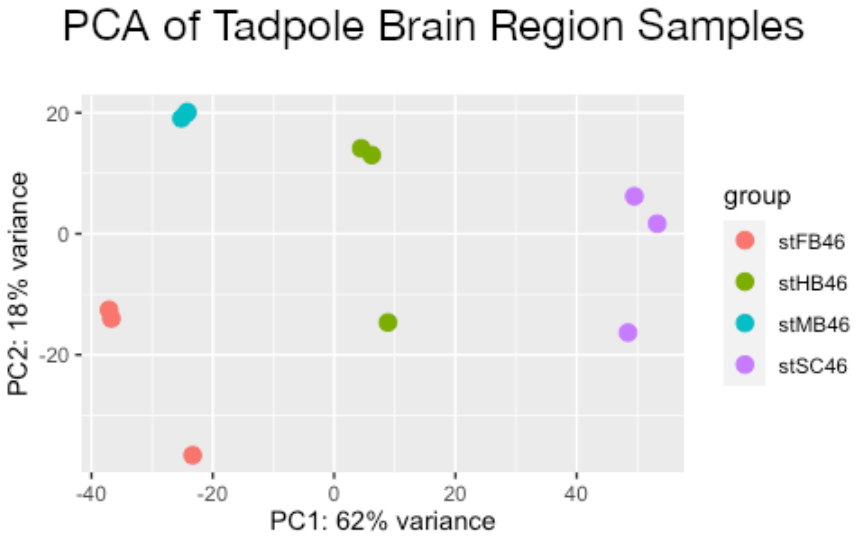


Figure 4.1: Principal component analysis of samples collected from the *X. laevis* stage 46 brain in the forebrain (stFB46), hindbrain (stHB46), midbrain (stMB46), and spinal cord (stSC46).

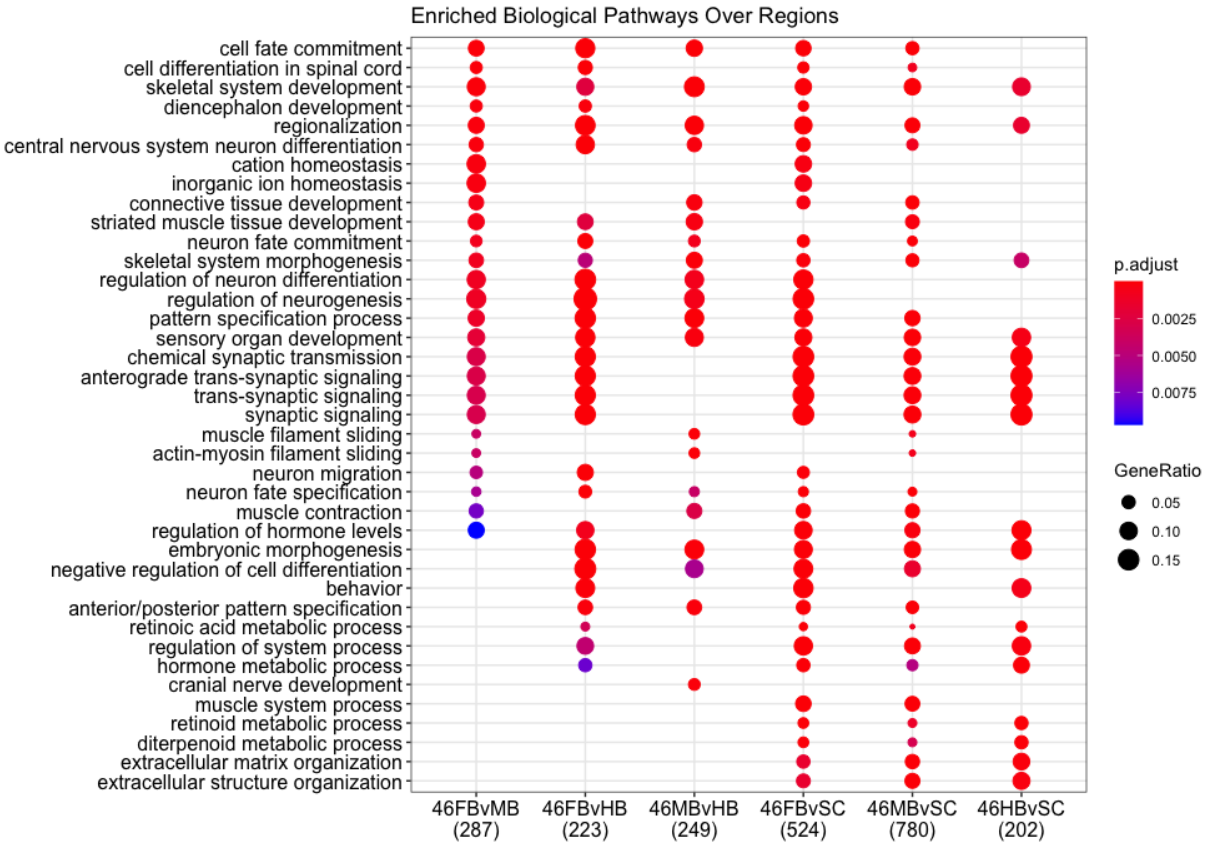


Figure 4.2: GO biological pathways enriched in the set of genes differentially expressed between two regions of the tadpole stage 46 brain. Dot size corresponds to the proportion of differentially expressed genes associated with the biological pathway, and dot color to the adjusted p-value. The numbers along the x-axis correspond to the total number of genes differentially expressed in this comparison.

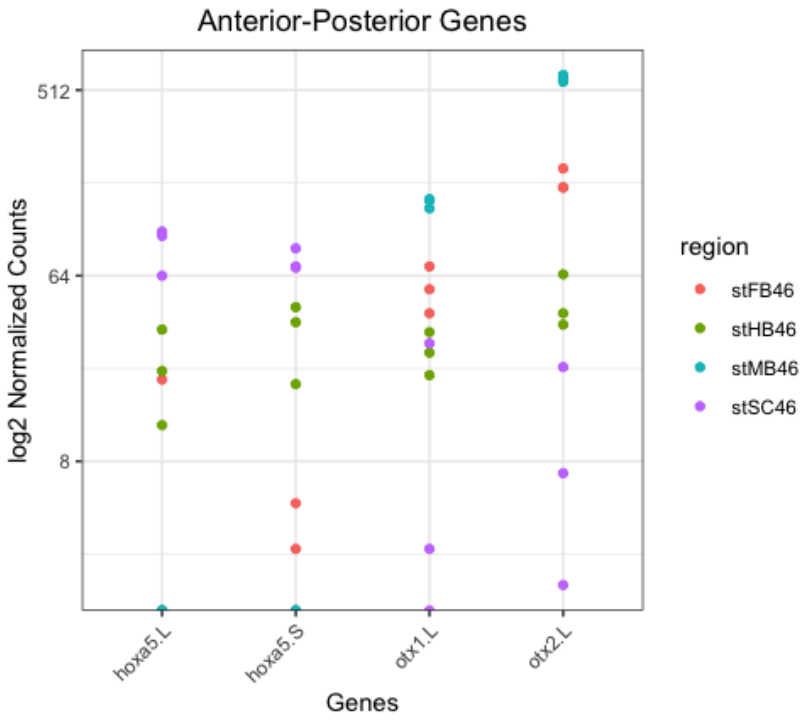


Figure 4.3: Normalized expression of *hox5*, expressed at the hindbrain-spinal cord region, and *otx1* and *otx2*, expressed at forebrain-midbrain region.

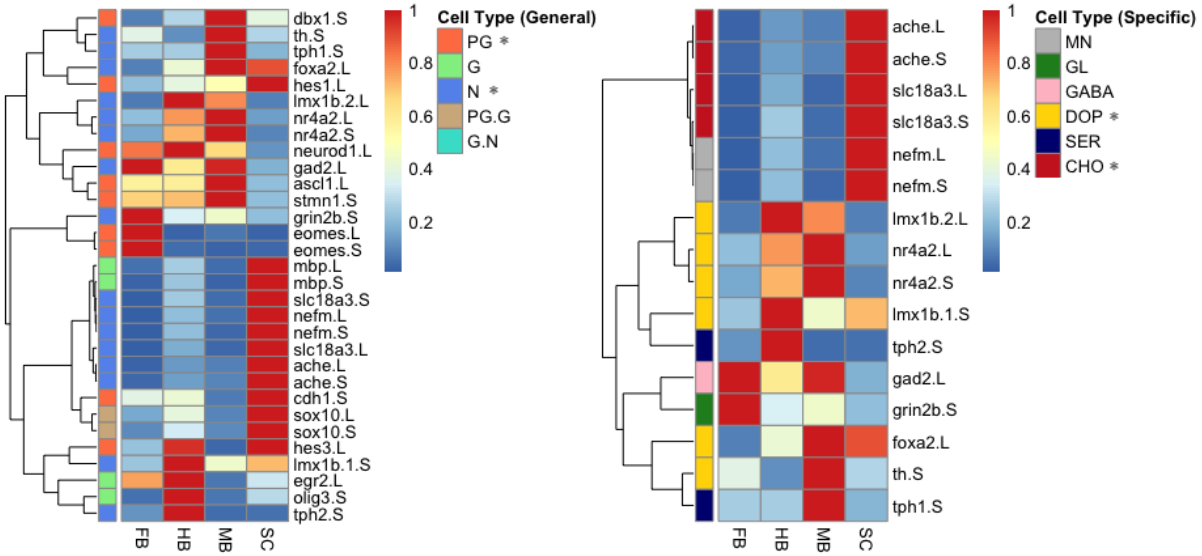


Figure 4.4: Left: Cell type marker expression over tadpole brain region at stage 46. PG = progenitor cell, G = glial cell, N = neuron, PG.G = progenitor cell/neuron marker, G.N = glial/neuron marker. **Right:** Neuron subtype marker expression over brain region at stage 46. MN = mature neuron, GL = glycinergic, GABA = GABAergic, DOP = dopaminergic, SER = serotonergic, CHO = cholinergic. Asterisks indicate significantly enriched cell types ($(p = 0.002$ for progenitor markers, $p < 0.001$ for neuron markers, $p < 0.001$ for both dopaminergic and cholinergic markers). Expression is normalized per-gene to the highest expression for gene.

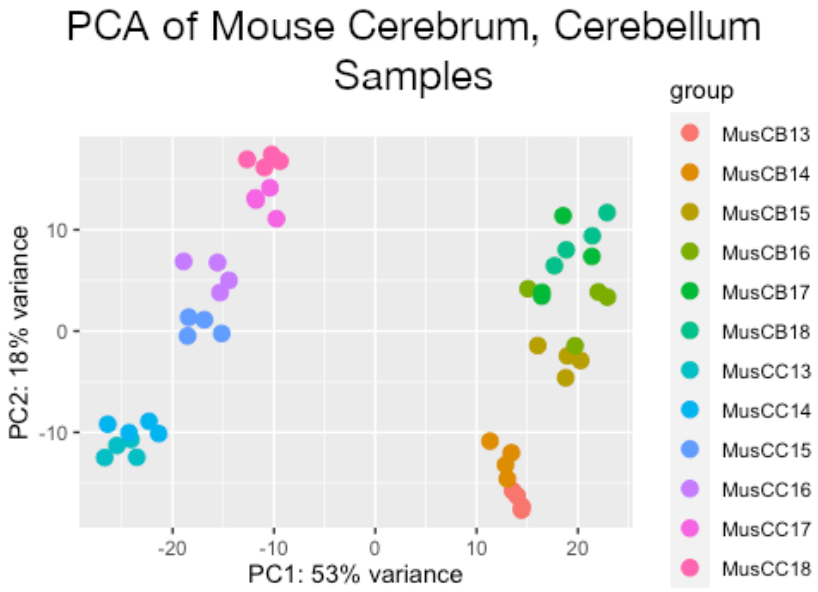


Figure 5.1: Principal component analysis of mouse cerebrum (CC) and cerebellum (CB) samples, collected at timepoints E13.5 – E18.5.

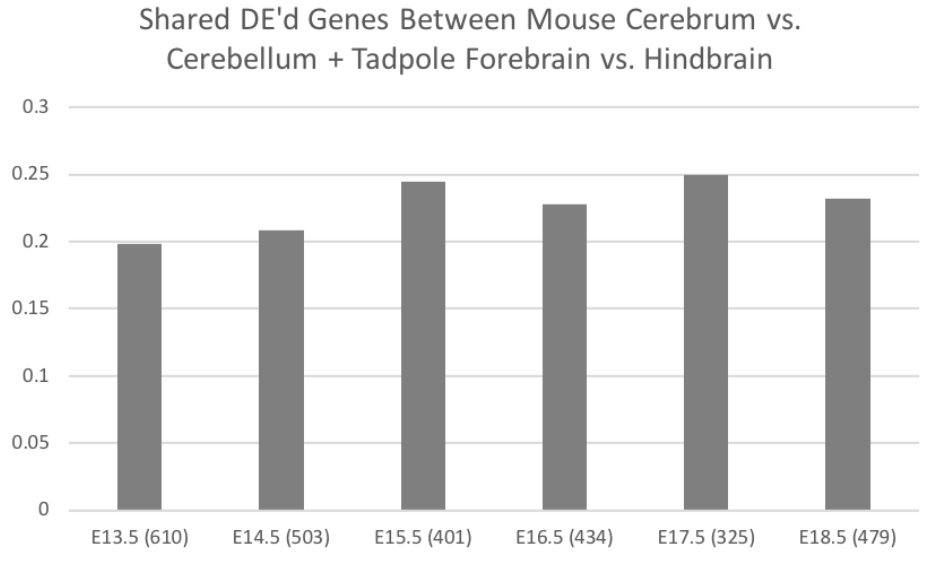


Figure 5.2: Proportion of genes differentially expressed in the mouse cerebrum vs. cerebellum at each timepoint that was also differentially expressed in the tadpole stage 46 forebrain vs. hindbrain. The total number of differentially expressed genes in at this timepoint is indicated in parentheses next to the timepoint on the x-axis. All timepoints in the mouse model shared a similar proportion of genes with the tadpole timepoint.

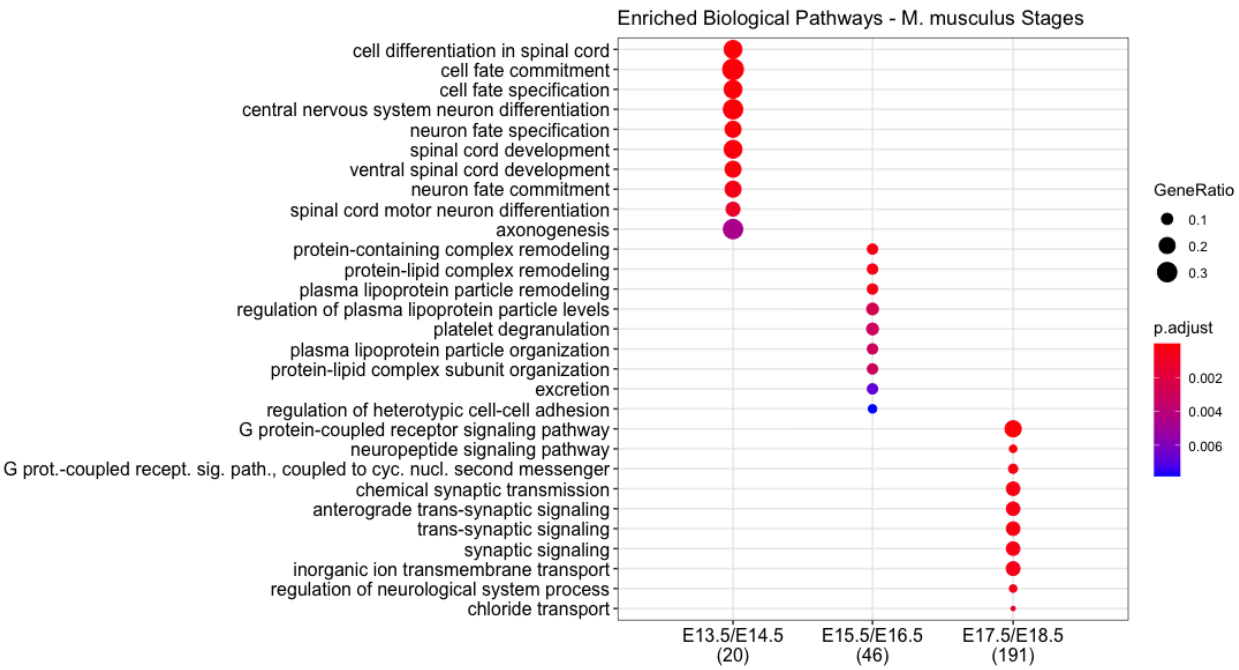
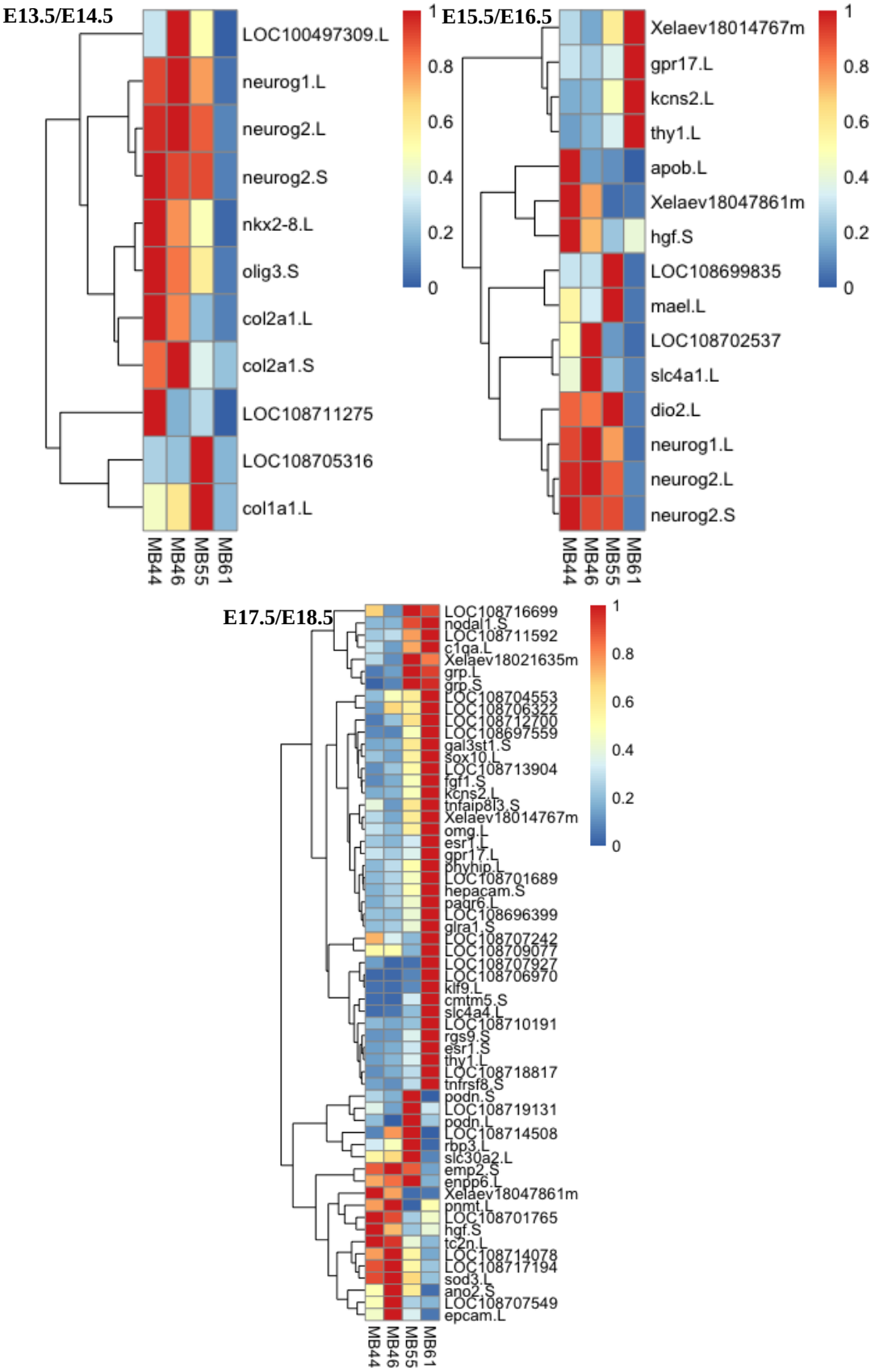


Figure 5.3: GO biological pathways enriched in the set of genes upregulated in both the mouse cerebrum and cerebellum at timepoints from E13.5 to E18.5. Distinct pathways are upregulated at each developmental period. Dot size corresponds to the proportion of differentially expressed genes associated with the biological pathway, and dot color to the adjusted p-value. The numbers along the x-axis correspond to the total number of genes differentially expressed in this comparison.

Figure 5.4: Expression of genes upregulated in the mouse cerebrum and cerebellum at E13.5/14.5 (top left), E15.5/E16.5 (top right), and E17.5/E18.5 (bottom) over *Xenopus* midbrain development. E13.5/E14.5 genes were overrepresented in genes upregulated at stages 44 and 46 ($p < 0.0001$), while E17.5/E18.5 genes were overrepresented in genes upregulated at stage B61 ($p < 0.0001$). Expression is normalized per-gene to the highest expression for gene.


E13.5/E14.5
LOC100497309.L
neurog1.L
neurog2.L
neurog2.S
nkx2-8.L
olig3.S
col2a1.L
col2a1.S
LOC108711275
LOC108705316
col1a1.L
MB44
MB46
MB55
MB61
1
0.8
0.6
0.4
0.2
0
E15.5/E16.5
Xelaev18014767m
gpr17.L
kcns2.L
thy1.L
apob.L
Xelaev18047861m
hgf.S
LOC108699835
mael.L
LOC108702537
slc4a1.L
dio2.L
neurog1.L
neurog2.L
neurog2.S
MB44
MB46
MB55
MB61
1
0.8
0.6
0.4
0.2
0
E17.5/E18.5
LOC108716699
nodal1.S
LOC108711592
c1qa.L
Xelaev18021635m
grp.L
grp.S
LOC108704553
LOC108706322
LOC108712700
LOC108697559
gal3st1.S
sox10.L
LOC108713904
fgf1.S
kcns2.L
tnfaip8l3.S
Xelaev18014767m
omg.L
esr1.L
gpr17.L
phyhip.L
LOC108701689
hepacam.S
paqr6.L
LOC108696399
glra1.S
LOC108707242
LOC108709077
LOC108707927
LOC108706970
klf9.L
cmtm5.S
slc4a4.L
LOC108710191
rgs9.S
esr1.S
thy1.L
LOC108718817
tnfrsf8.S
podn.S
LOC108719131
podn.L
LOC108714508
rbp3.L
slc30a2.L
emp2.S
enpp6.L
Xelaev18047861m
pnmt.L
LOC108701765
hgf.S
tc2n.L
LOC108714078
LOC108717194
sod3.L
ano2.S
LOC108707549
epcam.L
MB44
MB46
MB55
MB61
1
0.8
0.6
0.4
0.2
0

References

"Neural Markers." Abcam, 11 June 2020, www.abcam.com/neuroscience/neural-markers-guide.

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England), 31(2), 166–169. https://doi.org/10.1093/bioinformatics/btu638

Beck, C. W., Christen, B., & Slack, J. M. (2003). Molecular pathways needed for regeneration of spinal cord and muscle in a vertebrate. Developmental cell, 5(3), 429–439. https://doi.org/10.1016/s1534-5807(03)00233-8

Blum, M., & Ott, T. (2018). Xenopus: An Undervalued Model Organism to Study and Model Human Genetic Disease. *Cells Tissues Organs,205*(5-6), 303-313. doi:10.1159/000490898

Bowes, J. B., Snyder, K. A., Segerdell, E., Jarabek, C. J., Azam, K., Zorn, A. M., & Vize, P. D. (2009). Xenbase: Gene expression and improved integration. *Nucleic Acids Research,38*(Suppl_1). doi:10.1093/nar/gkp953

Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W., & Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science,360*(6392). doi:10.1126/science.aar5780

Calegari, F., & Huttner, W. B. (2003). An inhibition of cyclin-dependent kinases that lengthens, but does not arrest, neuroepithelial cell cycle induces premature neurogenesis. *Journal of Cell Science*, *116*(24), 4947 LP – 4955. https://doi.org/10.1242/jcs.00825

Costa-Silva, J., Domingues, D., & Lopes, F. M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. *Plos One,12*(12). doi:10.1371/journal.pone.0190152

D'Amico, L. A., Boujard, D., & Coumailleau, P. (2011). Proliferation, migration and differentiation in juvenile and adult Xenopus laevis brains. *Brain Research*, *1405*, 31–48. https://doi.org/https://doi.org/10.1016/j.brainres.2011.06.032

Denver, R. J., Hu, F., Scanlan, T. S., & Furlow, J. D. (2009). Thyroid hormone receptor subtype specificity for hormone-dependent neurogenesis in Xenopus laevis. Developmental biology, 326(1), 155–168. https://doi.org/10.1016/j.ydbio.2008.11.005

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2012). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15-21. doi:10.1093/bioinformatics/bts635

Hikosaka, A., Takaya, K., Jinno, M., & Kawahara, A. (2007). Identification and expression-profiling of Xenopus tropicalis miRNAs including plant miRNA-like RNAs at metamorphosis. *FEBS Letters*, *581*(16), 3013–3018. https://doi.org/10.1016/j.febslet.2007.05.054

Huang, L. C., Keuren-Jensen, K. V., Cline, H. (2016). Transcriptome profiles of neuronal activity-induced proliferating neural progenitor cells and newly-differentiated immature neurons in optic tectum of xenopus. Poster presented at: Neuroscience 2016. Society for Neuroscience. San Diego, CA.

Holzer, G., & Laudet, V. (2013). Chapter Fourteen - Thyroid Hormones and Postembryonic Development in Amniotes. In Y.-B. B. T.-C. T. in D. B. Shi (Ed.), *Animal Metamorphosis* (Vol. 103, pp. 397–425). https://doi.org/https://doi.org/10.1016/B978-0-12-385979-2.00014-9

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., Gil, L., Gupta, P., Hammond-Kosack, K. E., Haskell, E., Hunt, S. E., Jaiswal, P., Janacek, S.H., Kersey, P. J., Langridge, N., Maheswari, U., Maurel, T., McDowall, M. D., Moore, B., Muffato, M., Naamati, G., Naithani, S., Olson, A., Papatheodorou, I., Patricio, M., Paulini, M., Pedro, H., Perry, E., Preece, J., Rosello, M., Russell, M., Sitnik, V., Staines, D. M., Stein, J., Tello-Ruiz, M. K., Trevanion, S. J., Urban, M., Wei, S., Ware, D., Williams, G., Yates, A.D., & Flicek, P. (2019). Ensembl Genomes 2020—enabling non-vertebrate genomic research. *Nucleic Acids Research*, *48*(D1), D689–D695. https://doi.org/10.1093/nar/gkz890

Kakebeen, A. D., & Wills, A. E. (2019). More Than Just a Bandage: Closing the Gap Between Injury and Appendage Regeneration . In Frontiers in Physiology (Vol. 10, p. 81). https://www.frontiersin.org/article/10.3389/fphys.2019.00081

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nature methods, 12(4), 357–360. https://doi.org/10.1038/nmeth.3317 HISAT2

Kinsella, R. J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., Kersey, P., & Flicek, P. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database*, *2011*. https://doi.org/10.1093/database/bar030

Kornack, D. R., & Rakic, P. (1998). Changes in cell-cycle kinetics during the development and evolution of primate neocortex. Proceedings of the National Academy of Sciences, 95(3), 1242 LP – 1246. https://doi.org/10.1073/pnas.95.3.1242

Leinonen, R., Akhtar, R., Birney, E., Bower, L., Cerdeno-Tárraga, A., Cheng, Y., Cleland, I., Faruque, N., Goodgame, N., Gibson, R., Hoad, G., Jang, M., Pakseresht, N., Plaister, S., Radhakrishnan, R., Reddy, K., Sobhany, S., Ten Hoopen, P., Vaughan, R., Zalunin, V.,

Cochrane, G. (2011). The European Nucleotide Archive. Nucleic acids research, 39(Database issue), D28–D31. https://doi.org/10.1093/nar/gkq967

Lin, G., & Slack, J. M. W. (2008). Requirement for Wnt and FGF signaling in Xenopus tadpole tail regeneration. Developmental Biology, 316(2), 323–335. https://doi.org/https://doi.org/10.1016/j.ydbio.2008.01.032

Liu, L., Zhu, W., Liu, J., Wang, S., & Jiang, J. (2018). Identification and differential regulation of microRNAs during thyroid hormone-dependent metamorphosis in Microhyla fissipes. *BMC Genomics*, *19*(1), 507. https://doi.org/10.1186/s12864-018-4848-x

Lorena Pantano (2019). DEGreport: Report of DEG analysis. R package version 1.18.1. http://lpantano.github.io/DEGreport/

Love MI, Huber W, & Anders S (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology, 15, 550. doi: 10.1186/s13059-014-0550-8.

Miyata, K., & Ose, K. (2012). Thyroid Hormone-disrupting Effects and the Amphibian Metamorphosis Assay. Journal of toxicologic pathology, 25(1), 1–9. https://doi.org/10.1293/tox.25.1

Nieuwkoop , P.D. & Faber, J. (1994). Normal table of Xenopus laevis (Daudin): A systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis. Garland Science/Taylor & Francis, New York and London.

Pratt, K. G., & Khakhalin, A. S. (2013). Modeling human neurodevelopmental disorders in the Xenopus tadpole: From mechanisms to therapeutic targets. *Disease Models & Mechanisms,6*(5), 1057-1065. doi:10.1242/dmm.012138

Roberts, E. C., Shapiro, P. S., Nahreini, T. S., Pages, G., Pouyssegur, J., & Ahn, N. G. (2002). Distinct cell cycle timing requirements for extracellular signal-regulated kinase and phosphoinositide 3-kinase signaling pathways in somatic cell mitosis. Molecular and cellular biology, 22(20), 7226–7241. https://doi.org/10.1128/mcb.22.20.7226-7241.2002

Santini, S., Boore, J. L., & Meyer, A. (2003). Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. Genome research, 13(6A), 1111–1122. https://doi.org/10.1101/gr.700503

Schilling, T. F., & Knight, R. D. (2001). Origins of anteroposterior patterning and Hox gene regulation during chordate evolution. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 356(1414), 1599–1613. https://doi.org/10.1098/rstb.2001.0918

Segerdell, E., Ponferrada, V. G., James-Zorn, C., Burns, K. A., Fortriede, J. D., Dahdul, W. M., Vize, P. D., & Zorn, A. M. (2013). Enhanced XAO: The ontology of Xenopus anatomy

and development underpins more accurate annotation of gene expression and queries on Xenbase. Journal of Biomedical Semantics, 4(1), 31. https://doi.org/10.1186/2041-1480-4-31

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., Fukui, A., Hikosaka, A., Suzuki, A., Kondo, M., van Heeringen, S. J., Quigley, I., Heinz, S., Ogino, H., Ochi, H., Hellsten, U., Lyons, J. B., Simakov, O., Putnam, N., Putnam, N., Stites, J., Kuroki, Y., Tanaka, T., Michiue, T., Watanabe, M., Bogdanovic, O., Lister, R., Georgiou, G., Paranjpe, S. S., Kruijsbergen, I., Shu, S., Carlson, J., Kinoshita, T., Ohta, Y., Mawaribuchi, S., Jenkins, J., Grimwood, J., Schmutz, J., Mitros, T., Mozaffari, S. V., Suzuki, Y., Haramoto, Y., Yamamoto, T. S., Takagi, C., Heald, R., Miller, K., Haudenschild, C., Kitzman, J., Nakayama, T., Izutsu, Y., Robert, J., Fortriede, J., Burns, K., Lotay, V., Karimi, K., Yasuoka, Y., Dichmann, D. S., Flajnik, M. F., Houston, D. W., Shendure, J., DuPasquier, L., Vize, P. D., Zorn, A. M., Ito, M., Marcotte, E. M., Wallingford, J. B., Ito, Y., Asashima, Ueno, M. N., Matsuda, Y., Veenstra, G. J. C., Fujiyama, A., Harland, R. M., Taira, M., & Rokhsar, D. S. (2016). Genome evolution in the allotetraploid frog Xenopus laevis. Nature, 538, 336. https://doi.org/10.1038/nature19840

Sharma, P., & Cline, H. T. (2010). Visual activity regulates neural progenitor cells in developing xenopus CNS through musashi1. Neuron, 68(3), 442–455. https://doi.org/10.1016/j.neuron.2010.09.028

Sjöstedt, E., Zhong, W., Fagerberg, L., Karlsson, M., Mitsios, N., Adori, C., Oksvold, P., Edfors, F., Limiszewska, A., Hikmet, F., Huang, J., Du, Y., Lin, L., Dong, Z., Yang, L., Liu, X., Jiang, H., Xu, X., Wang, J., Yang, H., Bolund, L., Mardinoglu, A., , Zhang, C., von Feilitzen, K., Linskog, C., Pontén, F., Luo, Y., Hökfelt, T., Uhlén, M., Mulder, J. (2020). An atlas of the protein-coding genes in the human, pig, and mouse brain. Science, 367(6482), eaay5947. https://doi.org/10.1126/science.aay5947

Sunkin, S. M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T. L., Thompson, C. L., Hawrylycz, M., & Dang, C. (2012). Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Research, 41(D1), D996–D1008. https://doi.org/10.1093/nar/gks1042

The University of Heidelberg. (2019). Mouse RNA-seq time-series of the development of seven major organs. European Nucleotide Archive. https://www.ebi.ac.uk/ena/data/view/PRJEB26869

Thompson, C. L., Ng, L., Menon, V., Martinez, S., Lee, C. K., Glattfelder, K., Sunkin, S. M., Henry, A., Lau, C., Dang, C., Garcia-Lopez, R., Martinez-Ferre, A., Pombero, A., Rubenstein, J., Wakeman, W. B., Hohmann, J., Dee, N., Sodt, A. J., Young, R., Smith, K., Nguyen, T.N., Kidney, J., Kuan, L., Jeromin, A., Kaykas, A., Miller, J., Page, D., Orta, G., Bernard, A., Riley, Z., Smith, S., Wohnoutka, P., Hawrylycz, M., Puelles, L., Jones, A. R. (2014). A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain. Neuron, 83(2), 309–323. https://doi.org/10.1016/j.neuron.2014.05.033

Thuret, R., Auger, H., & 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. *Biology Open*, *4*(12), 1772 LP – 1781. https://doi.org/10.1242/bio.013391

Vernon, A. E., & Philpott, A. (2003). The developmental expression of cell cycle regulators in Xenopus laevis. *Gene Expression Patterns*, *3*(2), 179–192. https://doi.org/https://doi.org/10.1016/S1567-133X(03)00006-1

Wen, L., He, C., Sifuentes, C. J., & Denver, R. J. (2019). Thyroid Hormone Receptor Alpha Is Required for Thyroid Hormone-Dependent Neural Cell Proliferation During Tadpole Metamorphosis . *Frontiers in Endocrinology* , Vol. 10, p. 396. Retrieved from https://www.frontiersin.org/article/10.3389/fendo.2019.00396

Yaoita, Y., & Nakajima, K. (2018). Developmental gene expression patterns in the brain and liver of Xenopus tropicalis during metamorphosis climax. *Genes to Cells*, *23*(12), 998–1008. https://doi.org/10.1111/gtc.12647

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biology, 11(2), R14. https://doi.org/10.1186/gb-2010-11-2-r14