Evolution and Development of spinal cord stem cells and cell type diversity

Brigid Leung

St. Hilda's College

Submitted for the degree of Doctor of Philosophy in Zoology

Hilary Term 2021



Evolution and Development of spinal cord stem cells and cell type diversity

Brigid Leung

St. Hilda's College

Submitted for the degree of Doctor of Philosophy in Zoology

Hilary Term 2021

Abstract

How did spinal cord in vertebrates evolve to have its current morphology and function? To answer this I have looked at spinal cord organogenesis in different lineages. I use lamprey, a basally diverging vertebrate, as a model for the vertebrate ancestor and amphioxus, a cephalochordate, as an out-group for comparison. I have approached the question from three dimensions, the anterior-posterior (AP) axis, dorsal-ventral (DV) axis, and the medial-lateral (ML) axis, with my main focus on DV patterning. When comparing gene expression, I found that expression patterns are more similar between lamprey and gnathostomes than between amphioxus and vertebrates. Building on this foundation, some candidate genes with defined neural cell type expression were chosen to study the effects of signalling pathway manipulation on spinal cord patterning. Despite conserved gene expression patterns in lamprey, the effects of signalling pathway manipulation are very different from those observed in gnathostomes, with amphioxus even more diverged. This suggests that the roles of signals evolved divergently along different chordate lineages. In addition to the gene expression and signal patterning levels, I have also investigated the non-coding sequence level. I found an enhancer upstream of lamprey *Nkx2.2* which seems to be a vertebrate innovation. Overall, my work shows that gnathostomes, agnathans, and cephalochordates have evolved differently to pattern their spinal cord as demonstrated in the lack of cell type boundaries and different functions of signalling pathways in lamprey and differences in neuroanatomy, molecular pattern, and cis-regulatory elements in amphioxus.

Table of Contents

Abstract……1

Chapter 1: General Introduction

1.1 The evolution of the vertebrate nervous system……5

1.2 Spinal cord developmental anatomy……7

1.3 Patterning the vertebrate spinal cord: The AP axis……9

1.4 Patterning the vertebrate spinal cord: The DV axis……13

1.5 Patterning the vertebrate spinal cord: The ML axis, progenitor cell maintenance, and differentiation……16

1.6 Outgroups and evolutionary origins: spinal cords, cell types, and development in other chordates……18

1.7 Nerve cell types in adult lamprey and amphioxus spinal cords……20

1.8 AP spinal cord patterning in lampreys, sea squirts, and amphioxus……21

1.9 DV spinal cord patterning in lampreys, sea squirts, and amphioxus……23

1.10 ML spinal cord patterning in lampreys, sea squirts, and amphioxus: Differences in cell number……25

1.11 Conservation and innovation in chordate spinal patterning……26

1.12 A role for gene duplication in spinal cord evolution……26

1.13 Future perspectives: evolutionary change in spinal cord patterning……30

1.14 Aims of this thesis……32

Chapter 2: The anatomy of gnathostome, lamprey, and amphioxus spinal cords as visualized through developmental gene expression

2.1 Introduction……33

2.2 Methods

2.2.1 *Lampetra planeri* husbandry……37

2.2.2 *Branchiostoma floridae* husbandry……37

2.2.3 Molecular cloning……37

2.2.4 Whole mount in situ hybridization……37

2.2.5 Sectioning……38

2.2.6 Imaging……38

2.2.7 Phylogeny analysis……38

2.3 Results and Discussion

2.3.1 Defining spinal cord in amphioxus……39

2.3.2 Difficulty in comparing molecular pattern between amphioxus and vertebrates……39

2.3.3 Motor neuron development in chordates……40

2.3.4 Intermediate neural tube markers evolution in chordates……51

2.3.5 Dorsal neural markers……59
2.3.6 Signalling pathway gene expression……64
2.4 Summary, conclusion and future direction……72
Chapter3: Signal pathway manipulation effects on gene expression in the spinal cord
3.1 Introduction……75
3.2 Methods
3.2.1 Small pharmacological molecules treatment……79
3.3 Results and Discussion
3.3.1 The impact of manipulating Hh signaling on lamprey spinal cord patterning….81
3.3.2 The impact of inhibiting BMP signaling on lamprey spinal cord patterning…….86
3.3.3 The impact of manipulating RA signaling on lamprey spinal cord patterning….89
3.3.4 The impact of inhibiting Notch signaling on lamprey spinal cord patterning…..91
3.3.5 The impact of inhibiting Wnt and Fgf signaling on lamprey spinal cord patterning……95
3.3.6 *Pax2/5/8* expression in amphioxus when Wnt, Hh, and BMP activity are manipulated……104
3.3.7 The impact of inhibiting Notch signaling on amphioxus spinal cord patterning……106
3.3.8 The impact of Hh signaling inhibition on amphioxus spinal cord patterning….108
3.3.9 The impact of Wnt manipulation on amphioxus spinal cord patterning……..110
3.3.10 The impact of excess RA on motor neurons markers……113
3.3.11 The impact of Fgf inhibition on amphioxus spinal cord patterning……115
3.4 Summary, conclusion and future direction……117
Chapter4: Discovering cis-regulatory elements in lamprey
4.1 Introduction……123
4.2 Methods
4.2.1 ATAC and library preparation……125
4.2.2 ATAC-seq computational analysis……126
4.2.3 Chick neural tube in ovo electroporation……127
4.2.4 Lamprey whole mount electroporation……127
4.2.5 Lamprey whole mount immunohistochemistry……128
4.2.6 Synteny analysis……128
4.2.7 Conserved non-coding sequences BLAST against lamprey genome……128
4.2.8 Enhancer sequence alignment and motif database search……128
4.3 Results and Discussion
4.3.1 Vertebrate NK cluster synteny evolution……129

4.3.2 Chromatin landscape around lamprey genes……131
4.3.3 Phylogenetic foot-printing in the lamprey genome……139
4.3.4 Validating potential cis-regulatory elements……149
4.4 Summary, conclusion and future directions……152
Chapter 5: General Discussion……155
Acknowledgement……165
References……166
Supplementary Data
Chapter 2 supplementary data
S2.1 Amphioxus *Vacht, Netrin*, and *Gbx* gene expression characterization……218
S2.2 Phylogenetic trees and alignments……221
S2.3 Lamprey and Amphioxus probe sequences……231
Chapter 3 supplementary data
S3.1 Amphioxus drug treated data……251
Chapter 4 supplementary data
S4.1 Conserved non-coding sequences in 100 vertebrates but not lamprey……256
S4.2 *Nkx2.2, Nkx6, Pax6,* and *Pax7* conserved non-coding sequences in vertebrates including lamprey……259
S4.3 pTK-eGFP plasmid sequence……268

Chapter 1: General Introduction (Adapted from Leung and Shimeld, 2019)

1.1 The evolution of the vertebrate nervous system

Vertebrates develop complex nervous systems, with a level of sophistication that far exceeds that of invertebrates. The only challengers to this claim might be cephalopod molluscs, and some of the larger arthropods. Many of the complexities shown by vertebrates are considered innovations (Holland et al., 2013, Shimeld and Holland, 2000), and evolutionary biologists have long been interested in when and how vertebrate neural complexity evolved. Most studies on this topic have focused on the brain or the neural crest. This is not surprising, since the brain is the centre of characteristics such as upper level neural processing and memory, while the neural crest has long been recognised as a pivotally important tissue in vertebrate evolution, giving rise to many structures considered to be vertebrate innovations (Gans and Northcutt, 1983). There is, however, more to the vertebrate nervous system than the brain and neural crest. Nervous systems are typically divided into central (CNS) and peripheral (PNS) components. The vertebrate PNS derives primarily from neural crest cells, with cranial neurogenic placodes also playing a role. The evolution of neural crest and placodes has been extensively reviewed elsewhere (Patthey et al., 2014, Green et al., 2015) and will not be further considered here. The CNS is often dichotomised into the brain and spinal cord. While this makes sense from anatomical and functional perspectives, whether the separation is justified on developmental and evolutionary grounds is much more debatable. Aspects of spinal cord organisation and development extend into the hindbrain, including patterning by Hox genes, suggesting the separation may be better thought of as into a posterior, Hox-expressing territory (hindbrain and spinal cord) and anterior Hox-free territory (midbrain and forebrain). Recent work also shows deep and fundamental differences may exist in the gene regulatory networks (GRNs) that regulate the formation of the anterior and posterior neural cells very early in development (Metzis et al., 2018), and some authors have even speculated there may be separate evolutionary origins for anterior and posterior CNS components, traceable to separate anterior and posterior nervous systems in an early ancestor of the majority of extant animals (Arendt et al., 2016).

There is clearly still much to learn about the similarities and differences between anterior and posterior CNS in vertebrates and other animal groups, but regardless of the outcome the spinal cord has received relatively little attention from evolutionary biologists. This is in stark contrast to developmental biologists, who have dissected spinal cord development with impressive detail. This reflects a combination of its

relative simplicity and its experimental tractability for gene manipulation and imaging. It has resulted in the spinal cord becoming something of a ‘model organ system’ for understanding core developmental processes such as morphogen patterning and the balance between proliferation and differentiation (Gouti et al., 2015). This depth of knowledge represents an opportunity to explore the relationship between genetic change at the gene and genome level, developmental processes, and phenotypic consequences in terms of nervous system complexity. Key phenotypic changes in this regard might include increased cell number, new cell types, innovation in cell patterning, and mechanisms for axon navigation. I shall first summarise how vertebrate spinal cords develop. I will then relate this to genetic change, including gene duplication, with comparison to early diverging vertebrates such as lampreys, and to the vertebrates’ nearest invertebrate relatives, tunicates and amphioxus (Figure 1.1).

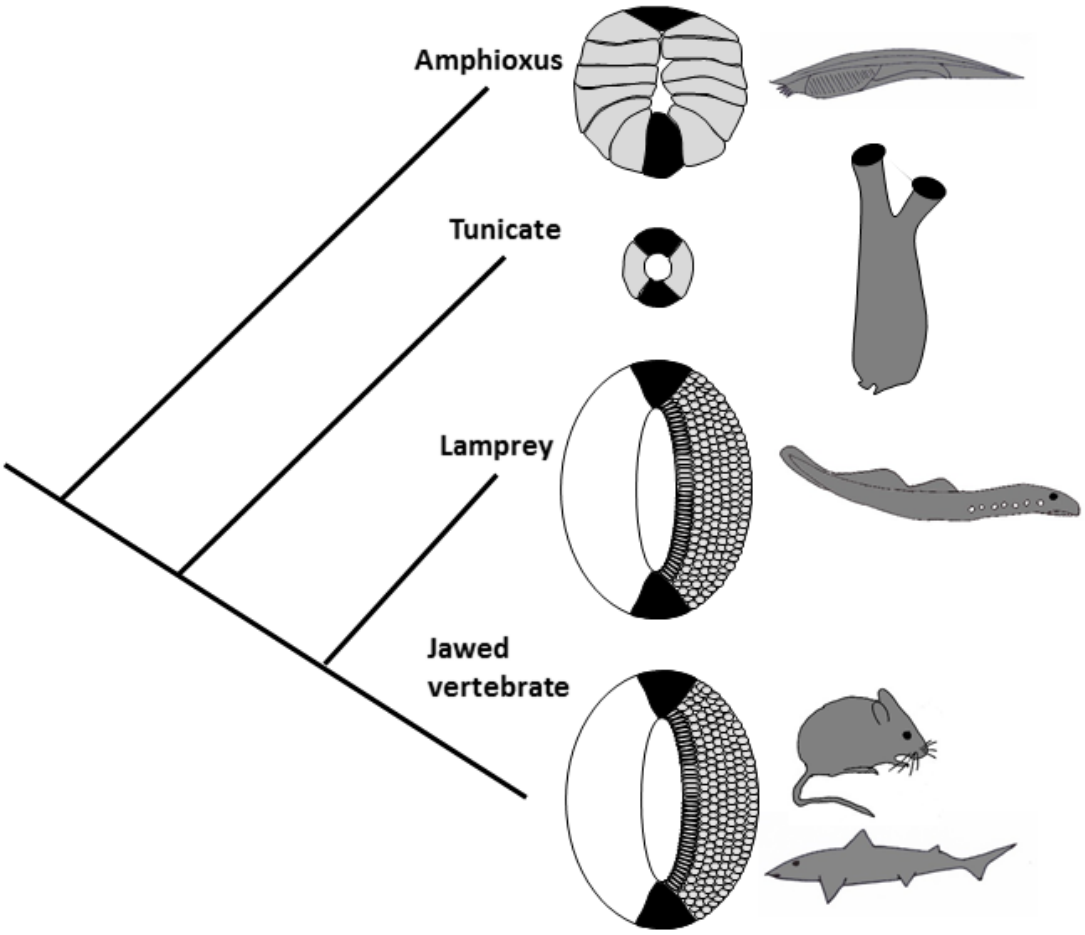


Figure 1.1. Chordate phylogeny and developing spinal cord anatomy.

Amphioxus is a member of the earliest diverging chordate lineage, and has a relatively simple spinal cord region that is just one cell thick. Tunicates, such as the sea squirt *Ciona*, have highly reduced spinal cord regions with just four cells in cross section and no neurons. Vertebrates have much more complex spinal cords, with more cells across the dorsal ventral axis, and multiple cell layers across the medial lateral axis. Lamprey and jawed vertebrate spinal cord diagrams are schematic, and differences are discussed more below. Diagrams adapted from (Lara-Ramirez et al., 2019).

1.2 Spinal cord developmental anatomy

The adult spinal cord of vertebrates serves a number of functions. Most obviously, it relays signals between the periphery and the brain. Put simplistically, sensory information is received via the dorsal nerve roots and passed on to the brain, while motor instructions are passed from the brain to the motor neurons which connect to the muscles. The spinal cord may also manage reflex actions like escape or withdraw

responses, and house central pattern generators managing rhythmic motor activity. Many different types of cell are needed to perform these functions, including various interneurons. While there are some important differences in spinal cords between vertebrate groups, for example, in motor control circuits in limbed vs finned vertebrates, cell types and their distributions appear fundamentally conserved (Goulding, 2009).

The developmental origins of spinal cord cell types has been intensively studied in model systems, especially in chick and mouse. Across the DV axis the embryonic spinal cord is divided into multiple discrete zones, each giving rise to a defined suite of cell types. Proliferative progenitor cells are sited ventricularly (by the lumen of the spinal cord) in stem cell zones that produce new cells through an extended period of development. Cells born from the same progenitor cell zone at different developmental times may have different fates, leading to ML diversity of cell types at the same DV level. The spinal cord is also patterned along the AP axis. Before considering the evolution of these axes, I will first summarise the molecular control of their development as understood from model vertebrates.

In mouse, the structures of the CNS along the AP axis are identified by their anatomical structure and also by Hox collinearity, with the spinal cord starting posterior to rhombomere 8 at the *Hox6* expression level, and this seems conserved in other jawed vertebrates (Nolte and Krumlauf, 2007). However, such anatomical and Hox expression characters are less clear in lamprey. The physical boundaries between rhombomeres are not observable during embryo development and the boundary between hindbrain and spinal cord is less clear. Lamprey *HoxL6* expression is expressed throughout the CNS posterior to the forebrain, which does not help identify the spinal cord (Cohn, 2002). In my studies, I have made use of other rhombomere markers and surrounding landmarks to identify the lamprey hindbrain-spinal cord boundary. Based on the lamprey *Hox1-4* and *Krox20* expression, individual rhombomeres can be identified. Lamprey *Hoxζ4* is expressed in rhombomere 7 and the spinal cord (Parker et al., 2019a). Although the boundary between hindbrain and spinal cord has not been clearly identified anatomically (and it remains unclear whether rhombomere 8 exists as a discrete entity in lamprey), I consider that the lamprey spinal cord begins immediately posterior to rhombomere 7 as the neural tube reduces in diameter at this point. For amphioxus, the exact extent of the region homologous to the vertebrate spinal cord is even less clear. From an anatomical point of view, there is an enlarged area called the cerebral vesical at the

anterior of the CNS, followed by a thinner neural tube along the rest of the AP axis. From a gene expression point of view, amphioxus *Hox6* begins its neural tube expression at the level between the somite 5 and somite 6 boundary (Pascual-Anaya et al., 2012, Schubert et al., 2006). The cerebral vesicle is proposed to be homologous to forebrain and midbrain (Candiani et al., 2012, Albuixech-Crespo et al., 2017), and neural tube posterior to the somite 5/6 boundary is likely homologous to spinal cord. However, the region in between should be comparable to the vertebrate hindbrain and possibly anterior spinal cord, but shows little evidence of similar morphological or anatomical differentiation, leading some to refer to the whole neural tube posterior to the cerebral vesicle as the 'rhombospinal region' (Holland and Holland, 1999). I return to this subject at the beginning of the Chapter 2 Results and Discussion section.

1.3 Patterning the vertebrate spinal cord: the AP axis

The spinal cord grows from anterior to posterior, in common with other axial tissues such as the somitic and intermediate mesoderm. New cells are born in the posterior growth zone, the tailbud, and as these cells continue to proliferate the tailbud grows posteriorly, leaving some cells behind in the forming spinal cord (Akai et al., 2005). As developmental time passes, these cells become progressively distant from the tailbud and may start to differentiate. AP identity in spinal cord cells is regulated by Hox gene expression. Establishment and maintenance of Hox gene expression in the spinal cord is a complex affair, involving regulation by multiple external signals, by interactions between Hox genes, and by change in chromatin state through developmental time. There is insufficient space here to detail all of these, and the reader is referred to a recent review for a detailed exploration of these issues (Deschamps and Duboule, 2017). However, some key generalisations can be extracted (Figure 1.2).

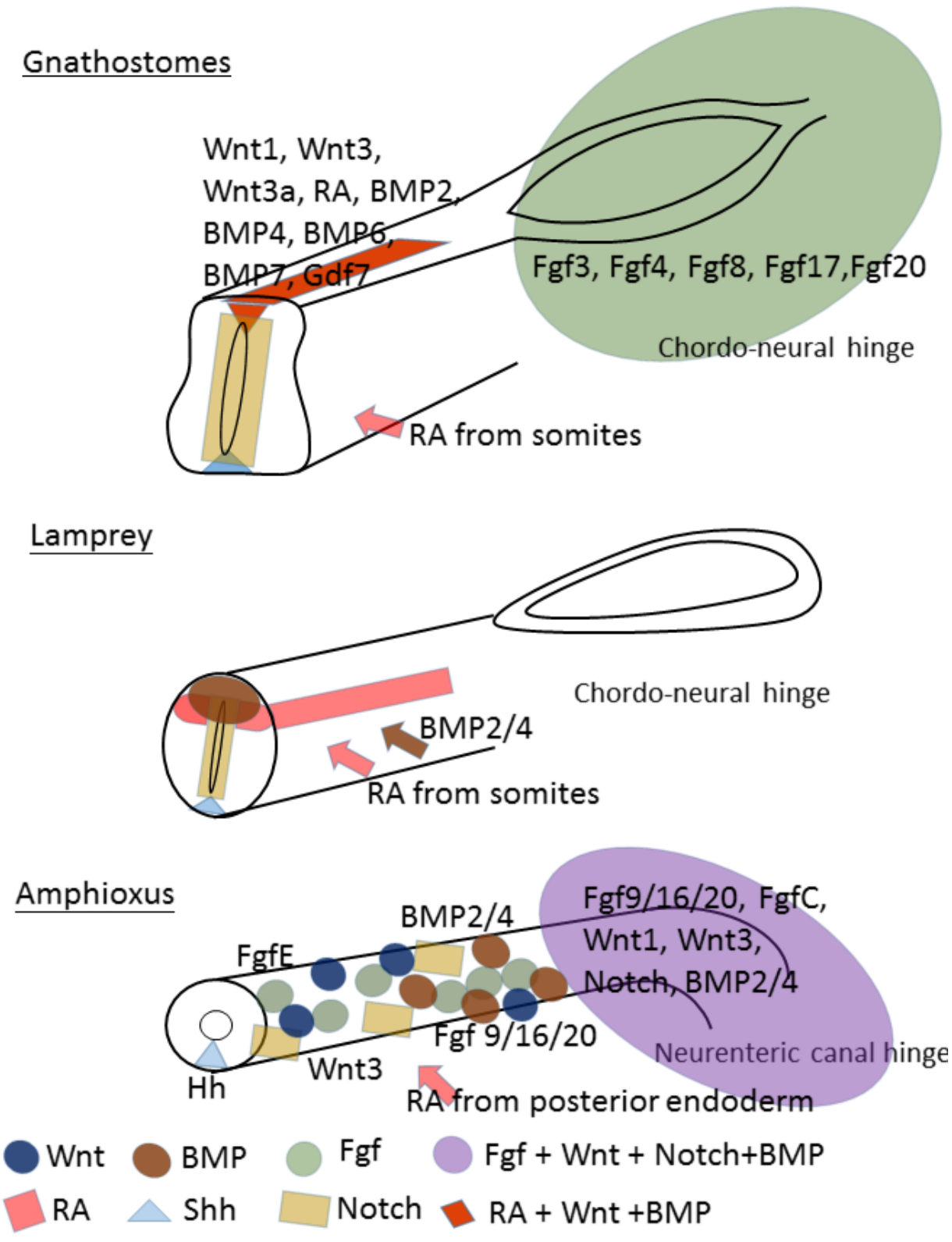

Gnathostomes
Wnt1, Wnt3,
Wnt3a, RA, BMP2,
BMP4, BMP6,
BMP7, Gdf7
Fgf3, Fgf4, Fgf8, Fgf17,Fgf20
Chordo-neural hinge
RA from somites
Lamprey
Chordo-neural hinge
BMP2/4
RA from somites
Amphioxus
BMP2/4
FgfE
Fgf9/16/20, FgfC,
Wnt1, Wnt3,
Notch, BMP2/4
Fgf 9/16/20
Neurenteric canal hinge
Wnt3
Hh
RA from posterior endoderm
Wnt
BMP
Fgf
Fgf + Wnt + Notch+BMP
RA
Shh
Notch
RA + Wnt +BMP

Figure 1.2. Signal sources in spinal cord patterning.

Schematic diagrams of gnathostome, lamprey and amphioxus spinal cords showing relevant expression of signalling molecules. Details and data sources as follows.

Hh in the floor plate in chordates. References: Mouse (Bai et al., 2002); Xenopus (Koide et al., 2006); Chick and zebrafish (Al Oustah et al., 2014); Lizard (Sun et al., 2018); Skate (Gillis and Hall, 2016); Lamprey (Sugahara et al., 2011); Amphioxus (Hu et al., 2017, Shimeld, 1999).

BMP in the roofplate in gnathostomes, dorsal neural tube and somites in lamprey, tailbud in amphioxus. References: Mouse *BMP2*, *BMP6*, *BMP7*, and *Gdf7* are expressed in roofplate (Butler and Dodd, 2003, Augsburger et al., 1999, Duval et al., 2014, Namm et al., 2015), *BMP4* in the whole spinal cord E9.5 then restricted to dorsal ectoderm by E10.5 (Andrews et al., 2017, Hu et al., 2004). Chick *BMP4*, *BMP7,* and *Gdf7* are in the roofplate (Chizhikov and Millen, 2004, Nitzan et al., 2016, Andrews et al., 2017), *BMP6* expressed through out spinal cord except roofplate (Andrews et al., 2017). Lizard BMP4 in the roof plate (Sun et al., 2018). Zebrafish *BMP4* and *BMP6* in the dorsal epidermis (Dick et al., 1999, Merkes et al., 2015, Thisse and Thisse, 2004), *BMP7a* in the tailbud (Schmid et al., 2000), *BMP7b* in the pronephric duct, dorso-ventral tail fin (Shawi and Serluca, 2008), *Gdf7* in the head cartilage, dorsal aorta and pronephric duct (Davidson et al., 1999). Lamprey *BMP2/4* in dorsal neural tube and somites (McCauley and Bronner-Fraser, 2004). Amphioxus *BMP2/4* in the tailbud (Panopoulou et al., 1998).

Canonical Wnt in the roof plate in gnathostomes and tailbud in amphioxus, no data on lamprey. References: Mouse *Wnt1*, *Wnt3,* and *Wnt3a* in the roofplate (Shinozuka et al., 2019, Parr et al., 1993); Chick *Wnt1* and *Wnt3a* (Alvarez-Medina et al., 2008); Zebrafish *Wnt1* and *Wnt3* in the roof plate (Duncan et al., 2015, Tanaka et al., 2012), *Wnt3a* in the tailbud (Fior et al., 2012). Lamprey only has data only on non-canonical *Wnt5* which is expressed across the DV spinal cord (Guerin et al., 2009). Amphioxus *Wnt1* in the tailbud, *Wnt3* in the tail fin and distributed along spinal cord (Somorjai et al., 2018).

Notch in the ventricular zone in vertebrates and in the tailbud in amphioxus. References: Mouse (Machado et al., 2014); Chick (Hämmerle and Tejedor, 2007); Zebrafish ventricular zone expressing *notch* and *delta* (Appel et al., 2001, Kim et al., 2008); Lamprey (Lara-Ramirez et al., 2019, Guerin et al., 2009); Amphioxus (Holland et al., 2001b).

RA from vertebrate somites and roof plate in gnathostomes, dorsal neurons in lamprey (Castillo et al., 2010), and from posterior endoderm in amphioxus beginning from neurula stage (Cañestro et al., 2000).

FGF in the tailbud of gnathostomes and amphioxus, no data on lamprey. References: Mouse *Fgf3*, *Fgf4*, *Fgf8* and *Fgf17* in the tailbud (Wahl et al., 2007). Chick *Fgf8* in the posterior neural plate and presomitic mesoderm (Bertrand et al., 2000). *Xenopus Fgf8* and *Fgf20* in the tailbud (Lin and Slack, 2008). Zebrafish *Fgf8* in the tailbud and posterior somites (Fior et al., 2012). Amphioxus *Fgf9/16/20* patchily expressed in the neural tube, *FgfE* in a different patch of the neural tube, *FgfC* in the tailbud (Bertrand et al., 2011).

Wnt signalling plays a role in posterior specification in many bilaterians (Niehrs, 2010), and in vertebrates is involved in early regulation of the Hox complexes (Deschamps and Duboule, 2017). It has also long been known from cell culture, explant and signalling pathway manipulation studies that Hox genes differ in their sensitivity to signalling molecules like Retinoic Acid (RA) and Fibroblast Growth Factor (FGF). RA is secreted from newly formed somites, while FGF (and other signalling molecules like the Tgfβ family member Gdf11) are produced by the tailbud (Diez del Corral et al., 2003, Liu et al., 2001). Since both tailbud and somitogenesis progress synchronously from anterior to posterior, all spinal cord cells should in principle move through developmental time from a high FGF, to a high RA environment. Exactly when and how these signals regulate Hox expression though has been tricky to decipher. Under RA treatment in chick embryos, the anterior boundaries of neural tube expression *Hoxb1, Hoxb3, Hoxb4* and *Hoxb5* extend, but their expression is not affected by FGF treatment; similarly, FGF treatment posteriorises the expression of *Hoxb6, Hoxb7, Hoxb8* and *Hoxb9,* but these are not sensitive to RA in the same assay (Bel-Vialar et al., 2002). However, others showed that *Hoxc6* and *Hoxc9* expression in explants respond to both RA and FGF treatment, with *Hoxc6* expression increasing when treated with RA and reducing when treated with FGF, whereas *Hoxc9* expression responded reciprocally (Liu et al., 2001).

Such differences in observation could be explained by studies using different methods and systems (for example whole embryo in vivo vs explants in vitro), but developmental time also plays a critical role. Tailbud derived FGF regulates proliferation (Diez del Corral et al., 2003, Diez Del Corral and Morales, 2017), but cells that spend more time in the tailbud appear to show progressive opening of chromatin across Hox clusters (Soshnikova and Duboule, 2009). How exactly this works needs more study, but it suggests a model in which developmental time regulates Hox cluster chromatin state, which is then acted on by other signals. Anterior spinal cord cells derive from progenitors that have fewer Hox genes accessible for transcription, while more posterior cells derive from progenitors that have more Hox genes accessible. Accessible Hox genes might then be regulated by signals like RA, and FGF. The homeobox gene *Cdx* is a good candidate for mediating FGF/Gdf11 regulation of Hox (Gaunt, 2017, Neijts et al., 2017), and RA probably directly regulates many Hox genes through RA responsive elements (RAREs) identified in Hox gene enhancers (Krumlauf, 2016). In addition some studies suggest the Hox genes may cross regulate, for example *Hoxc6* and *Hoxc9* appear to cross

repress each other and hence establish the boundary between the thoracic and lumbar regions of the spinal cord (Dasen et al., 2003).

1.4 Patterning the vertebrate spinal cord: the DV axis

Spinal cord DV patterning has become a model for understanding how morphogen signals may pattern a field of cells, with experimental approaches and mathematical models integrating signalling, downstream GRN interactions and developmental time. For a detailed exploration of this the reader is referred to recent reviews (Briscoe and Small, 2015, Gouti et al., 2015). Here, I will draw out the key points, focusing on aspects that allow comparison to other animals and hence evolutionary insight (Figure 1.2).

The spinal cord of amniote embryos is organised into a series of DV progenitor (or stem cell) zones (Figure 1.3). Ventrally these are named p0-p3, and pMN, with the first four producing interneurons and pMN producing motor neurons of various sorts, plus some other cells. As a generality, these cells are involved in controlling motor output (reviewed by (Goulding, 2009)). Dorsally these are dP1-6, which primarily produce interneurons involved in managing sensory input from the PNS. Many of the target subpopulations of interneuron innervation are known, as are the neurotransmitters they express. The expression of numerous transcription factor genes have also been mapped across these progenitor zones and the differentiated cells they produce. There is insufficient space to document this here: key aspects are shown in Figure 1.3, and the reader is referred to recent reviews for more detail (Alaynick et al., 2011, Goulding, 2009, Lai et al., 2016).

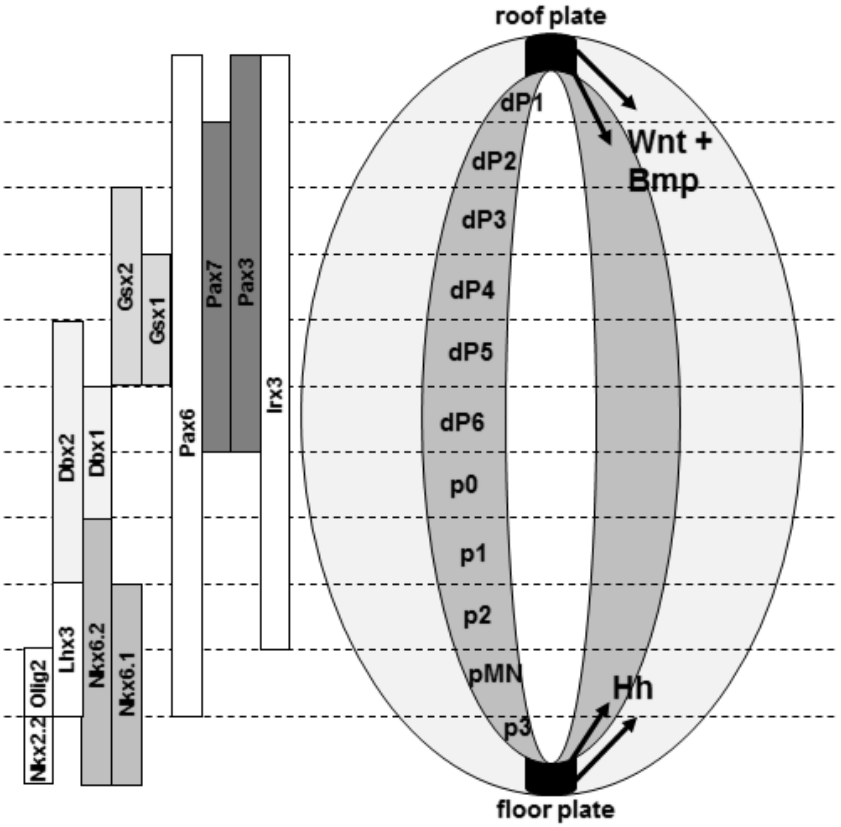


Figure 1.3. DV patterning of the amniote spinal cord.

To the right is a schematic cross-section of a spinal cord, with signalling molecules and sources shown on one side, and the 11 progenitor zones shown on the other. To the left are the domains of expression of some of the transcription factor genes involved in DV patterning of the progenitor zones. Paralogue pairs are a similar shade of grey. Not all known genes are shown, and it should be noted that many of the domains of expression are dynamic, particularly in early development when pattern is being established, and that their strength of expression may vary within their domain of expression. For a more detailed recent exploration, see Lai et al (2016).

In gnathostomes, the ventral spinal cord is dependent on the morphogen Shh secreted from the floorplate and notochord, evidenced by failure of ventral spinal cord patterning in *Shh*$^{-/-}$ mice (Chiang et al., 1996). Other factors may also have contributions to ventral patterning including Eph/Ephrin signalling (Laussu et al., 2017), and *Chordin*, which is expressed in the notochord and inhibits Bmp signalling (Sasai et al., 1994), but Hedgehog (Hh) signalling mediated by Shh has a dominant role. Pairs of cross repressive transcription factor genes lie downstream of the Shh signal. For example, *Nkx2.2* and *Olig2* cross-repress to establish the boundary between the p3 and pMN zones; *Nkx6.1* and *Dbx2* cross-repress to establish the boundary between p2 and p1; *Nkx6.2* and *Dbx1* cross repress to establish the

boundary between p1 and p0 (Sun et al., 2003, Briscoe et al., 2000); *Irx3* and *Olig2* cross-repress to form p2/pMN boundary (Novitch et al., 2001); *Sp8* and *Pax6* class I progenitor zone markers cross-repress *Nkx2.2* class II p3 zone marker to form pMN/p3 boundary (Li et al., 2014b). While these mechanisms are not fully resolved, it would appear these transcription factors generally act by interacting with Groucho/TLE repressors, and that as well as repressing each other they also repress a broad swathe of other genes that are normally expressed in other DV compartments (Kutejova et al., 2016). Receipt of Shh signalling is presumed to be quantitative, mediated by a balance between active and repressive forms of Gli proteins. To add additional complexity, the three Gli proteins involved (Gli1, Gli2 and Gli3) have different abilities to produce active and repressive forms, and their genes are also differently regulated by Hh signalling. All this creates a dynamic regulatory system that mathematical modelling suggests has interesting properties, including bi-stable outputs and the ability to scale across differently sized fields of cells (Uygur et al., 2016, Briscoe and Small, 2015).

Views on patterning of the gnathostomes dorsal spinal cord are more controversial. The roof plate produces multiple signalling molecules including Wnt1 and Wnt3, and Tgfβ members including Bmp2, Bmp4, Bmp7 and Gdf7 (Figure 1.2). Tozer and colleagues (Tozer et al., 2013) proposed that the gradient and exposure time of BMPs from the roof plate is responsible for patterning dorsal spinal cord cells. However, others have argued (Andrews et al., 2017) that the nature of the BMPs (rather than just their concentration) is responsible for cell type specification. BMP4 upregulates pSmad1/5/8 more efficiently than BMP7, and ectopic expression of *Bmp4* increases the dI1 and dI2 populations while ectopic *Bmp7* and *Gdf7* increases roof plate, dI1 and dI3 populations. In addition to Tgfβ signals, Wnt signalling is also involved in dorsal spinal cord patterning. The *Wnt1/Wnt3a* double homozygous mutant shows loss of expression of dI1 and dI2 markers *LH2, Islet1, Math1* and *Ngn1,* and gain of expression of dI3 markers *Pax2, Lim1/2* and *Mash1* in the dorsal spinal cord, while having no effect on roof plate expression of *Bmp4, Gdf7* and *Msx1* (Muroyama et al., 2002). This suggests a role for Wnt signalling in dorsal patterning, though Wnt also affects cell proliferation (see below). In summary, while there are clearly roles for both Tgfβ and Wnt signalling in development of the dorsal spinal cord progenitor cell zones, exactly which signal does what and how, how they inter-relate, and how this connects to proliferation, needs much more study.

To make things even more complex, Wnt signalling also induces the expression of *Gli3,* which generally represses targets of Shh signalling (Alvarez-Medina et al., 2008; Yu et al., 2008). Thus, dorsal and ventral signals are not independent entities, but potentially interact within receiving cells. Furthermore, AP and DV patterning are also interlinked. Posterior signals may block DV patterning from commencing: for example FGF regulates the expression of *Nkx1.2* in posterior cells, and this prevents cells from responding to Shh signals until *Nkx1.2* levels have dropped sufficiently (Sasai et al., 2014). RA also upregulates some progenitor cell zone genes (Novitch et al., 2003), and in vitro study shows that *Shh* expression depends on RA concentration (Okada et al., 2004). These interactions coordinate the progression from a relatively naïve tailbud progenitor, to an AP- and DV-specified progenitor, and eventually on to a differentiated cell.

1.5 Patterning the vertebrate spinal cord: the ML axis, progenitor cell maintenance, and differentiation

As AP and DV patterning progress, the spinal cord develops in an 'inside to outside' manner: undifferentiated proliferative cells lie ventricularly, forming progenitor cell zones with DV pattern as discussed above (Figure 1.3). When these cells divide, one or both daughter cells may migrate laterally to differentiate. Cells born later in development migrate through fields of earlier-differentiated cells to differentiate, meaning cells born in the same progenitor cell zone but at different times may migrate into different environments and hence adopt different fates. In this way, multiple cell types can form over time from the same progenitor cell population. The best studied example of this is the pMN zone, where motor neurons of different types, then later oligodendrocytes, are successively born from the same progenitor cell zone (Sockanathan et al., 2003).

Expression of SoxB1 genes marks neural stem cells, and the SoxB1 gene *Sox2* is sufficient to maintain progenitor cells at the ventricular zone (Graham et al., 2003). *Foxp2* and *Foxp4* are required and sufficient for neural differentiation, opposing SoxB1 activity, and with *Foxp*-expressing cells repressing the expression of junction protein genes such as *N-cadherin*, detaching from neuroepithelium and migrating laterally (Rousso et al., 2012). Notch signalling controls which cells undertake this transition, with sustained Notch signalling needed to maintain progenitor cells in the progenitor zones (Appel et al., 2001, Wettstein et al., 1997). From a cell-to-cell interaction point of view, Isomura and colleagues have shown through single-cell

bioluminescence imaging of progenitors that Notch-Dll signalling can entrain an intracellular *Hes1* oscillating negative feedback loop (Isomura et al., 2017). It would be interesting to know what breaks the feedback loop to direct the cell to differentiation.

What regulates Notch signalling, the continued maintenance of progenitor cells, and the balance between rates of proliferation and differentiation is not well understood, but many factors have been implicated, including several of the signals that pattern AP and DV axes. As discussed above, FGF maintains proliferative progenitors in the tailbud (Mathis et al., 2001), possibly via Notch signalling (Akai et al., 2005), while RA promotes differentiation (Diez del Corral et al., 2003). Shh signalling in the ventral neural tube is sufficient and required for *Jag2* expression, a Notch ligand. Jag2 activity prevents premature differentiation of oligodendrocytes and inhibits differentiation of motor neuron progenitors (Rabadan et al., 2012). Notch signalling may also impact on the receipt of Shh signalling, via effects of the localisation of the Hh receptor Ptc in the primary cilium (Kong et al., 2015).

Bmp6 also speeds up cell cycle transition (Andrews et al., 2017; Le Dreau et al. 2014), indicating Bmp may also be involved in ML development (Andrews et al., 2017, Le Dreau et al., 2014). Furthermore Shh and Wnt signalling control different stages of the cell cycle through regulation of cyclin D1, E, A, B, with Shh being epistatic to Tcf3/4 (a Wnt-regulated transcription factor) when regulating cyclin D1 (Alvarez-Medina et al., 2008). Cell cycle regulation in neural tube development is more extensively reviewed elsewhere (Molina and Pituello, 2017, Saade et al., 2018). In addition another ventral signal, Sema3B, may also regulate progenitor maintenance (Arbeille et al., 2015).

While still not fully pieced together, these data suggest regulation of proliferation is complex, involves inputs from both AP and DV signalling pathways, and may vary along both AP and DV axes as well as through developmental time at the same axial location. As a generalisation though, jawed vertebrates use these mechanisms to maintain a population of ventricular progenitors, which continue to divide and produce new differentiating cells. This allows embryos to make more cells and different types of cells, both components of a complex CNS. It may also allow them to set progenitor cell populations aside with the potential for these to persist into the adult and engage in the repair of damaged spinal cords (reviewed by (Grandel and Brand, 2013)).

1.6 Outgroups and evolutionary origins: spinal cords, cell types and development in other chordates

Amongst living animals, cyclostomes (lampreys and hagfishes) represent the earliest diverging lineage of vertebrates. Two other living lineages share phylum Chordata with the vertebrates: cephalochordates (amphioxus) and tunicates (including sea squirts). Their evolutionary relationships are shown in Figure 1.1, and they offer a window into the timing and genetic basis of evolutionary change in spinal cord patterning.

Lampreys are moderately well studied and in many respects similar to jawed vertebrates, but with important differences. For example, they appear to have a different history of gene and genome duplication, and lack myelination and hence are assumed to have differences in glial cell development (Smith et al., 2013, Smith et al., 2018). The hindbrain is segmented into rhombomeres with many similarities to those of jawed vertebrates, including patterns of Hox expression and neural crest organisation (Parker et al., 2014a, Parker et al., 2016, Parker et al., 2019b). Adult spinal cords are quite well studied as a model for understanding reflex circuitry and regeneration (Herman et al., 2018, Shifamn and Selzer, 2015), and found to have a variety of identified cell types (Figure 1.4: discussed more below).

The sea squirt *Ciona* is very well studied, but like others in its lineage the spinal cord region of its CNS is secondarily reduced and without neurons or a tailbud growth zone (Figure 1.1), with only four rows of ependymal-glial cells in posterior neural tube (Crowther and Whittaker, 1992). This derived status limits its use in understanding ancestral spinal cord patterning. Amphioxus has been quite well studied and offers significant insight into the pre-vertebrate condition. Limitations are that most studies have focused on embryos, with adult spinal cord cell types not well described, and that most embryogenesis studies focus on the anterior CNS, and as discussed earlier it is questionable whether this region is comparable to the spinal cord, or instead homologous to the hindbrain.

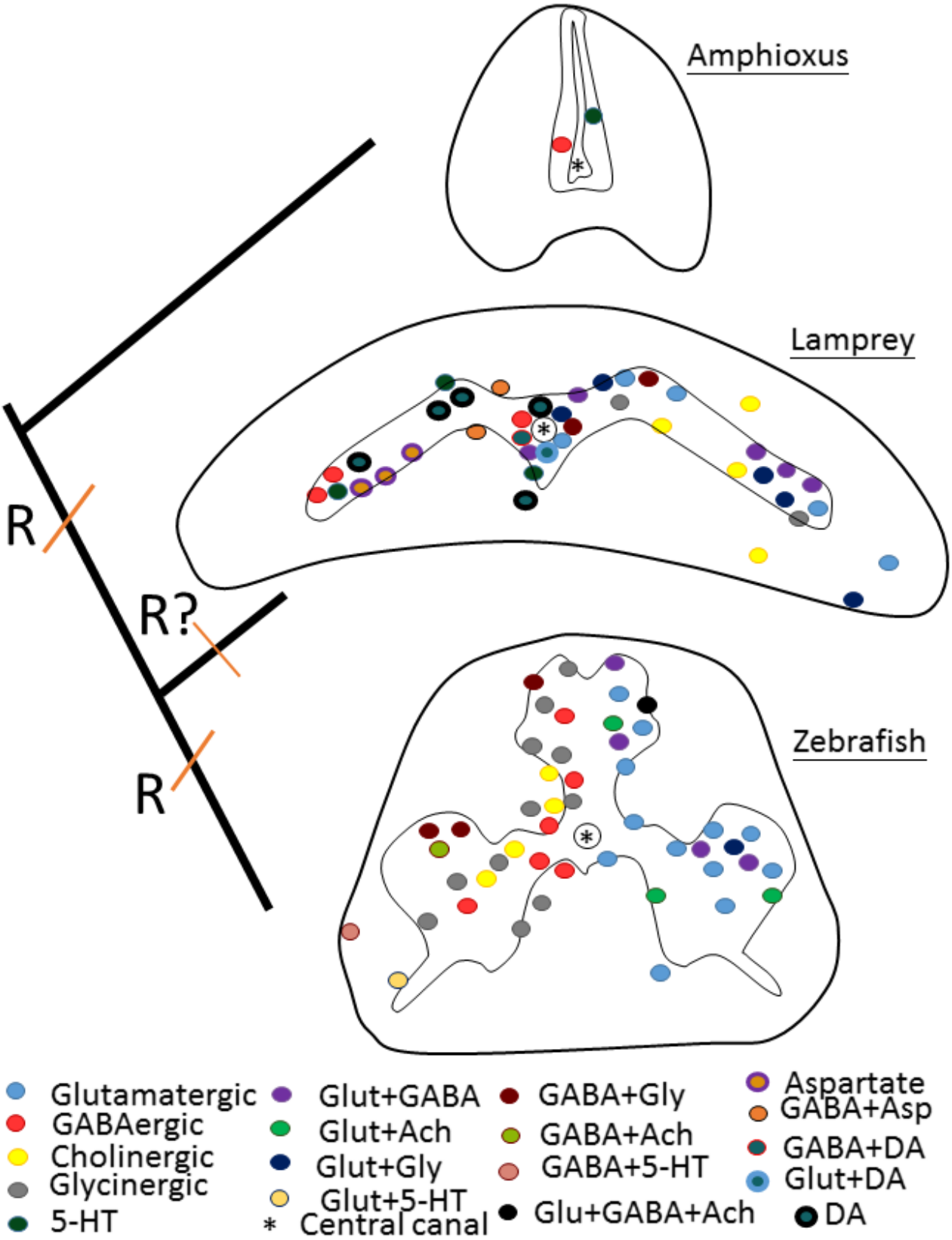

Amphioxus
Lamprey
Zebrafish
R
R?
R
Glutamatergic
GABAergic
Cholinergic
Glycinergic
5-HT
Glut+GABA
Glut+Ach
Glut+Gly
Glut+5-HT
* Central canal
GABA+Gly
GABA+Ach
GABA+5-HT
Glu+GABA+Ach
Aspartate
GABA+Asp
GABA+DA
Glut+DA
DA

Figure 1.4. Cross sections of adult spinal cords showing anatomy and cell types

A simple phylogenetic tree on the left shows the evolutionary relationships of the taxa. The Rs are marking autotetraploidy in the vertebrate ancestor, allotetraploidy in the gnathostome ancestor, and unresolved duplication in the agnathan ancestor. Adult spinal cord cross-section anatomy and cell type are shown on the right. Each type of neuron is only marked in one of either hemisegment. The dots only represent the general distribution and presence of the types of neurons, and not the numbers of expressing cells. Tunicates are not shown, as their spinal cord regresses before they reach adulthood. Data sources: Amphioxus**;** GABAergic neuron (Anadón et al., 1998), serotonergic neuron (Candiani et al., 2001), dopaminergic neurons not found in spinal cord (Moret et al., 2004), no data on other types of neurons. Lamprey**;** glutamatergic, glutamate-glycine, glutamate-GABA neurons (Fernández-López et al., 2012), dopamine-GABA neurons (Barreiro-Iglesias et al., 2009b), dopamine neurons (Fernández-López et al., 2015), dopamine-glutamate neurons (Fernández-López et al., 2017), aspartate, aspartate-GABA neurons (Villar-Cerviño et al., 2014), glycine, glycine-GABA neurons (Villar-Cerviño et al., 2008), cholinergic neurons (Quinlan and Buchanan, 2008). Larval data instead of adult for serotoninergic and GABAergic neurons (Barreiro-Iglesias et al., 2009a, Cornide-Petronio et al., 2014). Zebrafish; (Pedroni and Ampatzis, 2019).

1.7 Nerve cell types in adult lamprey and amphioxus spinal cords

Traditionally, the function of neurons are inferred from the type of neurotransmitters they secrete (Table 1.1), with glutamatergic rhythm-generating neurons signalling to acetylcholinergic motor neurons, and to glycinergic commissural neurons to inhibit motor neurons on the contralateral side (Goulding, 2009, Kiehn, 2016). However, recent studies have shown co-expression of two neurotransmitters in a single neuron (Vaaga et al., 2014), with a comprehensive zebrafish study (Pedroni and Ampatzis, 2019) identifying a small population of neurons with three neurotransmitters, acetylcholine, GABA and glutamate. Hence, while neurotransmitter distributions can help illustrate the number of different cell types, relying on them alone as an indication of cell homology may be problematic.

Figure 1.4 shows schematic cross-sections of spinal cords of amphioxus and lamprey, compared to a zebrafish. Lampreys show similarity to other vertebrates (Figure 1.4 and references therein), with a diversity of cell types as assessed by neurotransmitter complements. There are few data on neurotransmitter distributions in adult amphioxus, limited to studies showing GABAergic (Anadon et al., 1998), dopaminergic and serotonergic (Moret et al., 2004, Candiani et al., 2001) neuron distributions. Other studies have characterised amphioxus adult neurons by their immunoreactivity to a variety of neuropeptides including arginine vasopressin,

oxytocin, angiotensin II, cholecystokinin, urotensin, natriuretic peptide, FMRFamide, neuropeptide Y, calretinin and acetylated tubulin (Castro et al., 2003, Pestarino and Lucaroni, 1996, Castro et al., 2004, Castro et al., 2015, Uemura et al., 1994). However, these are not that informative for comparison across taxa as similar data from other species are not well developed. There are more data on embryonic and larval amphioxus (Candiani et al., 2012), showing that glutamatergic, serotoninergic, GABAergic, glycinergic and acetylcholinergic neurons can all be identified in the anterior CNS. However, none were identified more posteriorly. As considered more below, it is the posterior region that may be directly comparable to the spinal cord of vertebrates, with anterior regions potentially homologous to fore/mid brain and hindbrain. A possible explanation for the absence of neurotransmitter positive cells in posterior regions is that posterior cells had yet to differentiate and take on an identity. Irrespective, it means caution is needed in comparison between current understanding of neurotransmitter distributions, and more work is needed on the posterior larval and adult CNS of amphioxus.

Table 1.1. Common Types of neurotransmitters

Excitatory	**Inhibitory**	**Modulatory**
Glutamate	Glycine	Acetylcholine
Aspartate	GABA	Noradrenaline
Nitric Oxide	Serotonin	
	Dopamine	

1.8 AP spinal cord patterning in lampreys, sea squirts and amphioxus

Mechanisms of AP spinal cord patterning have not been well studied in lamprey embryos, although what is known is consistent with a general similarity to jawed vertebrates. Embryos grow from anterior to posterior via a tailbud, forming new somites and spinal cord over more than two weeks of development. RA appears to be synthesised by somite cells (Castillo et al., 2010) and Hox gene expression has been identified in the hindbrain and spinal cord, though posterior Hox genes have not been well studied (Parker et al., 2014a, Parker et al., 2019b, Pascual-Anaya et al., 2018).

Some *Ciona* Hox genes are expressed in the CNS, but the *ascidian* Hox cluster is fragmented (Ikuta et al., 2004, Sekigami et al., 2017) and this is not an ancestral state. Manipulation of RA and FGF signalling in *Ciona* can affect AP patterning, including Hox expression (Katsuyama et al., 1995, Pasini et al., 2012). Amphioxus Hox genes are expressed in a nested pattern in the CNS, and in the genome are located in a single canonical cluster (Garcia-Fernandez and Holland, 1994, Wada et al., 1999). There is anteriorisation of *Hox1*, *Hox2*, *Hox3*, *Hox4* and *Hox6* expression in the neural tube when embryos are treated with RA. *Hox14* is also responsive to RA even though it is only expressed in the posterior notochord, gut and cerebral vesicle at larval stage (Pascual-Anaya et al., 2012, Schubert et al., 2006). Several Wnt genes show tailbud expression (Schubert et al., 2001, Somorjai et al., 2018), though only limited manipulation of Wnt signalling has been conducted in amphioxus to date (Dailey et al., 2017) and we do not know whether Wnt regulates Hox. FGF inhibition alone does not seem to have an effect on *Hox1* expression but when coupled with an RA antagonist there is no *Hox1* expression in the neural plate (Bertrand et al., 2015). This study also found no evidence for opposing effects of RA and FGF on AP patterning in amphioxus, since inhibition of FGF caused loss of expression of typical downstream genes *ER81*/*Erm*/*Pea3* and *Sprouty*, but these remained unchanged when RA was inhibited. Similarly, FGF inhibition did not affect RA signalling downstream target genes, *Cdx* and *Xlox,* while RA inhibition caused reduction in expression (Osborne et al., 2009; Bertrand et al., 2015). In the larval stage of amphioxus, GABA-immunoreactive neurons exist in clusters along the neural tube posterior to the cerebral vesicle boundary. Exposure to BMS493, an RA antagonist, has a anteriorising effect on these cells with the gain of an extra cluster, while RA treatment has an posteriorising effect, with three GABA clusters lost (Zieger et al., 2018a). *Pax6* and *Prdm12* seem to have posteriorised expression in the neural tube under a high concentration of RA. When RA signalling is inhibited, *Prdm12* neural tube expression is reduced while *Pax6* remains the same (Carvalho et al., 2017).

A consideration when comparing the outcomes of these experiments to studies of vertebrates, including lampreys, is whether homologous regions of the CNS are being compared, particularly with respect to the hindbrain and spinal cord. The primary pigment spot of amphioxus develops at about the level of the boundary between somite 4 and somite 5, and many studies focus on cells anterior to this point. However, this region is within the domain of *Hox1* and *Hox3* expression, with the anterior boundary of *Hox4* mapping around the level of somite 6 (Wada et al., 1999).

If we consider Hox gene expression boundaries as landmarks between lineages, this region would be homologous to the vertebrate hindbrain and not the spinal cord. Spinal cord would be homologous to regions posterior to about somite 6/7. Hence some comparisons between vertebrate spinal cord and amphioxus CNS have to be made cautiously, acknowledging the lack of certainty around what exactly is, and is not, homologous.

1.9 DV spinal cord patterning in lampreys, sea squirts and amphioxus

Lamprey, amphioxus and sea squirt embryos all show DV polarisation of the spinal cord regions of their CNS. In most lineages orthologous signalling molecules are also expressed in similar positions as in jawed vertebrates. In lamprey the spinal cord is exposed to dorsal Bmp and Wnt signalling (McCauley and Bronner-Fraser, 2004, Guerin et al., 2009), with Hh localised to notochord and floor plate (Kano et al., 2010, Sugahara et al., 2011). Hh is also expressed by notochord cells in *Ciona* (Takatori et al., 2002), and by notochord and floor plate cells in amphioxus (Shimeld, 1999), while both Bmp and Wnt gene expression in amphioxus is consistent with a role in dorsal patterning (Panopoulou et al., 1998, Hu et al., 2017). The spinal cord is thus bounded across the DV axis by the expression of the same signals in all chordates.

Lamprey homologues of some spinal cord DV patterning genes have also been studied. The Hh receptor, *Patched* is expressed in the ventral-medial neural tube (Hammond et al., 2009) and a Gli gene that is probably orthologous to *Gli3* is expressed in the dorsal neural tube (Sugahara et al., 2011). These data imply lamprey spinal cord is patterned by Hh signalling, though this has yet to be validated experimentally. Hammond et al. (2009) also showed *Engrailed*, a V1 neuron marker in jawed vertebrates, is expressed in the ventral-medial neural tube, consistent with the presence of a homologous cell population. *Pax3/7*, a dl2-6 marker, is expressed in the dorsal neural tube (Kusakabe and Kuratani, 2005), while *OligA* is expressed in dorsal and ventral spinal cord (Lara-Ramirez et al., 2019), paralleling jawed vertebrate Olig gene expression in pMN and dl1-3 (Alaynick et al., 2011). Insightful studies confirmed that *Pax6β* and *Nkx2.2* are expressed in the spinal cord, and in addition showed a specific cell population in the ventral spinal cord expresses *Hb9*, a MN marker from the Mnx homeobox gene family (Yuan et al., 2018, Murakami et al., 2001, Derobert et al., 2002, Ravi et al., 2019). Although the spinal cord region has been overlooked in many other studies, hindbrain expression of other neural tube patterning genes has been demonstrated, including *Msx, Dbx, Wnt5, Gsh,* and *Isl*, and these are likely to extend into the spinal cord (Guerin et al., 2009, Shigetani et al.,

2002, Sugahara et al., 2011, Cerny et al., 2010). These data combined suggest many aspects of DV spinal cord patterning are conserved between lampreys and jawed vertebrates.

As considered above, amphioxus has a much simpler CNS than any vertebrate, with fewer cells and probably also fewer types of cells. ML progenitor zones are lacking (see Section 1.10), with evidence instead suggesting that individual cell types differentiate in situ across the DV axis. Nevertheless, studies have shown similarities in gene expression patterns to vertebrates, the most comprehensive being a study focused on midneurula stage (Albuixech-Crespo et al., 2017). This showed that:

- *Hh, Nkx2.1, Nkx6* and *Gsc* are expressed by the floorplate.
- *Six3/6*, *Lhx2/9b*, *Zic*, *Msx*, *Pax2/5/8*, and *Pax3/7* are expressed in the peripheral neural plate, which will become the dorsal neural tube as the plate folds.
- *Pou3f*, *Sim*, *FoxD*, *Meis*, *Lef*, *Lhx1/5*, *Hox3*, *Hox6*, *FoxB* are expressed in the internal neural plate, which will be the ventral neural tube.
- O*tx*, *Gbx*, *Fezf*, *Irx*, *Pax4/6*, *Six3/6*, *Nkx2.2*, *Meis*, *Rx*, *Hox1*, *Wnt3*, *Wnt7*, *Nova*, *Ebf/COE* are expressed in both internal and peripheral neural plate.

Broadly speaking, these patterns correspond to the relative expression of their orthologues in jawed vertebrates (though there are also many differences). However, the focus on the midneurula stage also means the neural plate at this time may be homologous to the brain only, as discussed above. At later developmental stages many of these genes mark scattered, isolated cells in the neural tube including *Islet* (Jackman and Kimmel, 2002), *Pax3/7* (Holland et al., 1999), *Evx* (Ferrier et al., 2001b), *OligA* (Beaster-Jones et al., 2008), *prdm12, engrailed* (Thélie et al., 2015), *Tlx* (Kaltenbach et al., 2009b), *Msx* (Sharman et al., 1999), *Mnx/Hb9* (Ferrier et al., 2001a), and *Err* (Schubert et al., 2006), with the exception of *IrxA* being expressed along the whole neural tube at late neurula stage (Beaster-Jones et al., 2008).

To date, there are insufficient functional studies for us to determine if and how Hh, Bmp and Wnt signalling might be involved in DV patterning of the amphioxus spinal cord. Wnt signalling has yet to be studied at the right developmental stages. Bmp signalling is known to regulate the patterning of epidermal neurons, though its impact on the spinal cord is not clear (Lu et al., 2012). Hh signalling has been studied with the smoothened inhibitor cyclopamine and no changes in neural patterning gene expression were observed (Ono et al., 2018). However, the study focus was gill

slit formation, and the drug treatment time only a 20 minute window, which might explain this outcome. Through generation of $Hh^{-/-}$ embryos using TALEN knockout, the role of Hh signalling in regulating left right asymmetry in amphioxus has also been studied (Hu et al., 2017). Again, neural patterning was not the focus of the work so few relevant genes or developmental stages were analysed. However, the study did show that *Err* expression became symmetrical in the brain and was lost from more posterior neural tube at the single stage shown. This hints at a role for Hh signalling in amphioxus neural patterning, though this needs more focused investigation with appropriate marker genes.

1.10 ML spinal cord patterning in lampreys, sea squirts and amphioxus: differences in cell number

As considered in Section 1.5, jawed vertebrate embryos build complexity into spinal cords via a ventricular progenitor cell population that produces new cells over a relatively long period of development. This yields spinal cords with multiple cell layers, and gives the potential for different cell types to be born at different times. Do other chordates operate with similar mechanisms? Lamprey embryos do, and resemble jawed vertebrates with a well-defined progenitor cell zone visible both by cell morphology and gene expression in brain and spinal cord (Guerin et al., 2009, Lara-Ramirez et al., 2019). The mechanism is conserved, since Notch signalling maintains progenitor cells at the ventricular zone, with inhibition of Notch by DAPT causing loss of progenitor cell markers *PCNA, OligA* and *HesB*, and premature differentiation as seen by gain of the differentiation markers *CoeA* and *CoeB,* (Lara-Ramirez et al., 2019).

With their highly reduced posterior CNS, sea squirt embryos do not have cells comparable to the ventricular progenitors of jawed vertebrates. Amphioxus also lacks a ventricular progenitor cell zone. Studies of cell proliferation have failed to identify more than a few scattered dividing cells (Bertrand and Escriva, 2011), while Notch signalling is lost early from the neural plate (Holland et al., 2001b, Rasmussen et al., 2007) and DAPT inhibition of Notch signalling has no clear effect on the CNS (Lu et al., 2012). This does not mean amphioxus has no proliferation in the CNS. Tailbud cells continue to produce new cells, scattered dividing cells in more anterior regions of the CNS show not all these cells are post-mitotic, and the anterior brain does eventually develop more of a layered structure. However, a ventricular proliferative layer is missing throughout the embryonic and larval spinal cord, and this matches the anatomy of the adult spinal cord which, judging from histological cross-sections,

remains one to two cells thick from embryo to adult (Bocina and Saraga-Babic, 2006, Bone, 1959, Bone, 1960, Candiani et al., 2012).

1.11 Conservation and innovation in chordate spinal patterning

Overall, the picture that emerges from these studies is that broad organisation of the neural plate shows conserved aspects between amphioxus and vertebrates. This includes AP patterning by Hox and DV organisation, as might be expected from evidence for more ancient similarities in basic nervous system development amongst bilaterally symmetrical animals (Holland et al., 2013). Markers of differentiated vertebrate cell types also sometimes show expression in subsets of amphioxus neurons. These may reflect conserved cell types, perhaps also functionally related and expressing similar neurotransmitters as discussed above, but this has not been experimentally demonstrated. The role of patterning signals is much less clear and in need of better experimental validation, especially those involved in the DV axis. Irrespective of the outcome of such work though, it is clear that amphioxus does not develop a ventricular progenitor cell layer as seen in vertebrates. Cells here are not divided into DV zones, rather cells are patterned at the level of individual differentiating cells.

1.12 A role for gene duplication in spinal cord evolution

It has long been thought that early vertebrate evolution was marked by extensive gene duplication, with broad comparisons of syntenic and paralogous gene organisation between jawed vertebrate and amphioxus genomes seen by most authors as showing this had occurred through two rounds of genome duplication, the '2R' hypothesis (Putnam et al., 2008). New genes offer the potential for new function (neofunctionalisation), and can diverge through splitting ancestral functions between duplicates (subfunctionalisation) or having duplicates maintaining broad ancestral functions with at least one becoming restricted to a more distinct function (specialization). How 2R gene duplication relates to evolutionary change in development and morphology has been hard to pin down, however, since both happened so long ago and intermediate lineages have gone extinct, meaning relating the timings of duplication, potential neofunctionalisation/subfunctionalisation/ specialisation and morphology is tricky. Recent observation suggest that specialisation occurs mostly in neural tissue, through gain in regulatory elements producing a more complex GRN (Marletaz et al., 2018).

Amphioxus and tunicates predate 2R and tend to have single copies of transcription factor family genes, except where secondary loss has occurred (Putnam et al., 2008, Paps et al., 2012). The position of lampreys with respect to 2R has also been debated. Lamprey genes often fail to resolve well in molecular phylogenetic analyses (Kuraku et al., 2009). These authors had analysed 55 gene families using the maximum likelihood method and could only conclude that there is a genome duplication before the divergence of the cyclostome lineage, but could not say anything definitive about subsequent lineage-specific changes. Additional gene family tree analyses like *Ngn, Coe,* and *Fox* also do not resolve the issue (Lara-Ramírez et al., 2015, Lara-Ramírez et al., 2017, Wotton and Shimeld, 2011). Escriva et al. (2002b) used neighbour-joining phylogenetic analysis on 33 gene families and came to a similar conclusion that one genome duplication happened before the split of cyclostomes and gnathostomes and that there is further gene-genome duplication and massive gene loss on separate vertebrate lineages. Others who analysed the Hox cluster and its relevant conserved non-coding elements had found lamprey orthologues are in many-to-many relationships with elephant shark and human orthologues, which could be interpreted as additional and/or independent duplications in the lamprey lineage after 2R (e.g. (Mehta et al., 2013, Escriva et al., 2002b)). More recently it was concluded that lampreys share the 2R duplications with jawed vertebrates after comparing the lamprey genome with a reconstructed vertebrate ancestor genome. The authors started with 61 extant species genomes to reconstructe an ancestral *Amniota* genome then extrapolated this to the pre-vertebrate genome taking into account chromosome duplication and fusion (Sacerdot et al., 2018). However, when comparing ancestral chordate linkage groups with lamprey scaffolds, Simakov et al. (2020) argued that the vertebrate ancestor had one round of whole genome duplication and the agnathan and gnathostome lineages have separate duplications/allotetraploidy. They found 6-8 lamprey chromosomes associated to each ancestral chordate linkage group. Furlong and Holland (2002) suggested that the vertebrate ancestor could be octaploid after 2R followed by delayed rediploidisation, which would also explain having 6-8 lamprey chromosomes associate to ancestral chordate linkage groups. It looks like different analysis methods have provided evidence for different pathways of genome evolution.

Irrespective of the precise timing and mechanism, it is clear lampreys do have more copies in many transcription factor gene families than amphioxus or sea squirts. Many of the jawed vertebrate paralogues are expressed by different cell populations in the spinal cord (Figure 1.3 and Table 1.2), as assessed by in situ hybridisation

(summarised by (Alaynick et al., 2011, Lai et al., 2016)) and more recently by single cell sequencing (Delile et al., 2019). Some of this could represent subfunctionalisation of original broader expression patterns into subdomains, however, some must also be neofunctionalisation and/or specialisation, for example the paralogues *Nkx6.1* and *Nkx6.2*, and *Dbx1* and *Dbx2,* overlap but are functionally distinct, setting up the interneuron progenitor zones p0, p1 and p2 (Alaynick et al., 2011). This would not be possible in amphioxus, with only one copy of each, and the duplication and resolution of *Nkx6* and *Dbx* into interacting pairs may have allowed the evolution of new interneuron domains. Reiterating the earlier section on gnathostome DV spinal cord patterning, *Nkx6.1 and Nkx6.2* cross repress with *Dbx2* and *Dbx1* respectively in chick (Briscoe et al., 2000), which could not be the case in lamprey because there is only one orthologue of *Dbx*. Also in chick, *Nkx2.2* and *Olig2* cross repress each other (Sun et al., 2003). Chapter four of this thesis attempted to test if this is the case in lamprey.

Table 1.2. Paralogues performing different roles in gnathostomes

Gene family	Paralogue expression domain(s)	Reference
Neurog	*Neurog2* dP2-pMN	(Timmer et al., 2002, Kriks et al., 2005, Lee et al., 2003)
	Neurog1 dP2 *Neurog3* p3	
Msx	*Msx1* roofplate-dP5 *Msx2* roofplate-dP1	(Timmer et al., 2002)
Pax3/7	*Pax3* dP1-dP6 *Pax7* dP2-dP6	(Moore et al., 2013, Gard et al., 2017)
Pax2/5/8	*Pax2* dI4, dI6-V1 *Pax8* dI4, dI6-V1	(Batista and Lewis, 2008)
Gata2/3	*Gata2* p2, V2 *Gata3* V2	(Nardelli et al., 1999)
Gsx	*Gsx2* dP3-dP5 *Gsx1* dP4-dP5	(Kriks et al., 2005)
Dbx	*Dbx2* dP5-p1 *Dbx1* dP6-p0	(Timmer et al., 2002)
Nkx6	*Nkx6.2* p1-floorplate *Nkx6.1* p2-floorplate	(Briscoe et al., 2000)
Olig	*Olig3* dP1-3, p0,p2,p3,V3 *Olig2* pMN	(Liu et al., 2014, Fu et al., 2002)
Barhl	*Barhl1* dI1 *Barhl2* dI1	(Ding et al., 2012, Bulfone et al., 2000)
Islet1/2	*Islet1* dI3,MN *Islet2* MN	(Song et al., 2009, Hutchinson and Eisen, 2006)
Lhx1/5	*Lhx1* dI2, dI4, dI6-V1 *Lhx5* dI2, dI4, dI6-V1	(Pillai et al., 2007)
Lhx2/9	*Lhx2* dI1 *Lhx9* dI1	(Wilson et al., 2008)
Evx1/2	*Evx1* dI1, V0 *Evx2* V0	(Juárez-Morales et al., 2016, Moran-Rivard et al., 2001)
Tlx1/3	*Tlx3* dI3, dI5 *Tlx1* dI5	(Qian et al., 2002, Borromeo et al., 2014, Logan et al., 1998)

1.13 Future perspectives: evolutionary change in spinal cord patterning

Comparing spinal cord development across the three chordate lineages (Figure 1.1) reveals both conservation and change. The general AP and DV polarisation of the spinal cord is conserved, with signalling molecules expressed by homologous tissues in similar positions (Figure 1.2). From this we can infer the common ancestor of vertebrates had tailbud mediated posterior growth, a DV axis framed by cells expressing Hh, Bmp and Wnt signals, and an AP axis framed by RA and FGF. Downstream from these signals conservation is less apparent, and differences between lineages lead me to hypothesise a set of evolutionary innovations that underlie vertebrate spinal cord development.

First, considering the AP axis, RA has a role in regulating Hox in both amphioxus and *Ciona*, showing this was likely present in the chordate common ancestor. Wnt signalling has a more ancient role in posterior specification and expression of Wnt ligand genes in amphioxus is consistent with a role in Hox regulation (though this has yet to be experimentally demonstrated). FGF has a role in posterior regulation in both amphioxus and *Ciona*, showing this is also likely ancestral, though it is not yet clear how much of an influence it had on the Hox genes. In this context it may be notable that comparison of vertebrate and amphioxus Hox cluster regulation suggests an important difference in the regulatory regions that lie outside the cluster, with vertebrate Hox being bipartite regulated while amphioxus Hox uses a single long-range cis-regulatory element on the anterior side (Acemel et al., 2016). This suggests aspects of FGF regulation of Hox may be vertebrate-specific. Understanding whether and how FGF regulates Hox in lamprey and amphioxus would resolve this.

Second, patterning downstream of DV signals looks to be more elaborate in vertebrates. This suggests differences in the GRNs responsible for patterning and differentiation of neural cells. This might reflect new roles for divergent paralogues following 2R duplication, incorporation of new genes into GRNs, and/or new regulatory connections between genes already involved in spinal cord development in the common ancestor (Marletaz et al., 2018). Understanding the regulatory connections between signals and targets in lamprey and amphioxus, and between transcription factors in the DV GRN, would answer this.

Third, progenitor cell maintenance and ML organisation look quite different between vertebrates and other chordates. Lampreys are similar to jawed vertebrates, with ventricular progenitor cells divided into progenitor zones with proliferation under

Notch regulation, and giving rise to subset populations of neurons and glia patterned by DV signals. These were therefore present in some form in the common ancestor of vertebrates, however *Ciona* and amphioxus show no ventricular progenitor cell layer in the spinal cord, and hence no progenitor cell zones. This suggests that a major innovation in vertebrate spinal cord evolution was the origin of a cohesive ventricular progenitor cell pool, able to produce cells over a long period of development. Understanding this third change means comparing how progenitor cell maintenance and proliferation are controlled in vertebrates and other chordates; challenging at present since neither is well understood.

Furthermore, as discussed above in Section 1.5, proliferation and DV patterning are linked in vertebrates. A simple hypothesis arising from this is that, primitively, DV signals such as Hh directly regulated differentiation of early neural plate cells into specific cell types. In vertebrates this has evolved by insertion of a progenitor cell stage with new regulatory connections between patterning and proliferation. Evolution of the DV GRN would parallel this, incorporating new paralogues and new connections, and perhaps generating a system with the feedback and cross-repressive properties allowing the sudden state changes and scaling necessary to divide a spinal cord into discrete zones of cells.

While these hypotheses are speculative, they make predictions about ancestral regulatory connections and are hence testable by dissecting gene regulation in multiple lineages. With the recent development of methods for manipulating genes and regulatory elements in amphioxus and lamprey (Parker et al., 2014a, Liu et al., 2013, Feng et al., 2014, Li et al., 2014a, Nikitina et al., 2009, Square et al., 2015, Zu et al., 2016), plus application of genome-wide methods to identify regulatory landscapes (Marletaz et al., 2018), this is rapidly becoming tractable.

1.14 Aims of this Thesis

This thesis aims to establish how the central nervous system of vertebrates has evolved into such a complex organ by focusing on the most conserved part, the spinal cord. I have already approached this question from the adult neuroanatomical point of view (Figure 1.4). I am also going to approach it from an evolution and development perspective by:

1.Determining if gnathostome spinal cord cell type developmental marker genes are conserved in lamprey and amphioxus. If not, then what markers are occupying the spinal cord?

2.Testing whether and how the common deuterostome signalling pathways Hh, BMP, RA, Notch, Fgf, and Wnt are patterning the spinal cord of lamprey and amphioxus, and comparing results to gnathostomes to see if the signals function similarly in agnathan and cephalochordate development.

3.Finding out how cell type marker genes are regulated by looking for open chromatin sites and conserved non-coding elements, then performing enhancer reporter assay to test enhancer properties.

4.Establishing how some of the genes of interest had evolved using synteny and phylogeny analyses.

Chapter 2: The anatomy of gnathostome, lamprey, and amphioxus spinal cords as visualised through developmental gene expression

(The majority of the data shown in this chapter were generated by myself, with some additional data generated through a collaboration with Guang Li and Qiongqiong Ren in Xiamen. Where figures contain data generated through this collaboration, provenance is identified in the figure legend. Some of the data and ideas generated as part of this collaboration were included in Ren et al. (2020) on which I am co-author. Other data are unplublished.)

2.1 Introduction

In gnathostome spinal cord development, a lot of data have been amassed describing which genes are expressed in which cells and through which periods of development (i.e. Figure 1.3, more extensively described in (Lai et al., 2016)). I am focusing on the earlier stage of development where transcription factors are acting in the progenitor cells, rather than the more differentiated states where neuron migration and axon guidance happen. This is because the technological tools to collect these other data are still relatively new in amphioxus (Hua et al., 2018) and non-existent in lamprey, making informative comparative study very challenging. In this chapter, I describe relevant transcription factor gene expression, plus expression of some other relevant genes (for example encoding signalling proteins) in amphioxus and lamprey, and compare it to that of other vertebrates using published data.

I hypothesise that the expression pattern of these lamprey genes would be similar to gnathostomes because lampreys have similar spinal cord embryonic anatomy to the rest of vertebrates (for example they all have a ventricular zone and a mantle zone) and they have diverged after at least one (Simakov et al., 2020) round of whole genome duplication so should have similar complexity of the gene regulatory network for spinal cord organogenesis. Lamprey genes (*OligA, OligB, PtchB, AxinA, Nkx2.2, Nkx2.1, Nkx6A, Pax6β, Pax7, Dbx, MsxA*) were chosen because they should provide information on whether there are spatially conserved homologous cell populations in lamprey developing spinal cord. My data suggest that lamprey spinal cord molecular pattern is evolutionarily conserved with the rest of the vertebrates.

On the other hand, amphioxus only has a single layer of cells in its neural tube throughout development meaning cells migrating laterally as they differentiate does not apply. Amphioxus has conclusively diverged before the first round of whole genome duplication so I hypothesise that amphioxus spinal cord organogenesis has a

very different, maybe simpler development because it lacks the extra gene duplicates, and associated functions, in the gene regulatory network from vertebrates. The genes chosen to compare between lamprey and gnathostomes are known markers of progenitor cell populations. However, some amphioxus orthologues are not expressed in neural tube at all, or only in the cerebral vesicle so some of the putative gnathostome differentiated cell markers (*Netrin, Nkx6, Gbx, Lhx3, Vacht, Pax2/5/8, IrxA, Islet, Hb9/Mnx, Axin, Dbx, Zic, SoxB1a*) are also used for comparison. This explains why different genes in amphioxus and lamprey were used in chapter three to test signalling manipulation effects. My data suggest that amphioxus and vertebrate spinal cord cell types have evolutionarily diverged. Table 2.1 indexes which gene is shown in which figures and their phylogenetic trees.

In addition to the above, Wnt family expression is characterised in lamprey for comparison to a published amphioxus Wnt expression atlas (Somorjai et al., 2018). This complements functional analyses described in chapter three, since one of Wnt's many functions in gnathostomes is patterning the spinal cord.

Table 2.1. Index for genes of interest expression and phylogenetic analysis

Genes	Gnathostome references	Lamprey figure no. and references	Amphioxus figure no. and references	Phylogenetic tree figure no. or references
Olig	(Fu et al., 2002, Liu et al., 2014)	2.1, (Lara-Ramirez et al., 2019), 2.2	2.3, (Ren et al., 2020)	(Lara-Ramirez et al., 2019)
Nkx6	(Vallstedt et al., 2001)	2.4	2.5, (Albuixech-Crespo et al., 2017)	Supp. Fig. 2.8
Mnx/Hb9	(Bello et al., 2012)	(Yuan et al., 2018)	2.6, (Ferrier et al., 2001a)	Supp. Fig. 2.6
Nkx2.2	(Fu et al., 2002)	2.7, (Sugahara et al., 2011, Yuan et al., 2018)	(Albuixech-Crespo et al., 2017)	Supp. Fig. 2.7
Nkx2.1	(Anderson et al., 2002)	2.8, (Sugahara et al., 2011)	(Albuixech-Crespo et al., 2017)	Supp. Fig. 2.7
Dbx	(Fogarty et al., 2005)	2.9 (Guerin et al., 2009)	2.10	Supp. Fig. 2.9
Pax4/6	(Bel-Vialar et al., 2007)	2.11, (Murakami et al., 2001)	(Albuixech-Crespo et al., 2017, Glardon et al., 1998)	Supp. Fig. 2.4
Pax2/5/8	(Burrill et al., 1997)	(McCauley and Bronner-Fraser, 2002, Murakami et al., 2001)	2.12(Kozmik et al., 1999)	(McCauley and Bronner-Fraser, 2002)
IrxA	(Houweling et al., 2001)	Not done	2.13(Kaltenbach et al., 2009a)	(Irimia et al., 2008)
SoxB1a	(Tanaka et al., 2004)	(York et al., 2019, Cattell et al., 2012)	2.14(Holland et al., 2000)	(Meulemans and Bronner-Fraser, 2007; Cattell et al., 2012)
Pax3/7	(Andrews et al., 2017, Hoover et al., 2008)	2.15, (Kusakabe et al., 2011)	(Holland et al., 1999, Barton-Owen et al., 2018)	(McCauley and Bronner-Fraser, 2002)
MsxA	(Misra et al., 2014)	2.16, (Cerny et al., 2010)	(Sharman et al., 1999)	(Cerny et al., 2010)
Zic	(Merzdorf, 2007)	(Castillo et al., 2010)	2.17, (Gostling and Shimeld, 2003)	(Gostling and Shimeld, 2003)
Ptch	(Matsumoto et al., 2007)	2.18, (Hammond et al., 2009)	(Ren et al., 2020)	(Hammond et al., 2009)
Axin	(Yu et al., 2007)	2.19	2.20, (Jing et al., 2016)	Supp. Fig. 2.5
Vacht	(Clovis et al., 2016)	Not done	Supp. Fig. 2.1, (Candiani et	(Candiani et al.,

			al., 2008)	2008)
Netrin	(Kennedy et al., 2006, Wang et al., 1999, Yin et al., 2000)	Not done	Supp. Fig. 2.3,(Shimeld, 2000)	(Shimeld, 2000)
Gbx	(Waters et al., 2003)	Not done	Supp. Fig. 2.4, (Castro et al., 2006)	Supp. Fig. 2.10 (Castro et al., 2006)

2.2 Methods

2.2.1 *Lampetra planeri* husbandry

L. planeri (Brook lamprey) embryos were collected at the New Forest National Park, Southampton, with the permission from the Forestry Commission. The embryo cultures were kept at 16°C in filtered river water.

2.2.2 *Branchiostoma floridae* husbandry

Full methods detailed in Li et al. (2012). Briefly, *B. floridae* (amphioxus) were reared in buckets of sand and seawater at 22°C. Heat-shock spawning was induced according to (Li et al., 2013, Li et al., 2015). In vitro fertilisation was performed to obtain embryos by pipetting 20µL of sperm on to a tightly clustered group of 300-500 eggs. Excess sperm was washed away by replacing half of the artificial sea water in the petri dish at a time, for three times. For ease of visualising fertilised eggs with the naked eye, a dark background and sideways light from a torch was used. The embryo cultures were kept at 29°C in artificial sea water.

2.2.3 Molecular cloning

When lamprey embryos had grown to stage 24 to 27, RNA was extracted using TRIzol Reagent according to the manufacturer's protocol. cDNA was synthesized from the RNA using SuperScript III reverse transcriptase. Primers of genes of interest were designed base on the transcript sequences. Amplicons were cloned into PCRII vector and sequenced. Some clones were synthesised commercially (Twist Bioscience) with SP6 and T7 sequences flanking the gene ends. The amplicon sequences, gene accession number, and relevant primers used are in Supplementary Table 2.1 and 2.2 for lamprey and amphioxus respectively. Most amphioxus clones were gifted from Guang Li, with others from Linda Holland, Sky Yu, and Peter Holland.

2.2.4 Whole mount In situ hybridization

The lamprey protocol was adapted from Lara-Ramírez et al. (2015). Briefly, when the embryos had reached the desired stages, they were fixed in 4%PFA overnight, bleached in hydrogen peroxide and methanol, treated in proteinase K, refixed for a short period of time then hybridized with riboprobe which was synthesized by T7 or Sp6 RNA polymerase from the cloned gene sequence. After hybridization in formamide based hybridization buffer at 65°C for two days, the embryos were washed, treated with RNAse A, then blocked with heat treated sheep serum together with anti-DIG antibody overnight. The antibody was washed off and embryos

proceeded to the colorimetric reaction with NBT/BCIP in NTMT (at pH9.5). Once sufficient colour had developed, the reaction was halted by washing with PBT and fixed in PFA. The embryos were then dehydrated in graded methanol and cleared in BABB for imaging.

Amphioxus whole mount in situ hybridisation followed Yu and Holland (2009) with slight modifications: 48-well plate and in situ hybridisation baskets (P1000 pipette tips melted into 55micron mesh) were used to scale up data production. The glutamate treatment step was skipped because of its inconsistency in quenching protease K action.

2.2.5 Sectioning

Whole mount embryos were sectioned to observe expression at cellular resolution. Embryos that were cleared in BABB were switched to ethanol, then to Xylene Substitute (Sigma-Aldrich: 265-068-8). Paraffin wax was melted in 65°C then poured into a mould on a cold surface. Embryos were positioned with the posterior end towards the bottom as the wax was setting, then more hot paraffin filled the mould and a cassette was placed on top. After the paraffin had cooled, a microtome was used to create 10micron thick sections. The ribbon of sections were smoothened out in a 40°C waterbath then attached onto glass slides. The slides were deparaffinised with HistoClear (National Diagnostics: NDIHS-200) then mounted in glycerol for imaging. Alternatively, embryos were also sectioned directly after BABB infiltration by razor blade under dissecting microscope on a glass slide with the help of a pair of forceps to stabilise the embryo.

2.2.6 Imaging

Whole mount lamprey embryos were imaged using Leica MZ16F microscope with Leica DFC300 FX camera at a magnification of 50X.
Sectioned lamprey embryos and whole mount amphioxus embryos were imaged with Zeiss Axioskop2 Plus with Moticam 580 camera at a 200X magnification.

2.2.7 Phylogeny analysis

Protein sequences of different species were collected from NCBI database, while lamprey sequences were extracted from the 2017 *Petromyzon marinus* genome assembly hosted by the Stowers site (https://genomes.stowers.org/sealamprey) using both BLAST searches and predicted gene annotation. Some homeobox gene sequences were extracted from homeodb (Zhong and Holland, 2011;

http://homeodb.zoo.ox.ac.uk/). Multiple sequence alignment was done with Clustal Omega on the EMBL-EBI site. Default parameters were used with Clustal chosen as output format (https://www.ebi.ac.uk/Tools/msa/clustalo/). The alignment was downloaded as a text file and converted into a .meg file using MEGA7. To help avoid long branch attraction, only the homeodomain was used for alignment and tree construction. For other genes, alignments were trimmed by eye. The trees are rooted and have *Saccoglossus kowalevskii,* a hemichordate, as an outgroup. The trees were constructed with default settings in MEGA7 with the maximum-likelihood method and Jones-Taylor-Thornton model, and nearest-neighbor-interchange heuristic method for tree inference options (Kumar et al., 2015).

2.3 Results and Discussion

2.3.1 Defining spinal cord in amphioxus

Many previously studied neural marker genes are different in amphioxus compared to vertebrates with most of the same genes not expressed in the spinal cord region, so more genes were explored hoping to capture spinal cord development and the definition of "spinal cord" was expanded to include the hindbrain since the primordial spinal cord-hindbrain boundary is not well defined. If we classify the neural identity of the hindbrain and spinal cord by Hox collinearity then spinal cord should start around somite 5 or 6 at the late neurula stage and shift to the posterior pigment spot (which is past half way of its body length) as it ages to the larval stage (Schubert et al., 2006). However, anatomically there is a clear cerebral vesical which resembles the primordial brain followed by a thinner neural tube resembling a spinal cord. Moreover, judging from my own *Vacht* amphioxus data at late larval stage (Supp. Fig. 2.1), motor neurons are still only present at the hindbrain level which makes one wonder how the larva moves its tail. It is likely that the motor neurons at hindbrain level extend axons to the tail, though this has not been described. With such different neural anatomy from the vertebrate spinal cord, I think it is reasonable that we consider the hindbrain homologous region of amphioxus as part of the spinal cord in the current context. A possible explanation would be that the hindbrain, being a vertebrate novelty, was built upon a spinal cord-like region.

2.3.2 Difficulty in comparing molecular pattern between amphioxus and vertebrates

In general, gnathostome progenitor neuron markers are conserved in lamprey, based on the spatio-temporal expression pattern of *Olig, Hb9/Mnx, Nkx2.2, Pax6, SoxB1a, Pax7,* and *Zic* (References in Table 2.1). Comparing spatio-temporal expression between vertebrates and amphioxus is more difficult because of their differences in

neuroanatomy and developmental event sequence. There are overlapping progenitor markers and differentiated markers in the same neurons e.g., referring to double in situ hybridisation amphioxus data from Ren et al. (2020), *OligA, Mnx,* and *Vacht* are expressed together in the same cells in 7-somite to 14-somite stage while in gnathostomes *Olig2* is in ventricular progenitor cells at an early stage (Fu et al., 2002) and *Mnx* and *Vacht* are in differentiated neurons in the mantle zone in later stages (Bello et al., 2012, Clovis et al., 2016). A potential explanation could be that through the separation of progenitor and differentiated cell populations in vertebrates, the gene regulatory network for neuron differentiation, which in the chordate ancestor had been confined in a neuron within the same developmental timeframe, had split across developmental stages as the cells migrate laterally. It is also possible that the chordate ancestor and vertebrate ancestor had different gene regulatory networks which happen to share a few of the same transcription factors, as suggested by Ren et al. (2020) on noting that *Err* was not an amphioxus motor neuron marker while *Err3* is marking motor neurons in mice (Friese et al., 2009). Regardless, I have attempted to compare their expression, function, and evolution in detail below.

2.3.3 Motor neuron development in chordates

In amphioxus, all three paralogues of *Olig* are expressed in a similar pattern in motor neurons in the neural tube, with *OligA* and *OligC* having posterior neural tube expression while *Olig B* does not (black arrowheads pointing at expression Figure 2.3 C, F, G; (Ren et al., 2020)). Double in situ hybridisation would be needed to prove all three amphioxus *Olig* paralogues are expressing in the same motor neurons. In chick and mice, *Olig3* and *Olig2* expression do not overlap. *Olig3* is downstream of *Nkx2.2* and involved in V3 (postmitotic) interneuron differentiation, but not in oligodendrocytes since it remains in grey matter (Liu et al., 2014). *Olig3* is also required for *Islet* expression in dI3/4 (Müller et al., 2005). *Olig1* is involved in oligodendrocyte maturation (Lu et al., 2002). *Olig2* antagonises further differentiation of progenitor motor neurons by repressing *Hb9/Mnx* (Lee et al., 2005); also *Nkx2.2* and *Olig2* co-repress *Ngn* for oligodendrocyte differentiation (Zhou et al., 2001). We do not know if *Olig* represses *Hb9/Mnx* (Figure 2.6) in amphioxus but we do know that *Mnx* is not required upstream of *Olig* because the five pairs of neurons and faint posterior spinal cord *OligA* expression remains in *Mnx* knockout embryos (Ren et al., 2020). Judging by their synteny, it is likely that the three amphioxus *Olig* genes arose by tandem duplication after amphioxus diverged from the olfactores and the one orthologue in the vertebrate ancestor has duplicated in 2R (possibly with gene loss, since *Olig1* and *Olig2* are closely linked) giving rise to

three paralogues with divergent functions. This is supported by my lamprey data showing *OligA* and *OligB* both being expressed in dorsal and ventral strips of cells in the spinal cord (arrows on Figure 2.1 and 2.2), which suggests vertebrate ancestral *Olig* was functioning in V3, dI1-3 interneurons, pMN motor neuron and ensheathing glia cell differentiation. In Ren et al. (2020) and figure 2.3, posterior neural tube expression of amphioxus *Olig* paralogues, which do not overlap with other motor neuron markers, could potentially fulfill the function of interneuron (and possibly glial) differentiation, though this has not been further examined. *Mnx* suppresses interneuron identity development in fish and chick (Seredick et al., 2012, Tanabe et al., 1998), while *Olig2* is downstream of *Nkx6.1* and *Nkx6.2* but upstream of *Mnx/Hb9(Mnr2), Lim3*, and *Islet1/2*(Vallstedt et al., 2001, Novitch et al., 2001). Ren et al. (2020) showed that *Mnx* is upstream of *Islet, Lhx3* and downstream of *Olig* and *Nkx6* but we have yet to find out or not whether *Nkx6* is upstream of *Olig* in amphioxus.

According to Ren et al., (2020), in amphioxus, *OligA, OligB,* and *OligC* are all downstream of Hh, the subset of motor neurons which co-express *OligA* and *Nkx6* are lost in the *Hh* knockout and *Nkx6* is not co-expressed in ectopic *OligA* neurons in the *Ptch* knockout. Therefore, within those subsets of motor neurons, amphioxus *Nkx6* is upstream of *OligA* because if it is the other way around *Nkx6* should also be present in those ectopic *OligA* neurons. An *Olig* triple knockout and an *Nkx6* knockout would be able to give a more definitive answer because a potential caveat of such deduction would be that *OligA* is required but not sufficient to induce *Nkx6* in those subsets of motor neurons. Also, older stages of *Olig* gene characterization would be helpful to see if it matches *Nkx6* in early larval stages (black arrowheads Figure 2.5 H). Given the limitations of lamprey experimental tools, I could only show that *Nkx6A* (Figure 2.4) is expressed within the spatio-temporal vicinity of *OligA/OligB,* so that they can potentially interact with each other in the motor neuron differentiation gene regulatory network. Lamprey *Mnx/Hb9* does seem to remain in the mantle zone at the motor neuron level of the spinal cord (Yuan et al., 2018). A broader temporal range of lamprey *Mnx/Hb9* expression characterization would show whether, unlike gnathostomes, *Mnx/Hb9* and *Islet* occupy the same space in the mantle zone due to the lack of lateral motor column (lamprey has no limbs nor fins). In chick, *Islet* and *Hb9* cells occupy separate spaces in the lateral motor column (Bello et al., 2012).

Whether *Olig* being downstream of *Nkx2.2* is conserved in lamprey and amphioxus would be difficult to tell without additional experiments. In lamprey, *Nkx2.2* and *OligA/OligB* (arrows on Figure 2.1 C', C", 2.2C, C', 2.7A, E') have similar expression spatio-temporally near the spinal cord floorplate. According to Albuixech-Crespo et al., (2017), in amphioxus, *Nkx2.2* is restricted to the cerebral vesicle but its tandem duplicate *Nkx2.1* is expressed in close enough vicinity spatio-temporally for it to be upstream of *OligA, OligB* and *OligC* (Figure 2.3). Amphioxus *Nkx2.2* could potentially mark the primordial hypothalamus (Albuixech-Crespo et al., 2017), because in vertebrates *Nkx2.1* is required for hypothalamus development (Orquera et al., 2019). My lamprey data (Figure 2.8) and Sugahara et al. (2016) also suggest that *Nkx2.1* is marking the hypothalamus in the vertebrate ancestor. Sugahara et al. (2016) identified lamprey hypothalamus from the co-expression of *Hh2* and *Nkx2.1*, plus it being rostral to the diencephalon.

In chick embryos, Islet1 and Lhx3 form a complex to bind to the enhancer of *Islet1* and *Lhx3*. This positive feedback loop promotes motor neuron maturation in the spinal cord and represses the formation of V2 neurons through LMO4 (Erb et al., 2017). The enhancer sequence of *Islet1* to which this complex binds is conserved in lamprey (Kim et al., 2015). In amphioxus, *Islet* and *Lhx3/4* are expressed in the same motor neurons at the same stages (Ren et al., 2020) but whether they form a protein complex and autoregulate is unknown, and the conserved vertebrate enhancer sequence of *Islet* is not found in amphioxus. I have searched for the enhancer sequence from (Kim et al., 2015) in the updated Marletaz et al. (2018) amphioxus genome and it remains undetected. There are no expression data for lamprey *Lhx3*. From in person communication with Guillaume Poncelet, I learnt that lamprey *Islet* expression is similar to *Olig,* which has dorsal and ventral strips of cells along the spinal cord. This makes it plausible that *Islet* is also involved in motor neuron differentiation in lamprey.

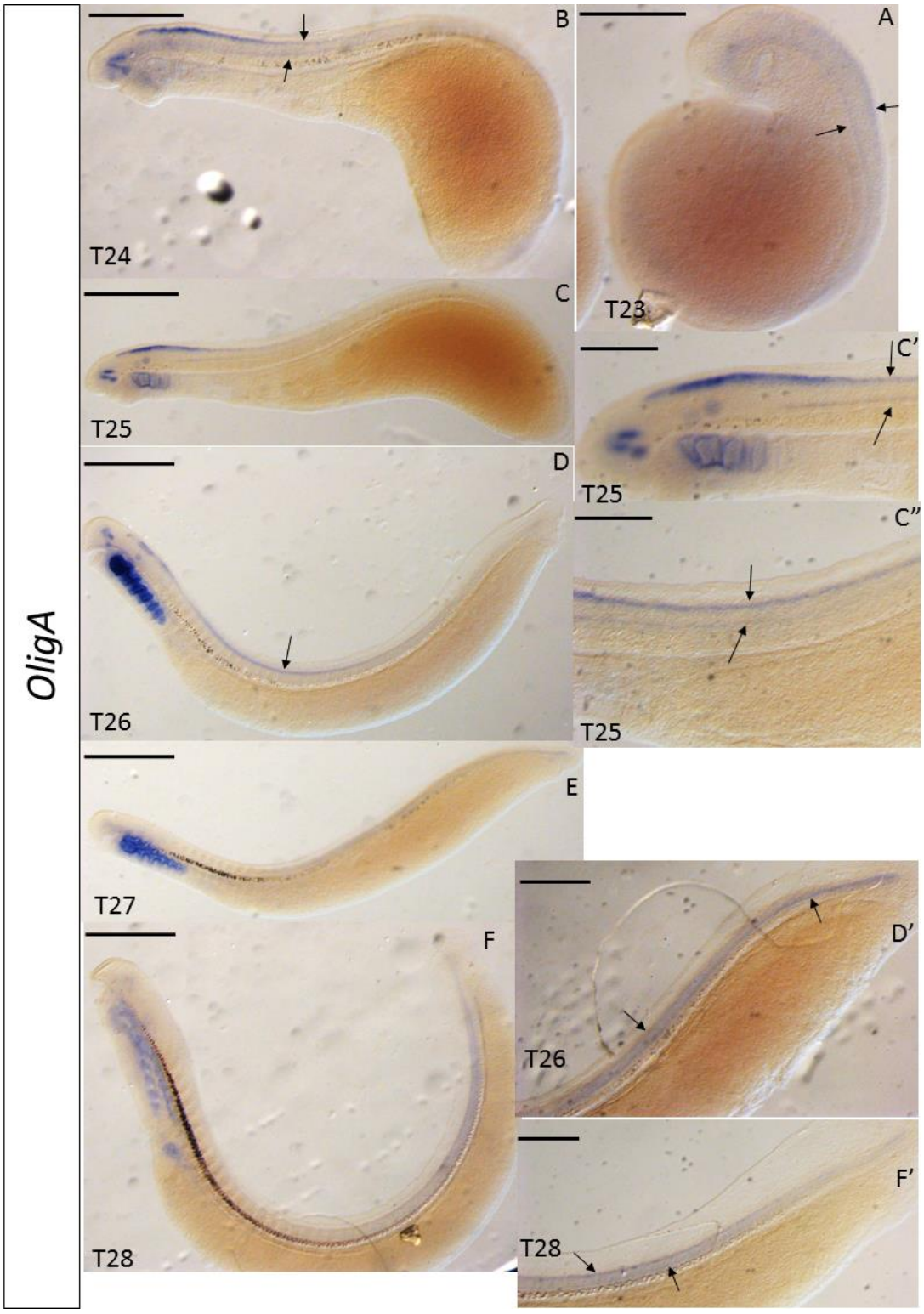


Figure 2.1. Lamprey *OligA* expression characterization. Black arrows point at expression in dorsal and ventral spinal cord cells from T23 to T28. In older embryos, *OligA* expression fades from anterior to posterior direction. Stages of embryos are at the bottom left corners. Panel C' and C" are the zoomed-in photo of T25 embryo (C). D' is the zoomed-in tail of T26 embryo (D). F' is the zoomed-in tail of T28 embryo (F). Scale bars: 500μm(A-F); 200μm (C', C", D', F').

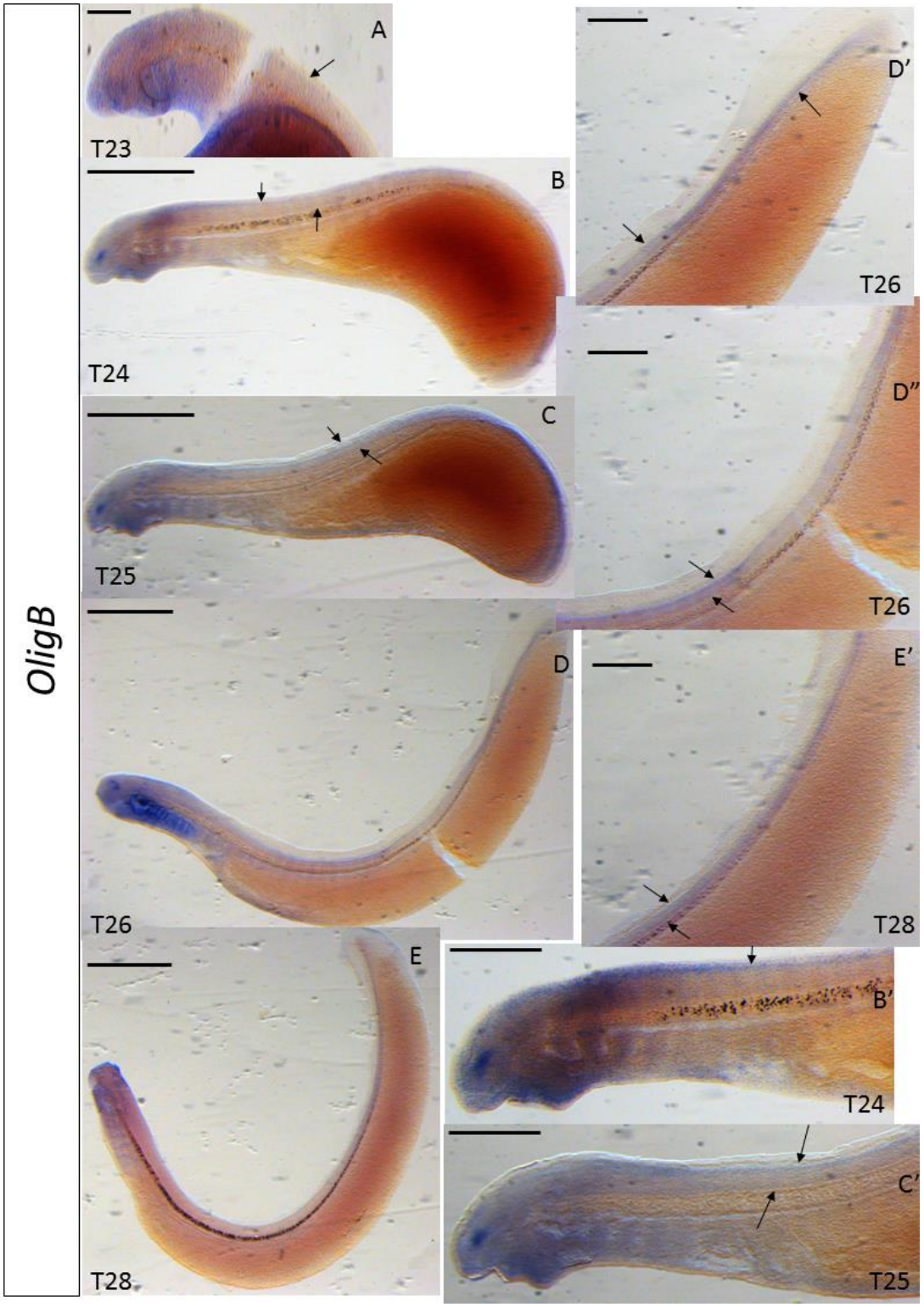


Figure 2.2. Lamprey *OligB* expression characterization. Black arrows point at expression in dorsal and ventral spinal cord cells from T23 to T28. In older embryos, its expression fades. Stages of embryos are at the bottom corners. B' and C' are the zoomed-in head of T24 (B) and T25 (C) embryos respectively. D' and E' are the zoomed-in tail of T26 (D) and T28 (E) embryos respectively. D" is the zoomed-in trunk of T26 (D) embryo. Scale bars: 100µm (A); 500µm (B-E); 200µm (B', C', D', D", E')

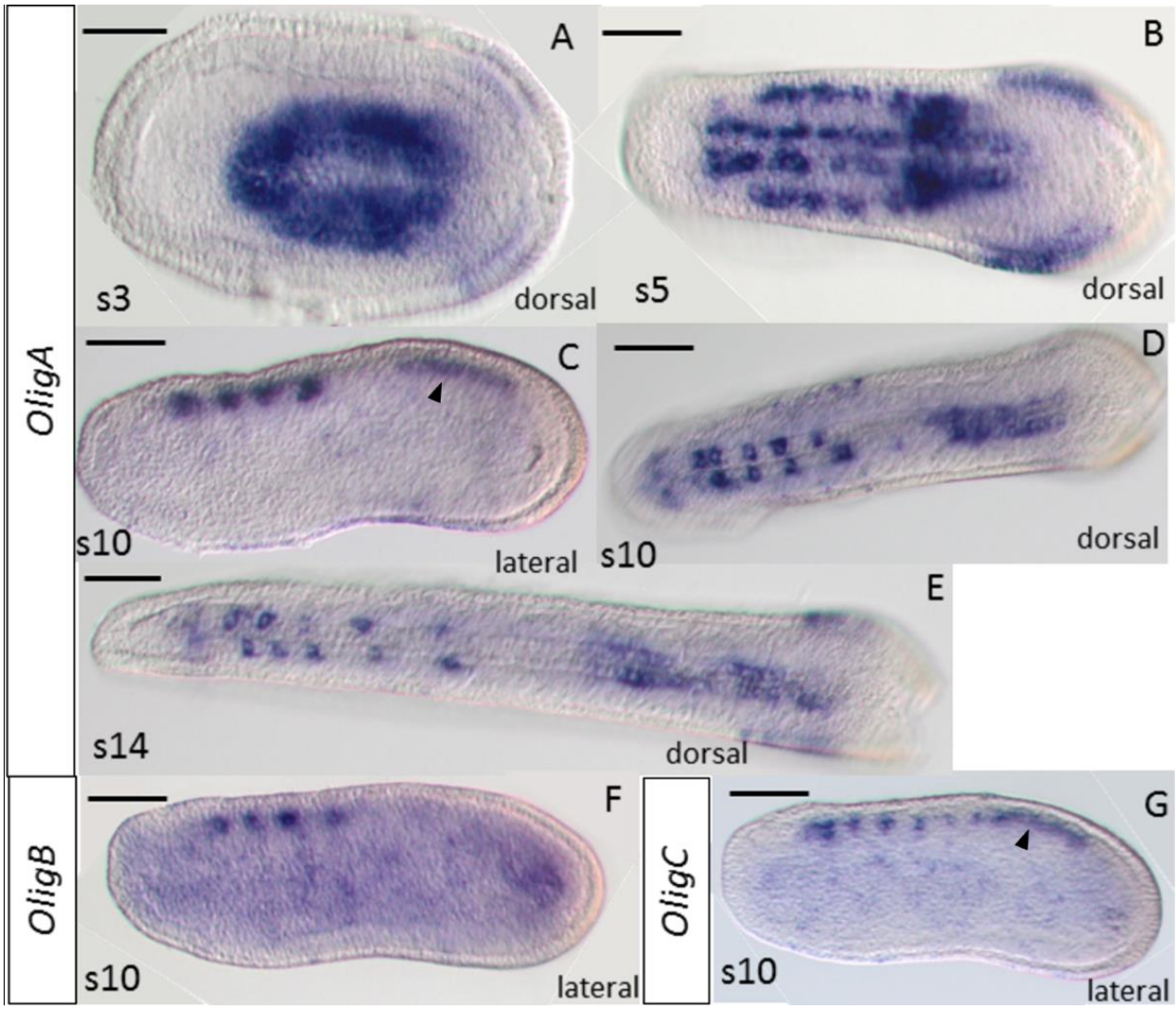


Figure 2.3. Amphioxus *OligA, OligB,* and *OligC* expression characterisation. At s10 stage, all three paralogues have similar expression in the spinal cord with the exception of *OligB* not having posterior spinal cord expression (A, F, G). Black arrowheads pointing at posterior spinal cord expression in C and G. Embryo stages are at bottom left corner while bottom right shows angle of perspective. All scale bars are 50µm. All photos from this figure were taken by Qiongqiong Ren and team (2020) at Guang Li's lab in Xiamen, China.

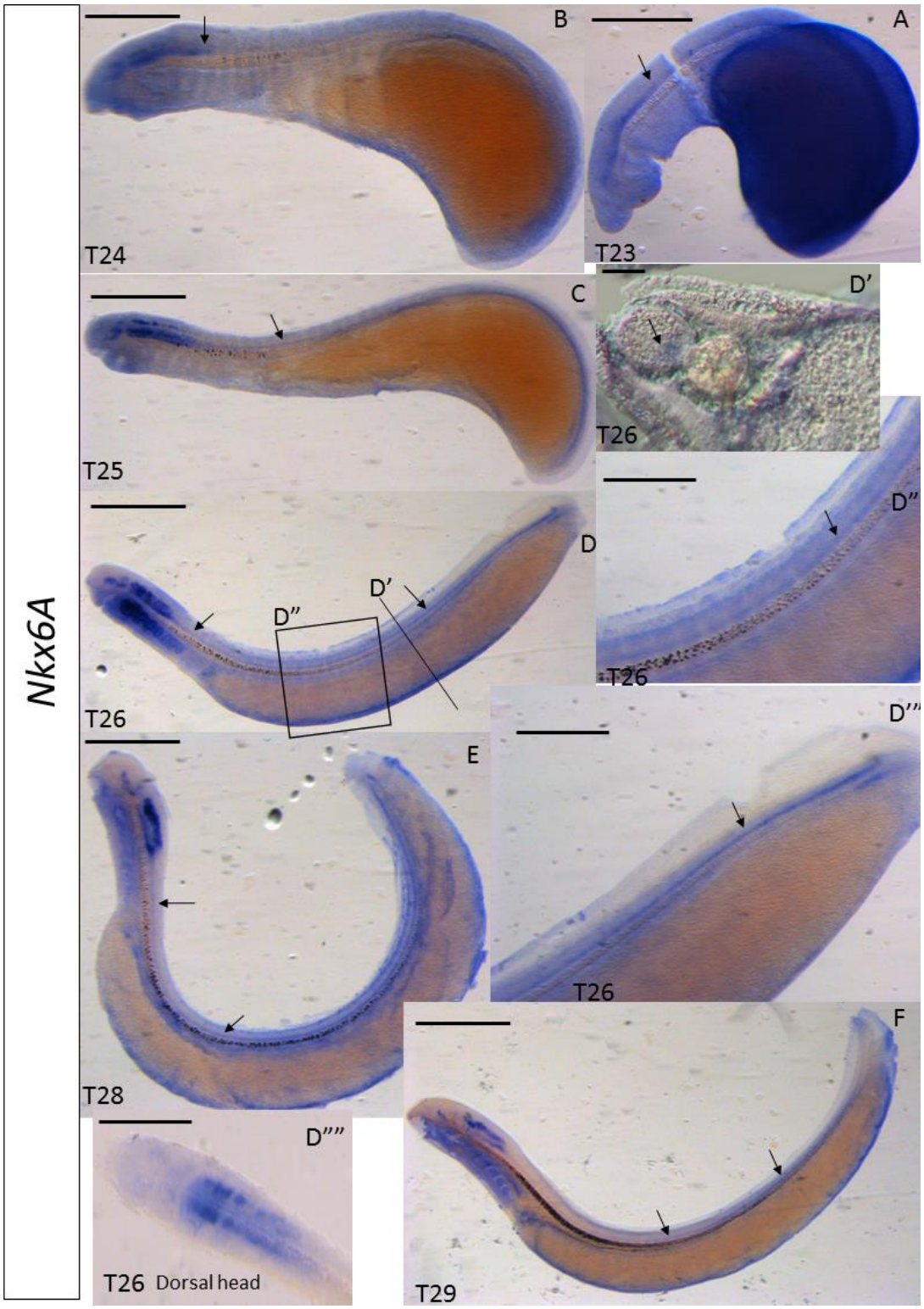


Figure 2.4. Lamprey *Nkx6A* expression characterization. Black arrows pointing to ventral spinal cord expression from T23 to T29. The transverse cross-section photo in D' is showing the trunk level of a T26 embryo. Stages of embryos are at bottom left corner. At the bottom left of the figure, there is a dorsal view of a T26 lamprey head to depict hindbrain expression from a different perspective. Scale bar: 500μm (A-F); 50μm (D'); 200μm(D", D'", D"").

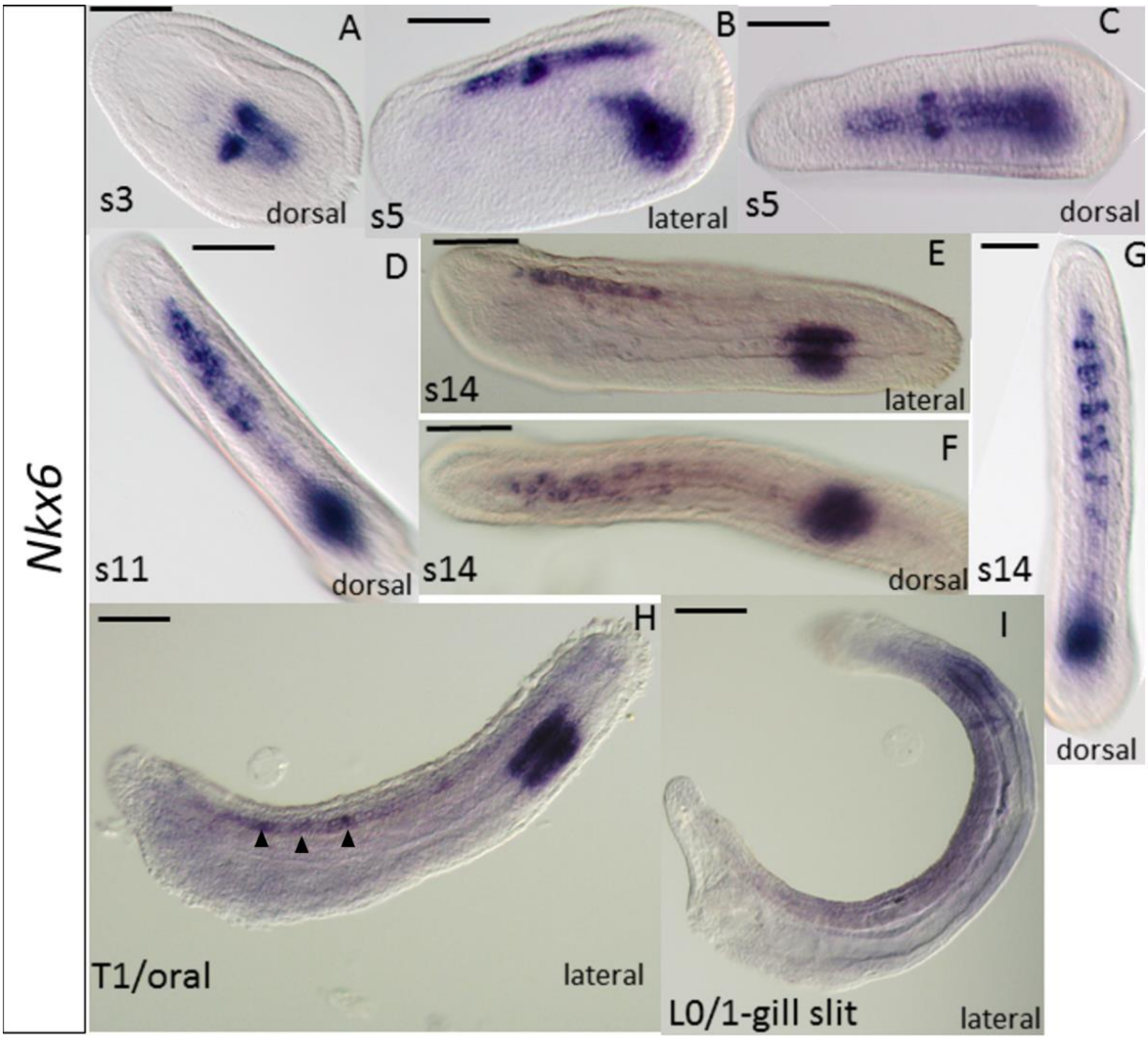


Figure 2.5. Amphioxus *Nkx6* expression characterization. In early stages, there is specific neural plate expression with more diffuse medial neural plate expression. As the embryos age, there are specific neurons with expression in the spinal cord which then fades as they reach the 1-gill slit stage while posterior endoderm expression remains. Black arrowheads pointing at neural tube expression that could potentially be compared to *Olig* expression at the same stage (H). Embryo stages are at bottom left corner while bottom right shows angle of perspective. Images E, F, H, and I are those taken by me, others were taken by Qiongqiong Ren and team. Scale bars on top left corners are 50µm.

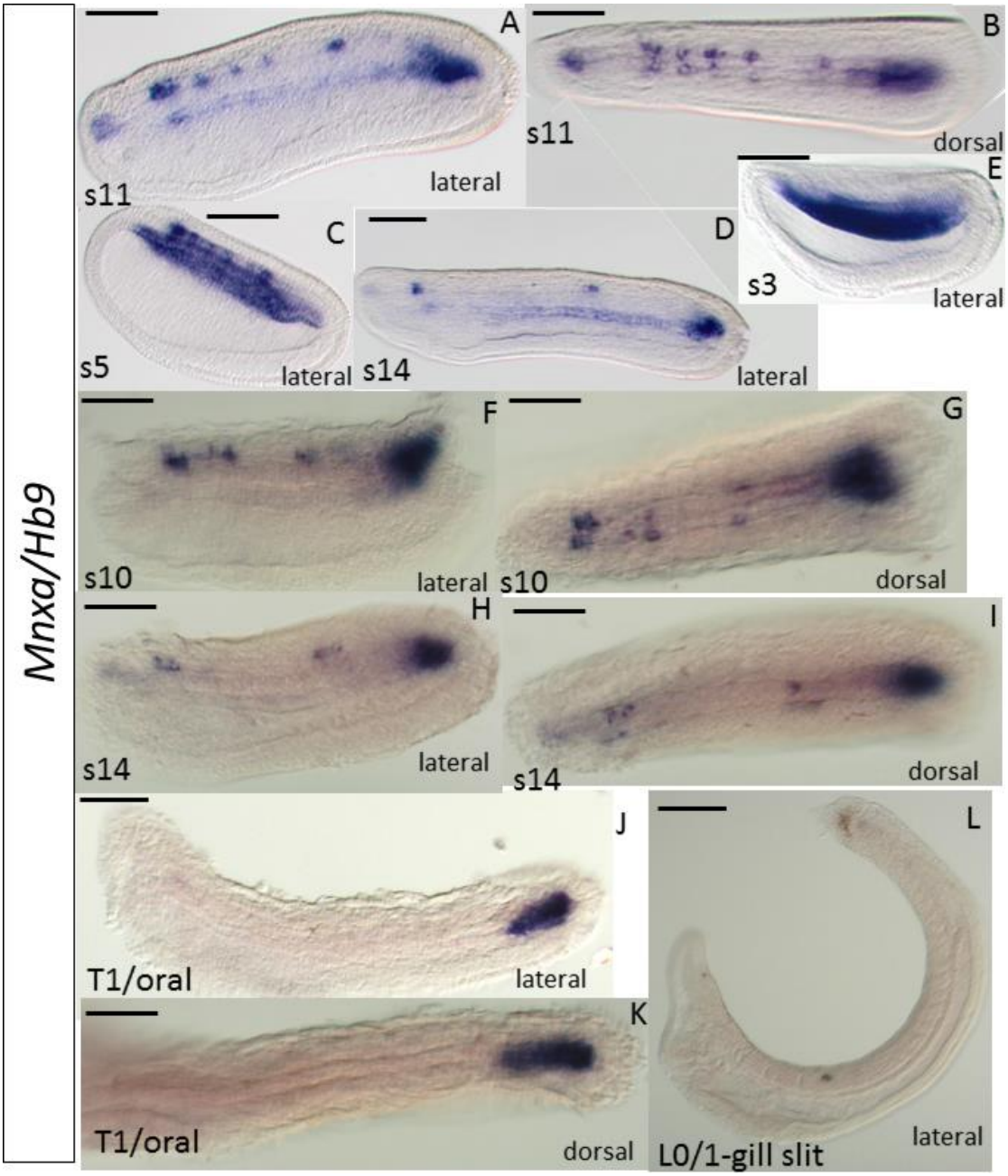


Figure 2.6. Amphioxus *Mnxa/Hb9* expression characterization. In early stages expression is mainly in the mesendoderm, then specific sites in the neural plate. As the embryos age, mesendoderm expression regresses into tailbud only, while the number of specific neurons expressing *Mnxa* in the spinal cord reduced and were eventually all gone by the oral stage. Tailbud expression remains at the 1-gill slit stage. Embryo stages are at bottom left corner while bottom right shows angle of perspective. Images F-L are taken by me, others were taken by Qiongqiong Ren and team. Scale bars on top left corners are 50μm.

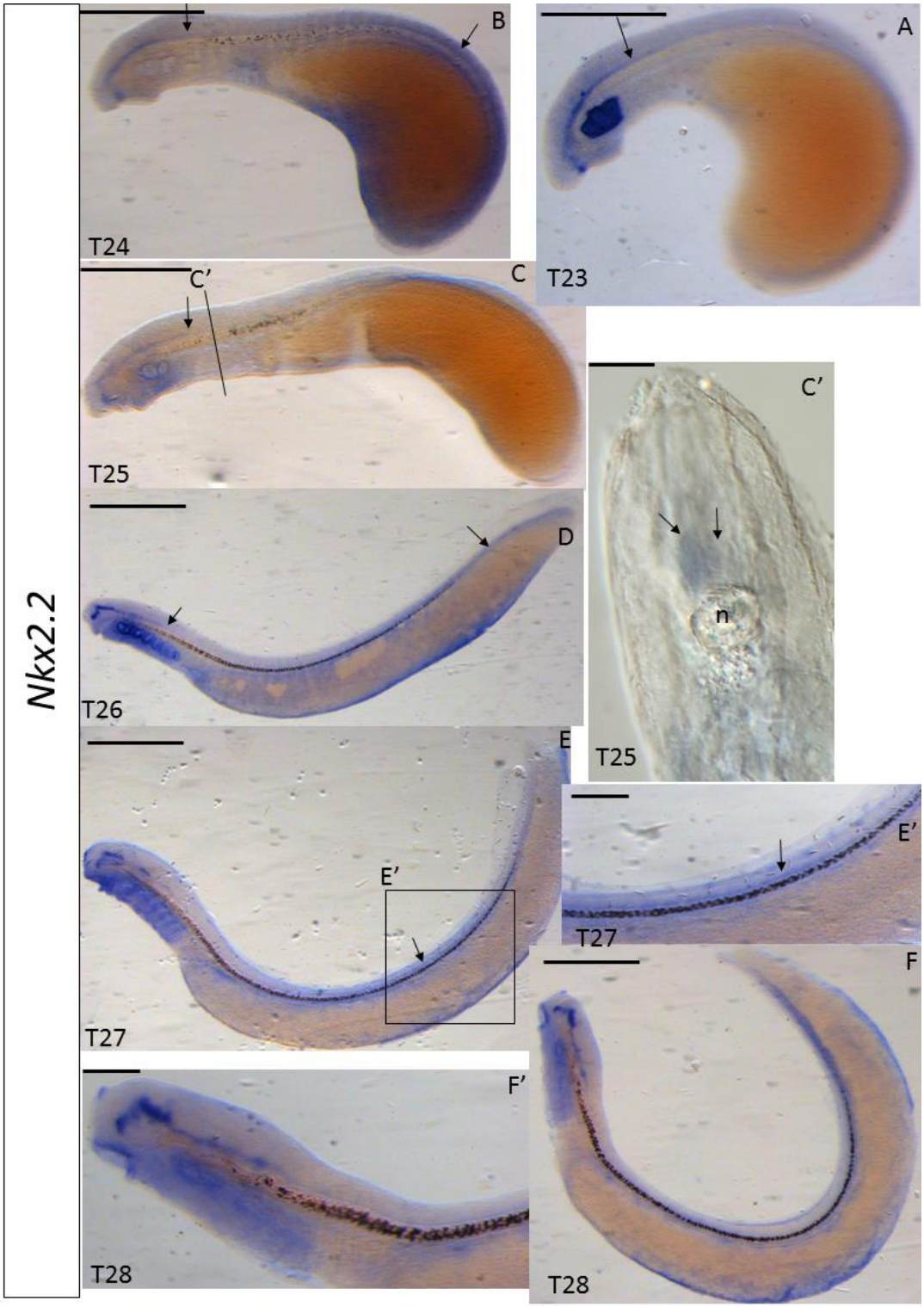


Figure 2.7. Lamprey *Nkx2.2* expression characterization. Black arrows highlighting ventral spinal cord expression from T23 to T27. The black line on the T25 embryos shows the level of transverse cross-section photo next to it. E' is the zoomed-in trunk of T27 embryo (E). F' is the zoomed-in head of T28 embryo (F). Stages of embryos are at the bottom left corners. Scale bars: 500µm (A-F); 200µm (E', F'); 50µm(C').

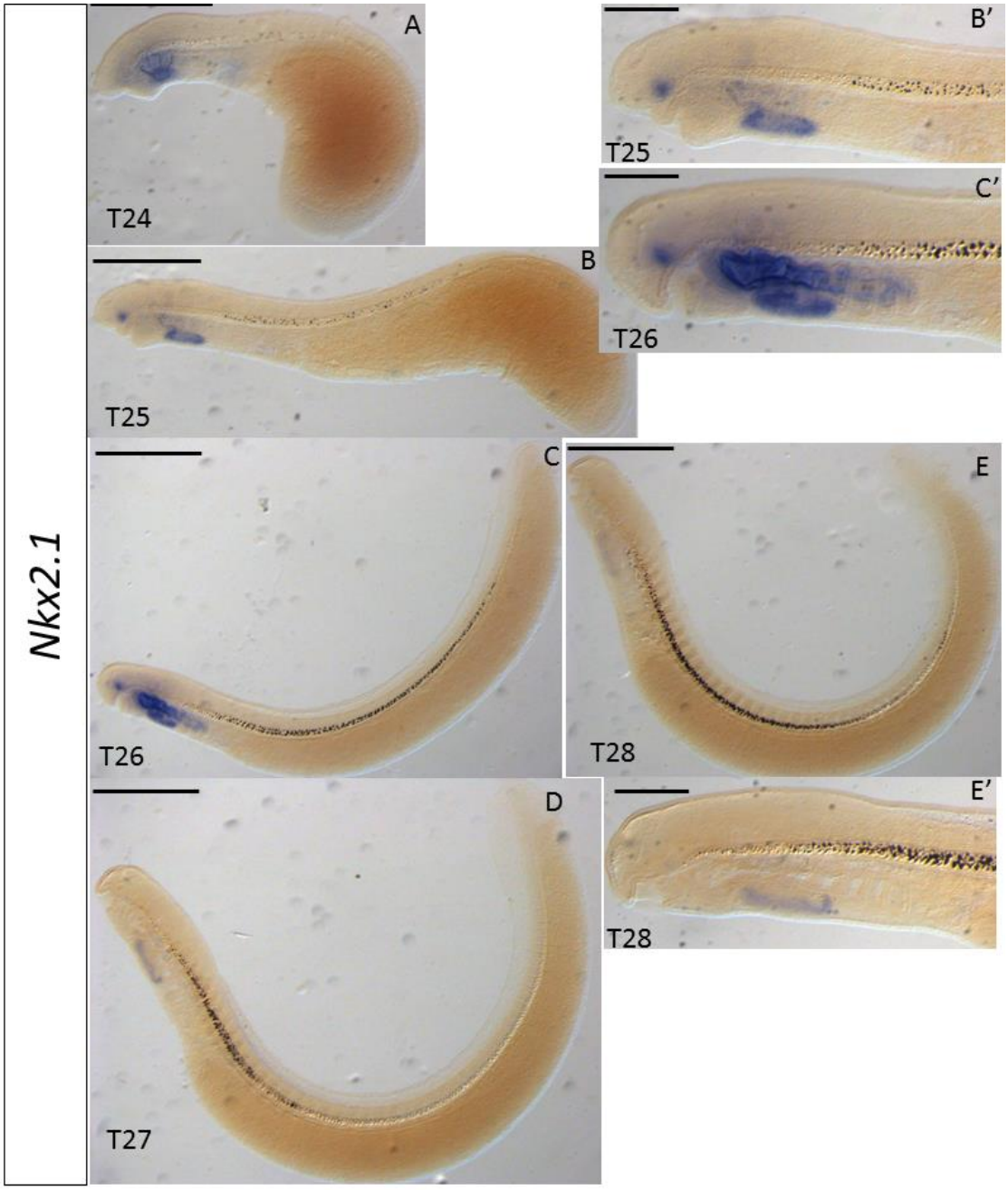


Figure 2.8. Lamprey *Nkx2.1* expression characterization. Endostyle expression persists through T24 to T28. Hypothalamus expression is only in T25 and T26. Stages of embryos are at the bottom left corners. B', C', and E' are the zoomed-in head of B, C, and E respectively. Scale bars: 500μm (A-E); 200μm (B', C', E').

2.3.4 Intermediate neural tube markers in chordates

Here I use the term intermediate to describe the central region, dl4-p2, of the gnathostome neural tube DV axis, although for some genes there might be fainter expression extending beyond that limit, e.g. for *Pax6* (Figure 2.11) and *Irx* (Houweling et al., 2001). Transcription factors in the intermediate regions are responsible for boundary maintenance by cross-repressing ventral and dorsal markers; other transcription factors are responsible for maintaining the progenitor state of the cells in the ventricular zone. *Pax6* is performing both functions (discussed more below). Thélie et al. (2015) argue that *Prdm12* specifies p1 interneurons by cross-repressing *Dbx1* and *Nkx6* in Xenopus. However, Yildiz et al. (2019) argue that, in zebrafish, *Prdm12* acts independently of *Nkx6* to maintain p1 interneurons. In mice, Gard et al. (2017) found that *Pax3* and *Pax7* repress *Dbx1* and *Prdm12* to set the p0 and p1 interneuron boundaries. In lamprey, *Dbx* seems to be expressed in the p0 area (Figure 2.9: arrows on D'-D"') while in amphioxus, *Dbx* is expressed in the posterior endoderm, tailbud, and maybe part of the mouth, but not in specific neurons (Figure 2.10). In the hemichordate, *Saccoglossus kowalevskii*, *dbx* is expressed in an ectodermal ring and the prosome throughout the ectoderm and in specific anterior scattered cells (Lowe et al., 2003). In the red flour beetle Tribolium, *dbx* is expressed in specific cells in the head epidermis and is responsible for bristle formation and is also expressed in a specific segmental pattern in the trunk epidermis (Posnien et al., 2011). There are also *dbx* expression data in *Novocrania anomala,* a brachiopod, but it is unclear what structure it is expressed by in the ectoderm, and *dbx* is expressed in the brain of *Lineus ruber,* a nemertean (Martín-Durán et al., 2018). In the deuterostome ancestor *dbx* was therefore probably expressed in ectoderm tissue. Amphioxus *dbx* expression might therefore have diverged from other deuterostomes. The only common property between lamprey and amphioxus *Dbx* expression is that both are in the sites of cell proliferation with the lamprey gene being in the ventricular zone of the neural tube and the amphioxus one in the tailbud. It is possible that the amphioxus *Dbx* tailbud expression could be involved in the AP patterning of the neural tube. The idea that as the tail elongates posteriorly, the neurons which are exposed to Dbx later specify into certain types of neurons is possible. But, currently, there is no evidence that that is the case. Also, amphioxus *Dbx* does not appear to have a cross-repressive property like in gnathostomes because both amphioxus *Dbx* and *Nkx6* are expressed at a similar location in the posterior endoderm (Figure 2.10, 2.5). This thesis focuses on DV patterning instead of AP patterning of the spinal cord so I shall not delve any deeper into speculation on how the tailbud could function.

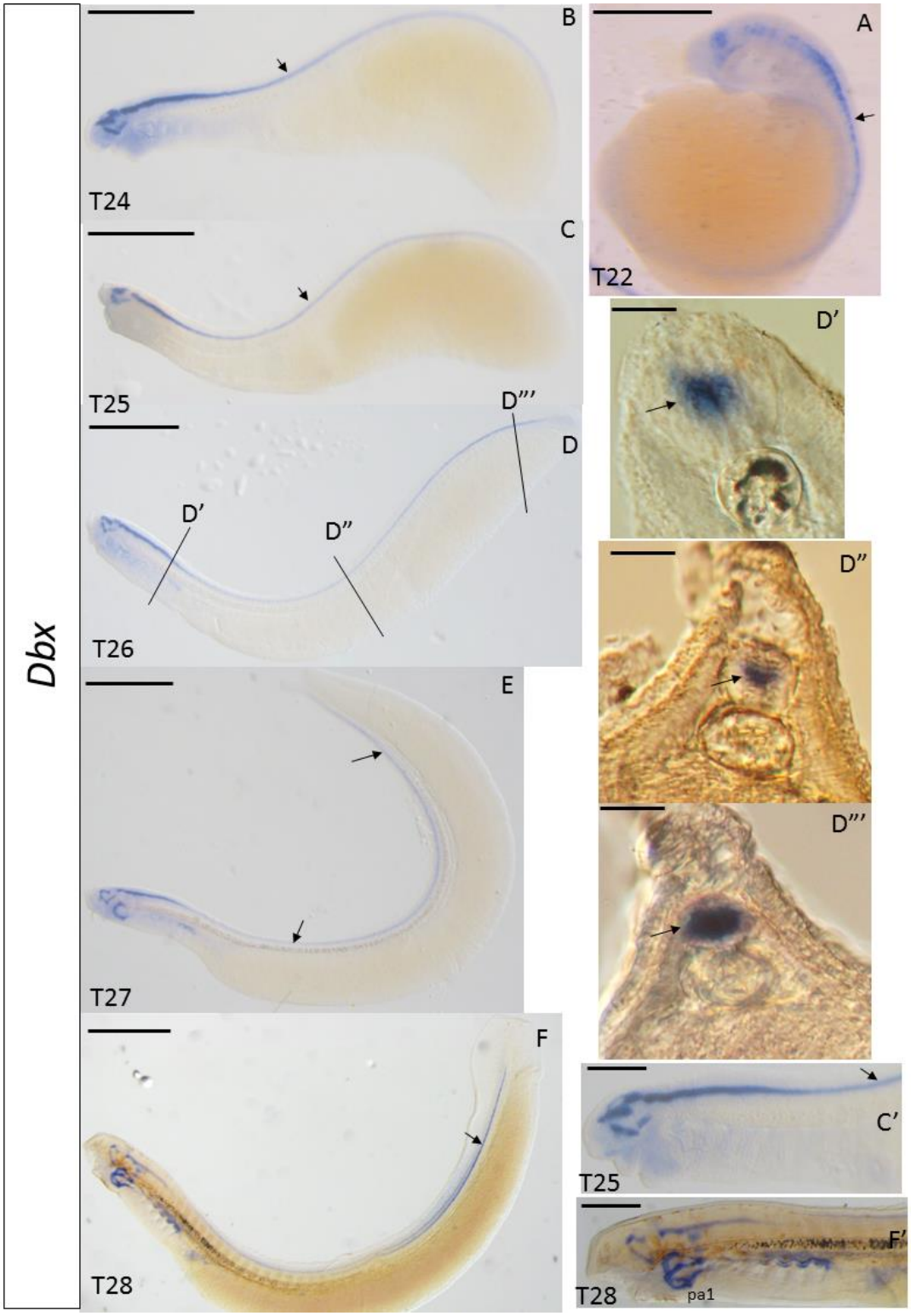


Figure 2.9. Lamprey *Dbx* expression characterization. Black arrows pointing at intermediate (around dl5-p1) spinal cord expression from T22 to T28. Transverse cross-section photos are of T26 embryo (D) at different level as indicated in D', D", and D'". C' and F' are the zoomed-in head of T25 (C) and T28 (F) embryos respectively. Stages of embryos are at the bottom left corners. Pa1: 1st pharyngeal arch. Scale bars: 500μm (A-E); 50μm (D', D", D'"); 200μm (C', F').

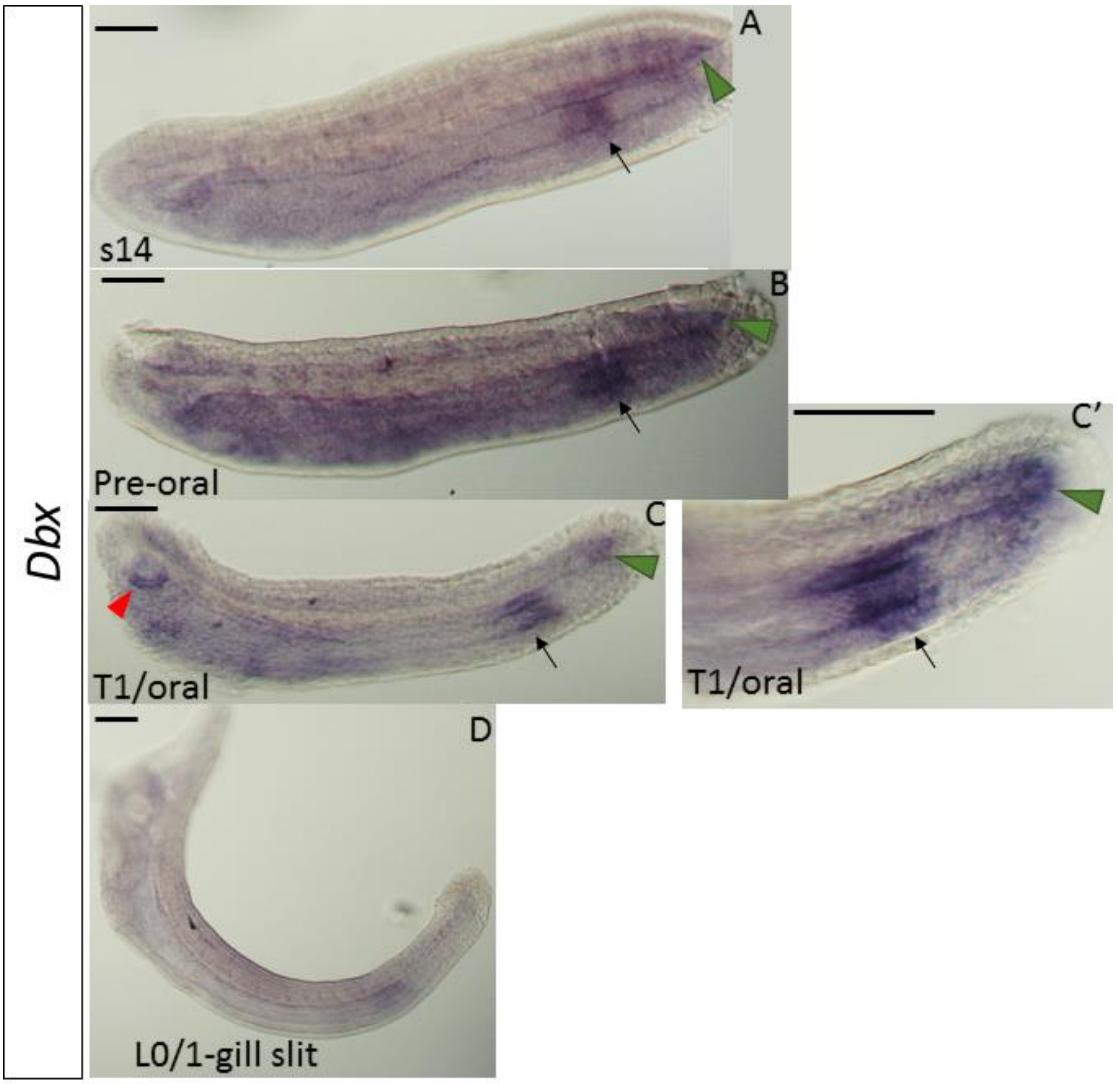


Figure 2.10. Amphioxus *Dbx* expression characterization. Posterior endoderm expression lasts from s10 to 1-gill slit stage (black arrows). There is tail-bud expression (green arrowheads) from s14 to oral stage with a zoomed-in photo of the tail (C') at oral stage. Red arrowhead showing partial mouth expression. All photos are showing lateral view. Stages of embryos are marked in bottom left corners. Scale bars are 50μm.

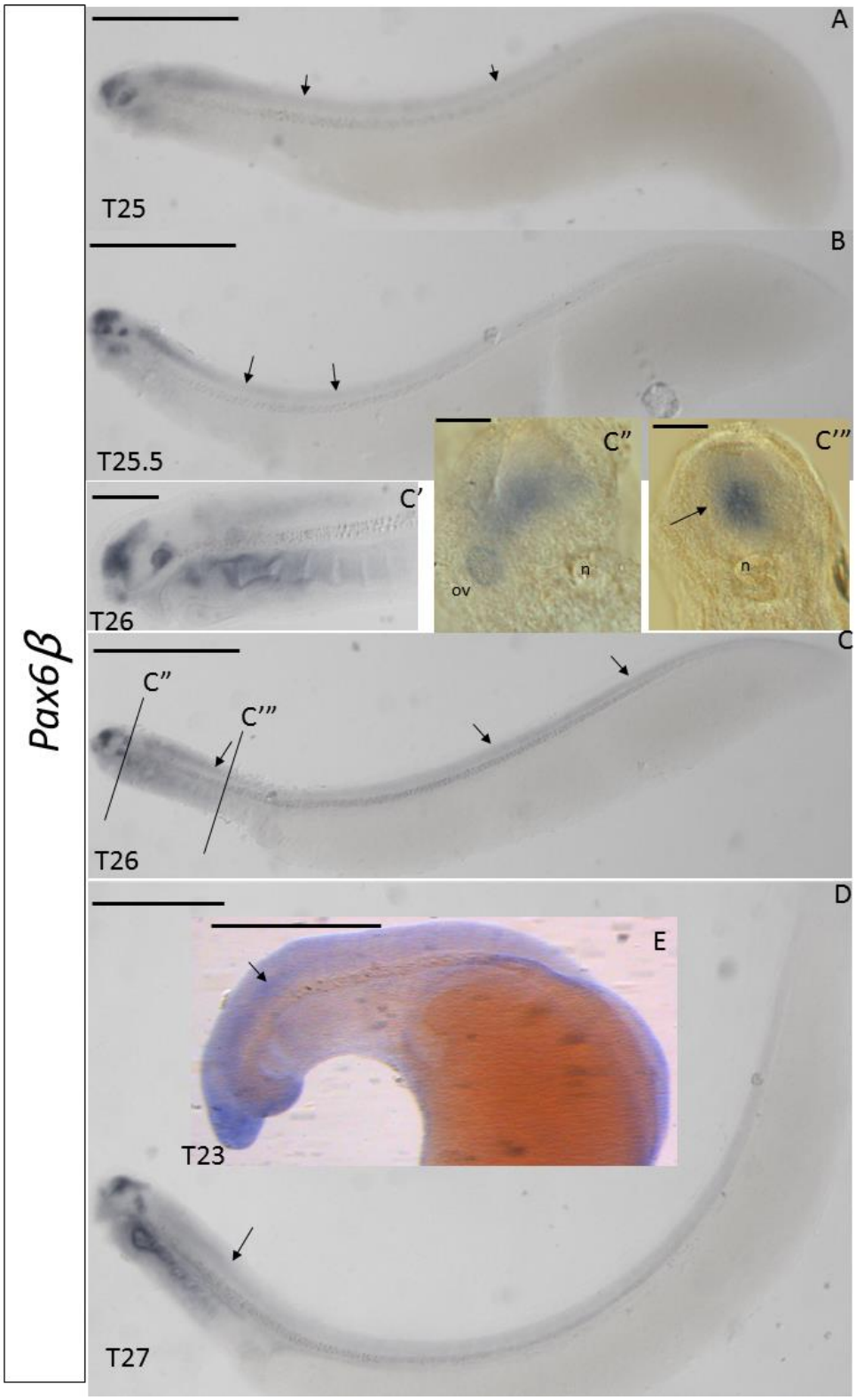


Figure 2.11. Lamprey *Pax6β* expression characterization. Black arrows point at spinal cord expression. Transverse cross-section photos (C", C'") is at the level shown in C. C' is the zoomed-in head of T26 (C) embryo. Stages of embryos are at the bottom left corners. Scale bars: 500μm (A-E); 200μm (C'); 50μm (C", C'"). ov: otic vesical; n: notochord.

Amphioxus *Pax4/6* remains expressed in the cerebral vesicle throughout development (Glardon et al., 1998). In addition to eye, nasohypophysial plate, and brain (Murakami et al., 2002), lamprey *Pax6*β is also expressed in the spinal cord around dI2-p2 (Figure 2.11, arrows). Ravi et al. (2019) found three paralogues of *Pax6* in lamprey, but it is yet to be seen where the other two paralogues are expressed. There is only one copy of *Pax6* in chick and human unless we consider *Pax4* to be a *Pax6* paralogue, however, irrespective of this *Pax4* is expressed in pancreas but not CNS in vertebrates (Djiotsa et al., 2012). In mice spinal stem cells, *Pax6* is induced by Notch1 to maintain proliferation by upregulating cyclin D1 (Wang et al., 2018). In chick spinal cord, *Pax6* is cross-repressing *Nkx2.2* to maintain the p3/pMN boundary (Briscoe et al., 2000). Combined, these data show *Pax6* neofunctionalised in the vertebrate ancestor, from only expressing in the cerebral vesicle as in amphioxus evolving to also expressing in the spinal cord in vertebrates, where it is functioning in progenitor cells for cell type boundary maintenance.

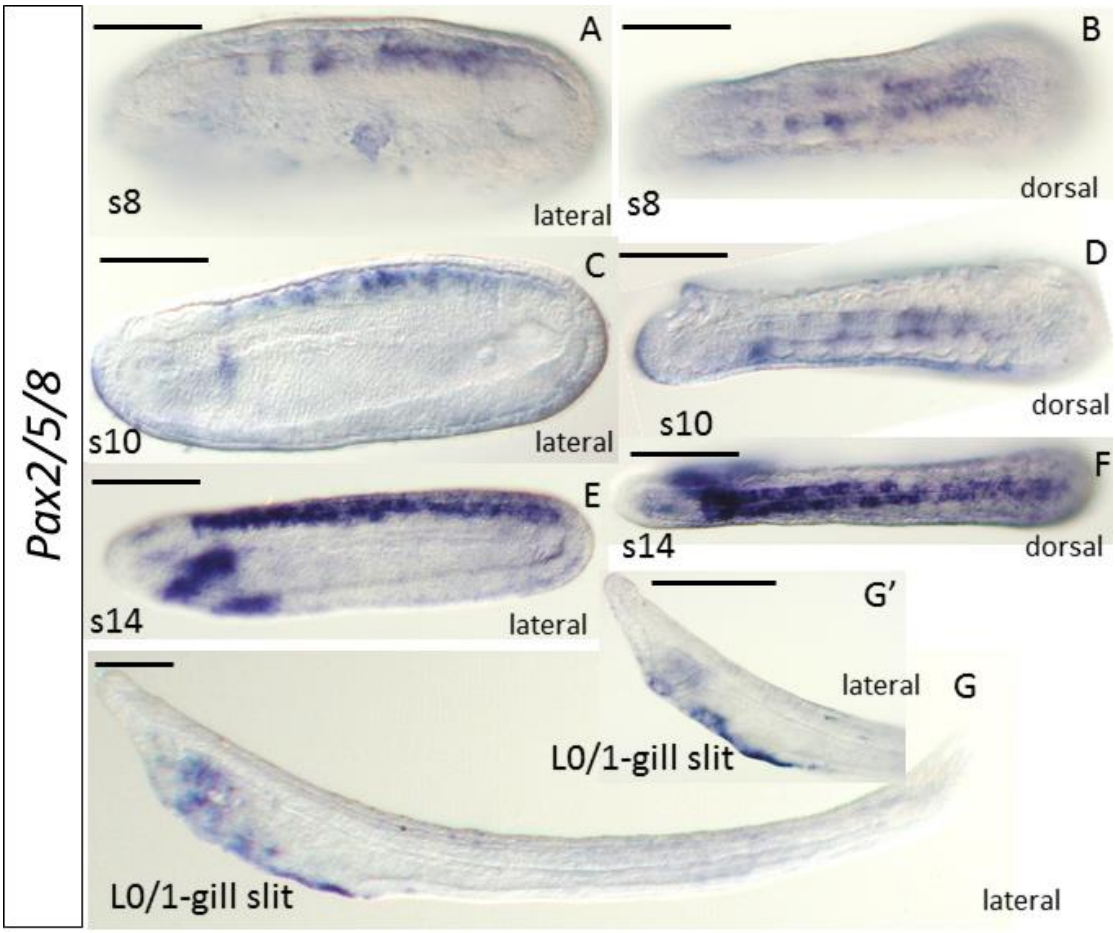


Figure 2.12. Amphioxus *Pax2/5/8* expression characterization. Specific cells in the spinal cord express *Pax2/5/8.* It is also expressed in pre-oral endoderm at s14 with expression reduced at the 1-gill slit stage. Embryo stages are at bottom left corner while bottom right shows angle of perspective. G' is the zoomed-in head of 1-gill slit embryo (G). Scale bars are 50μm.

Lamprey *Pax2/5/8* is expressed at the mid-hind brain boundary and broadly in hindbrain and spinal cord (McCauley and Bronner-Fraser, 2002, Murakami et al., 2001). In gnathostomes, *Pax2, Pax5* and *Pax8* are expressed in mid-hind brain boundary with only *Pax2* being expressed in around dl4 to v2 (in the mantle zone) and by GABAergic neurons in adult spinal cord (Pfeffer et al., 2000, Pfeffer et al., 2002, Burrill et al., 1997, Larsson, 2017, Ohyama et al., 2006). Amphioxus *Pax2/5/8* is expressed in spinal cord neurons in the dorsal and lateral neural tube but not floorplate (Figure 2.12; (Kozmik et al., 1999)). In addition to gene duplication, *Pax2, Pax5* and *Pax8* have multiple isoforms with different functions in chordates (Short et al., 2012, Fabian et al., 2015).

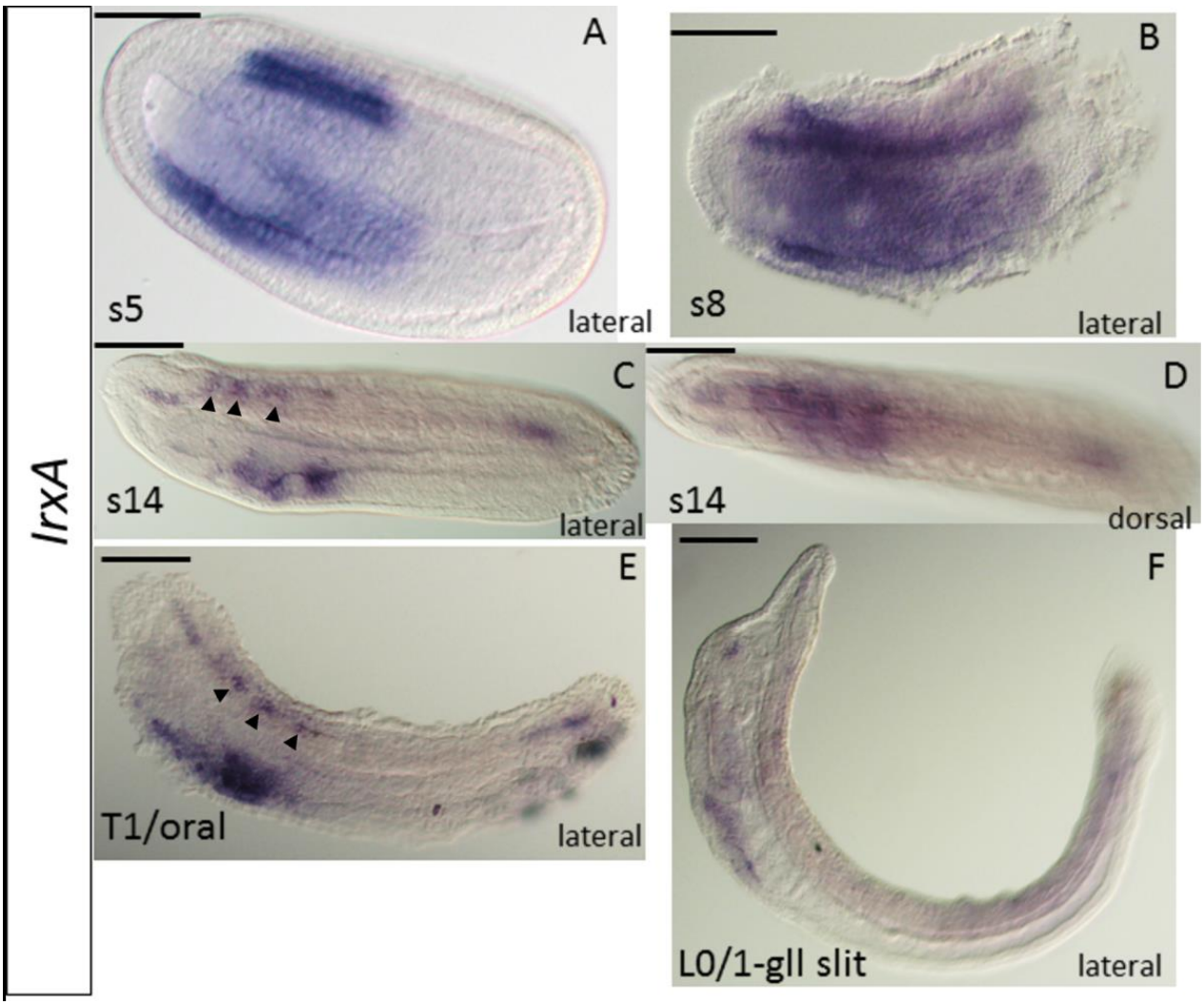


Figure 2.13. Amphioxus *IrxA* expression characterization. *IrxA* is expressed in the neural plate in early stages and restricted to specific neurons in spinal cord in later stages, then the number of expressing neurons reduced in older stage. Pre-gill slits endoderm also has expression in these stages. There is also tailbud and cerebral vesical expression in s14 and oral stages. Black arrowheads pointing at spinal cord expression. Embryo stages are at bottom left corner while bottom right shows angle of perspective. Images B-F are those taken by me, A was taken by Qiongqiong Ren and team. All scale bars are 50μm.

There are three paralogues of *Irx* in amphioxus and six in vertebrates. In mice, *Irx1* and *Irx2* are expressed in the floorplate, notochord, and around p0/v0. *Irx3* and *Irx5* are expressed in the ventricular zone and by motor neurons in the mantle zone. *Irx6* is only expressed in the lateral motor neuron. *Irx4* is not expressed in spinal cord at all (Houweling et al., 2001). In amphioxus, all three *Irx* genes in a cluster have similar expression in the neural tube at spinal cord level (black arrowheads in Figure 2.13; (Kaltenbach et al., 2009a)). In chick, *Irx3* represses motor neuron marker

Mnr2/Hb9/Mnx (Briscoe et al., 2000). In zebrafish, the *olig2* motor neuron progenitor marker represses the Irx gene *iro3.* In addition to ventral interneurons, *iro3* is also expressed in postmitotic motor neurons (Lewis et al., 2005). There are no data on other potential *Irx* functions in spinal cord. Analysis in lamprey would provide more information on how *Irx* paralogue functions diverged in the vertebrate ancestor.

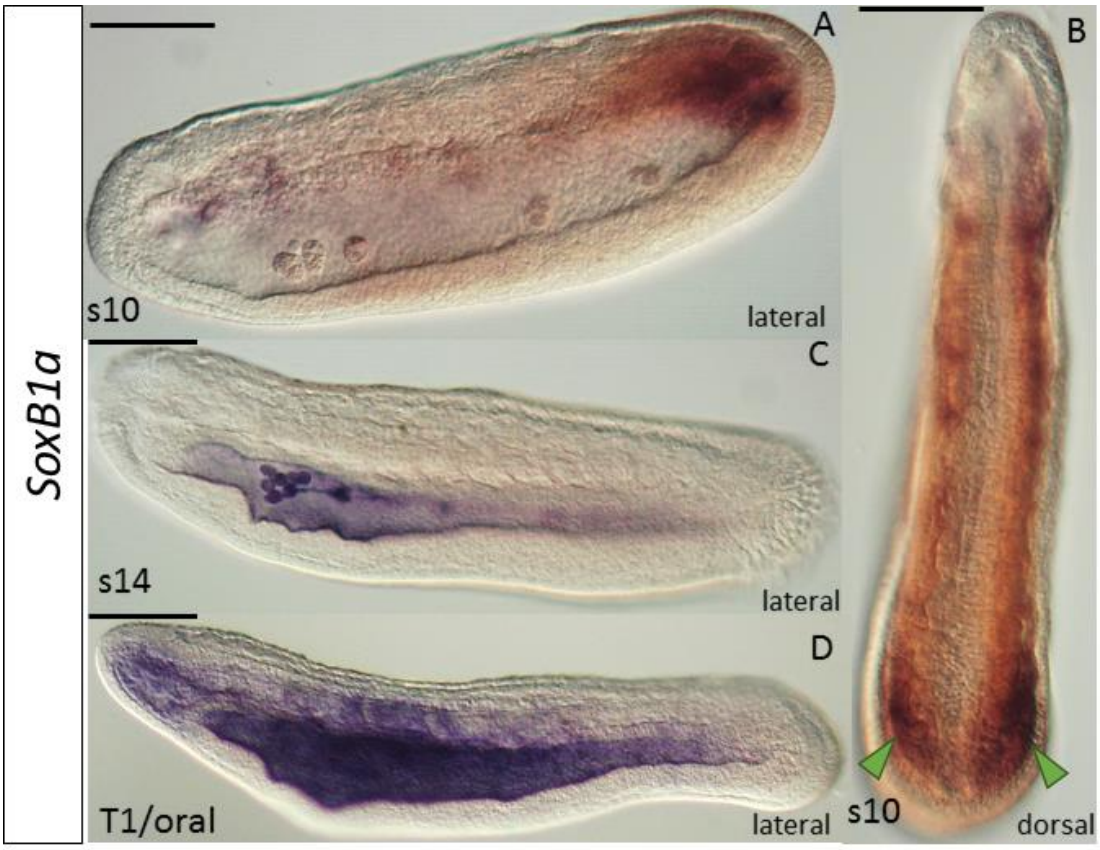


Figure 2.14. Amphioxus *SoxB1a* expression characterization. In s10 stage, it is expressed mostly in presomitic mesoderm. From s14 to oral stage, it is expressed in the endoderm. Embryo stages are at bottom left corner while bottom right shows angle of perspective. Green arrowheads are pointing at presomitic mesoderm expression. The endoderm staining could potentially be antibodies trapping artefact, even though *SoxB1c* is also expressed in endoderm (Meulemans and Bronner-Fraser, 2007). All scale bars are 50μm.

SoxB1 (Sox1/2/3) is known to maintain stem cell identity in the ventricular zone of the spinal cord in mouse (Tanaka et al., 2004). Amphioxus *SoxB1a*, previously known as *AmphiSox1/2/3*, is not expressed in the neural tube (Figure 2.14; (Holland et al., 2000)). *SoxB1b* is expressed in the neural plate but there are no data beyond early neurula. *SoxB1c* is expressed throughout the neural tube from the mid-neurula stage to fainter expression in the anterior neural tube in 1-gill slit stage. From the authors' phylogenetic analysis of the SoxB family, amphioxus *SoxB1c* has closer affinity to

vertebrate *Sox1, Sox2*, and *Sox3* genes (Meulemans and Bronner-Fraser, 2007, Cattell et al., 2012). This would suggest that vertebrate *Sox1, Sox2*, and *Sox3*, including lamprey *SoxB1a* and *SoxB1b*, had subfunctionalised from the chordate ancestor *SoxB1c.* Also, the aforementioned vertebrate orthologues are expressed throughout the AP axis of ventricular CNS which supports the idea that the vertebrate SoxB1 family had evolved from the chordate ancestor *SoxB1c* (York et al., 2019, Cattell et al., 2012, Tanaka et al., 2004).

2.3.5 Dorsal neural markers

Dorsal neural genes sometimes have dual roles. They are used in dorsoventral patterning of the neural tube and sometimes also in neural crest cell specification. Here I focus on their spinal cord patterning function. In chick and mice, *Msx1* is expressed in the roofplate and p2 interneurons. *Msx2* is in dI1 and *Msx3* in the dorsal half of the spinal cord (Misra et al., 2014). In lamprey, *MsxA* is only expressed in the upper and lower lips (black arrows in Figure 2.15 B', C'); as for *MsxB,* in addition to upper and lower lip, it is also expressed in branchial arches but not spinal cord (Cerny et al., 2010). Amphioxus *Msx* is expressed in the dorsal and lateral neural tube (Sharman et al., 1999). Judging from different species *Msx* expression, it is likely that neural tube expression is conserved between amphioxus and gnathostomes while the function of *Msx* in lamprey (Figure 2.15) has diverged.

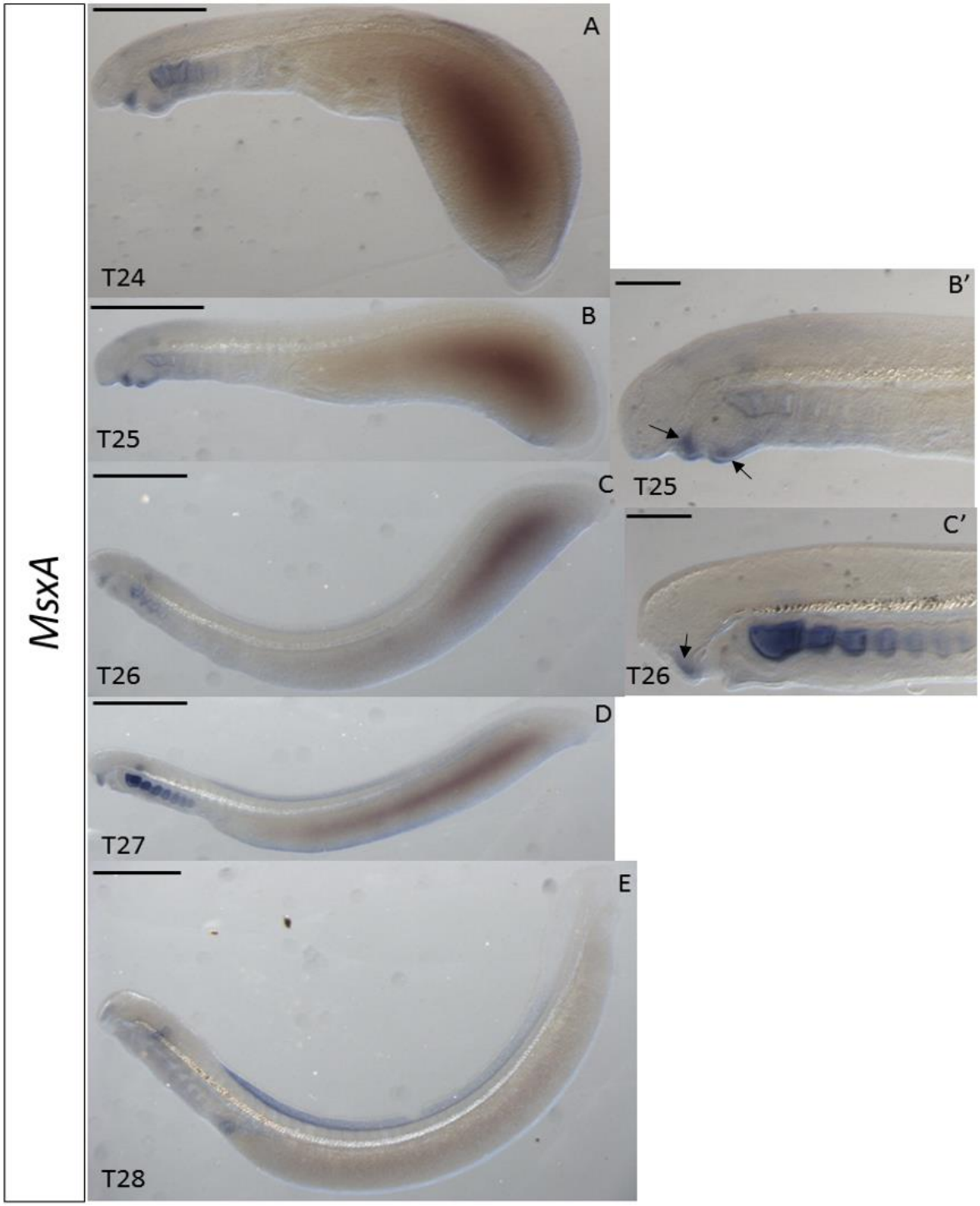


Figure 2.15. Lamprey *MsxA* expression characterization. *MsxA* is not expressed in spinal cord within my stages of interest. It is expressed in the upper and lower lip (black arrows in B' and C'). Stages of embryos are at the bottom left corners. The dorsal strip of blue which can be seen in T27 and T28 embryos is dorsal fin trapping of anti-dig antibody, not expression of transcript. Trapping is also seen in the pharyngeal pouches, especially in stage T26 (C') and T27 (D) embryos. B' is the zoomed-in head of T25 (B) embryo. Scale bars: 500µm (A-E); 200µm (B', C').

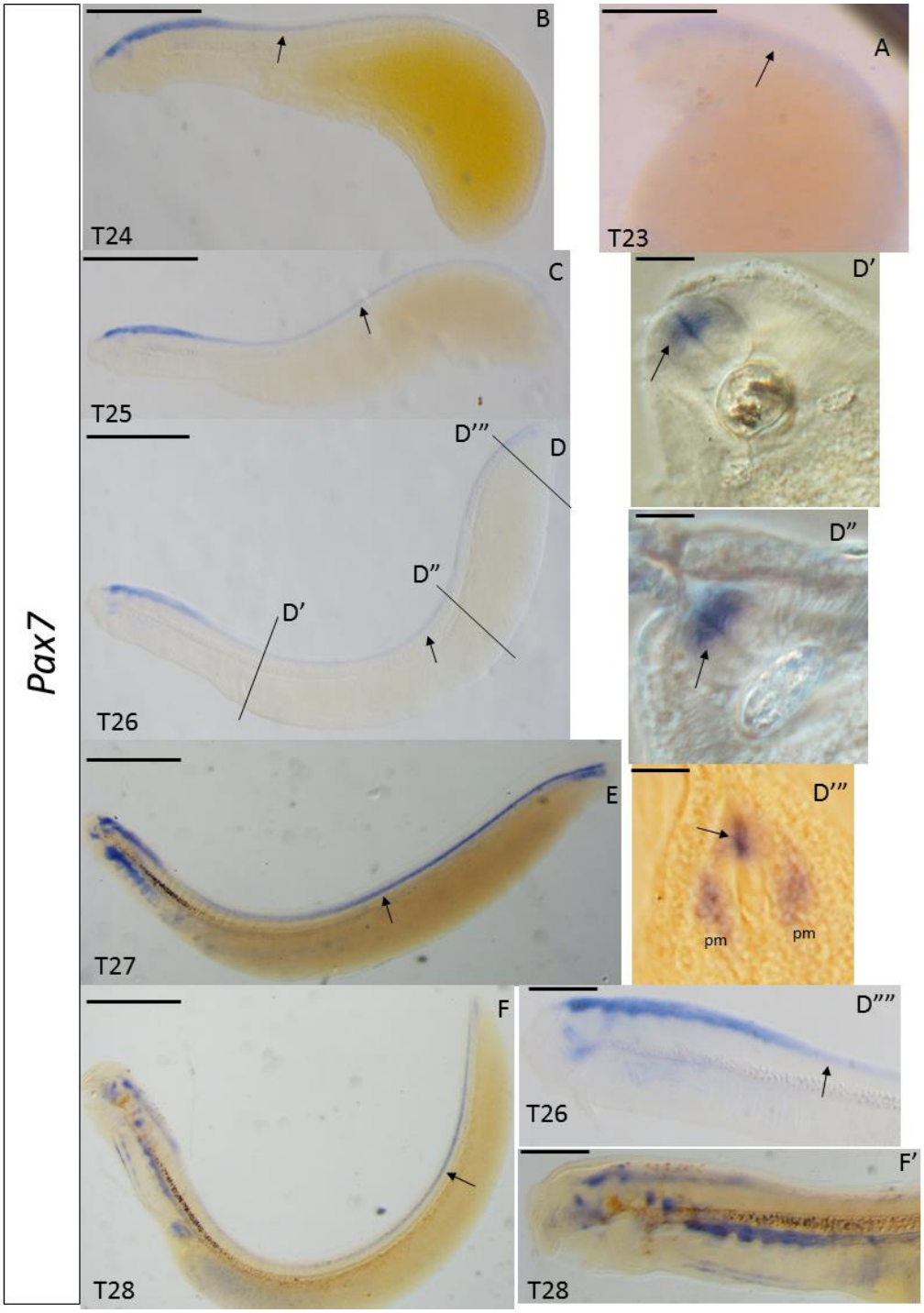


Figure 2.16. Lamprey *Pax7* expression characterization. Black arrows point to dorsal spinal cord expression. Level of transverse cross-section of T26 embryo are indicated in black lines. Stages of embryos are in the bottom left corners. D"" and F' are the zoomed-in head of T26 (D) and T28 (F) embryos. T23 embryo (A) has its head fell off but dorsal spinal cord expression can still be observed. Scale bars: 500µm (A-F); 50µm (D', D", D'"); 200µm (D"", F').

Barton-Owen et al. (2018) found two paralogues of *Pax3/7* in amphioxus. Gene expression characterisation in *Branchiostoma lanceolatum* seems to be missing the cerebral vesicle expression which was described in *B. floridae* as described by Holland et al. (1999) and confirmed by Ren et al. (2020). Regardless, *Pax3/7* is not present in amphioxus spinal cord. In chick, mouse, and lamprey, *Pax3/7* is expressed in the dorsal spinal cord (arrows on Figure 2.16; (Hoover et al., 2008, Andrews et al., 2017)). *Pax3/7* has hence probably evolved spinal cord expression in the vertebrate ancestor. *Pax3/7* is expressed at the dorsolateral domains of ectoderm in branchiopoda, *N. anomala*; there is a broad ubiquitous low level of expression in nemertean, *L. ruber*; it is expressed at the ventral ectoderm midline and also lateral larval expression domains in Annelid, *O. fusiformis* (Martín-Durán et al., 2018). These outgroup data do not clarify whether spinal expression is gained in vertebrates or lost in amphioxus.

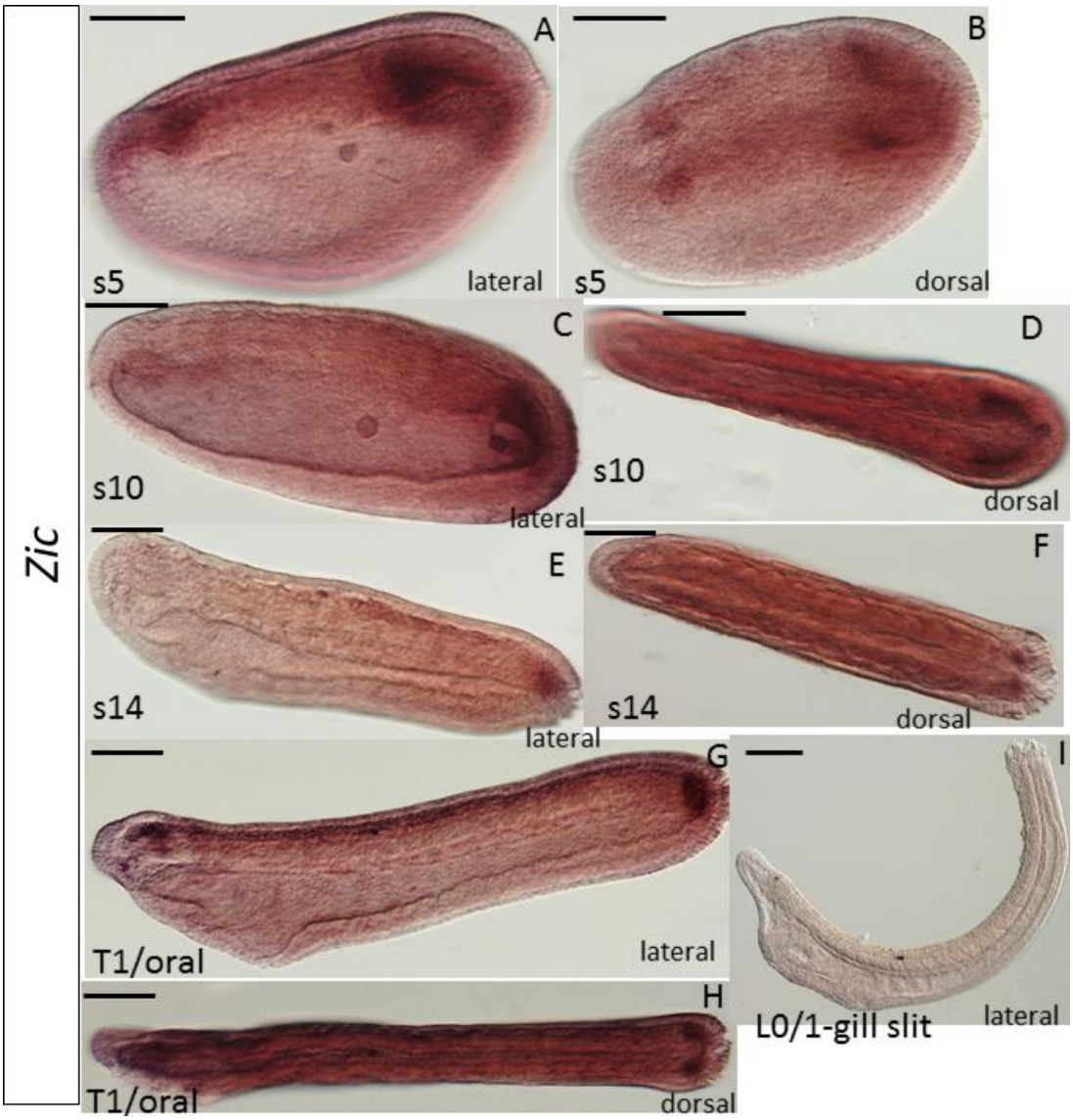


Figure 2.17. Amphioxus *Zic* expression characterization. *Zic* is expressed in presomitic mesoderm from s5 to oral stage and ceased to be expressed by the 1-gill slit stage. At s5 stage, the first pair of somites also express *Zic.* Embryo stages are at bottom left corner while bottom right shows angle of perspective. All scale bars are 50μm.

Amphioxus *Zic* expression (Figure 2.17) in T1/oral stage shows expression throughout the spinal cord. Gostling and Shimeld (2003) have cross-section amphioxus data showing that the expression is in the dorsal neural tube and dorsolateral somites which is similar to lamprey *ZicA/Zic1* and mouse *Zic1* and *Zic2* expression (Castillo et al., 2010, Kusakabe et al., 2011, Merzdorf, 2007). *Zic3* seems to have subfunctionalised/specialised into dorsal-lateral somites only and *Zic5* into roofplate only. *Zic4* is similar to *Zic1* and *Zic2* but has subfunctionalised to exclude the roofplate.

2.3.6 Signalling pathway gene expression

Patched is the downstream receptor of Hh. Both Ptch1 and Ptch2 autoregulate with ligand-dependent antagonism of Hh, and double knockout of the two *Ptch* genes led to ventralisation of the neural tube in mice (Holtz et al., 2013). *Ptch* can express more broadly and initiate earlier than the subset of cells receiving Hh signals, so at least in early development has some Hh-independent regulation in vertebrates and amphioxus (Goodrich et al., 1996, Ren et al., 2020). In both lamprey and mice, *Ptch* is expressed in floorplate (Figure 2.18), and faint ventral ventricular zone expression is also detected mice (Matsumoto et al., 2007). In zebrafish, *Ptch2* is expressed throughout the ventricular zone of spinal cord and sclerotome (Jacobs and Huang, 2019, Shen et al., 2013). There is evidence in gnathostomes showing that Notch primes neural tube cells to be responsive to Shh/Ptch signalling (Stasiulewicz et al., 2015, Jacobs and Huang, 2019). In *B. floridae, Ptch* is expressed in the notochord and posterior neural tube from mid-neurula to 1-gill slit stage (Ren et al., 2020). In *Branchiostoma belcheri, Ptch* is expressed mainly in the endoderm with posterior neural tube expression begining at the much later 2-gill slits stage (Lin et al., 2009). How *Ptch* is affected by different signalling pathways shall be explored more in chapter three.

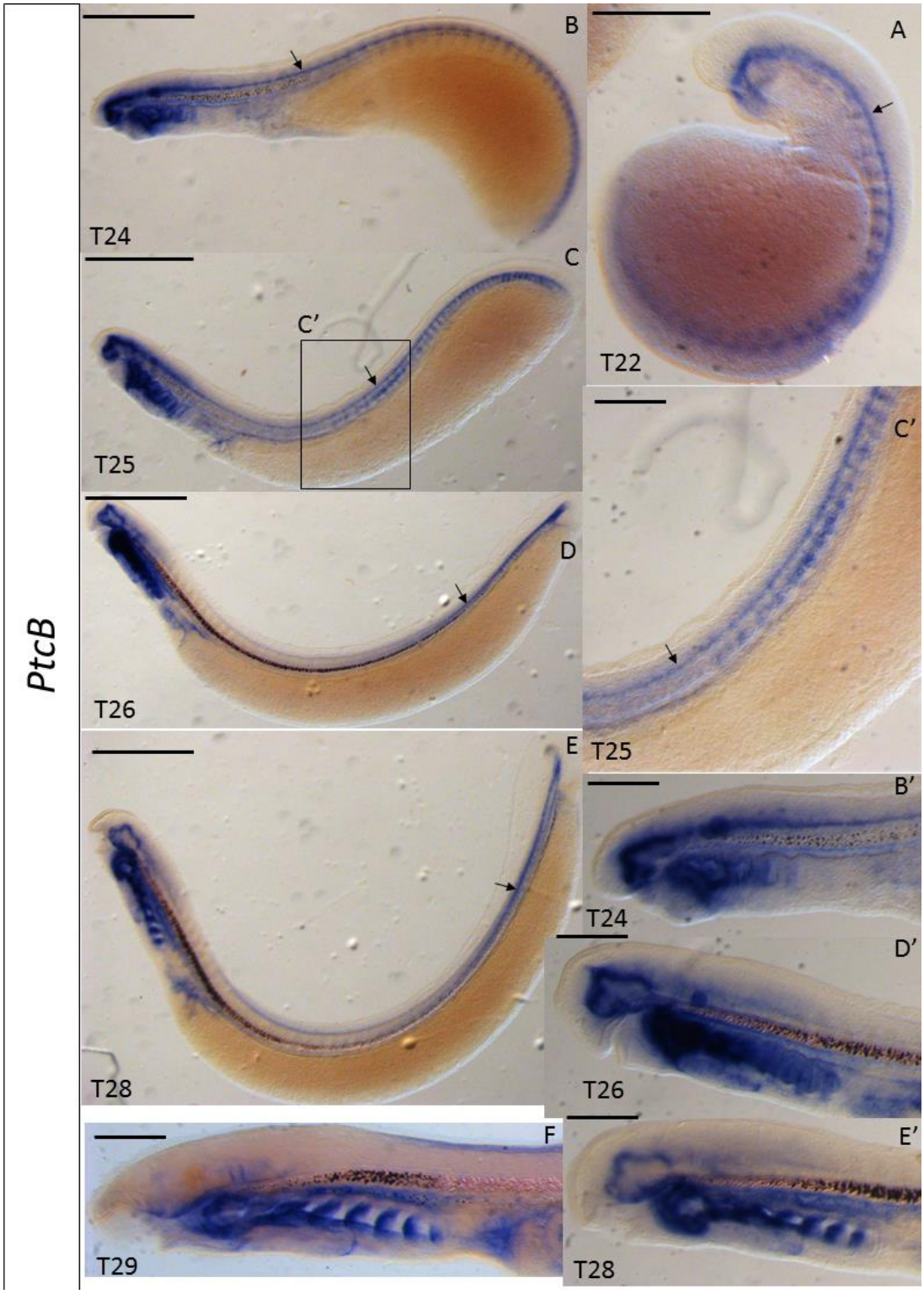


Figure 2.18. Lamprey *PtcB* expression characterization. Black arrows point at ventral spinal cord expression. It is also expressed in tail somites. Expression seems to fade from the hindbrain to spinal cord as the embryos age. Stages of embryos are at the bottom left corners. C' is the zoomed-in trunk of T25 (C) embryo. B', D', and E' are the zoomed-in head of T24 (B), T26 (D), and T28 (E) embryos respectively. Scale bars: 500μm (A-E); 200μm (C', B', D', E', F)

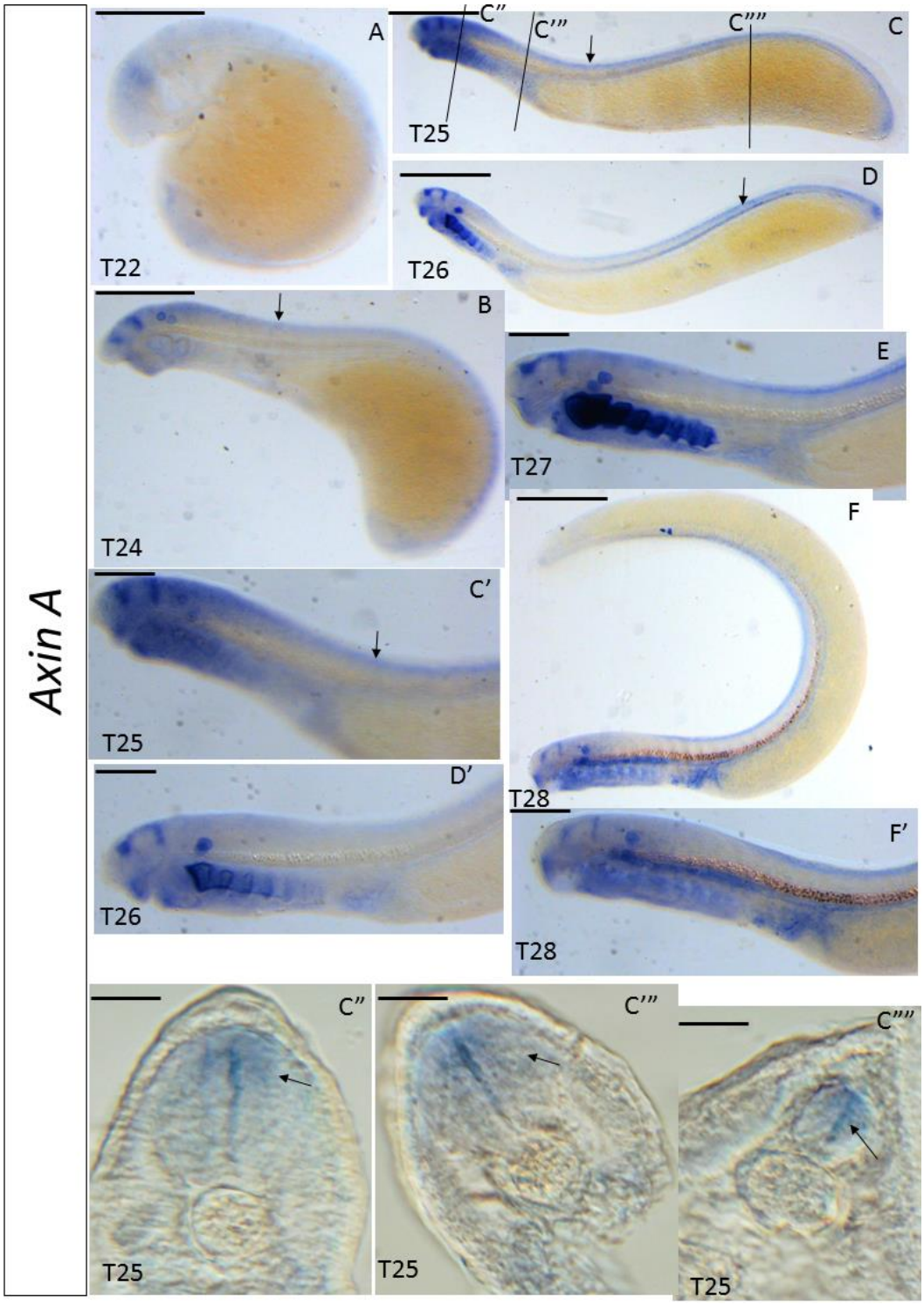


Figure 2.19. Lamprey *AxinA* expression characterization. Black arrows point at dorsal spinal cord expression. Transverse cross-section photos are showing T25 embryo expression at different level (C", C'", C""")marked by black lines on the top photo of T25 (C) lamprey. Expression seems to fade at embryos age to T28 (F, F'). Stages of embryos are at the bottom left corners. Scale bars: 500µm (A-D,F); 200µm (C', D', E, F'); 50µm (C", C'", C""").

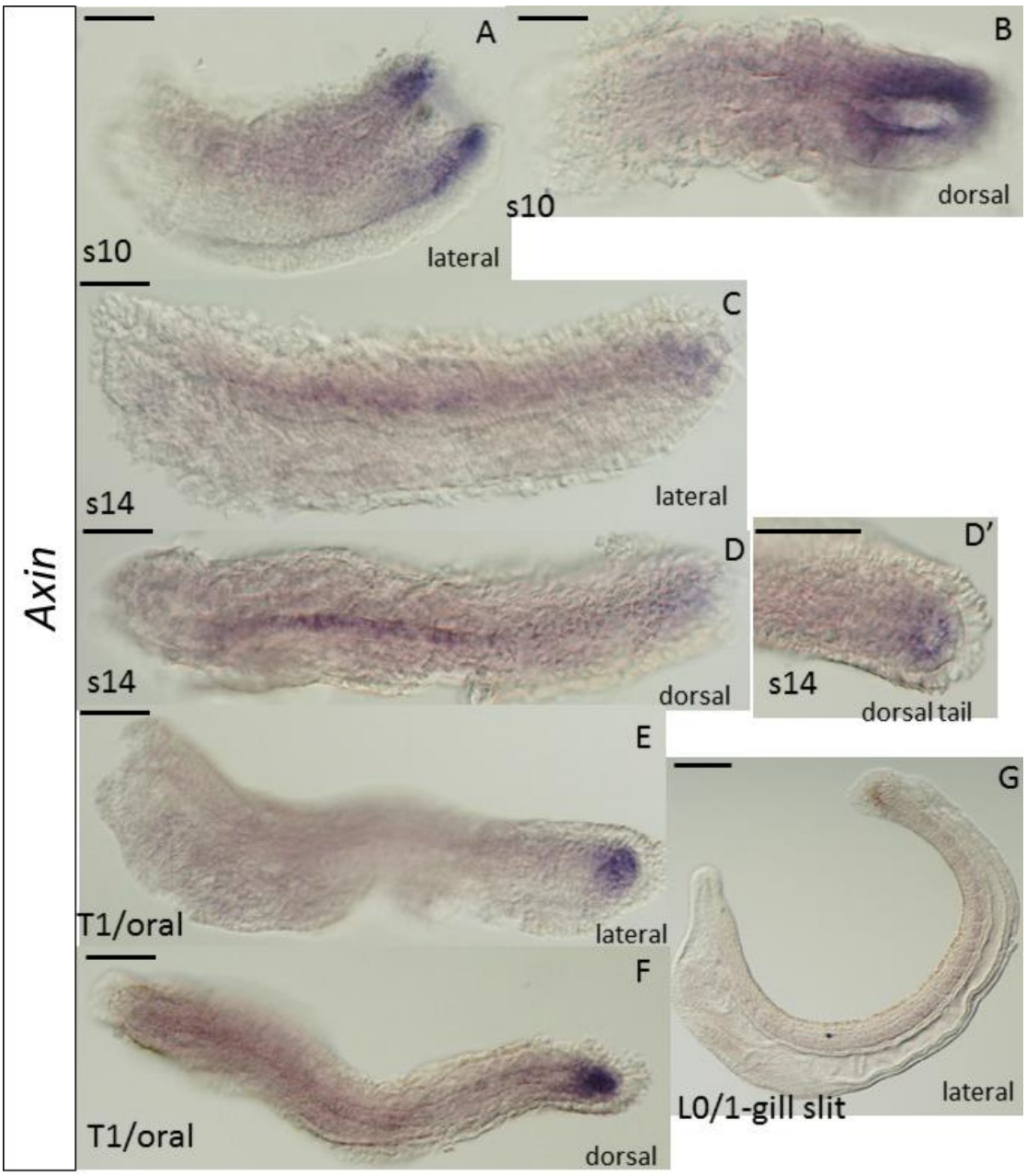


Figure 2.20. Amphioxus *Axin* expression characterization. Mainly expressed in tailbud from s10 to 1-gill slit stage with s14 stage also showing notochord expression. Embryo stages are at bottom left corner while bottom right shows angle of perspective. Some of the embryos have a kinked shaped and epidermal tissue has been lost, which was probably due to laboratory technicalities like the embryos were fixed at an angle against other embryos, overdigestion with protease K and/or being desiccated sometime during experiment or storage. All scale bars are 50µm.

Axin is downstream of canonical Wnt signalling and forms part of a negative feedback loop. Axin regulates β-catenin at the neuroepithelium to control the cell cycle, delamination, and survival (Yu et al., 2007). In amphioxus, *Axin* is expressed faintly in notochord and strongly in the tailbud in different stages (Figure 2.20). Lamprey *AxinA* is expressed in the dorsal ventricular zone of the central nervous system (Figure 2.19C-C"') with strong expression in the fore-mid brain boundary and mid-hind brain boundary, upper and lower lip mesenchyme, and also tailbud (Figure 2.19). Eckei et al. (2016) characterised a comprehensive temporal range of chick *Axin2* expression showing dorsal spinal cord expression from HH15-27. The dorsal spinal cord expression in lamprey and chick would be consistent with the canonical pathway being a cell fate determinant, downstream of Wnt1 and Wnt3 in the roofplate (Figure2.21; (Parr et al., 1993, Bulfone et al., 2004). However, as suggested by other studies, canonical and non-canonical pathways overlap (Grumolato et al., 2010, van Amerongen et al., 2012), which would mean *Axin* in amphioxus, lamprey, and chick being involved in cell proliferation, cell movement, cell patterning, and cell survival.

To understand the dynamics of Wnt signaling in lamprey, I also characterised the expression of nine of the 23 lamprey Wnt genes. Only *Wnt1* and one copy of *Wnt5* and *Wnt7* expression patterns had been reported in lamprey so far (Guerin et al., 2009, Sugahara et al., 2016). The expression of those that are expressed in the spinal cord are included here. Lamprey *Wnt1, Wnt5* and *Wnt6* are expressed in the roof plate of the lamprey spinal cord which means they could be involved in patterning the dorsal spinal cord cell populations. In gnathostomes, only *Wnt1* is well described as patterning dl1 and dl2 cell types (Muroyama et al., 2002). However, a more recent study shows that *Wnt1* is also required in proliferation of the ventricular zone (Shinozuka et al., 2019). The role of Wnt5 is not well studied in gnathostomes, though Dickson (2005) suggest that it might be involved in axon guidance along the anterior-posterior axis of the spinal cord because of its function in *Drosophila* (Yoshikawa et al., 2003). Schubert et al. (2002) characterised *Wnt6* in chick, where it is expressed in epidermal ectoderm. It is conceivable that Wnt6 could have a role in patterning the spinal cord though this is not discussed in the study. *Wnt1, Wnt5,* and *Wnt6* in amphioxus are mostly expressed in the proliferative tailbud region. It is possible that in addition to stem cell maintenance, they might also play a role in patterning the neural tube that forms in this region. Functional studies using amphioxus knockout lines would need to be carried out to confirm such a hypothesis. Mouse *Wnt7* expression has been characterised in the ventral two-thirds of the

ventricular zone but its function here has not been studied. Lamprey *Wnt7_6* differs in expression pattern from gnathostomes and seems to be expressed in the dorsal one-third of the spinal cord and throughout the ependymal cell layer. Expression of the other five lamprey *Wnt7* paralogues is still unknown. Amphioxus *Wnt7* is more likely to be involved in neuron differentiation as it is expressed in the spinal cord but not the tailbud. Comparing to the *Wnt* expression in mouse by Summerhurst et al. (2008), very rarely do *Wnts* have this ubiquitous CNS expression like *wnt2, wnt11, wnt5_5, wnt9* and *wnt16* in lamprey (Figure 2.22), maybe with the exception of mouse *Wnt2b.* Even though *Wnt* expression is well characterised in hemichordate (Darras et al., 2018), amphioxus (Somorjai et al., 2018), and gnathostomes (Summerhurst et al., 2008), their respective functions in each species remain poorly understood. While the lack of specificity of *Wnt* expression visible in figure 2.22 might suggest that the vertebrate ancestor Wnts have yet to subfunctionalise and specialize to pattern a larger more complex CNS, the conserved expression seen in figure 2.21 throughout vertebrates would suggest that functions like dorsal spinal cord patterning and ventricular zone planar cell polarity/stem cell maintenance have been well conserved since the vertebrate ancestor. A potential explanation for the widespread mRNA distribution of some lamprey *Wnt* genes compared to the relatively restricted expression of lamprey *AxinA* could be that not all Wnts are functional across their expression range: for example some genes may not produce functional protein, or may only produce proteins in certain areas due to posttranscriptional regulation. Another explanation could be that other *Axin* paralogues which I have not tested might have a more ubiquitous expression.

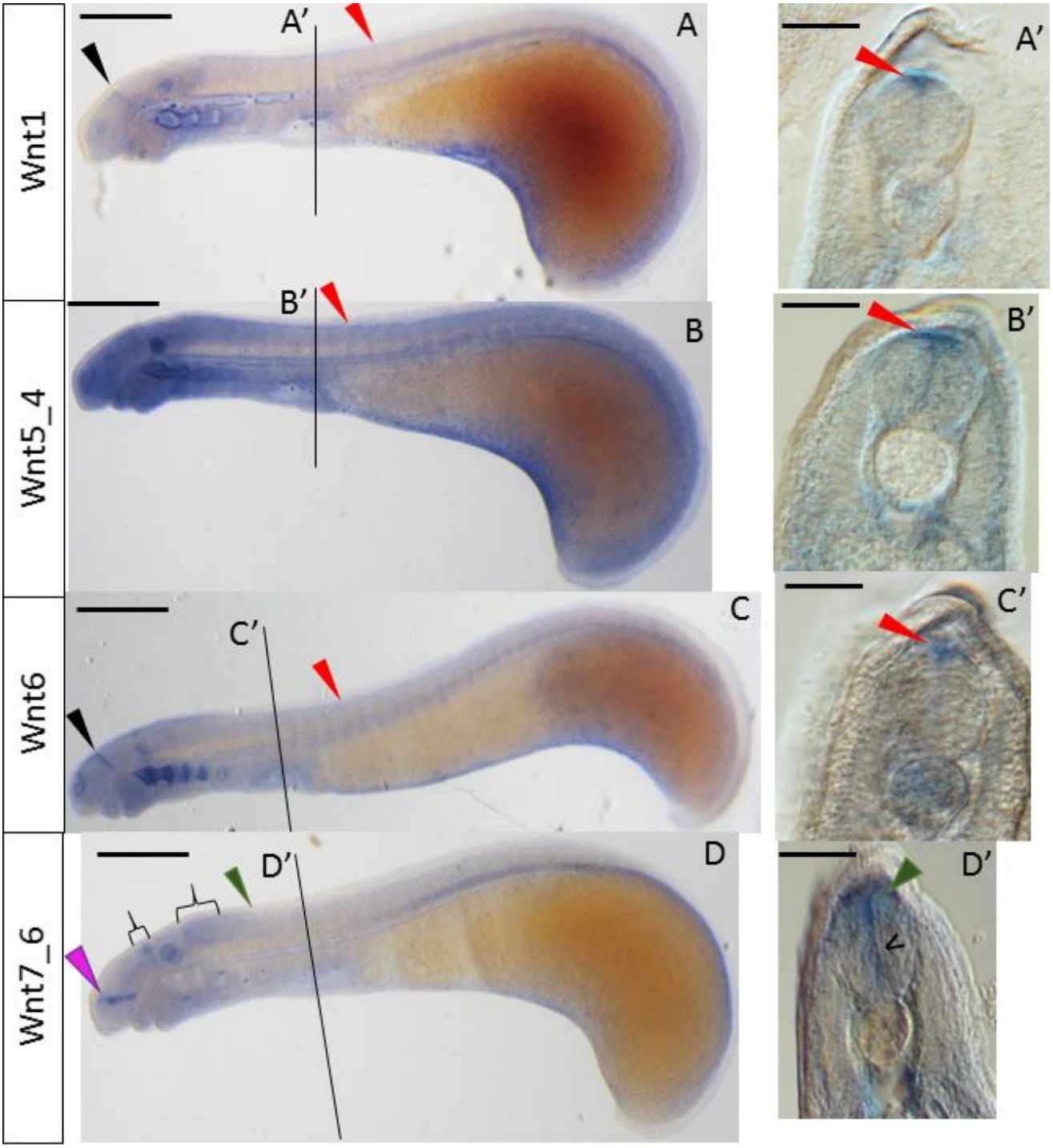


Figure 2.21. Lamprey *Wnt1, Wnt5_4, Wnt6,* and *Wnt7_6* expression at T25. Black arrow heads mark the mid-hind-brain boundary expression. Red arrow heads marking dorsal spinal cord expression. Green arrow heads mark dorsal spinal cord expression. Magenta arrow head marks zona limitans intrathalamica (ZLI) expression. Black brackets show the hindbrain expression in segments. "V" is pointing to ventricular zone expression. Transverse cross-section photos correspond to the black lines on the embryos on the left. Scale bars: 500µm (A-D); 100µm (A'-D').

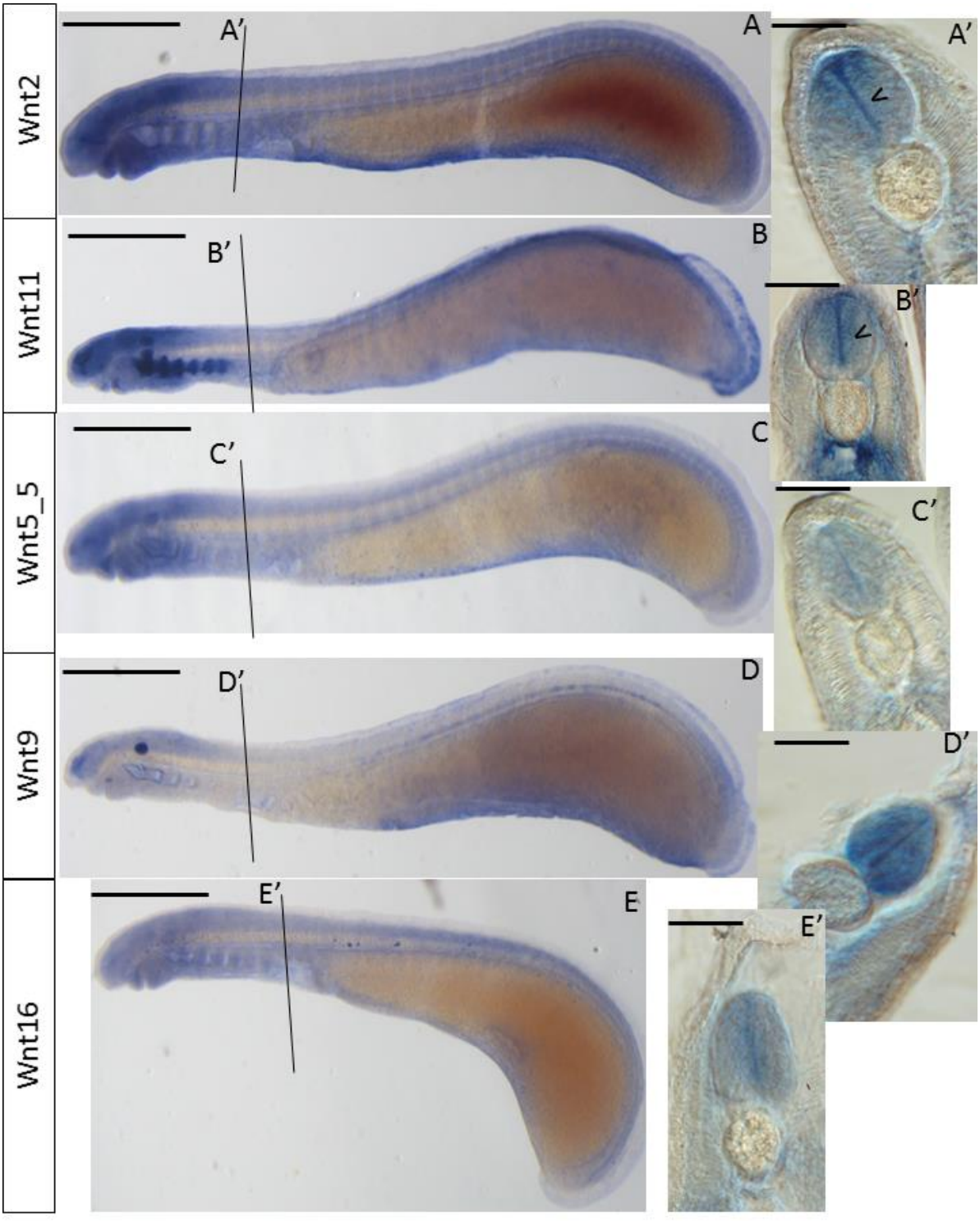


Figure 2.22. Lamprey *Wnt2, Wnt11, Wnt5_5, Wnt9,* and *Wnt16* expression at T25. All of these Wnts have general ubiquitous CNS expression. Level of transverse cross-section photos on the right corresponds to the black line on the left. "V" marking ventricular zone expression. Scale bars: 500µm (A-E); 100µm (A'-E').

2.4 Summary, conclusions and future direction

Before generating these data, Kusakabe and Kuratani (2005), who focused on the myotome, characterised lamprey *Pax7* from T22 to T28 using *Lethenteron japonicum* while Modrell et al. (2014), who focused on cranial sensory ganglia using *Petromyzon marinus,* characterised lamprey *Pax7* from T20 to T25. I characterised lamprey *Pax7* in *L. planeri* and my results are the same as these other two lamprey species. Lamprey *OligA* had only been previously shown in *L. planeri* at stage T26 (Lara-Ramirez et al., 2019). I have contributed stages from T23 to T28 showing that *OligA* starts expressing in earlier spinal cord patterning stages and anterior CNS expression fades in older embryos. For lamprey *Pax6β,* Murakami et al. (2001) used *L. japonica* and characterised stages from T19 to T26 while Derobert et al. (2002) used *Lampetra fluviatilis* showing expression of stages T25 and T27. Both studies focused on brain development. My results show that *L. planeri* expression is the same as the two other published lamprey species. Yuan et al. (2018) showed lamprey *Nkx2.2* expression from stage T22 to T25 using *P. marinus,* while Sugahara et al. (2011) showed a T26 head of *L. japonicum*. My *L. planeri* data recapitulate that of the other two published lamprey species. Guerin et al. (2009) characterised lamprey *Dbx* at stage T25 and T26 focusing on the head with *L. fluviatilis.* My *L. planeri* data recapitulates the head expression of Guerin et al. (2009) data and also confirmed that *Dbx* expression is in the spinal cord from stages T22 to T28. In addition, I reported the expression of genes previously not studied, *OligB, Nkx6, AxinA, Wnt2, Wnt5_4, Wnt6, Wnt9, Wnt11,* and *Wnt16.* Generally, existing literature studying relevant gene expression has been on other species of lamprey and has focused on the brain or non-neural organs instead of the spinal cord. My data confirms that *L. planeri* expression for my genes of interest is the same as the other lamprey species and also expands on the range of genes and temporal stages being reported. This confirms that these genes are involved in the developmental timeframe of spinal cord patterning. *Nkx2.2, Nkx6A, Dbx, Pax6β, Pax7, PtchB, AxinA,* and *Wnt1* lamprey spinal cord expression are within the same region as gnathostomes. Lamprey *OligA* and *OligB* are similar to gnathostome *Olig2* and *Olig3* spinal cord expression with both lamprey orthologues present in dorsal spinal cord while only *Olig3* is expressed in dorsal spinal cord. These data also provided the foundation for deciding which genes to test for in the following signal manipulation chapter.

OligA expression in *B. floridae* has been reported by Beaster-Jones et al. (2008). Ren et al. (2020) characterised the other two amphioxus paralogues of *Olig,* showing that

they have similar expression in the neural tube. *Nkx6* expression has been characterised in *B. floridae* by Albuixech-Crespo et al. (2017). However, their study focused on the mid-neurula stage. My data showed that *Nkx6* neural tube expression extends to early larval stages but not to older larval stages. *SoxB1a* has been described as expressed in the posterior neural tube (Holland et al., 2000). However, my data, showing dorsal views of the same stage of *B. floridae,* revealed that the expression is at the presomitic tissue and not the neural tube. I also report older developmental stages to show that it is not expressed in the neural tube later in development. Amphioxus *Dbx* expression data has not been reported before. *Zic* expression has been characterised by Gostling and Shimeld (2003) in *B. floridae*. My data recapitulates theirs and also establishes Zic posterior neural tube expression. *B. floridae Vacht* had been previously studied by Candiani et al. (2008). My data for *Vacht* confirms theirs, showing it is not expressed in posterior larval neural tube and hence helping define the amphioxus spinal cord region for this thesis. My amphioxus *Mnx/Hb9, Pax2/5/8, IrxA, Axin, Netrin,* and *Gbx* data show the same expression as previously published data (Table 2.1).

Out of the limited number of spinal cord expressing transcription factors I have analysed, the number of orthologues in amphioxus compared to lamprey seems to vary between 1:1-4. This is not enough evidence to support either hypothesis of one autotetraploidisation in the vertebrate ancestor with subsequent tetraploidisation and gene losses; or octaploidisation followed by delayed rediploidisation. Other methods that compare many species genomes would provide more convincing evidence to solve this decades long question. My data fit the hypothesis that lamprey would be similar to gnathostomes while amphioxus would be more different to vertebrates. The increase in number of orthologues in the vertebrate ancestor could have played a role in increasing the size of the spinal cord and changed the way that spinal cord neurons are being patterned: in this context I note that some lamprey spinal cord expressing genes are not expressed in amphioxus spinal cord at all, e.g. *Nkx2.2, Dbx, Pax6, Pax7, Axin*. The way that spinal cord neurons specialised will be explored more in the following chapter.

In general, with amphioxus transcription factor genes, it is much harder to define whether they are progenitor markers or differentiation markers (or both) because these processes are temporally but not spatially separate and genes can be co-expressed in the same cells within the whole developmental timeframe. However, some of the gnathostome spinal cord interneuron markers like *Dbx, Pax4/6, Pax3/7*

seem to be a vertebrate novelty because they are not expressed in the amphioxus spinal cord.

In addition to *Nkx6* (Dasen, 2018), *Olig* also seems to be well conserved across chordate motor neurons. Other genes like *Msx* seem to have diverged in function exclusively in the cyclostome lineage because both amphioxus and gnathostome *Msx* genes are expressed in the spinal cord. The rest of the transcription factors studied here, *Zic, Irx, SoxB1, Pax2/5/8,* and *Nkx2,* have at least one amphioxus paralogue which is expressed in the spinal cord, but whether they are in the same types of cells as gnathostomes would require further studies. To identify whether the putative gnathostome interneuron markers have conserved functions in amphioxus, double in situ/co-expression experiments would be a starting point as it would identify whether expression overlaps in the same cells or is mutually exclusive. Single cell RNA-seq might also achieve this aim. Knockout functional studies are needed to decipher gene regulatory networks for GABAergic or glutamatergic neuron differentiation.

Chapter3: Signal pathway manipulation effects on gene expression in the spinal cord

3.1 Introduction

As described in chapter two, lamprey spinal cord gene expression patterns are similar to gnathostomes. In this chapter, I focus on lamprey *Nkx2.2, Nkx6A, Dbx, Pax7, OligA, PtchB,* and *AxinA,* and whether they are regulated by the same signalling pathways as in the gnathostome spinal cord. Given that the lamprey genes have similar expression patterns as gnathostomes, I hypothesise that their gene regulation by patterning signals would also be similar. Such a hypothesis is drawn because similar anatomy, gene expression patterns, and number of orthologues might suggest that their gene regulatory network would also be similar. However, my data suggest otherwise. Some genes which were known to be downstream of certain signals in gnathostomes were not affected when the signalling pathway was manipulated in lamprey, and I instead discovered these lamprey genes were downstream of other signals for gene induction and or maintenance. Some of the transcription factors in gnathostomes have cross-repressive effects, but that also does not seem to be the case in lamprey because when one of the transcription factors is ectopically expressed, it overlaps with other transcription factor expression.

As discussed in chapter two, amphioxus markers that separate between progenitor and differentiated neurons are not well defined and their expression overlaps spatio-temporally in amphioxus motor neurons. Amphioxus *Pax2/5/8, Nkx6, Netrin, Vacht, IrxA,* and *Gbx* are chosen to be studied in this chapter because they are expressed in the developing spinal cord within the experimental timeframe, from five-somite/early-neurula stage to 14-somite /early larval stage (Figure 3.4). From published amphioxus Hox (Pascual-Anaya et al., 2012) and brain molecular regionalization data (Albuixech-Crespo et al., 2017), amphioxus neuroanatomy, gene expression patterns, and number of orthologues in key gene families being different to vertebrates, I hypothesise that gene regulation by signals would also be very different from vertebrates. Concordant with my hypothesis, signalling pathway manipulation seems to affect the expression of most of these genes. A potential explanation could be that the gene regulatory networks of the signals are interconnected, so manipulating one pathway would inevitably change other pathways. Alternatively, signalling manipulation could be changing positional/temporal cues and hence identity of neurons. It could also be a combination of the above explanations.

For general background on gnathostome spinal cord patterning and evolutionary comparisons, please refer to Chapter one. All the signal manipulation cell biology relevant to this chapter is summarized in figures 3.1 and 3.2.

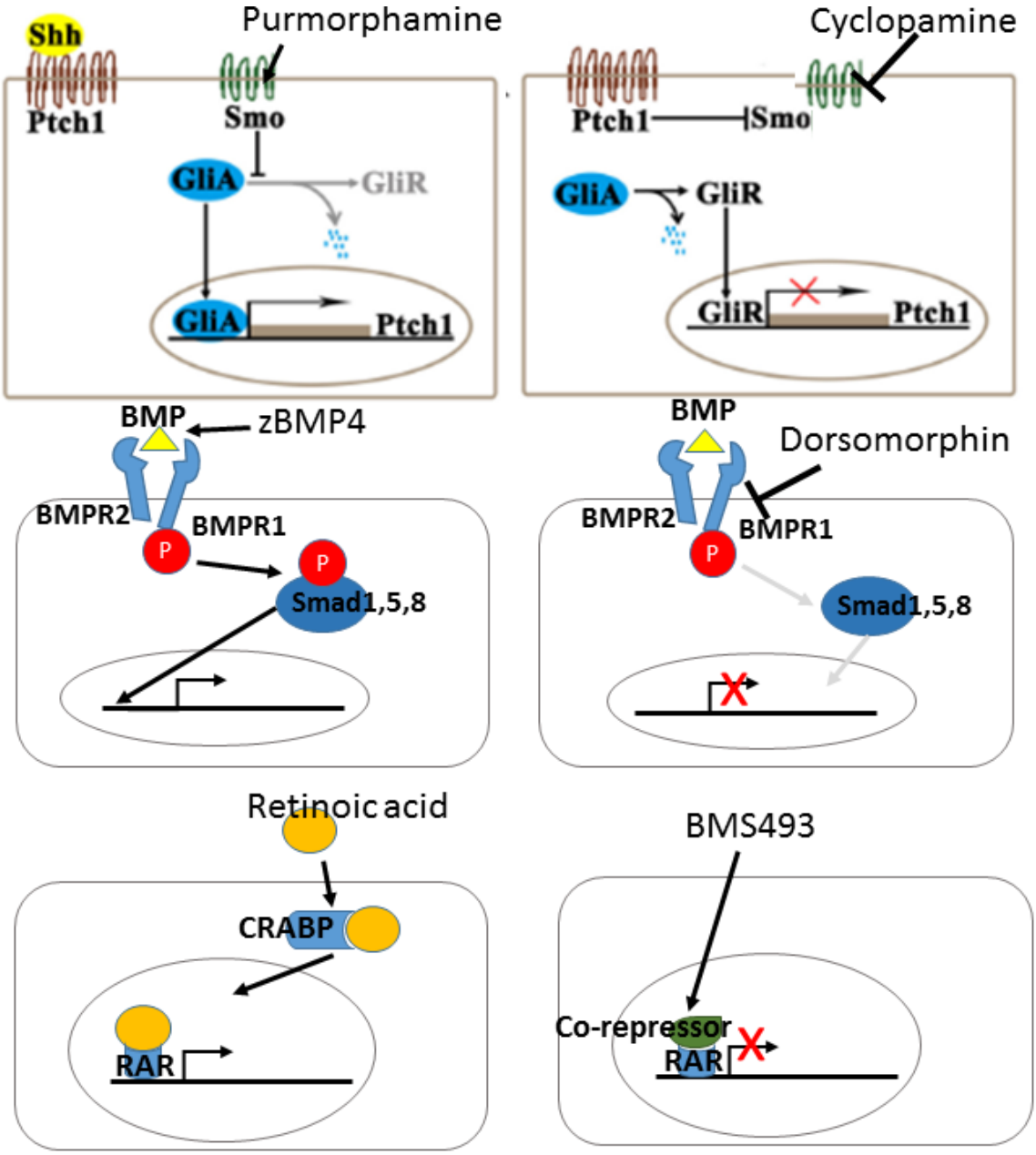


Figure 3.1. Pharmacological treatment pathway manipulation of Hh, BMP, and RA. The top row is showing Hh pathway manipulation using small pharmacological molecules. When Hh is present, it binds to Patched (Ptch) which allows Smoothened (Smo) to be actively repressing the cleavage of Gli activator protein and hence downstream transcription can occur. Purmorphamine overactivates while cyclopamine inhibits Smo. When Smo is inhibited, Gli activator is cleaved to become Gli repressor so

downstream transcription is repressed.

The middle row is showing BMP pathway manipulation using zebrafish BMP4 protein and the small pharmacological molecule, dorsomorphin. In the presence of BMP, BMP receptor one and two (BMPR1 and 2) bind to extracellular BMP protein and change their intracellular conformation and get phosphorylated, which then phosphorylates Smad1/5/8 which acts as a transcription factor for downstream signalling. When there is excess exogenous BMP, the pathway would be overactivated. Dorsomorphin inhibits BMP receptor one, which halts downstream intracellular transduction, resulting in no downstream transcriptional activation.

The bottom row is showing RA pathway manipulation. Endogenous retinoic acid normally diffuse into a cell and binds to cellular retinoic acid binding protein (CRABP) which transports it into the nucleus. Inside the nucleus, RA binds to retinoic acid receptor (RAR) together acting as transcription factors. Excess exogenous RA overactivates the RA pathway while BMS493 is a pan-RAR inverse antagonist, which causes conformational change to RAR and stops it from being a transcription factor.

Schematics of signalling pathway transduction are overly simplified for ease of understanding of how exogenous protein or small pharmacological molecules are acting on different pathways.

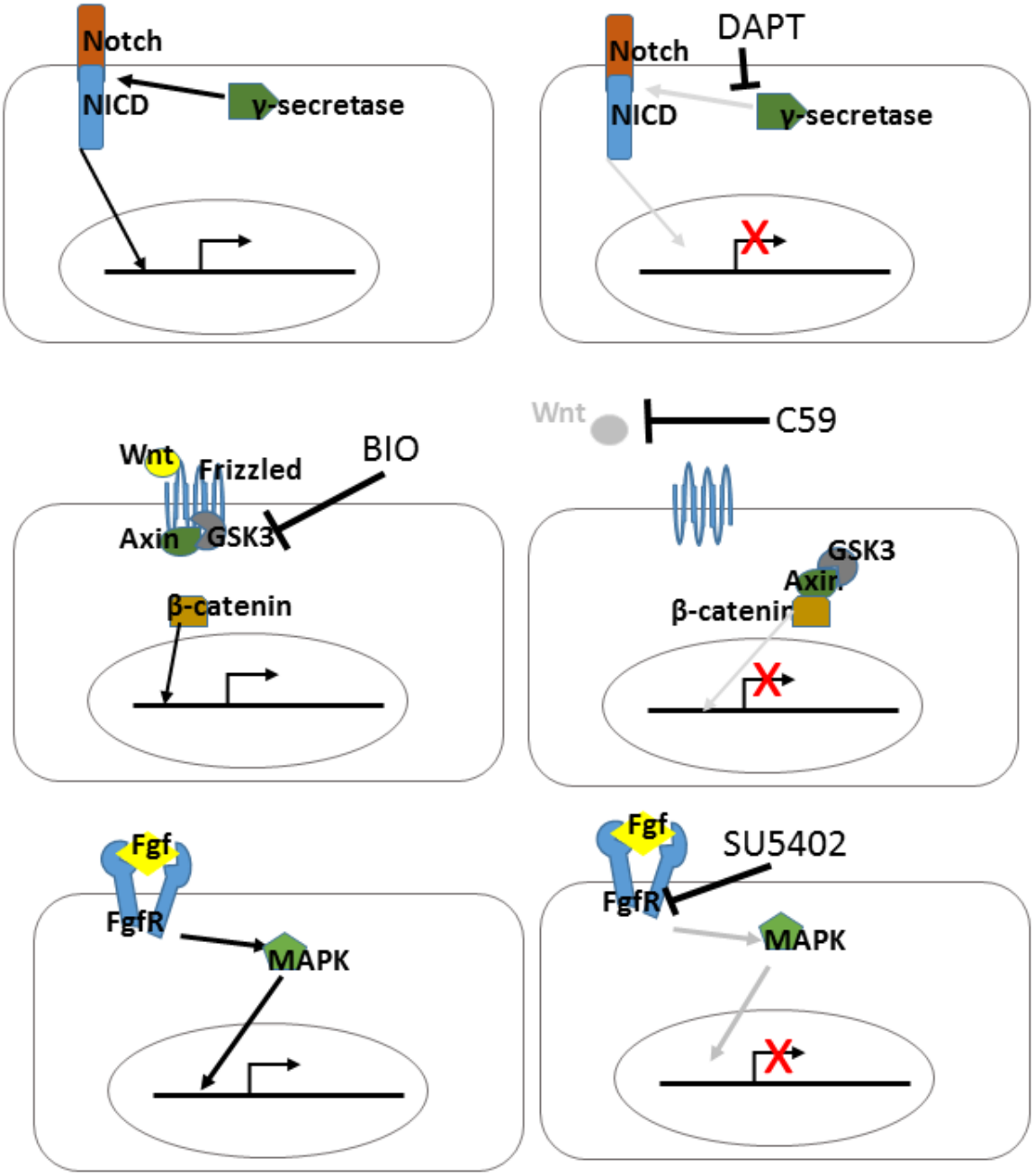


Figure 3.2. Pharmacological treatment pathway manipulation of Notch, Wnt, and Fgf. The top row is showing Notch pathway manipulation. Notch pathway is a cell-to-cell interaction. Delta from a neighbouring cell binds to the extracellular part of notch and causes conformation changes that allows γ-secretase to bind and cleave the notch intracellular domain (NICD) which enters the nucleus and acts as transcription factor. DAPT is a small pharmacological molecule which inhibits the γ-secretase so NICD remains at the plasma membrane and hence there is no downstream transcription.
The middle row is showing canonical Wnt pathway manipulation. When Wnt binds to Frizzled receptor, Axin and Gsk3 are recruited intracellularly to form a complex, which allows β-catenin to diffuse into the nucleus to act as transcription co-factor. BIO is a small pharmacological molecule that

inhibits Gsk3 which means β-catenin is constantly released from the complex and overactivates the canonical Wnt pathway. C59 is a small molecule that inhibits Porcupine, a protein at the endoplasmic reticulum, which prevents the release of all Wnts. Without the presence of Wnt, Gsk-Axin complex phosphorylates β-catenin so it is degraded. This means downstream transcription is prevented. Since C59 is blocking all Wnt release the planar-cell-polarity pathway, calcium dependent pathway, and canonical pathway are all halted. C59 was chosen because all Wnt pathways could potentially be involved in the spinal cord development. Future experiments could use inhibitors specific to certain Wnt pathways to futher dissect the gene regulatory network.

The bottom row is showing Fgf pathway manipulation. When Fgf binds to a pair of Fgf receptors (FgfR), there is a chain of intracellular transduction events which leads to Ras/mitogen-activated protein kinase (MAPK) entering the nucleus acting as a transcription factor. SU5402 is a small pharmacological molecule which antagonises FgfR by competitively binding to the ATP-binding site. So the transduction events cannot happen, which then leads to no downstream transcription.

Schematic of signalling pathway transduction are simplified for ease of understanding of how exogenous protein or small pharmacological molecules are acting on different pathways.

3.2 Methods

3.2.1 Small pharmacological molecule treatments

All small pharmacological molecules were dissolved in DMSO as stock solutions. The volume needed to make up the working concentrations (listed in table 3.1) were added to culture medium. Their respective control experiments were treated with the same concentration of DMSO. For example, 5mg of cyclopamine solid was dissolved in 1.21mL of DMSO to make up 10mM of stock solution, then 400µL of the stock solution is added to 40mL of filtered river water to make up 100µM of working concentration to rear lamprey embryos in; for the control experiment 400µL of DMSO was added to the 40mL of filtered river water. For zBMP4 treatment, the same volumes of 1mM HCl 0.2%BSA were used as control experiments because the zBMP4 was dissolved in 1mM HCl 0.2%BSA. Lamprey embryos were reared in 6cm or 3cm diameter petri dishes with volumes of 40mL or 5mL respectively. Amphioxus embryos were reared in 32-well plates with 2mL for each well. The drugs were mixed into the culture medium by gently pipetting. The treatment timeframes are summarised in figures 3.3 and 3.4. Please refer to chapter two methods for husbandry and in situ hybridisation protocols.

Table 3.1. Drug treatment. Working concentration were decided based on published methods for drug treatment in lamprey and amphioxus. Some drugs were not previously used on lamprey so embryos of similar size were referenced, like zebrafish (#), skate(@); for amphioxus, similar size embryos referenced here was sea urchin (*).

Drug	Stock	Lamprey	Citation	Amphioxus	Citation	Pathways
Cyclopamine	10mM	100µM 24°C	(Sugahara et al., 2011)	1.75µM	(Ono et al., 2018)	Shh↓
Purmorphamine	10mM	10µM	(Winata et al., 2009)#	1.75µM	Guessed based on Ono et al. (2018)	Shh↑
Retinoic acid	50uM	0.01µM	(Kuratani et al., 1998)	1µM	(Koop et al., 2014)	RA↑
BMS493	10mM	0.1µM	(Stafford and Prince, 2002) #	1µM	(Carvalho et al., 2017)	RA↓
C59	10mM	20µM	(Onimaru et al., 2016) @	0.5µM	(Cui et al., 2014) *	Wnt↓
BIO	10mM	-		0.5µM	(Onai et al., 2012)	Wnt↑
SU5402	10mM	50µM	(Jandzik et al., 2014)	50µM	(Bertrand et al., 2011)	Fgf↓
DAPT	10mM	100µM	(Lara-Ramirez et al., 2019)	50µM	(Lu et al., 2012)	Notch↓
Dorsomorphin	5mM	5µM	(Paul et al., 2008)#	20µM	(Kozmikova et al., 2013)	BMP↓
zBMP4	200ng/uL	-		250ng/mL	(Lu et al., 2012)	BMP↑

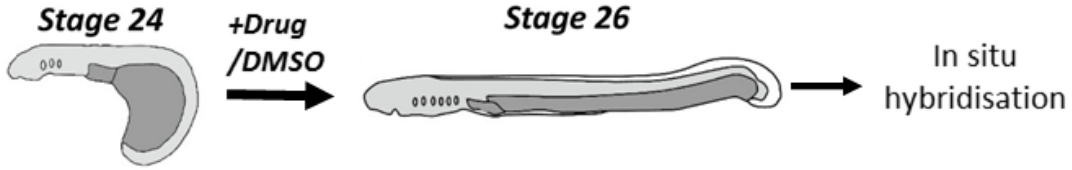


Figure 3.3. Lamprey pharmacological treatment schedule. Stage T24 lamprey embryos are transferred from normal river water rearing culture to drug treated river water rearing culture. When they reach T26, embryos are harvested and fixed for downstream experiments. Drawings adapted from Tahara, 1988.

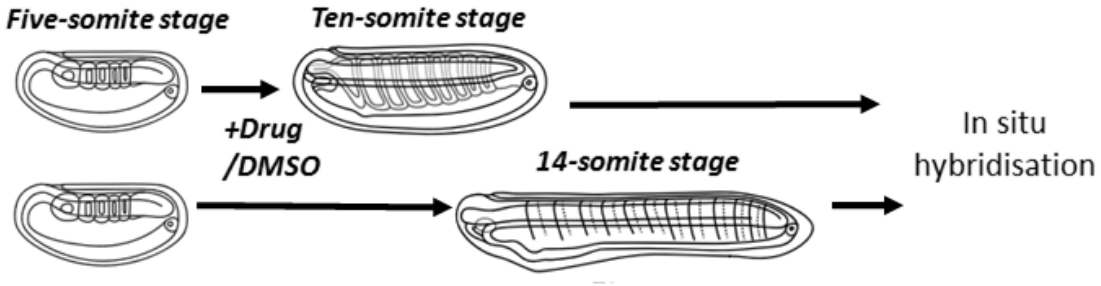


Figure 3.4. Amphioxus pharmacological treatment schedule. Five-somite stage embryos reared in artificial sea water are treated with drugs and left to grow to ten-somite stage. Part of the culture are harvested and fixed while remaining embryos are left to grow to 14-somite stage then fixed. Drawings adapted from Carvalho et al., 2020.

3.3 Results & Discussion

3.3.1 The impact of manipulating Hh signalling on lamprey spinal cord patterning

Considering the effects of cyclopamine and purmorphamine on lamprey development, we can postulate how Shh affects patterning of the developing spinal cord. Figure 3.5A-C show *Nkx2.2* expression changes following Hh manipulation. Cyclopamine-treated embryos have slightly lower *Nkx2.*2 expression at the ventral tail spinal cord (Figure3.5 A, A', B, B' black arrows). The purmorphamine treated embryos have ectopic expression throughout the whole spinal cord (Figure 3.5 C, C', B, B'; compare boxed region). Figure 3.5 J, K, L show *Pax7* expression at the trunk spinal cord level. The cyclopamine and DMSO treated embryos have the same expression pattern, but the purmorphamine treated embryos have greatly reduced expression as emphasized in the inset photos of transverse sections at the same level in the trunk. In Figure 3.6J-O, *AxinA* expression in cyclopamine has reduced expression compared to DMSO-treated embryos while the purmorphamine-treated embryos have completely lost *AxinA* expression in the head and dorsal CNS (compare L and O to regions arrowed in K and N). The *Nkx6A* and *Dbx* expression in Figure

3.5D-I and the *OligA* expression in Figure 3.6A-C may seem like they have expression variation under Hh manipulation but upon close examination, the spatial expression along the DV axis in the spinal cord is the same as in the controls. The variation is in colour intensity. This could reflect differences in transcript number and qPCR would test this. Such differences could also in principle be a consequence of technical variation. To provide a level of confidence that reduced or loss of expression is an experimental result and not a consequence of technical variation, I performed controls and the other experiments in parallel, under the same conditions and staining times, including testing overstaining: In some instances, increasing staining time resulted in BCIP/NBT solidifying into yellow clumps on the surface of the embryos and not intensifying the purple/blue colour with the antibody colorimetric conjugate. In other instances, staining for longer would raise the background staining level too high, though the clearing step with BABB helps with this by leaching out background staining.

Over-activation of the Hh pathway with purmorphamine ventralises the p3 cell marker *Nkx2.2* and consequentially reduced partially or fully the dorsal cell type markers *Pax7* and *Axin,* respectively. However, it caused no change in *Nkx6A*(p2-3), *Dbx*(dI6-p0) or *OligA*(dI1-3,pMN) expression. This suggests that cell type identity does not have set boundaries in developing lamprey because the ventralised area of *Nkx2.2* overlaps with the *Nkx6A, Dbx*, *Pax7* and *OligA* expression. This led me to believe that other cell type markers are maintained by signals other than Shh, though it is difficult to judge whether there might be subtle shifts in expression boundaries with *OligA, Nkx6A* and *Dbx* without transverse sections and double in situ hybridisation. Cyclopamine treated embryos showed a reduction in *PtchB* expression in lamprey fore and mid-brain but not in spinal cord (Figure 3.6 D, E, G, H: compare arrowed and bracketed regions). This is consistent with Sugahara et al. (2011) which depicted loss of *GliA* (downstream of *PtchB*) expression in brain when lamprey embryos were treated with cyclopamine. My data seem to suggest that Hh inhibition has little effect on spinal cord gene expression. From the results of both up and down regulation of the Shh pathway, I conclude that other signalling pathways are also important for the maintenance of DV cell type markers in lamprey. As a side note, both cyclopamine and purmorphamine treated embryos have ectopic *PtchB* expression in somites and spinal cord at trunk level (Figure 3.6 G-I).

In chick embryos, *Nkx6.1* expression in the ventral neural tube is lost in *Shh* and *Gli*2 homozygous mutants while foregut expression is not affected (Cai et al., 2000). In mouse *Shh* mutant embryos, *Nkx2.2* expression in neural tube is lost (Pabst et al.,

2000). In chick explants and zebrafish embryos, Shh is sufficient to induce *Nkx2.2* (Soula et al., 2001, Schäfer et al., 2005). Lamprey mimics the overexpression of *Nkx2.2* when Shh signalling is upregulated, and also its ectopic expression overlaps with *Nkx6.1* and *Pax7* expression in chick explants (Briscoe et al., 2000), but my cyclopamine-treated lamprey data show only slightly reduced *Nkx2.2* expression instead of it being completely lost like in *Shh* knockout mice (Figure 3.5 A, B). In the dorsal spinal cord, chick and lamprey develop similarly when *Shh* is overexpressed as *Pax3* and *Pax7* expression are repressed (Figure 3.5 K, L; (Lin et al., 2016)). In *Shh* mutants, *Dbx1* and *Dbx2* are expressed in the floorplate instead of the p0/p1 area (Pierani et al., 1999), whereas cyclopamine treated lamprey embryos show no effects on *Dbx* expression in spinal cord (Figure 3.5 G, H; arrows). *Dbx* also cross-represses with *Nkx6* (Vallstedt et al., 2001, Gribble et al., 2007). In gnathostomes, Shh is required (Qi et al., 2003) and sufficient (Danesin et al., 2006) for *Olig2.* While there is no evidence of a Shh relationship with ventral subpopulations of *Olig3* cells, the dorsal *Olig3* population is known to be instructed by both Wnt/beta-catenin and BMP signals (Zechner et al., 2007). Comparing this to lamprey, *OligA* is affected by the inhibition of BMP signalling. This leads us onto the next section on the impact of BMP signalling.

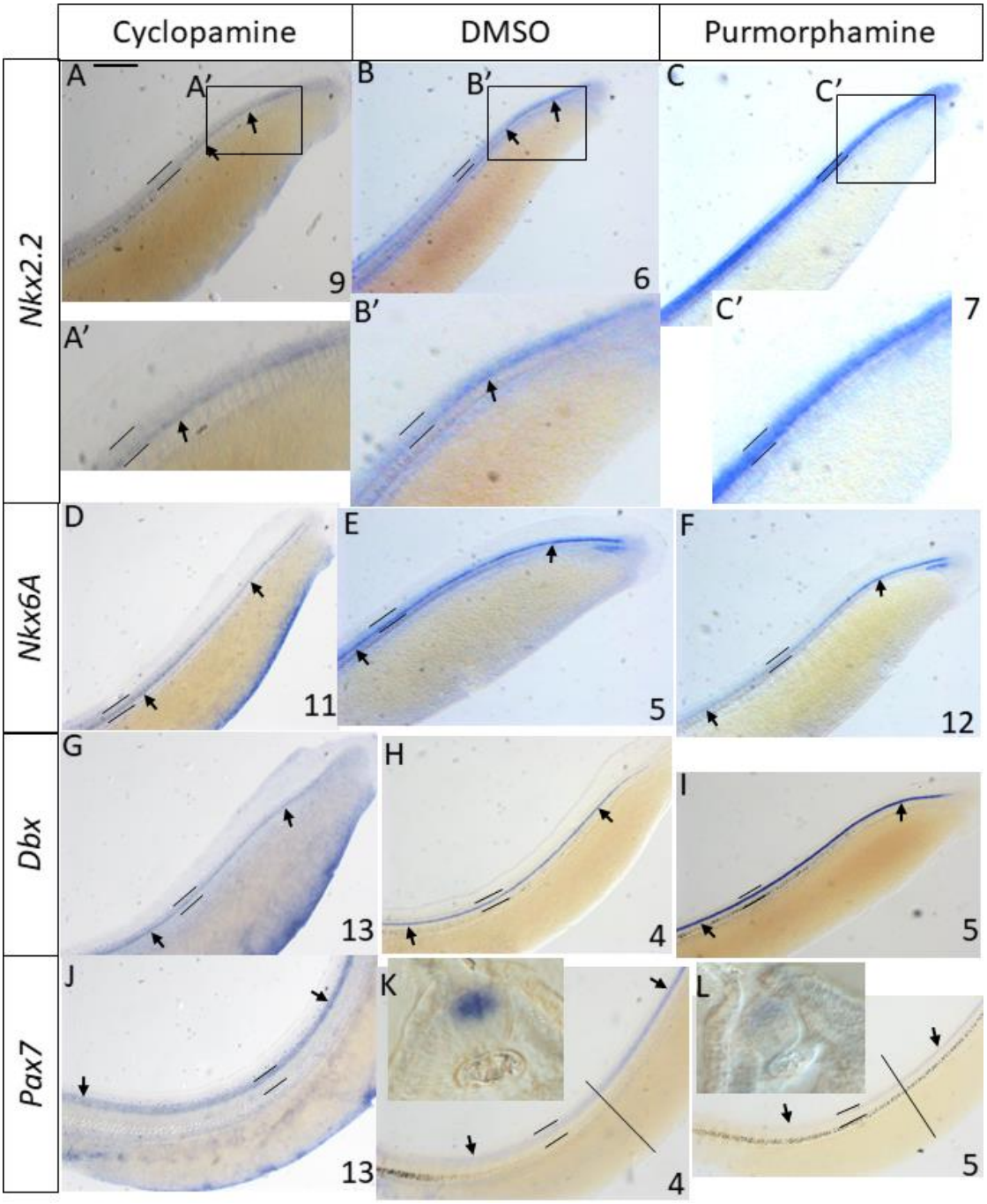


Figure 3.5. Effects of Hh manipulation on *Nkx2.2, Nkx6A, Dbx,* and *Pax7* expression. The photos are lateral views of T26 lamprey tail expression of the respective genes except the last row which is showing trunk expression. Transverse sections of the tails are at the corners of the lateral view images for *Nkx2.2* and *Pax7* DMSO and purmorphamine treated embryos. Cyclopamine inhibits Hh signalling while purmorphamine overactivates it. DMSO treated embryos act as controls. The N number is at the bottom right of each image. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The top left black bar on the *Nkx2.2* expression of cyclopamine treated embryo photo is the scale bar for 200µm.

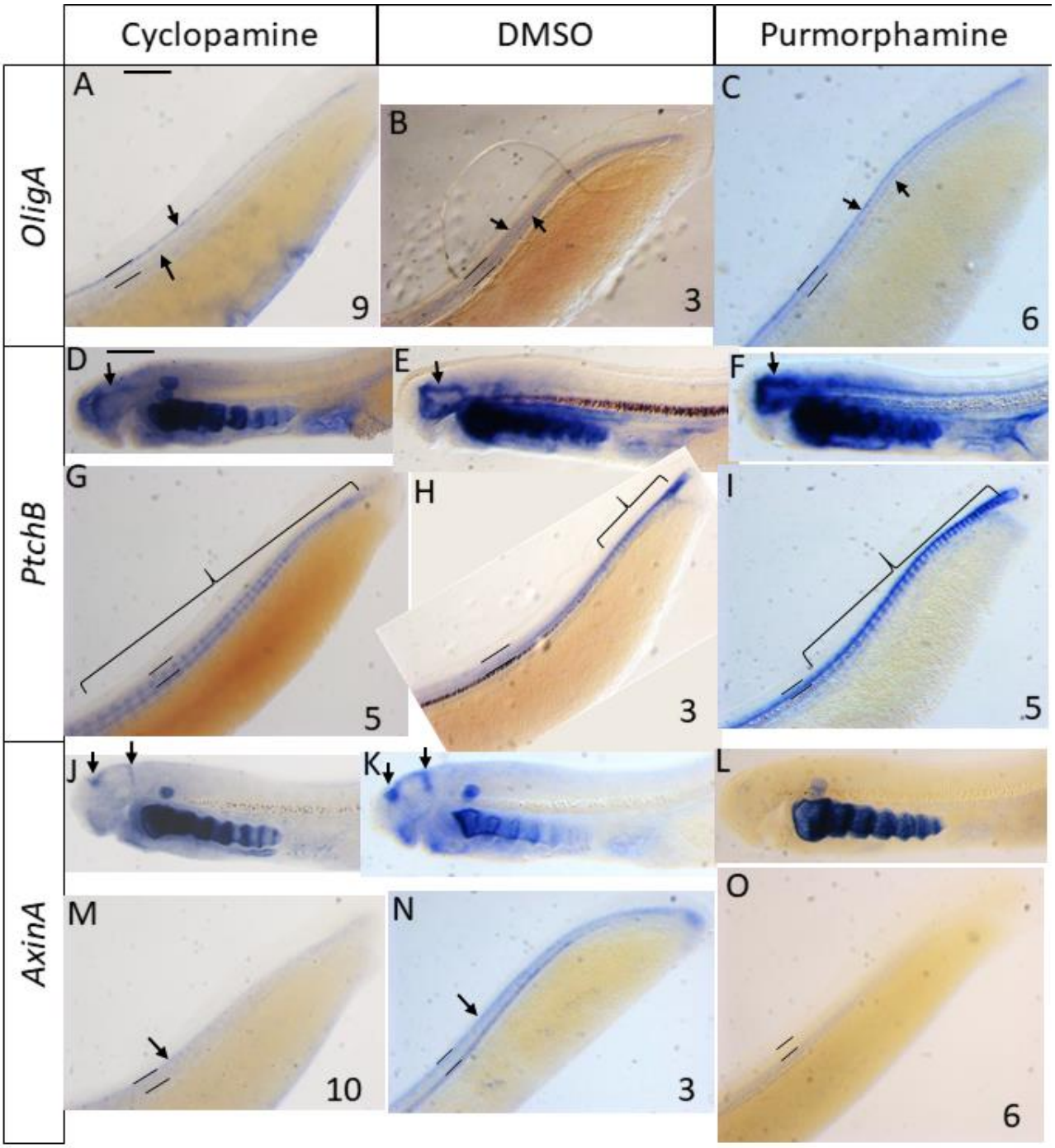


Figure 3.6. Effects of Hh manipulation on *OligA, PtchB,* and *AxinA* expression. Lateral views of T26 lamprey tail expression for all three genes with addition of head expression for *PtchB* and *Axin.* The N number is at the bottom right of each image. Cyclopamine inhibits the Hh signalling while purmorphamine overactivates it. DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The black brackets highlight the extent of *PtchB* somite expression. The top left black bars on the *OligA* and *PtchB* expression of cyclopamine treated embryo photo are the scale bars for 200µm.

3.3.2 The impact of inhibiting BMP signalling on lamprey spinal cord patterning

Out of all the genes tested, dorsomorphin treatment seems to only have an effect on *Nkx2.2, OligA, AxinA,* and *PtchB* (Figure 3.7A, B, 3.8 A-J). In the treated embryos, *Nkx2.2* ectopic expression covers the whole spinal cord as highlighted in the transverse section inset in Figure 3.7 A, B. In Figure 3.8, *OligA* and *AxinA* expression are shown to be lost when BMP is inhibited (Figure 3.8 A-F, I, J; compare arrowed region and between black lines) while *PtchB* has ectopic trunk expression at the ventral spinal cord and somites (Figure 3.8 G, H; compare region between black lines and brackets). The lack of gnathostome-like cell type boundaries is also supported by these data, as when *Nkx2.2* expression is ventralised in the spinal cord while *Nkx6A, Dbx,* and *Pax7* remains the same and occupy space overlapping with the ectopic *Nkx2.2* expression (Figure 3.7). Perhaps Nkx2.2 is only repressing *Pax6* but not the other transcription factors (Briscoe et al., 2000). Contrary to my original hypothesis, when BMP is inhibited *Pax7* (dI2-6) dorsal cell type marker is not lost and *Dbx* (dI5-p1) intermediate cell type marker remains the same (Figure 3.7 E-H), while both dorsal and ventral stripes of OligA cells were lost instead of just the dorsal strip (Figure 3.8 A-F; compare arrowed region and between black lines), unlike in chick where only the dorsal population is affected (Zechner et al., 2007). This further supports that other mechanisms are at play in DV axis patterning. In chick, BMP over activation causes dorsalised expression of *Pax3* and *Pax7* at the expense of *Dbx1* and *Dbx2* (Timmer et al., 2002, Andrews et al., 2017). There are no gnathostome data showing effects of BMP inhibition on these transcription factors for direct comparison.

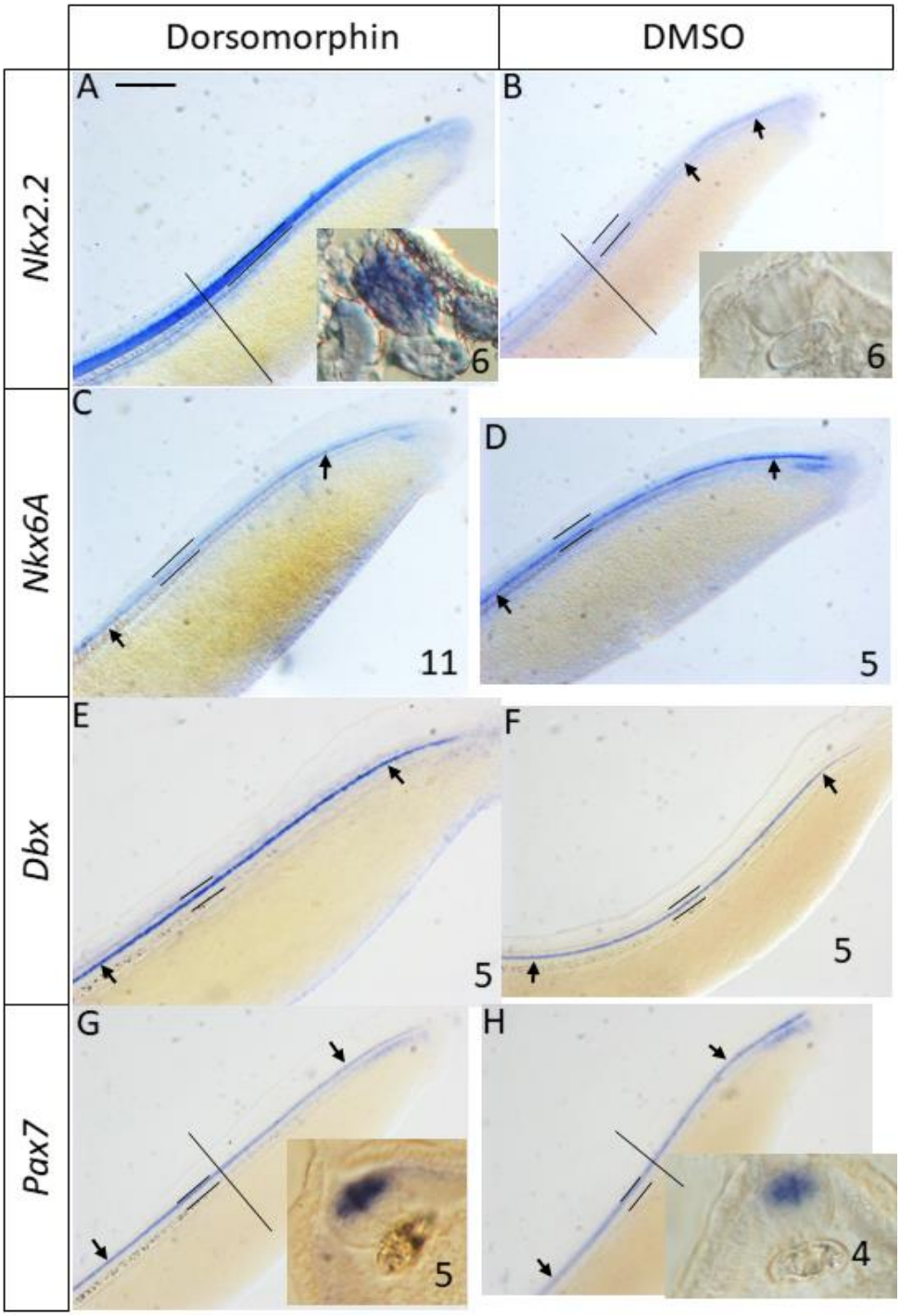


Figure 3.7. Effects of BMP inhibition on *Nkx2.2, Nkx6A, Dbx,* and *Pax7* expression. Lateral views of tail of T26 lamprey embryos. Inset images of *Nkx2.2* and *Pax7* are showing transverse sections at the tail level. The N number is at the bottom right of each image. Dorsomorphin inhibits the BMP signalling pathway and DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The top left black bar on the *Nkx2.2* expression of dorsomorphin treated embryo photo is the scale bar for 200µm.

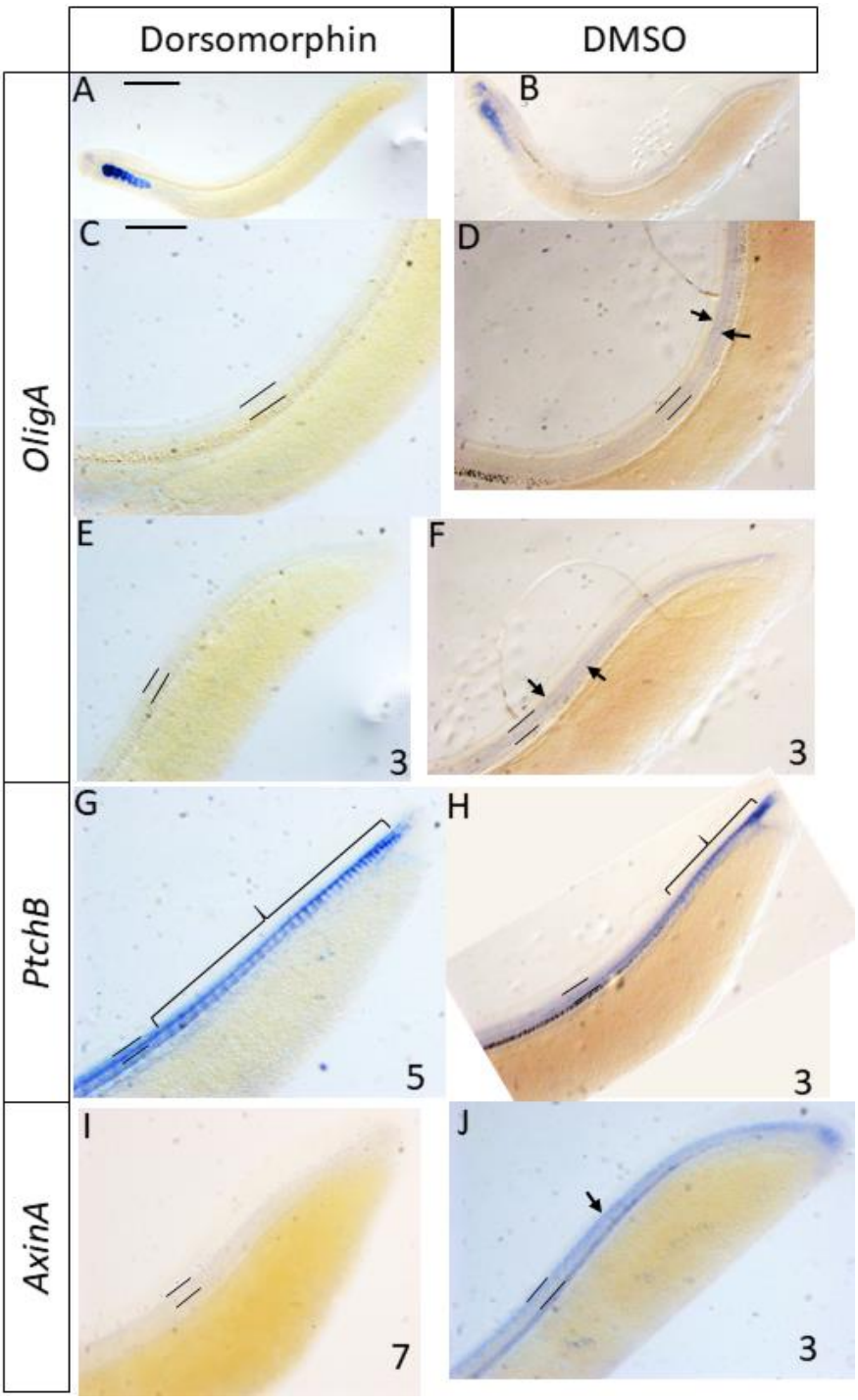


Figure 3.8. Effects of BMP inhibition on *OligA, PtchB,* and *AxinA* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. Dorsomorphin inhibits the BMP signalling pathway and DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The black brackets highlight the extent of *PtchB* somite expression. The top left black bars on the *OligA* expression of dorsomorphin treated embryo photos are the scale bars for 500µm for the whole embryo and 200µm for the zoomed in ones.

3.3.3 The impact of manipulating RA signalling on lamprey spinal cord patterning

The RA pathway is downregulated by BMS495 and upregulated by RA treatment. When the RA pathway is downregulated, there is a loss of expression of *Nkx2.2, Nkx6A, Dbx, Pax7, OligA, PtchB,* and *AxinA* (Figure 3.9, 3.10 compare left column BMS495 with central column DMSO). This shows that RA is required for all those cell identities, regardless of their DV position, and that other signalling pathways cannot compensate for the lack of RA. When RA is upregulated, there is a loss of expression for *Nkx2.2* (Figure 3.9 B, C)*, Nkx6A* (Figure 3.9 E, F), and *Axin* (Figure 3.10 H, I). There is a reduction in expression for *Pax7* with its presomitic mesoderm expression also lost, such that only a short section of tail spinal cord expression remains (Figure 3.9 K, L; region marked 'pm'). *PtchB* has ectopic expression in the trunk spinal cord and somite while tail spinal cord expression is reduced (Figure 3.10 E, F; compare bracketed regions). Looking at the *Pax7* and *PtchB* RA upregulated data alone, they suggest that other signals might be involved in the regulation of the posterior region of the tail spinal cord. Meanwhile, *Dbx* (Figure 3.9 H, I) and *OligA* (Figure 3.10 B, C) spinal cord expression remains the same. Taking the RA pathway manipulation results together, *Nkx2.2, Nkx6A, Pax7*, and *Axin* expression requires a specific endogenous level of RA activity because both up and down regulation caused the loss of their expression. However, *Dbx* and *OligA* spinal cord expression is only affected by down but not up regulation of RA, which suggests that in addition to the RA pathway, other mechanisms are involved in regulating *Dbx* and *OligA* expression.

Novitch et al. (2003) showed that in chick RA deficient embryos, unlike lamprey, *Nkx6.1* expression is ventralised. However, similar to lamprey, *Pax3* and *Pax7* expression are lost. In chick neural tube explant experiments by Pierani et al. (1999), *Dbx1* and *Dbx2* expression are induced by Shh and RA through independent pathways. In zebrafish, RA inhibition does not seem to affect *dbx* expression in significant ways (Gribble et al., 2007). RA deficient quails have much weaker *Nkx6.2* and *Olig2* but much stronger *Nkx6.1* expression, which expanded from pMN and p2 to the floorplate (Diez del Corral et al., 2003, Wilson et al., 2004). Excess RA represses *Nkx2.2* expression in zebrafish while inhibition of RA led to ventralisation of *Nkx2.2* at the expense of *olig2* expression (Schäfer et al., 2005). It seems like different vertebrates have their own mechanisms of regulating spinal cord transcription factors, although we have to be aware that developmental stages have always been difficult to compare across species and different experimental methods may produce different results. For example, drug treatment of embryos can precisely target certain developmental timeframes for manipulating a signal, while typical gene edited

embryos would be affected by that change through their whole development, unless a tissue-specific and time-dependent conditional knockout or knockin strategy is used.

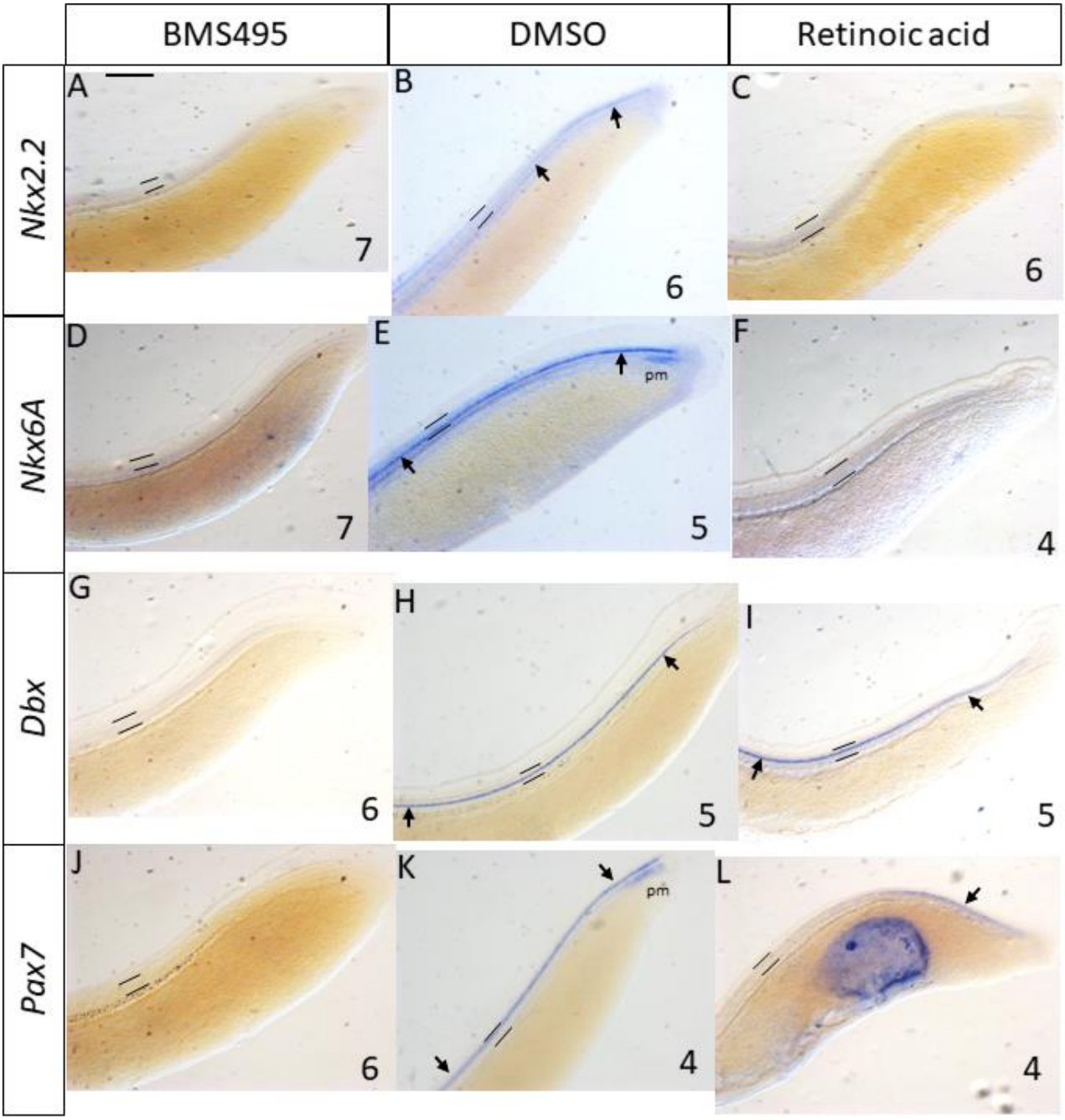


Figure 3.9. Effects of retinoic acid manipulation on *Nkx2.2, Nkx6A, Dbx,* and *Pax7* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. BMS495 inhibits the RA signalling pathway while exogenous RA overactivates it. DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. Presomitic mesoderm: pm. The top left black bar on the *Nkx2.2* expression of BMS495 treated embryo photo is the scale bar for 200µm.

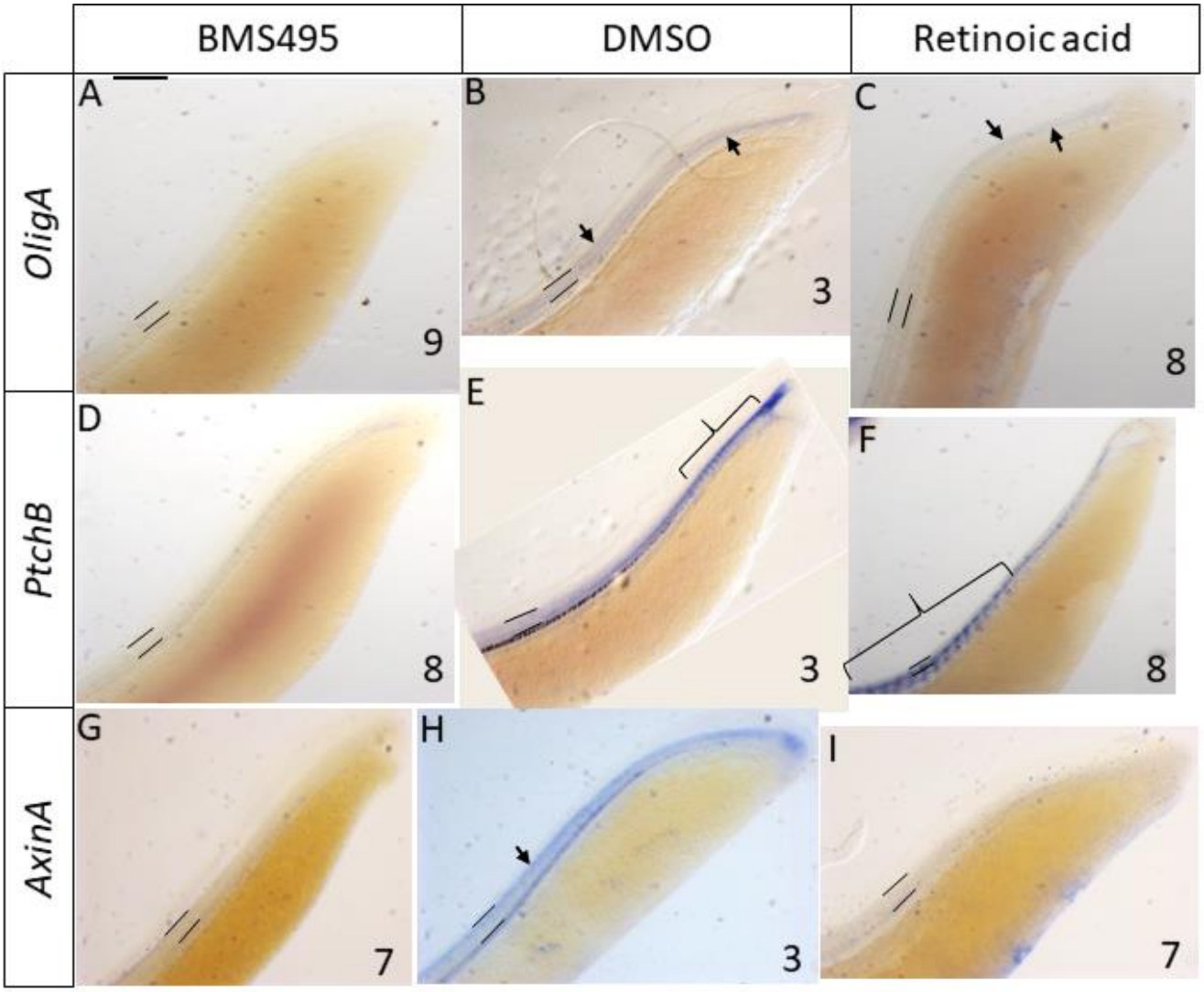


Figure 3.10. Effects of retinoic acid manipulation on *OligA, PtchB,* and *AxinA* expression. Lateral views of T26 lamprey embryo tail. The N number is at the bottom right of each image. BMS495 inhibits the RA signalling pathway while exogenous retinoic acid overactivates it. DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The black brackets highlight the extent of *PtchB* somite expression. The top left black bar on the *OligA* expression of BMS495 treated embryo photo is the scale bar for 200µm.

3.3.4 The impact of inhibiting Notch signalling on lamprey spinal cord patterning

I found that Notch pathway inhibition by DAPT causes loss of *Dbx* and *OligA* expression in spinal cord. Unlike the complete loss of expression in BMS495 treated embryos, the DAPT treated embryos maintain dotted *Dbx* expression (Figure 3.11 E-H; black arrows) and a short continuous dorsal/ventral strips of cells expressing *OligA*

at the tail spinal cord (Figure 3.12 C-F; black arrows). Combining these observations, a possible explanation could be that *Dbx* and *OligA* cell types need to be initiated by RA signals from posterior/tail somites (Castillo et al., 2010) then maintained throughout trunk spinal cord by Notch signals. Loss of expression of other cell markers like *Nkx2.2, Nkx6A,* (Figure 3.11 A-D; compare arrowed regions) *Pax7, PtchB,* and *Axin* (Figure 3.12 A, B, G-J; compare arrowed regions) in DAPT treated embryos shows that Notch signalling is also required for maintaining those genes. A likely mechanism for these results would be that these genes are expressed in progenitor cells and since Notch signalling maintains stemness in lamprey spinal cord (Lara-Ramirez et al., 2019) when Notch signalling is inhibited, the progenitor cells prematurely differentiated and lost the progenitor cell marker expression. Whether Notch signalling plays a role in initiating spinal cord transcription factor expression could be tested by treating embryos with DAPT at an earlier developmental stage.

DAPT treated lamprey embryos have similar phenotypes to conditional *Notch* knockout mice, as both have reduced expression in *Dbx1, Nkx2.2, Nkx6.1,* and *Olig2* (Yang et al., 2006). Li et al. (2016) showed that *Nkx6.1* in ventral neural tube regulates the expression of *Notch1* in stem cells. It would be interesting to see if this cross regulation between *Nkx6.1* and *Notch* is conserved in other chordates. There is no evidence from gnathostomes of the Notch pathway having an effect on *Pax3* or *Pax7* but I would predict the results to be the same as my lamprey data (loss of expression) given how conserved the effects of Notch signalling on spinal neural progenitor markers in vertebrates are.

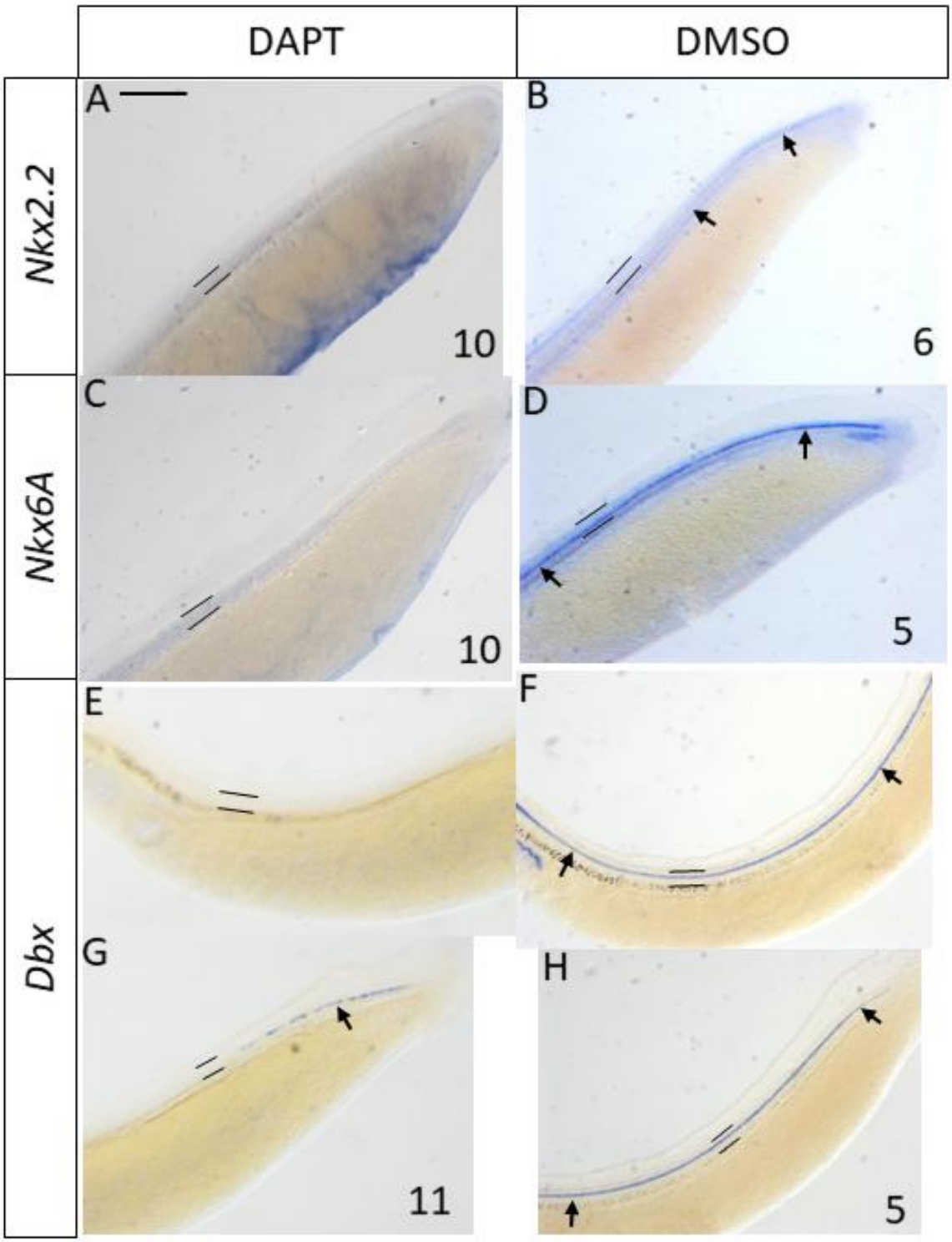


Figure 3.11. Effects of Notch inhibition on *Nkx2.2, Nkx6A,* and *Dbx* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. DAPT inhibits the Notch signalling pathway while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The top left black bar on the *Nkx2.2* expression of DAPT treated embryo photo is the scale bar for 200µm.

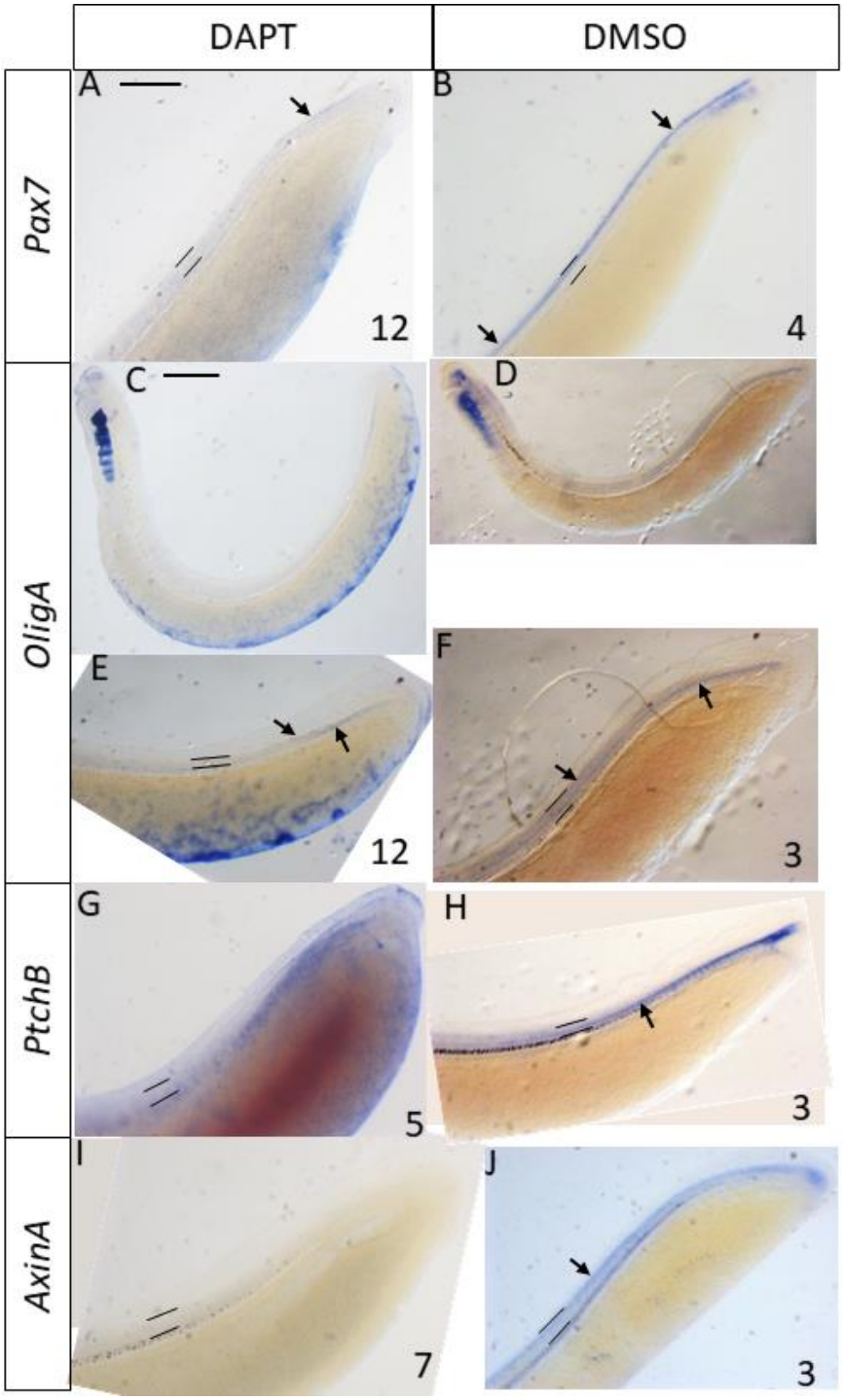


Figure 3.12. Effects of Notch inhibition on *Pax7, OligA, PtchB,* and *AxinA* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. DAPT inhibits the Notch signalling pathway while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The top left black bar on the *Pax7* expression of DAPT treated embryo photo is the scale bar for 200µm. The top left black bar on the *OligA* expression of DAPT treated whole embryo photo is the scale bar for 500µm.

3.3.5 The impact of inhibiting Wnt and Fgf signalling on lamprey spinal cord patterning

For the pan-Wnt inhibited lamprey embryos, there is a reduced expression of *Nkx2.2* in the tail spinal cord (Figure 3.13 A, B; compare boxed regions). Wnt/Tcf signals repress *Nkx2.2* expression in chick and mouse, contributing to the establishment of the p3/pMN boundary (Lei et al., 2006). Lamprey *Nkx6A, Pax7* and *OligA* expression remained the same in the spinal cord with the exception of loss of expression at the posterior-most region (Figure 3.13 C-F; Figure 3.14 A-F; compare arrowed regions). There is a general malformation or loss of tissue of the tailbud in embryos treated with C59. There is also a loss of presomitic mesoderm expression for *Nkx6A* and *Pax7* (Figure 3.13 E, F; Figure 3.14 A, B; regions marked 'pm'). Wnts have been known to regulate somite and tailbud formation in mouse, *Xenopus*, and zebrafish (Takada et al., 1994, Goto et al., 2017). This is also conserved in lamprey, seeing that the tailbud malformed into a cyst (most visible in the *PtchB* embryo in Figure 3.14 G). There are other morphological changes in C59-treated lamprey embryos. There is a reduction in size of forebrain (Figure 3.13.1, 3.14.1) and abnormal elongation of tail (Figure 3.13, 3.14; tails appear bulkier than the more tapered form in DMSO treated embryos). For *Dbx* and *Pax7* expression, the developmental abnormalities might make it difficult to judge whether there are shifts in expression boundaries along the DV axis. One possibility is that there is a reduction of dorsal neural tissue caused by C59, reducing the overall size of the spinal cord. This would make *Dbx* expression appear more dorsal than its control (Figure 3.13.1) and the *Pax7* expression appear thinner/more restricted dorsally when comparing to its control (Figure 3.14.1). In *Wnt1* and *Wnt3a* double knockout mice, there is a reduction of dorsal neural precursors while DV polarity remains the same (Ikeya et al., 1997). Their mice data recapitulates my observation in pan-Wnt blocked lamprey embryos. There is also a complete loss of *AxinA* expression and ectopic trunk expression of *PtchB* in the spinal cord and somites (Figure 3.14 G-J; arrows and bracketed regions).

SU5402 treated embryos in which Fgf signalling is inhibited have observable branchial arch malformations (Figure 3.16 A, B, G, H) as established previously by Jandzik et al. (2014). Similar to Wnt inhibited embryos, Fgf inhibited embryos also have ectopic trunk *PtchB* expression in the spinal cord and somites, which can be interpreted as posteriorisation (Figure 3.14 G, H; Figure 3.16 E, F; compare bracketed regions). There is also a generally higher level of *PtchB* expression in Fgf inhibited embryos than its control (Figure 3.16 E, F). The *PtchB* result matches observations in

other vertebrates, in which Wnts and Fgf are expressed in the tailbud and are responsible for AP patterning. However, Ribes et al. (2009) had shown that mouse caudal development relies on cross-talk between Shh, Fgf, and RA pathways while Stulberg et al. (2012) depicted cross-talk between Fgf and Wnt in zebrafish tailbud. This means all these signals could be interconnected, and as mentioned in the previous section, RA signalling activity modulates Hox collinearity and this is likely to be conserved in lamprey. The results of Wnt and Fgf signalling inhibition causing posteriorisation could be due to their connection to lamprey RA and Hox cluster chromatin openness. There is also evidence in mouse embryonic stem cells showing Fgf and Wnt signals induce *Cdx2* to clear away repressive H3K27me3 from *Hox1-9* chromatin domains, leading to branchial and thoracic identity of neurons (Mazzoni et al., 2013). A recent review by Saito et al. (2020) summarised different signal interactions in AP patterning and elongation mainly in chicks and mice. *Axin* is downstream of Wnt, which explains the complete loss of expression in C59 treated embryos but does not explain why *Axin* is also absent in Shh, RA, BMP, Fgf, and Notch manipulated embryos (Figure 3.6K-O, 3.8I-J, 3.10G-I, 3.12 I-J, 3.14I-J, 3.16G-J). A potential explanation is that all these signals are cross-regulating one another. This is similar to how Ptch is traditionally known to be downstream of Shh but in my cyclopamine treated embryos, *PtchB* expression remains while DAPT- and BMS495-treated embryos have a complete loss of *PtchB* expression in spinal cord. A possible explanation could be that RA signals primed the cells to be receptive to Shh signals (Ribes et al., 2009) while Notch signals maintain the cell types (Lara-Ramirez et al., 2019). Regardless, *PtchB* is still downstream of Shh in lamprey because Purmorphamine treated embryos have increased expression (Figure 3.6 E, F, H, I).
In chick, Wnt1 and Wnt3a are sufficient to dorsalise *Pax6* and *Pax7* expression at the expense of intermediate and ventral cell type markers like, *Dbx1, Dbx2, Nkx2.2, Olig2, FoxA2, Nkx6.2,* and *Nkx6.1* (Alvarez-Medina et al., 2008, Alrefaei et al., 2020). In zebrafish, when Wnt is inhibited with Dkk1, dorsal *Pax3* expression is reduced with a dorsal shift of *Dbx2* expression while *Nkx6.2* remains unchanged (Bonner et al., 2008). The authors also found that disruption in proliferation is not enough to induce patterning defects. This subject of spinal cord size and scale regulation is reviewed more recently by Kuzmicz-Kowalska and Kicheva (2020). In my data, when Wnt is inhibited by C59, the genes I studied remain unchanged DV-wise (noting that some of the loss of expression in the posterior most region is likely caused by tailbud developmental failure), which might suggest that Wnt behaves differently in different vertebrates. It is also possible that C59 inhibits the release of all Wnts, all three Wnt response pathways (canonical, PCP, Calcium dependent) are inhibited so the results

would be different from the zebrafish experiments which only inhibited canonical Wnt signalling, or chick in which only *Wnt1* and *Wnt3a* were overexpressed.

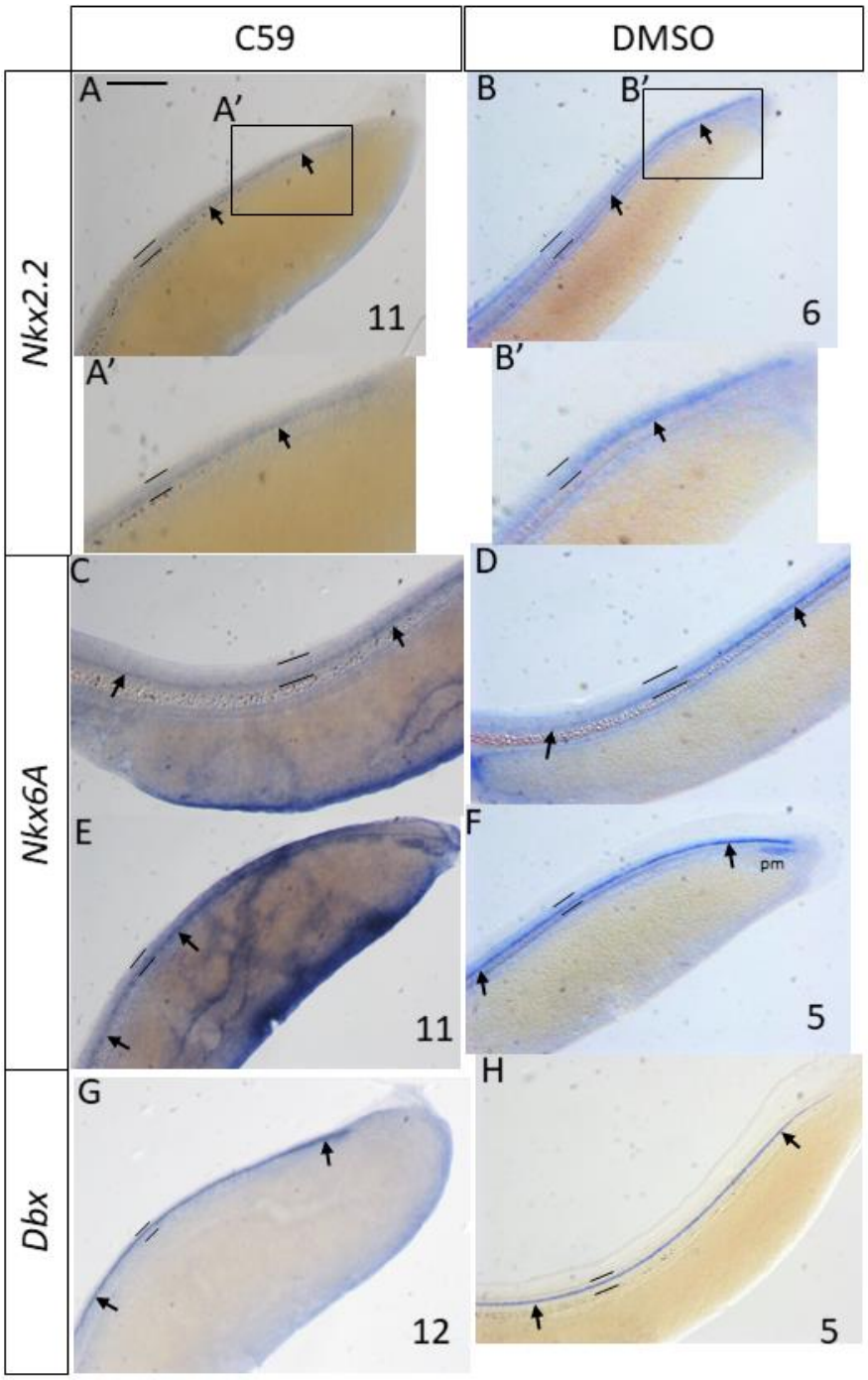


Figure 3.13. Effects of Wnt inhibition on *Nkx2.2, Nkx6A,* and *Dbx* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. C59 inhibits all Wnt pathway signalling, including PCP, calcium-mediated and canonical, while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The top left black bar on the *Nkx2.2* expression of C59 treated embryo photo is the scale bar for 200µm.

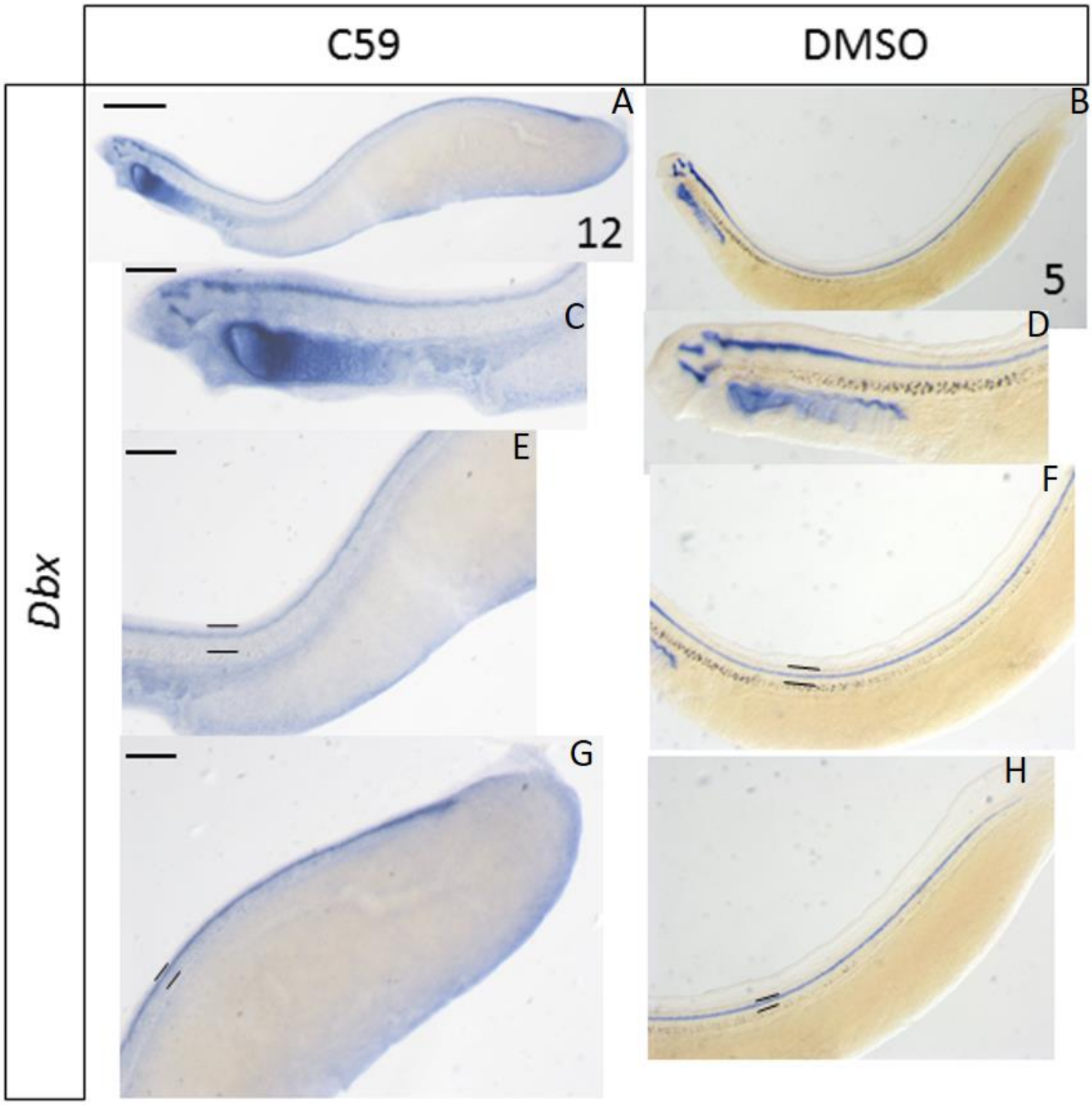


Figure 3.13.1. Effects of Wnt inhibition on *Dbx* expression. Lateral views of T26 lamprey embryos. The N number is at the bottom right of each image. C59 inhibits all Wnt pathway signalling, including PCP, calcium-mediated and canonical, while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. Scale bars: 500µm (whole embryo); 200µm (zoomed-in photos).

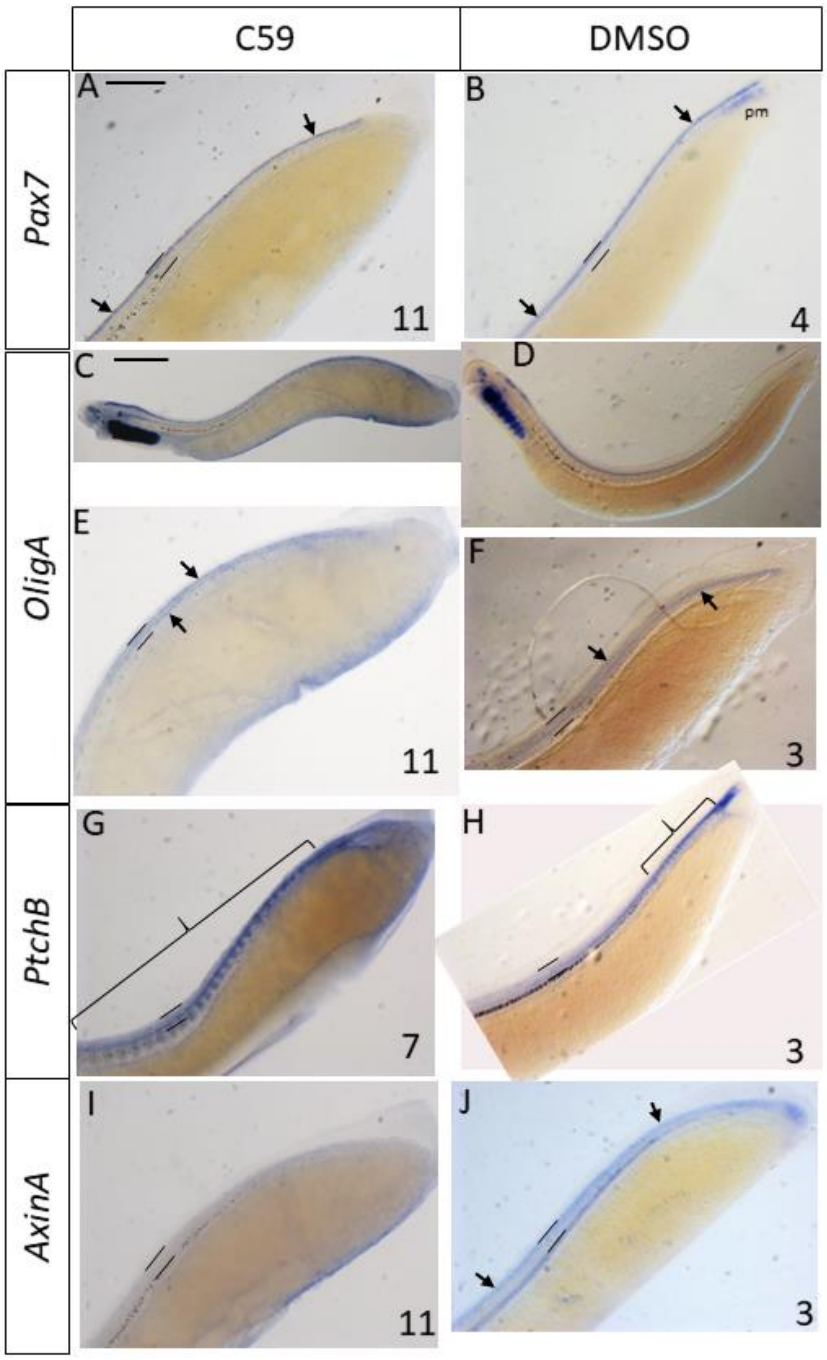


Figure 3.14. Effects of Wnt inhibition on *Pax7, OligA, PtchB,* and *AxinA* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. C59 inhibits all Wnt pathway signalling, including PCP, calcium-mediated and canonical, while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The black brackets highlight the extent of *PtchB* somite expression. The top left black bar on the *Pax7* expression of C59 treated embryo photo is the scale bar for 200µm for the embryo. The top left black bar on the *OligA* expression of C59 treated embryo photo is the scale bar for 500µm.

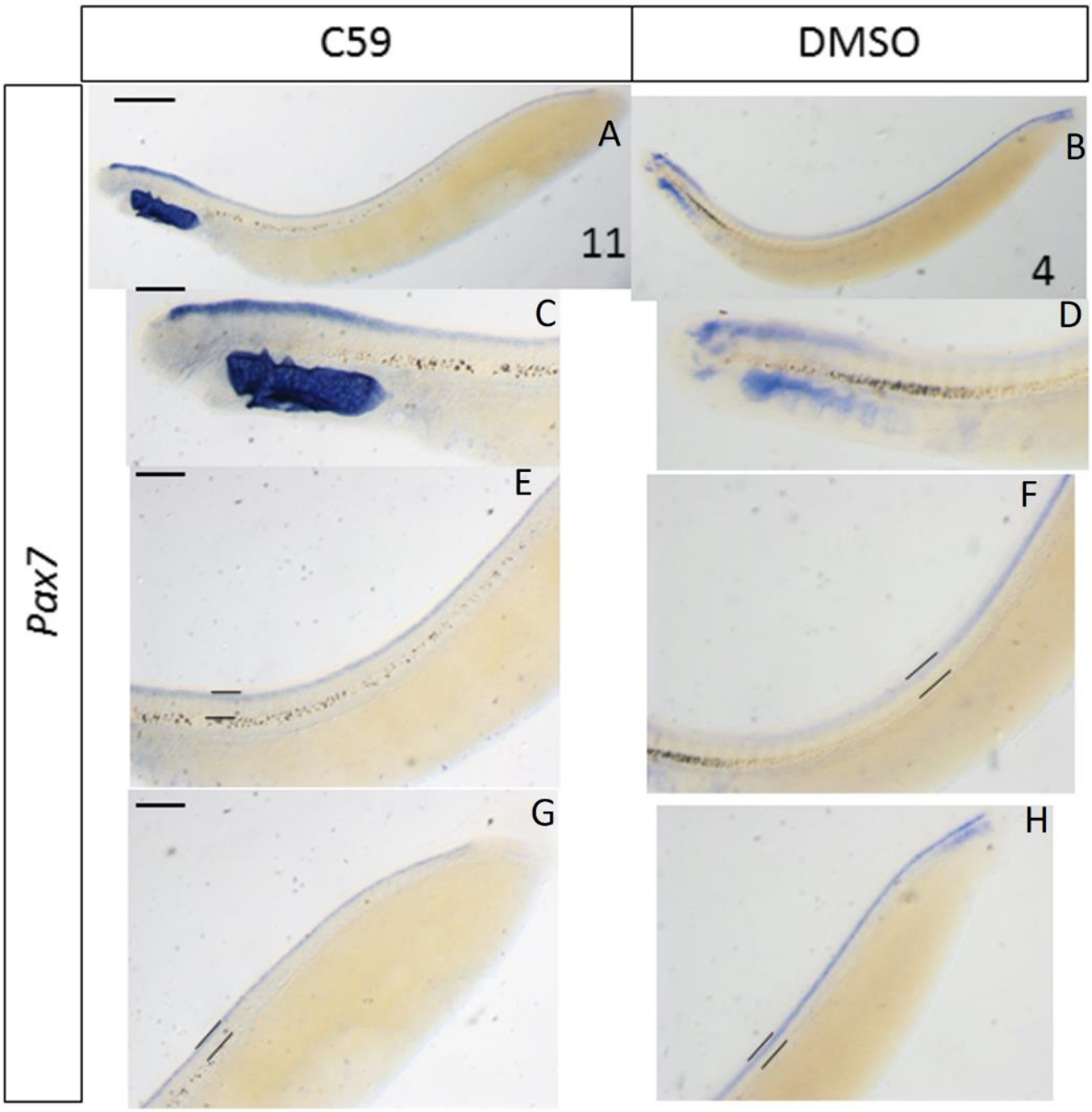


Figure 3.14.1. Effects of Wnt inhibition on *Pax7* expression. Lateral views of T26 lamprey embryos. The N number is at the bottom right of each image. C59 inhibits all Wnt pathway signalling, including PCP, calcium-mediated and canonical, while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. Scale bars: 500µm (whole embryo); 200µm (zoomed-in photos).

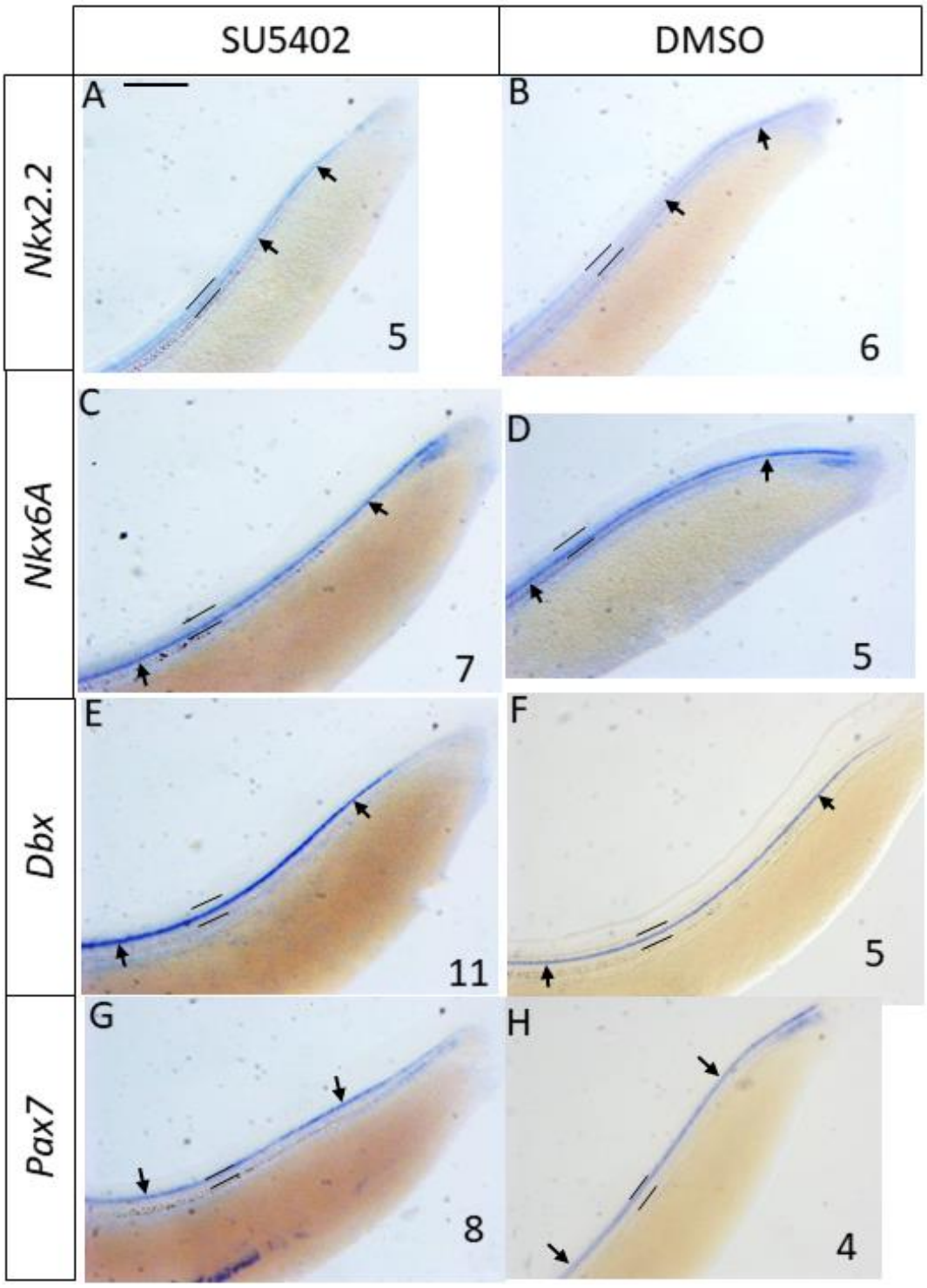


Figure 3.15. Effects of Fgf inhibition on *Nkx2.2, Nkx6A, Dbx,* and *Pax7* expression. Lateral views of T26 lamprey embryo tails. The N number is at the bottom right of each image. SU5402 inhibits the Fgf signalling pathway and DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The top left black bar on the *Nkx2.2* expression of SU5402 treated embryo photo is the scale bar for 200µm.

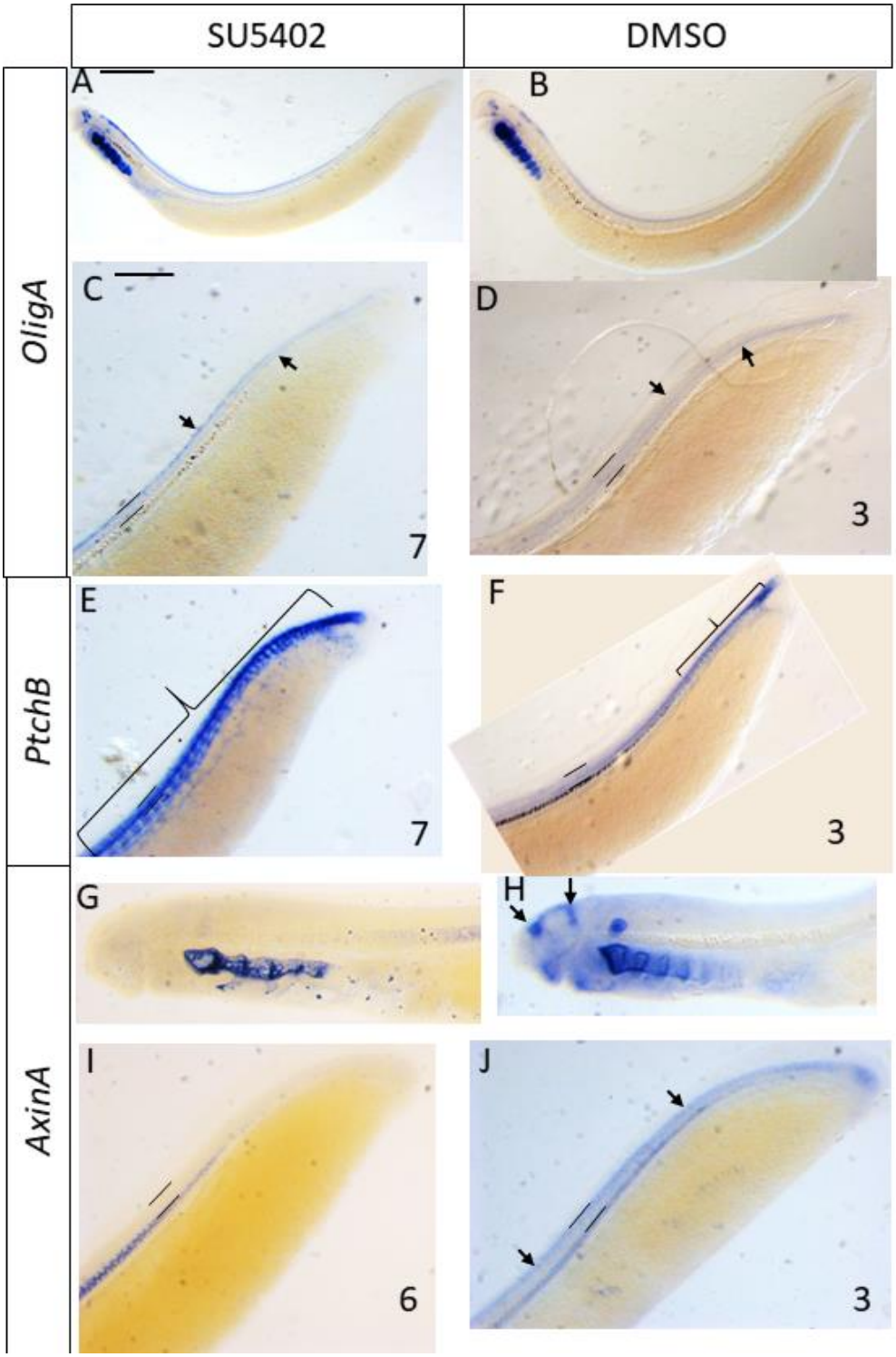


Figure 3.16. Effects of Fgf inhibition on *OligA, PtchB,* and *AxinA* expression. Lateral views of T26 lamprey embryos tails. The N number is at the bottom right of each image. SU5402 inhibits the Fgf signalling pathway and DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. The black arrows point out the gene expression. The black brackets highlight the extent of *PtchB* somite expression. The top left black bar on the *OligA* expression photo is the scale bar for 200µm.

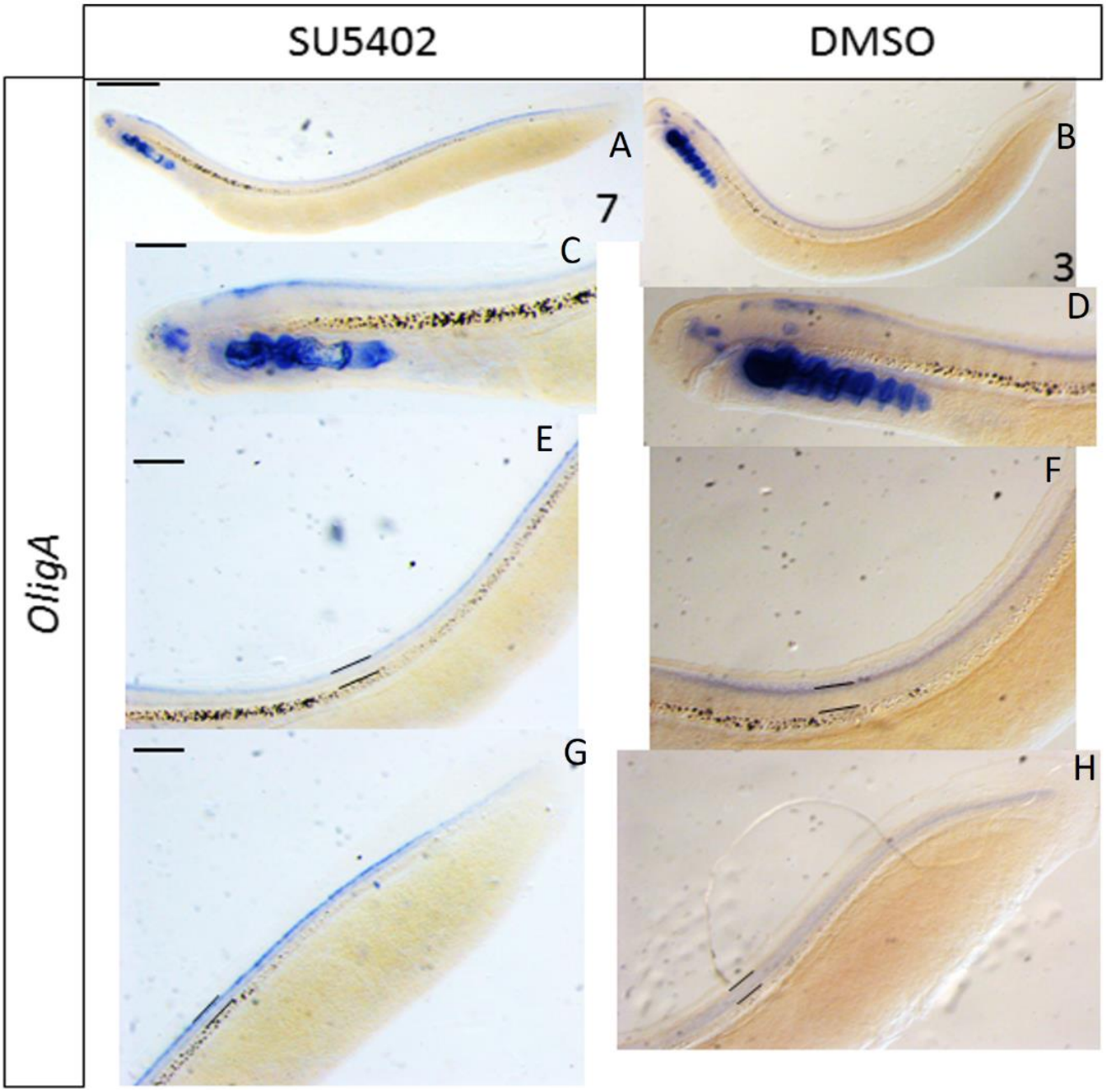


Figure 3.16.1. Effects of Fgf inhibition on *OligA* expression. Lateral views of T26 lamprey embryos. The N number is at the bottom right of each image. SU5402 inhibits Fgf pathway signallingl, while DMSO treated embryos act as controls. The pairs of parallel black lines demarcate the dorsal-ventral boundaries of the spinal cord. Scale bars: 500µm (whole embryo); 200µm (zoomed-in photos).

Fgf inhibition in lamprey has no effect on *Nkx2.2, Nkx6A, Dbx,* and *Pax7* (Figure 3.15 A-H). *PtchB* and *AxinA* expression changes had been discussed together with Wnt data in the above. Inhibition of chick Fgf signalling significantly increases the number of cells expressing *Nkx2.2* and *Olig2* while *Nkx6.1* remains the same (Morales et al., 2016, Sasai et al., 2014). The changes caused by electroporating a dominant negative

FgfR into one side of the neural tube are rather subtle and require comparison to the internal contralateral control side of the neural tube to measure the change in cell numbers. Such experiments cannot be repeated in lamprey for technical reasons and the current pharmacological inhibition study affects the whole embryo instead of just one side of the neural tube so subtle changes to cell number would not be picked up without making detailed cell counts across many embryos. The way chick *Olig2* is regulated could be conserved in lamprey and explain the stronger dorsal tail spinal cord expression of lamprey *OligA* (Figure 3.16.1; Figure 3.16 C, D). Fgf inhibition has no effect on both lamprey *Nkx6A* and chick *Nkx6.1.* Mouse dorsal spinal cord explant data suggest that Fgf2 treatment induces expression of *Olig1, Olig2,* and *Nkx2.2* in a Hh independent way (Chandran et al., 2003). Another study in chick showed that co-overexpression of Fgf8 and Hh is sufficient to induce *Olig2* and *Nkx2.2* (Farreny et al., 2018). No evidence on effects of Fgf manipulation on *Dbx, Pax3,* and *Pax7* has been reported in gnathostomes. After comparing within vertebrates, the following sections compare between vertebrates and cephalochordates.

3.3.6 *Pax2/5/8* expression in amphioxus when Wnt, Hh, and BMP activity are manipulated

In gnathostomes, of the *Pax2/5/8* family only *Pax2* is expressed in the intermediate level of the spinal cord (dI2,4,p0-2), in the lateral zone with some of the neurons overlapping with *Engrailed* and *Evx* expression (Burrill et al., 1997). Most of the *Pax2* neurons are GABAergic and Glycinergic, with a small number of them being Glutamatergic (Batista and Lewis, 2008). *Pax2* is also found to be co-localised with *neuropeptide Y*, nociceptin/orphanin, dynorphin, and galanin (Huang et al., 2008). In *Wnt1* and *Wnt3a* double knockout mice, the number of cells expressing *Pax2* doubles in the dorsal spinal cord (Muroyama et al., 2002). In *Smo-/-* embryos in which the Shh pathway is inhibited, ventral cells show ectopic expression of *Pax2* (Wijgerde et al., 2002). Misexpression of BMP7 is sufficient to decrease *Pax2* in dI4 (Andrews et al., 2017). Whether amphioxus *Pax2/5/8* is patterning GABA-Glycin-Glutamat-ergic neurons is unknown, and it would need double in situ hybridisation and/or scRNA-seq to confirm this. In contrast, to gnathostome *Pax2*, amphioxus *Pax2/5/8* expression is lost when Wnt is up or down regulated and when the Hh pathway is blocked (Figure 3.17 A-H). Similar to gnathostome *Pax2*, amphioxus *Pax2/5/8* has reduced expression when there is an excess of BMP4 (Figure 3.17 C, D, I, J). There are no vertebrate studies on the direct effect of upregulation of Shh, downregulation of BMP, inhibition of Notch, excess RA, or inhibition of Fgf on *Pax2* spinal cord

expression, so evolutionary comparison here cannot be easily drawn.

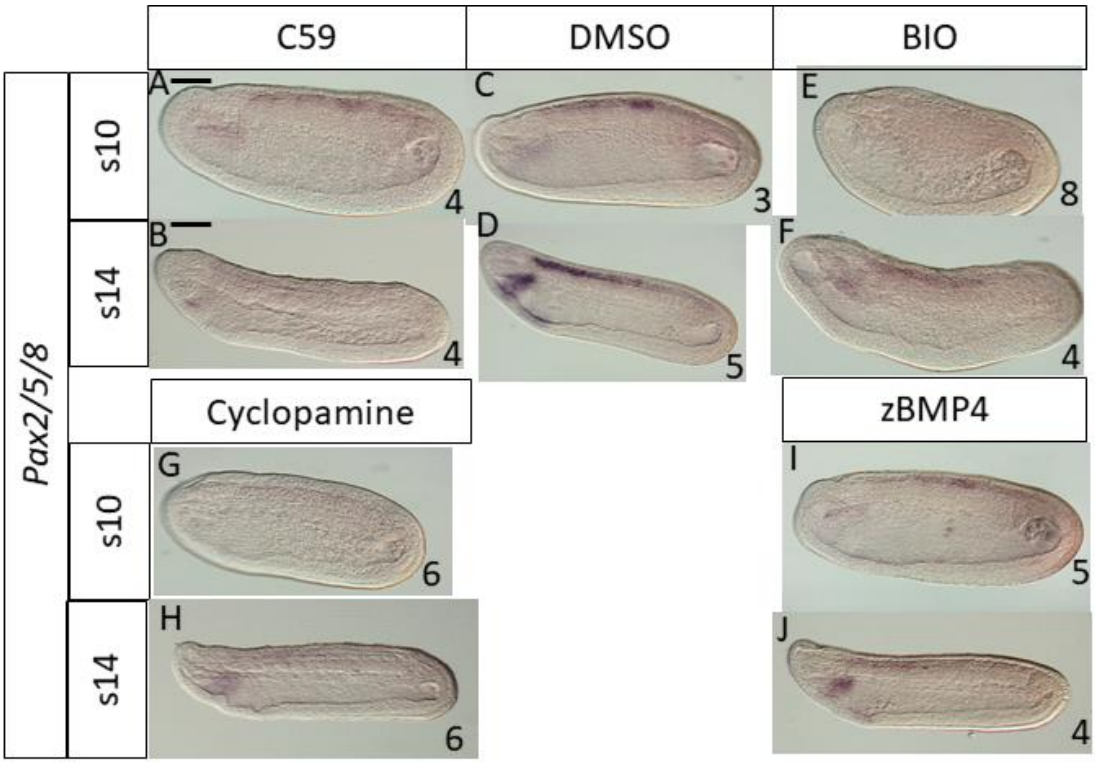


Figure 3.17. *Pax2/5/8* expression under the influence of signalling pathway manipulation. C59 inhibits all Wnt pathways. BIO upregulates the canonical Wnt pathway. Cyclopamine inhibits the Hh pathway. Exogenous zBMP4 overactivates the BMP pathway. DMSO treated embryos act as controls for all four treatments. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14). Scale bars: 50µm.

3.3.7 The impact of inhibiting Notch signalling on amphioxus spinal cord patterning

Phosphohistone 3 in amphioxus labels few and far apart cells in the neural tube/plate throughout development from the early neurula to early larval stage (Coppola et al., 2018, Bertrand and Escriva, 2011), which makes one wonder what signals are the dividing cells receiving to maintain progenitor identity. The Notch-delta pathway is maintaining stemness in the vertebrate CNS (Lara-Ramirez et al., 2019) and in amphioxus epidermal sensory neurons (Lu et al., 2012), but it remains dubious whether it is doing the same in the amphioxus central nervous system because *Delta* is only expressed in the presomitic mesoderm and a few cells in the neural plate (Rasmussen et al., 2007), whereas *Notch* is expressed throughout the neural tube/plate and tailbud (Holland et al., 2001a) which seems like a waste of *Notch* transcript in the neural tube if there is no delta ligand there to have cell-cell interaction. Nonetheless, my data suggest that the Notch pathway is likely to also be maintaining the progenitor cells in the CNS because there is a reduced expression of *Nkx6* (progenitor motor neuron marker; Figure3.18 E-H) and *Pax2/5/8* (potential progenitor inhibitory neuron marker; Figure 3.18 A-D), and an increase in expression of *Vacht* (differentiated motor neuron marker; Figure 3.18 I-L) and *Gbx* (a late stage motor neuron development marker in mice(Buckley et al., 2020); Figure 3.18 M-P), suggesting premature differentiation of neurons. There is a ligand-independent pathway for Notch signalling which involves endocytosing the Notch receptor and breaking it down but this has only been demonstrated in T cells (Steinbuck and Winandy, 2018). There is no evidence this mechanism is happening in neurons.

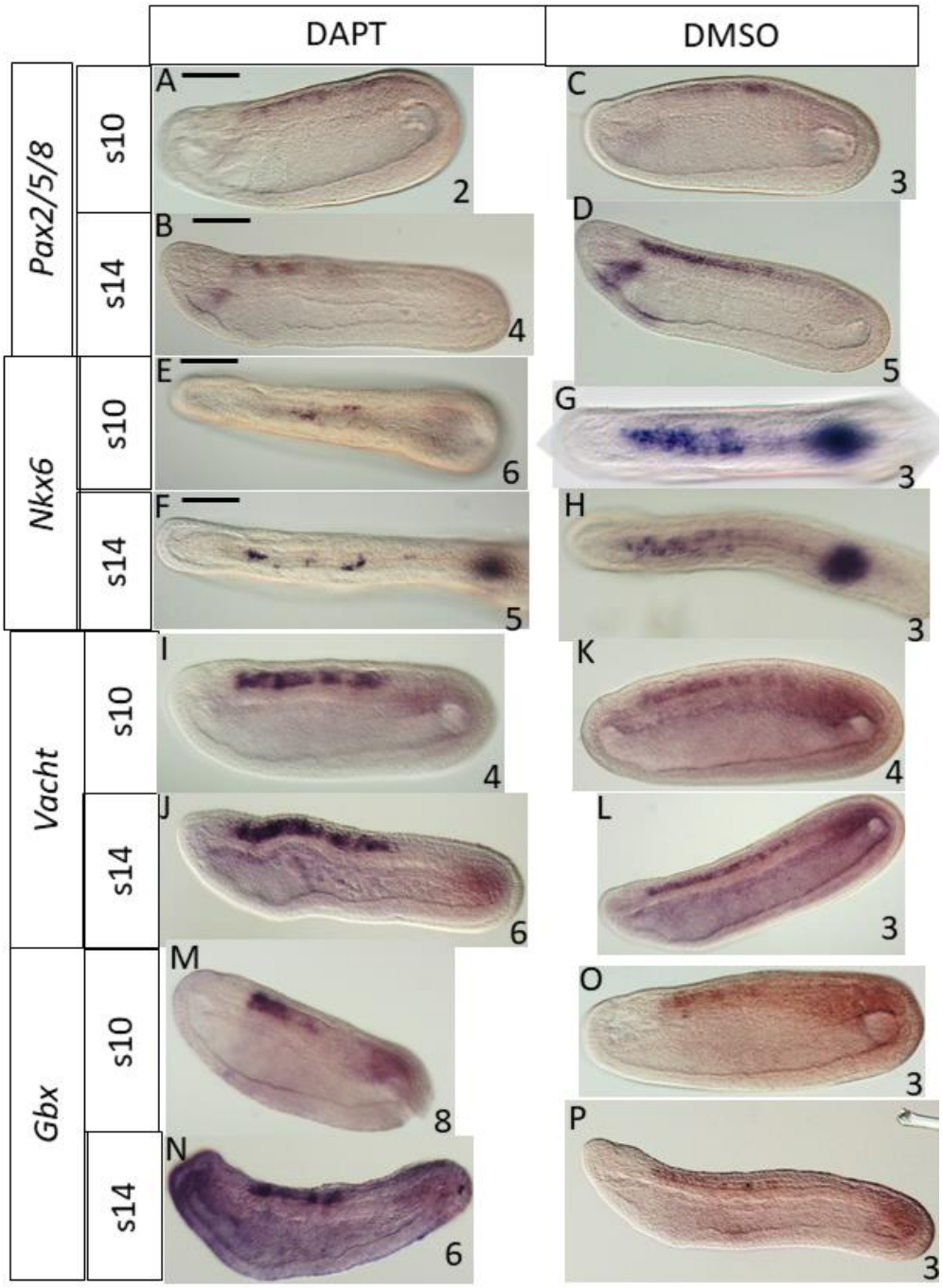


Figure 3.18. Effects of Notch inhibition on *pax2/5/8, Nkx6, Vacht,* and *Gbx* expression. DAPT inhibits the notch signalling. DMSO acts as the control. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14). *Nkx6* is shown in dorsal view. *Nkx6* DMSO-treated photo (G) is produced by collaborator Qiongqiong Ren, all other photos are produced by me. The rest of the images are in lateral views. Scale bars: 50μm.

3.3.8 The impact of Hh signalling inhibition on amphioxus spinal cord patterning

Irx3 is sufficient to induce interneurons at the expense of motor neurons in the early gnathostome spinal cord (Briscoe et al., 2000, Lewis et al., 2005). At a later developmental stage, *Irx3, Irx5,* and *Irx6* are expressed in the lateral motor neurons in addition to the interneurons at the ventricular zone (Houweling et al., 2001). So it seems that at the progenitor stage Irx represses motor neurons but in the differentiated stage *Irx* is expressed in motor neurons. In amphioxus, expression of all three paralogues, *IrxA, IrxB,* and *IrxC,* are not affected by *Hedgehog* knockout or *Patched* knockout in mid-neurula stage embryos (Ren et al., 2020), which shows that Irx is not downstream of the Hh pathway. However, in my data, *IrxA* expression is lost when embryos are treated with a Hh inhibitor which suggest that Irx is downstream of Hh at least from late neurula to early larval stages (Figure 3.19 A-D). In addition to the developmental timeframe difference, *Hh* knockouts deprive the embryo of Hh signal throughout its development, whereas my drug treatment experiment is targeting a specific developmental timeframe when spinal cord is developing. Moreover, as established before, Ptch could also have other ligands activating its pathway while cyclopamine is directly inhibiting Smo, since Smo is downstream of Ptch, cyclopamine treatment would give a more complete block of the signalling pathway. Irx in amphioxus does not seem to have cross-repressive activity with motor neuron identity because there is no gain in *Vacht* or *Nkx6* expression in cyclopamine treated embryos when *Irx* expression is lost (Figure 3.19 E-L). In fact, *Nkx6* expression is also completely lost when Hh is inhibited (Figure 3.19 E-H). Whether Irx is labelling interneurons or motor neurons, or a combination of both, is unknown.

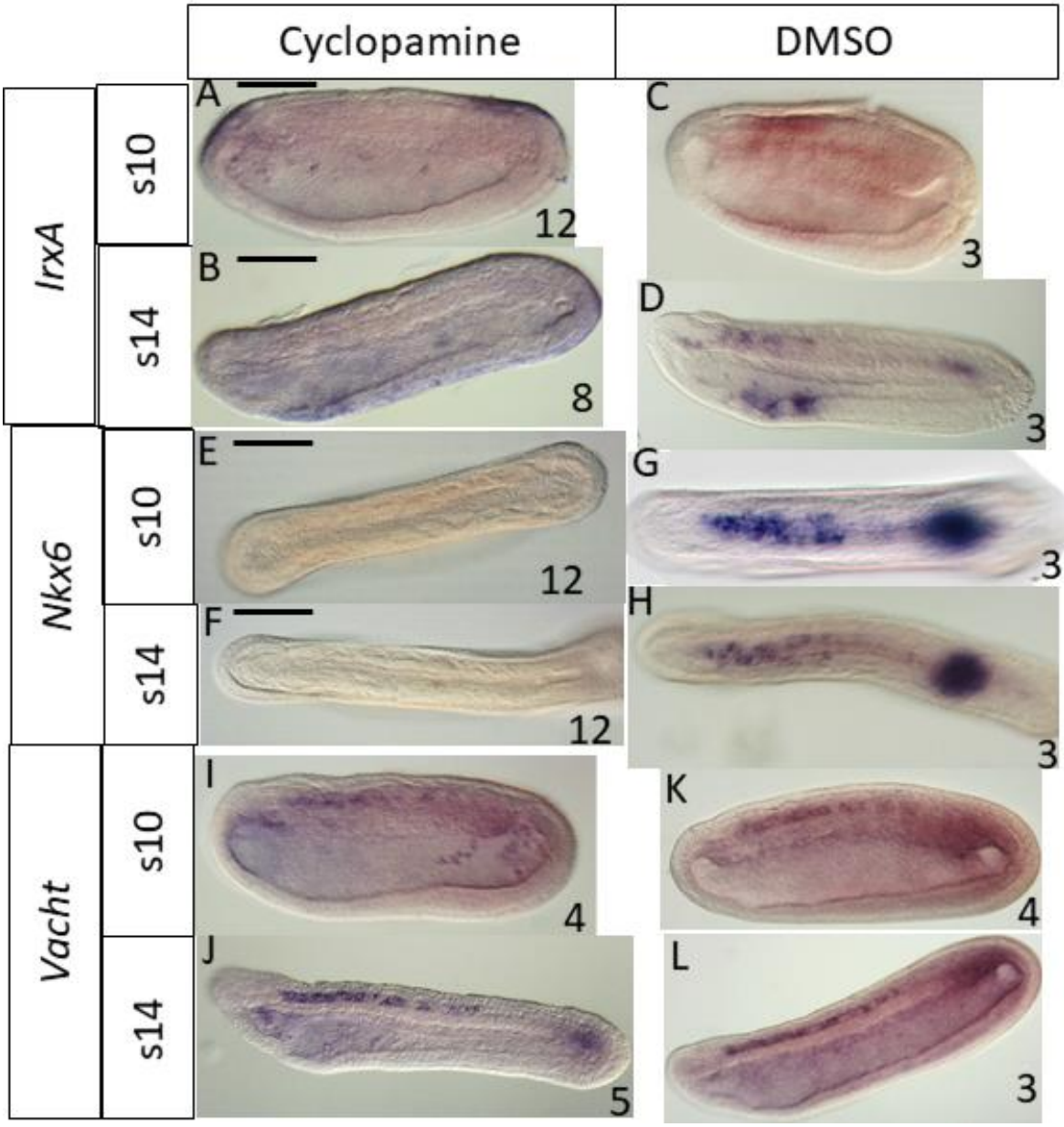


Figure 3.19. Effects of Hh inhibition on *IrxA, Nkx6,* and *Vacht* expression. Cyclopamine inhibits the Hh pathway while DMSO acts as the control. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14). *Nkx6* s10 is shown in dorsal view, the rest are lateral view. *Nkx6* DMSO-treated photo (G) is produced by collaborator Qiongqiong Ren, all other photos are produced by me. Scale bars: 50µm.

3.3.9 The impact of Wnt manipulation on amphioxus spinal cord patterning

C59 acts on Porcupine which prevents the release of all Wnts while BIO is an inhibitor of GSK3, which only upregulates the canonical Wnt pathway. *Gsk3β* is expressed ubiquitously in the mesendoderm of amphioxus throughout the neurula stage (Jing et al., 2016), which is in rather stark contrast to Wnt family expression which is specifically in tailbud and/or somites, and neural tube (Somorjai et al., 2018). It is possible that cells birthed from the tail bud experience Wnt signal briefly and maintain its downstream signal GSK3 throughout neurula to early larva in mesendoderm. From this understanding, my timeframe of BIO treatment would cause a ubiquitous upregulation of the canonical Wnt pathway in all of the mesendoderm, potentially misdirecting the cell identity such that they are closer to a tailbud phenotype (progenitor pool cell type) than they should be. This may result in failure of convergent extension during neurulation which is depicted in my BIO treated embryo morphology with them having a shorter body length, and ectopic *Netrin* expression in endoderm (Figure 3.20 A-B asterisks). In addition to morphological changes, my data also suggest that the positional identity in the neural tube is also dependent on Wnt pathway activity of surrounding tissues. More notable reduction in expression in the neural tube includes *Pax2/5/8* (Figure 3.17 C-F)*, Nkx6, IrxA, Gbx, Vacht* (Figure 3.21 C-F, I-L, O-R, U-X) expression in BIO treated embryos. Onai et al. (2009) have described similar posteriorisation/loss of expression of amphioxus neural tube markers like *Lim1/5* and *Otx* when Wnt is upregulated by Lithium chloride. Canonical Wnt signalling is also used for anterioposterior patterning in sea urchin, hemichordate, and tunicate (Darras et al., 2018, Khadka et al., 2018, Feinberg et al., 2019).

If upregulation of Wnt causes posteriorisation, then inhibition of the Wnt pathway should cause anteriorisation. Contrarily, C59-treated embryos do not show anteriorisation (Figure 3.21 A-D, G-J, M-P, S-V). There is a complete loss of expression for *Pax2/5/8* (Figure 3.17 A-D) and *Nkx6* (Figure 3.21 A-D). I would predict that there would be some loss of *Vacht* neurons in C59-treated embryos since *Nkx6* is upstream of *Vacht* in patterning a subset of motor neurons (Ren et al., 2020). However, that does not seem to be the case as *Vacht* expression in the neural tube remains the same at stage s14 regardless of the elongation and developmental malformation which are more noticeable in stage s10 (Figure 3.21 G, H). More experiments are needed to determine the role of *Nkx6* in the specification of amphioxus motor neurons. Refering back to C59-treated embryos not showing anteriorisation, a possible explanation would be that C59 is blocking all canonical and non-canonical

pathways of Wnt and some of the Wnts were used in neural cell type induction, like *Wnt2*, *Wnt3, Wnt4, Wnt5, Wnt6, Wnt7* and *Wnt8* which are expressed in particular neurons (Somorjai et al., 2018), instead of tailbud progenitor pool maintenance/ anterioposterior identity. In a way, this response to Wnt manipulation is similar to the tunicate which also posteriorises the CNS when Wnt is upregulated, but inhibition just led to loss of *Six1/2, Six3/6,* and *Hox12* expression without anteriorising neural markers, *Mytf, Gbe1,* and *Hox1* (Feinberg et al., 2019).

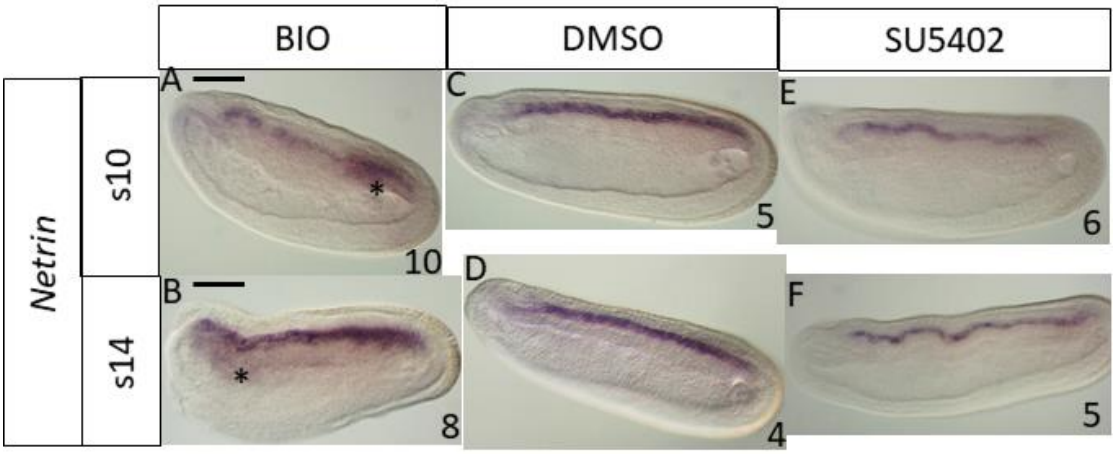


Figure 3.20. Convergent extension failure during neurulation as indicated by floorplate/notochord marker, *netrin*. BIO upregulates the canonical Wnt pathway. SU5402 inhibits the Fgf signalling. DMSO acts as the control. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14).

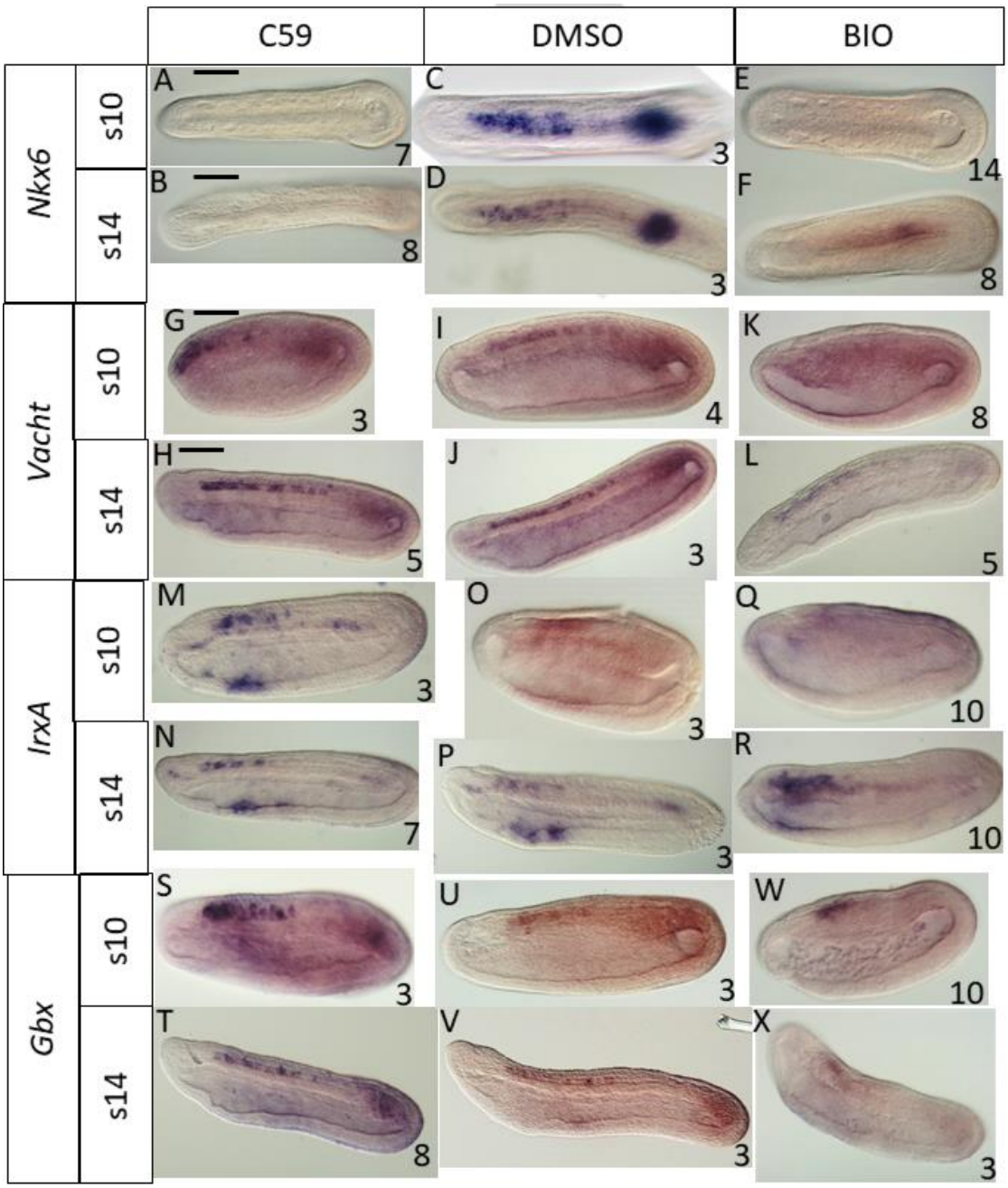


Figure 3.21. Effects of Wnt manipulation on *Nkx6, Vacht, IrxA,* and *Gbx* expression. C59 inhibits all Wnt pathways, including PCP, calcium-dependent, and canonical. BIO upregulates the canonical Wnt pathway. DMSO acts as the control. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14). Nkx6 is shown in dorsal view. Nkx6 DMSO-treated photo (C) is produced by collaborator Qiongqiong Ren, all other photos are produced by me. The rest of the images are in lateral views. Scale bars: 50µm.

3.3.10 The impact of excess RA on motor neuron markers

Several studies have confirmed the role of retinoic acid in anterior-posterior axis patterning in amphioxus branchial arches (Koop et al., 2014, Onai et al., 2009), neural tube (Zieger et al., 2018a), and epidermal sensory neurons (Zieger et al., 2018b, Schubert et al., 2004). Generally, *Hox* cluster expression is posteriorised by excess RA and anteriorised when RA is inhibited (Schubert et al., 2006, Marlétaz et al., 2006), with a less straight-forward response in *Parahox* genes (Osborne et al., 2009). RA is also self-regulated in amphioxus, as all three *cyp26* paralogues are upregulated when there is an excess of RA and down-regulated when the RA pathway is inhibited (Carvalho et al., 2017). Also *RAR* expression is posteriorised when RA is upregulated and anteriorised when RA is inhibited (Escriva et al., 2002a). My data suggest that RA is regulating motor neurons because of partial loss of *Nkx6* and *Vacht* neurons when there is an excess of RA, but not in the way that the neural tube is posteriorised because the pattern of *Vacht* and *Nkx6* expression loss is selective, with a few neurons lost without changing their expression boundaries (Figure 3.22).

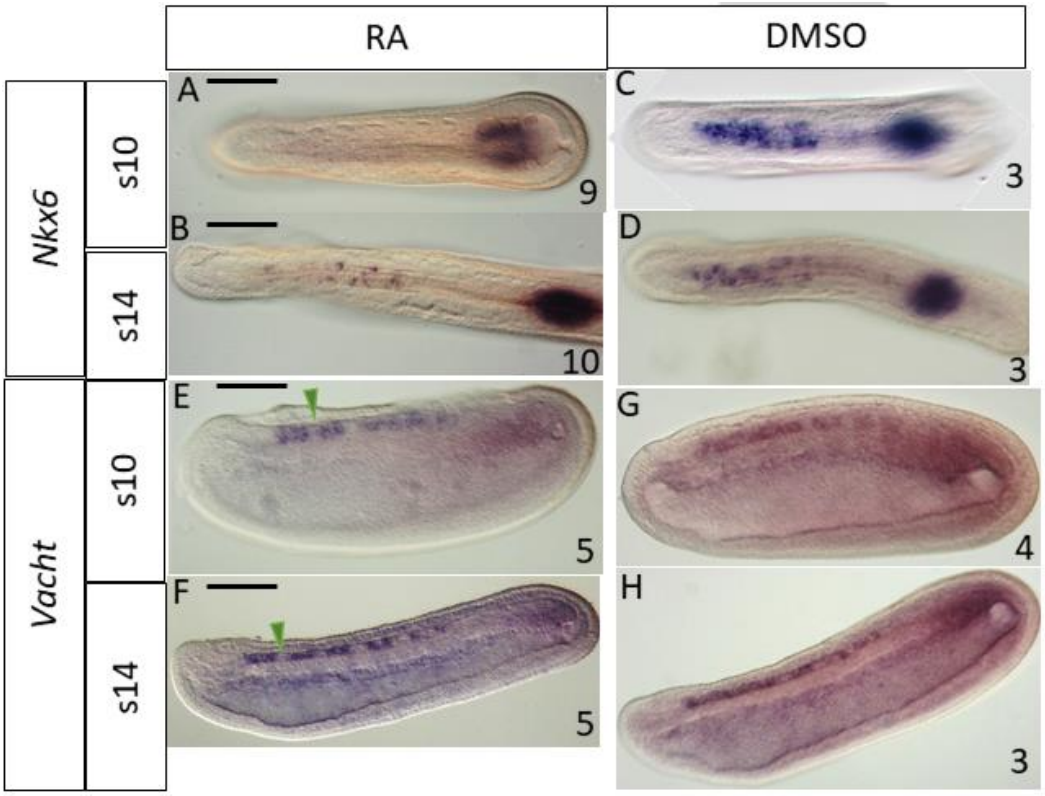


Figure 3.22. Effects of RA upregulation on *Nkx6* and *Vacht* expression. *Nkx6* at 14-somite stage is shown in dorsal view. The rest of the images are in lateral views. Exogenous retinoic acid overactivates the RA signalling pathway. DMSO-treated embryos act as controls. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14). Green arrowheads are pointing out the missing *Vacht* expression in the RA-treated embryos. *Nkx6* is shown in dorsal view. *Nkx6* DMSO-treated photo (C) was produced by collaborator Qiongqiong Ren, all other photos were produced by me. The rest of the images are in lateral views. Scale bars: 50µm.

3.3.11 The impact of Fgf inhibition on amphioxus spinal cord patterning

Bertrand et al. (2015) and Beaster-Jones et al. (2008) concluded that the RA and Fgf "clock and wavefront" only exists in the vertebrate lineage and not in cephalochordates. The fact that *Fgfr* is mainly expressed in mesendoderm (Bertrand et al., 2011) instead of the neural tube during neurulation could explain why there is no observable change in neural tube expression for *Vacht* and *IrxA* (Figure 3.23 E-L) while *Netrin* expression in floorplate remains with notochord expression gone (Figure 3.20 C-F). On the other hand, amphioxus Fgf could potentially have a complementary role to Notch signalling for maintaining progenitor cells in the neural tube because my *Nkx6* and *Gbx* data at stage s10 seems to recapitulate my Notch inhibition data, with partial loss of *Nkx6* expression and an increase in *Gbx* expression (Figure 3.23 A, C, M, O). However, such gene expression changes are not observed in the older Fgf inhibited s14 stage embryos. A potential explanation could be that *Nkx6* and *Gbx* are regulated by other signals in the older stage since *Fgf9/16/20* is expressed segmentally along the AP axis of the neural tube during mid-late neurula stage, but in pre-oral stage, *Fgf9/16/20* expression is restricted to the posterior neural tube. The other four Fgf paralogues are not expressed in the amphioxus spinal cord during mid-late neurula stage (Bertrand et al., 2011). Both the lack of "clock and wavefront" and changes in neural tube expression in older stage embryos would suggest that Fgf is not involved in anterior-posterior axis patterning in amphioxus. The amphioxus *Fgf* expressed in CNS neurons and neuroenteric canal could be giving positional cues to mesendoderm, which would explain the kinks in the three germ layers of the SU5402 treated embryos, most obvious in *Netrin* in situ hybridisation (Figure 3.20 C-F). It suggests that there could be a mismatch in convergence extension among the three germ layers during neurulation.

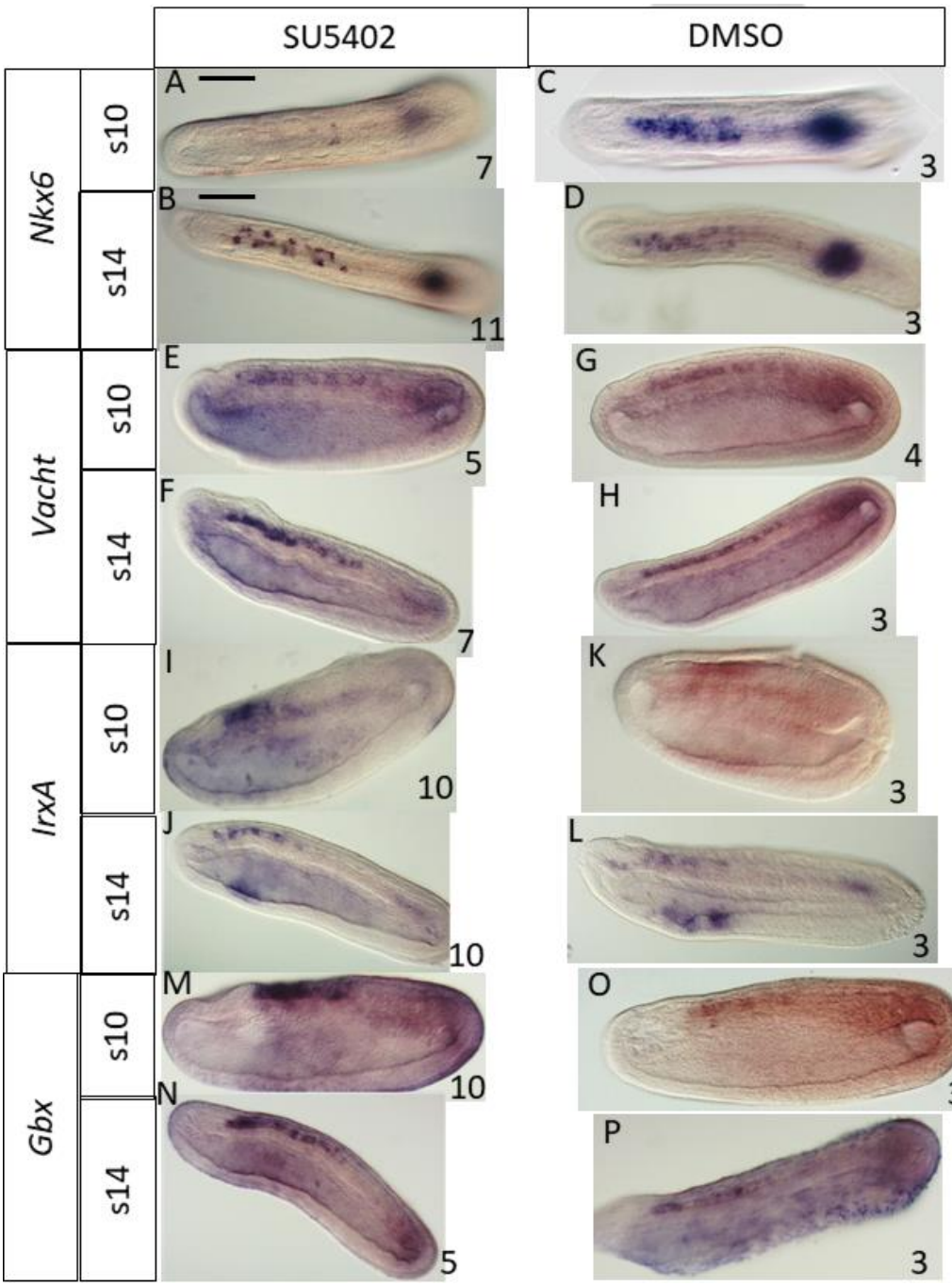


Figure 3.23. Effects of Fgf inhibition on *Nkx6, Vacht, IrxA,* and *Gbx* expression. *Nkx6* at 14-somite stage is shown in dorsal view. The rest of the images are in lateral views. SU5402 inhibits the Fgf pathway. DMSO treated embryos act as controls. The N number is at the bottom right of each image. Amphioxus embryo stages are indicated on the left as ten-somite stage (s10) or 14-somite stage (s14). *Nkx6* is shown in dorsal view. *Nkx6* DMSO-treated photo (C) is produced by collaborator Qiongqiong Ren, all other photos are produced by me. The rest of the images are in lateral views. Scale bars: 50μm.

3.4 Summary, conclusions and future direction

Table 3.2. Summary of gene expression changes under signal manipulation. First row is showing the up (↑) and down (↓) regulation of various signalling pathways. The table starts with seven lamprey genes followed by six amphioxus genes. ↑ means increased in expression or ectopic expression. ↓ means loss of expression or reduced in expression. – means no observable change. 'na' means not applicable, no data was collected. The symbols in green mean the image data are in the Supplementary section.

Drug / Lamprey	**Shh↓**	**Shh↑**	**RA↑**	**RA↓**	**Wnt↓**	**Wnt↑**	**Fgf↓**	**Notch↓**	**BMP↓**	**BMP↑**
Nkx2.2	↓	↑	↓	↓	↓	na	-	↓	↑	na
Dbx	-	-	-	↓	-	na	-	↓	-	na
Pax7	-	↓	↓	↓	-	na	-	↓	-	na
OligA	-	-	-	↓	-	na	-	↓	↓	na
PtchB	↓	↑	↑	↓	↑	na	↑	↓	↑	na
AxinA	↓	↓	↓	↓	↓	na	↓	↓	↓	na
Nkx6A	-	-	↓	↓	-	na	-	↓	-	na
Amphioxus										
Nkx6	↓	↓	↓	na	↓	↓	↓	↓	↓	↓
Pax2/5/8	↓	-	-	na	↓	↓	-	↓	-	↓
Netrin	↓	-	-	na	-	↑	↓	↓	-	↓
Vacht	-	↓	↓	na	-	↓	-	↑	-	-
IrxA	↓	-	na	na	-	↓	-	-	-	-
Gbx	-	na	na	na	-	↓	↑	↑	-	-

General results are summarized in Table 3.2. Across the results and discussion section of this chapter, I have mentioned that the equivalent experiments of signal manipulation effects on the genes studied here have often not been done in gnathostomes, leading to difficulty in comparing to outgroups. So, even though spinal cord is a relatively well-studied organ in developmental biology, more data are still needed. Below I will first list from the discussion above what is conserved, different, or not known between these lineages. I will then try to integrate these into a broader discussion of what this does and does not tell us about the evolution of spinal cord patterning.

A number of aspects are conserved between lamprey and gnathostomes:

- Hh is sufficient to induce ectopic *Nkx2.2*.
- Hh overactivation causes reduced expression of *Pax3* and *Pax7.*

- In RA-deficient chick embryos and RA-inhibited lamprey embryos, *Pax3* and *Pax7* expression are lost.
- RA-deficient quails have weaker *Nkx6.2* and *Olig2* expression, when RA is inhibited in lamprey there is a loss of *Nkx6A* and *OligA* expression.
- Excess RA represses *Nkx2.2* in zebrafish and lamprey.
- DAPT-treated lamprey embryos have similar results to conditional *Notch* knockout mice, as both have reduced expression in *Dbx, Nkx2.2, Nkx6,* and *Olig.*
- Wnt inhibition has no effect on zebrafish *Nkx6.2* and lamprey *Nkx6A* expression.
- Fgf inhibition has no effect on chick *Nkx6.1* and lamprey *Nkx6A.*

There are also many differences between lamprey and gnathostomes:

- Hh inhibition causes loss of chick *Nkx6.1* expression while having no observable effect on lamprey *Nkx6A* expression.
- Hh inhibited chick, *Dbx1* and *Dbx2* expression shifted ventrally from p0/p1 to floorplate while lamprey *Dbx* remains unaffected.
- Shh is sufficient and required for *Olig2* while lamprey *OligA* not affected by Hh manipulation.
- Overactivation of BMP causes dorsalisation of *Pax3* and *Pax7* in chick while BMP inhibition in lamprey has no effect on *Pax7.*
- BMP inhibition affects only the dorsal chick *Olig3* population while lamprey *OligA* dorsal and ventral populations were lost.
- In chick RA-deficient embryos, *Nkx6.1* is ventralised while lamprey *Nkx6A* is lost.
- RA inhibition in zebrafish does not affect *dbx* expression while in lamprey, *Dbx* expression is lost.
- In chick explant, RA induces *Dbx1* and *Dbx2* expression while in lamprey, excess RA does nothing to *Dbx* expression.
- RA-deficient quails have higher *Nkx6.1* expression while RA-inhibited lamprey has lost its *Nkx6A* expression.
- Inhibition of RA in zebrafish ventralises *Nkx2.2* at the expense of Olig2 while both *Nkx2.2* and *OligA* expression are lost in lamprey.
- In chick, Wnt1 and Wnt3a are sufficient to dorsalise *Pax7* expression while in lamprey Wnt inhibition has no effect on *Pax7* expression.
- Wnt-inhibited zebrafish have reduced *Pax3* and a dorsal shift of *Dbx2* while Wnt-inhibited lamprey has no change in *Pax7* and *Dbx.*

- Fgf inhibition increases chick *Nkx2.2* and *Olig2* expression while lamprey *Nkx2.2* and *OligA* remain the same.

Several important points of comparison between lamprey and gnathostomes remain unknown:

- Does lamprey *Dbx* cross-repress with *Nkx6A*, as it does in gnathostomes?
- Is gnathostome *Axin* expression affected by Hh or BMP signalling?
- Does gnathostome *Ptch* expression posteriorise in both up and down regulation of Hh, RA, and down regulation of BMP, Notch, Fgf?
- How does BMP manipulation affect *Dbx* and *Nkx6* expression in gnathostomes?
- Is RA required for gnathostome *Ptch* expression, like in lamprey?
- Does Notch manipulation have an effect on gnathostome *Ptch, Pax3* and *Pax7?*
- Is gnathostome *Dbx, Pax3* and *Pax7* affected by Fgf manipulation?

A small number of aspects are conserved between amphioxus and vertebrates:

- Excess BMP in chick and amphioxus is sufficient to reduce *Pax2* and *Pax2/5/8* expression.
- Notch inhibition causes premature reduced of progenitor marker and increase in differentiation marker of motor neurons.
- Up and down regulation of RA affect motor neuron development in different ways depending on the species.

There are also differences between amphioxus and vertebrates:

- In *Wnt1* and *Wnt3a* double knockout mice, the number of cells expressing *Pax2* doubles in the dorsal spinal cord while in amphioxus, Wnt up and down regulated embryos have lost expression of *Pax2/5/8.*
- When Hh pathway is inhibited in mice, there is a ventral ectopic expression of *Pax2* while in Hh-inhibited amphioxus there is a loss of *Pax2/5/8* expression.
- When Hh is inhibited in chick and zebrafish, *Irx3* expression expands ventrally while in amphioxus, *IrxA* expression is completely lost.
- *Irx3* represses motor neurons in chick but not in amphioxus.
- Wnt is mostly involved in AP patterning in amphioxus while Wnt is also involved in DV patterning in vertebrates.
- BMP(9) is sufficient to induce expression of *Vacht* in mice but not in amphioxus.

Finally, important aspects of comparison between amphioxus and vertebrates remain unknown, including:

- There are no vertebrate studies on the direct effect of upregulation of Shh, downregulation of BMP, inhibition of Notch, excess RA, or inhibition of Fgf on *Pax2* spinal cord expression, so evolutionary comparison here cannot be easily drawn.

Patterning by signals seems to be conserved between gnathostomes and lamprey. Hh induces the ventral spinal cord marker *Nkx2.2* (Figure 3.5B,C; (Soula et al., 2001, Schäfer et al., 2005)), while inhibiting dorsal markers, *Pax3* and *Pax7* (Figure 3.5 K, L; (Lin et al., 2016)). This seems to suggest Hh patterning is conserved between lamprey and gnathostomes. However, the lineages also have their differences. Hh manipulation has no effect on *Nkx6A, Dbx* and *OligA* in lamprey while their expression is affected in gnathostomes. In amphioxus, Hh inhibition led to loss of *Nkx6* expression (Figure 3.19 E-H) which is similar to gnathostomes.

From my data, lamprey BMP patterning seems to have nothing in common with gnathostomes. Both dorsal and ventral *OligA* cell populations were lost in lamprey when BMP was inhibited (Figure 3.8 A-F), with the ventral marker *Nkx2.2* being ectopically expressed (Figure 3.7 A-B) while dorsal and intermediate markers, *Pax7* and *Dbx,* were not affected (Figure 3.7 E-H). Whereas, in gnathostomes, there is only evidence for dorsal and intermediate spinal cord patterning effects (Zechner et al., 2007, Timmer et al., 2002, Andrews et al., 2017). Both Hh and BMP signal manipulation data suggests that transcription factor cross-repression is not observed in lamprey because the ectopically expressed *Nkx2.2* in both cases overlaps with the intermediate and dorsal markers. In gnathostomes, those markers would shift their expression along the DV axis accordingly.

Regarding RA patterning of the spinal cord, each vertebrate lineage seems to be different. *Dbx1* and *Dbx2* are induced by RA in chick, *dbx* in zebrafish is not affected by RA inhibition, *Dbx* in lamprey is lost on RA inhibition but not affected by RA overactivation ((Pierani et al., 1999, Gribble et al., 2007); Figure 3.9 G-I). RA deficient chick and quail have increased *Nkx6.1* expression but weaker *Nkx6.2* expression (Novitch et al., 2003, Wilson et al., 2004) while inhibition or overactivation of RA led to loss of *Nkx6A* expression in lamprey (Figure 3.9 D-E). RA inhibition in amphioxus also led to reduced *Nkx6* expression (Figure 3.22 A-D). In lamprey, both up and down

regulation of RA leads to loss of *Nkx2.2* expression (Figure 3.9 A-C) while in zebrafish, *Nkx2.2* expression is ventralised when RA is inhibited (Schäfer et al., 2005).

Notch signalling had been known to be involved in stem cell maintenance of the spinal cord and is active in mitotically active cells in chick, zebrafish, and lamprey (Akai et al., 2005, Kim et al., 2008, Lara-Ramirez et al., 2019). My lamprey data provide more evidence to support that Notch signalling is involved in stem cell maintenance in spinal cord because the genes I tested are progenitor cell markers and they all lost or reduced their expression when Notch is inhibited (Figure 3.11,12). *Dbx, Pax7,* and *OligA* remain at the very posterior end of the tail spinal cord in the Notch inhibited lamprey embryos, suggesting that other signals are involved in the induction of those markers (Figure 3.11E-H; 3.12A-F). Treating embryos with DAPT at an earlier developmental stage can test whether Notch is involved in induction of these gene markers. My amphioxus data also suggest that Notch is also involved in maintaining stem cells in the cephalochordate lineage because when Notch is inhibited, progenitor motor neuron marker expression is reduced while differentiated motor neuron markers increased in expression (Figure 3.18 E-L). Taking the Notch inhibition data together, Notch could have functioned in stem cell maintenance in spinal cord in the chordate ancestor and diverged to function in the ventricular progenitor zone in the vertebrate ancestor.

There is no direct evidence demonstrating a Wnt signalling relationship with *Nkx2.2.* There is evidence in mouse and chick that canonical Wnt signalling represses the Shh pathway in the ventral spinal cord (Alvarez-Medina et al., 2008, Lei et al., 2006) which is contrary to the reduced *Nkx2.2* expression in Wnt inhibited lamprey embryos (Figure 3.13 A, B) because Shh induces *Nkx2.2* and if Shh is no longer repressed then I would expect an increase in *Nkx2.2* expression. An explanation would be that in the mouse and chick studies, only canonical Wnts were studied while the C59 molecule I used inhibits all Wnts pathways. Future studies can use other pharmacological molecules which are specific to manipulating certain type of Wnt pathway. Regardless of the developmental failure of some dorsal neural tissue, the other markers do not seem to have shifted along the DV axis (Figure 3.13, 3.14) when pan-Wnt was inhibited, which concurs with the findings of Ikeya et al. (1997). Mouse and chick seem to have different Fgf2 patterning of spinal cord cells with mouse showing that Fgf induces *Olig1, Olig2,* and *Nkx2.2* while the Hh pathway was inhibited (Chandran et al., 2003). In the chick study, Fgf8 alone was not enough to induce *Olig2* or *Nkx2.2* while co-expression with Shh produced a lot more Olig2 and

Nkx2.2 positive cells than Shh alone (Farreny et al., 2018). When Fgf was inhibited, *Nkx2.2* and *Olig2* increased in expression while *Nkx6.1* remained the same (Morales et al., 2016). Different vertebrates seem to have their own Fgf patterning and interaction with other signals. My *Nkx6A* and *OligA* lamprey data recapitulates the chick Fgf inhibition study (Figure 3.15 C, D, 3.16 A-D, 3.16.1). Amphioxus *Nkx6* expression is reduced when Fgf is inhibited but only at the s10 stage, not the s14 stage (Figure 3.23 A-D).

In addition to the lamprey genes studied in this chapter, as depicted in chapter two *Pax2/5/8* and *Zic* are also good candidates for studying signalling manipulation on spinal cord patterning, but unfortunately I did not have the chance to study them. Lamprey *Pax6* was not studied in this chapter because of a technical issue with the probe. Currently, it is hard to tell whether different signals have a hierarchy or temporal sequence like *Dbx* and *OligA* maintaining small areas of expression near the tailbud in Notch-inhibited embryos (Figure 3.11, 12); or *Pax7* retaining a small section of tail spinal cord expression in RA-overactivated embryos (Figure 3.9). We also do not know if there are equal combinatorial effects on induction or maintenance of markers like *Nkx2.2, Nkx6A,* and *Axin*, because they are all affected by different signalling pathway manipulations.

In amphioxus, signal manipulation could be affecting major developmental events like neurulation, which has downstream effects that could cause problems in neuron specification (anatomically, I see kinked shape of mesendoderm and floorplate in drug-treated embryos, suggesting some aspects of morphogenesis may be compromised). Ideally, tissue-specific and time-dependent conditional knockout would be more accurate to study the effects of signal manipulation on spinal cord development in amphioxus and lamprey. I would then be able to observe the unadulterated effects of different signals on spinal cord without contributing factors like convergent extension and elongation problems affecting spinal neuron differentiation. Unfortunately, the tools and technology for these two species are not there. Single cell RNA-seq would give a better idea of whether the different signals are interconnected or cross-modulate. The comparison of drug treated zebrafish and amphioxus embryo RNA-seq from Gil-Galvez et al. (2020) seems to suggest that signalling pathways are cross-modulated. If single cell RNA-seq were done, we would obtain the perspective of in which cell types those transcripts were being up or down regulated in, instead of the tissue-level resolution provided by the whole mount in situ hybridization data.

Chapter4: Discovering cis-regulatory elements in lamprey

4.1 Introduction

Some of the genes studied in the previous chapters were within the NKL subclass of homeobox genes, which form part of the ANTP class (Zhong and Holland, 2011). The ANTP class also includes the HOXL subclass which contains the Hox genes, famous for their tight clustering and collinearity, and ParaHox genes, known for their roles in endoderm patterning in development (Holland, 2012). Within the NKL subclass, the protostome-deuterostome ancestor probably contained an NK cluster of nine genes because in *Magallana hongkongensis,* the Hong Kong oyster, all its NK cluster genes are on one chromosome (Li et al., 2020). The NK cluster is more broken up and dispersed in vertebrates with only three pairs of tightly linked genes in human (Wotton et al., 2009). Also within the NKL subclass, there are two genes, *Nkx2.1* and *Nkx2.2,* which belong in the micro-syntenic pharyngeal cluster. Other genes in the pharyngeal cluster include *Pax1/9, FoxA, mipol1, Msxlx,* and *slc25A21* (Ferrier, 2016). These other genes are not related to the NKL subclass but their positions in genomes are linked tightly with *Nkx2.1* and *Nkx2.2*, and they are all expressed in the developing pharyngeal slits in deuterostomes (Simakov et al., 2015). The NK cluster may function to pattern mesoderm in Drosophila. Drosophila *tin (Nkx4), bap (Nkx3), lbl (Lbx), lbe (Lbx), C15 (Tlx),* and *slou (Nkx1)* are all tightly linked on the right arm of the third chromosome and have a hierachical order in specifying the mesoderm tissue (Jagla et al., 2001). *Tin* is expressed and restricted to the dorsal mesoderm, which activates *bap* in segmental visceral mesoderm which contributes to the musculature of the midgut (Azpiazu et al., 1993). Downstream of the *tin* specified mesoderm tissue, *lbl* and *lbe* are responsible for the differentiation of the heart and the segmental border muscle (Jagla et al., 2001), *C15* is expressed in the alary muscles (Carmena et al., 1998), and *slou* is expressed in a subset of muscle founder cells (Cox and Baylies, 2005). The expression of *tin* becomes more restricted during development, from dorsal mesoderm to pericardial cells, and coordinates with *bap, lbl,* and *lbe* to specify different cardiac cells (Yin and Frasch, 1998, Zikova et al., 2003). Enhancer sharing within the NK cluster has been demonstrated in the beetle *Tribolium*. In *Tribolium*, *C15* replaces the *lbl* and *lbe* expression in the heart field because of an inversion causing the bypass of an insulator which normally suppresses *C15* expression (Cande et al., 2009). There is no other evidence in bilaterians that the NK cluster plays a role as a functional cluster, in the sense that the genes need to be in a cluster to function appropriately, since the expression of individual genes are usually divergent. Onychophoran *Euperipatoides rowelli* and annelid *Platynereis dumerilii NK1, NK3, NK4, NK5, Tlx, Msx, and Lbx* are expressed in segment polarity

patterns in ectoderm and mesoderm along the AP axis, but there is no evidence from other bilaterians that suggests that this could be an urbilaterian feature (Treffkorn et al., 2018, Saudemont et al., 2008). Work in this chapter updates knowledge of synteny evolution of the NK cluster in chordates to include lampreys, then describes the search for cis-regulatory elements related to all the lamprey genes considered in earlier chapters with the exception of the Wnts. Two methods were used to identify potential cis-regulatory elements, phylogenetic footprinting and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq).

As reviewed in chapter one, there is evidence for transcription factors cross-repressing each other to form cell type boundaries in gnathostomes. Even though my chapter three data suggest that cell type boundaries might not exist in lamprey, I would still like to identify whether *Nkx2.2* and *Olig*; *Nkx6* and *Dbx*; *Pax6* and *Nkx2.2* cross-repression is conserved in lamprey (Sun et al., 2003; Briscoe et al., 2000). If the same cis-regulatory element is found in lamprey then it would suggest that the regulation of that transcription factor is conserved. *Pax6* and *Nkx2.2* are known to be downstream of the Shh-Gli pathway but their exact regulation is still unclear (Soula et al., 2001, Ericson et al., 1997). To better elucidate the gene regulatory network of lamprey spinal cord development, this chapter aims to identify cis-regulatory elements for my genes of interest. Cross-species functional testing of lamprey conserved non-coding sequence would prove that it is a cis-regulatory element with a conserved function in another vertebrate species. Further, in vitro analysis can be done to identify what transcription factors are binding to the cis-regulatory element.

Phylogenetic footprinting compares non-coding orthologous regions of the genome across different species to look for conserved sequences, which usually means they have a slower evolutionary rate and, possibly, important cis-regulatory functions. For non-coding sequences with a faster evolutionary rate it is assumed their sequences could be changed through adaptation, or that the sequences are not under selection and have drifted apart. Yasuoka (2020) highlighted that the difficulty in identifying cis-regulatory elements in amphioxus when comparing to vertebrates may be because of the genome rearrangement in two rounds whole genome duplication, plus their phylogenetic distance. This author also suggested that comparison between cephalochordate species might give better results, which concurs with Yue et al. (2016) who proposed that comparing non-coding sequences between *Asymmetron lucayanum* and *B. floridae* is within the phylogenetic “goldilocks” distance for finding cis-regulatory elements. However, functional testing of enhancers

in amphioxus embryos has heterogeneous mosaic results (Guang Li and Tokiharu Takahashi, pers. comm.).

My approach was to make use of the 100 vertebrates basewise conservation track on the UCSC human GRch38/hg38 genome browser to look for non-coding conserved sequences related to my genes of interest. One conserved non-coding element near *Nkx2.2* from lamprey is tested in an enhancer reporter assay in chick embryos showing that enhancer function is conserved in chick and active throughout the neural tube.

To look for lamprey-specific cis-regulatory elements in spinal cord cells, I used ATAC-seq on dissected lamprey spinal cord cells to observe chromatin openness. However, in vivo experiments are needed to validate that chromatin openness near genes of interest is functional. Since I was not able to establish a reliable way of testing enhancers in lamprey embryos and the open chromatin site sequences are not conserved in other vertebrates, not much meaning can be derived from the ATAC-seq data other than overall chromatin landscape.

4.2 Methods

4.2.1 ATAC and library preparation

Spinal cords of embryos at stage T25 and 27 were dissected in 50uM MS222 in filtered river water. Figure 4.1 shows the sections of spinal cord tissue that were harvested. The explants were kept in L-15 medium with 10% fetal bovine serum at room temperature across the span of dissection, since the dissection of 11 to 14 embryos could last two to three hours, the dissected spinal cord cells are kept alive in the serum.

Spinal cord from 11 to 14 embryos, depending on how much tissue I managed to harvest from each embryo, was dissociated in 20mg collagenase in 200uL Hank's Balanced Salt Solution (HBSS) with 600uL 0.05% Trypsin mixture for 7-8mins at room temperature. The dissociating reaction was halted by diluting in HBSS, spinning down (at 500 xg for 5mins at 4°C), resuspended, passed through a cell strainer (40 micron; Falcon 352340) to remove cell clumps, spun down again (at 500 xg for 5mins at 4°C) and resuspended in 150uL of HBSS. This generated around 6,000-10,000 live cells per sample replicate. After washing the cells with PBS, they were lysed and tagmented using the Illumina Nextera kit, adjusting the volume to account for the smaller number of cells (Table 4.1) and reducing the 37°C incubation time from 30 to 20 minutes to avoid overtagmentation. NEB Q5 High-Fidelity 2X Master Mix was used to

amplify the tagmented DNA for 14 cycles, then quality control of the library was assessed with Agilent TapeStation and quantified with Qubit. The library was then assessed with KAPA Library Quantification Kit and sequenced with paired-end 40bp runs on Illumina NextSeq500 platform.

Table 4.1. Volume adjustments for fewer number of cells in the transposition reaction

	Illumina for 50,000 cells	**Reduced for 6,000-10,000 cells**
2X TD Buffer	25µL	10µL
Tn5 Transposase	2.5µL	1µL
Nuclease free water	22.5µL	9µL
Total volume	50µL	20µL

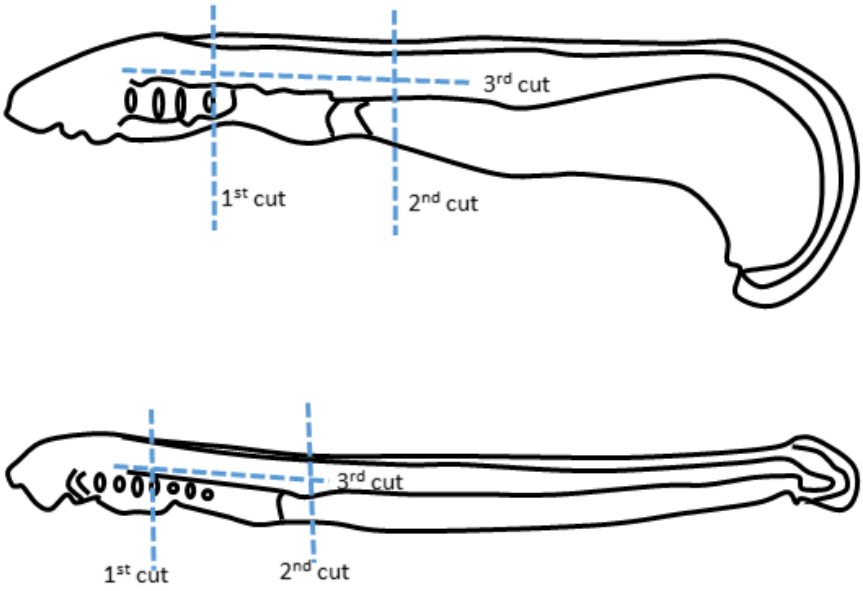


Figure 4.1. Lamprey spinal cord dissection. The top schematic is of a T25 embryo while the bottom is at T27. The blue dotted lines indicate the sequence of the dissection. The isolated spinal cord was then briefly cleaned from surrounding tissues with forceps. Only the anterior spinal cord is being extracted because the spinal cord tapers out towards the posterior and becomes difficult to dissect. Drawings adapted from Tahara, 1988.

4.2.2 ATAC-seq computational analysis

FastQC was used to examine data quality (Andrews et al., 2010). Trimmomatic (v0.36) was used to trim away the adaptor sequence of the paired-end reads (parameters: LEADING:20 TRAILING:20 MINLEN:30 ILLUMINACLIP:NexteraPE-PE.fa:2:30:10) (Bolger

et al., 2014). The *Petromyzon marinus* genome assembly was downloaded (Smith et al., 2018) and indexed using Bowtie2 (Langmead et al., 2008). The trimmed sequence reads were mapped back into the genome with Bowtie2 (parameters: --very-sensitive –X 2000). Such parameters were used to include as many reads as possible because the number of reads and data quality are not high to begin with. Duplicates were removed by MarkDuplicates (Picard v1.83) and peaks were called by MACS2 (parameters: --nomodel –shiftsize 100 –extsize 300) (Zhang et al., 2008). Such parameters were used to smooth the pileup signals. Peaks were visualized with Integrative Genomics Viewer (V.2.4.10). The raw datasets have been deposited in the European Nucleotide Archive under accession code: PRJEB47184 http://www.ebi.ac.uk/ena/data/view/PRJEB47184

4.2.3 Chick neural tube in ovo electroporation

Fertilised Bovans Brown eggs were obtained from Henry Stewart & Co (Norfolk, England) and incubated at 38°C. Solutions of plasmid DNA (2ug/uL) in TE buffer pH 8/ 0.01% Fast Green were injected into the lumen of HH10 chick neural tube from tailbud using a Nanojet II with capillary pulled needle, then electroporated with 1-mm diameter platinum electrodes using BTX ECM 830 with settings 18 volts, 5 pulses, 50 ms duration, and 950ms intervals. Method was adapted from Croteau and Kania (2011) with the voltage reduced to adjuste for the younger HH10 stage. An RFP expression vector (pCI-H2B-RFP, vector map on AddGene Cat.# 92398) was included as a control for successful introduction of DNA. A pTK-eGFP enhancer reporter construct (Chen and Streit, 2015) was the experimental DNA (full sequence in the supplementary data). Some embryos have both sides of the neural tube electroporated to give a stronger signal when screening embryos, while some have one side electroporated with the contralateral side acting as internal negative control. Chick embryos were collected at 16 to 24 hours following electroporation and imaged live in 1X PBS with Leica MZ16F microscope with Leica DFC300 FX camera and Leica EL6000 fiber optic fluorescence illuminator at a magnification of 50X.

4.2.4 Lamprey whole mount electroporation

Please refer to chapter two methods for lamprey husbandry. The lamprey electroporation method is adapted from zebrafish embryo electroporation protocol (Kera et al., 2010). After attempts of injecting plasmids into the neural tube lumen of lamprey, I realised that the method is not feasible because of how small the central canel is. So I tried whole embryo electroporation instead, which renders the step of fixing the embryo in low melting-point argrose obsolete. T25 to T27 embryos were

transported to a 100uL electroporation chamber containing electroporation buffer, pTK-GFP enhancer reporter construct (Chen and Streit, 2015) and RFP vector (pCI-H2B-RFP) as a control for successful introduction of DNA into the lamprey embryos. The working concentration of each plasmid is about 20ug/uL. Current was passed unidirectionally using BTX ECM 830 with the setting of 30V, 50ms pulse, 950ms pause, with 5 pulses. The 1-mm diameter platinum electrodes are placed parallel to either side of the embryo. Lamprey embryos were observed under a florescence microscope at 16 to 24 hours following electroporation then fixed in 4%PFA for 20minutes to allow amplification of GFP signals by immunohistochemistry.

4.2.5 Lamprey whole mount immunohistochemistry

Fixed embryos were treated in 1mL 30% H_2O_2 with 450uL PBT for an hour. After washing away the H_2O_2 with PBS, embryos were incubated in chick anti-GFP (Sigma-Aldrich, AB16901), with 10% Heat-Treated Sheep Serum, in PBT at 4 °C overnight. After washing away primary antibody with PBT, they were incubated in anti-chick AlexaFluor-488 (Invitrogen, A32931) in PBT at 4 °C overnight. Secondary antibody was then washed away with PBT before imaging with a Leica MZ16F microscope with Leica DFC300 FX camera and Leica EL6000 fiber optic fluorescence illuminator at a magnification of 50X.

4.2.6 Synteny analysis

Synteny diagrams of NK clusters were drawn referring to UCSC genome browser for human, the UCSC custom track hub created by Marletaz et al. (2018) for amphioxus and the Stowers Institute genome browser for lamprey (Smith et al., 2018).

4.2.7 Conserved non-coding sequence BLAST against lamprey genome

The 100 vertebrates basewise conservation track displayed in the UCSC human genome browser showing highly conserved region sequences were extracted using the table browser tool. The extracted sequences were then searched against the lamprey genome using the Stowers Institute *Petromyzon marinus* blastn suite (https://simrbase.stowers.org/blast/nucleotide/nucleotide) at its default settings with expected threshold: 0.1, word size: 11, max matches in a query range:0, match/mismatch scores:1,-2, gap costs: Existence5 Extension2

4.2.8 Enhancer sequence alignment and motif database search

Enhancer sequences were extracted from the UCSC genome browser using the BLAT tool or through Ensembl BLAST (in their default parameters) against the genomes of

different species. Extracted sequences were aligned with EMBL-EBI multiple sequence alignment tool Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Motif sequences were searched through the JASPAR database (Fornes et al., 2020).

4.3 Results & Discussion

4.3.1 Vertebrate NK cluster synteny evolution

Five of the genes tested in chapter two belong to the NKL subclass; *Nkx6, Dbx, Dlx* and *Msx* from the NK cluster and *Nkx2.2* from the pharyngeal cluster (Ferrier, 2016, Simakov et al., 2015). Although *Dbx* and *Dlx* were classified in the NKL subclass by Schierwater et al. (2008) with Baysian Inference and Maximum Likelihood analyses, later on Hui et al. (2012) put that in question with genomic linkage analysis when the protostome *Platynereis dumerilii* was studied. Building upon the work of chordate NK cluster evolution analysis by Wotton et al. (2009), I analysed the synteny of NK genes from the *Petromyzon marinus* genome 2018 assembly (Figure 4.2). The NK cluster seems to have undergone only one round of chromosomal duplication in lamprey because most genes only have two paralogues. Hmx, which has three paralogues, is likely due to tandem duplication instead of whole genome duplication based on their position in the phylogenetic tree, close proximity on the chromosome and conserved non-coding sequences between duplicates (Vasileios Papadogiannis, pers. comm.). While Smith et al. (2018) provide convincing evidence that the Hox cluster in lamprey has undergone two-rounds of duplication plus an extra chromosomal duplication to give six clusters, other clusters, like the NK, might have only been through one-round of duplication, or might have been through two duplications with subsequent loss of clusters which is consistent with the idea that lampreys have diverged from the vertebrate ancestor after the first duplication by autotetraploidisation followed by additional lineage specific genomic duplications (Simakov et al., 2020).

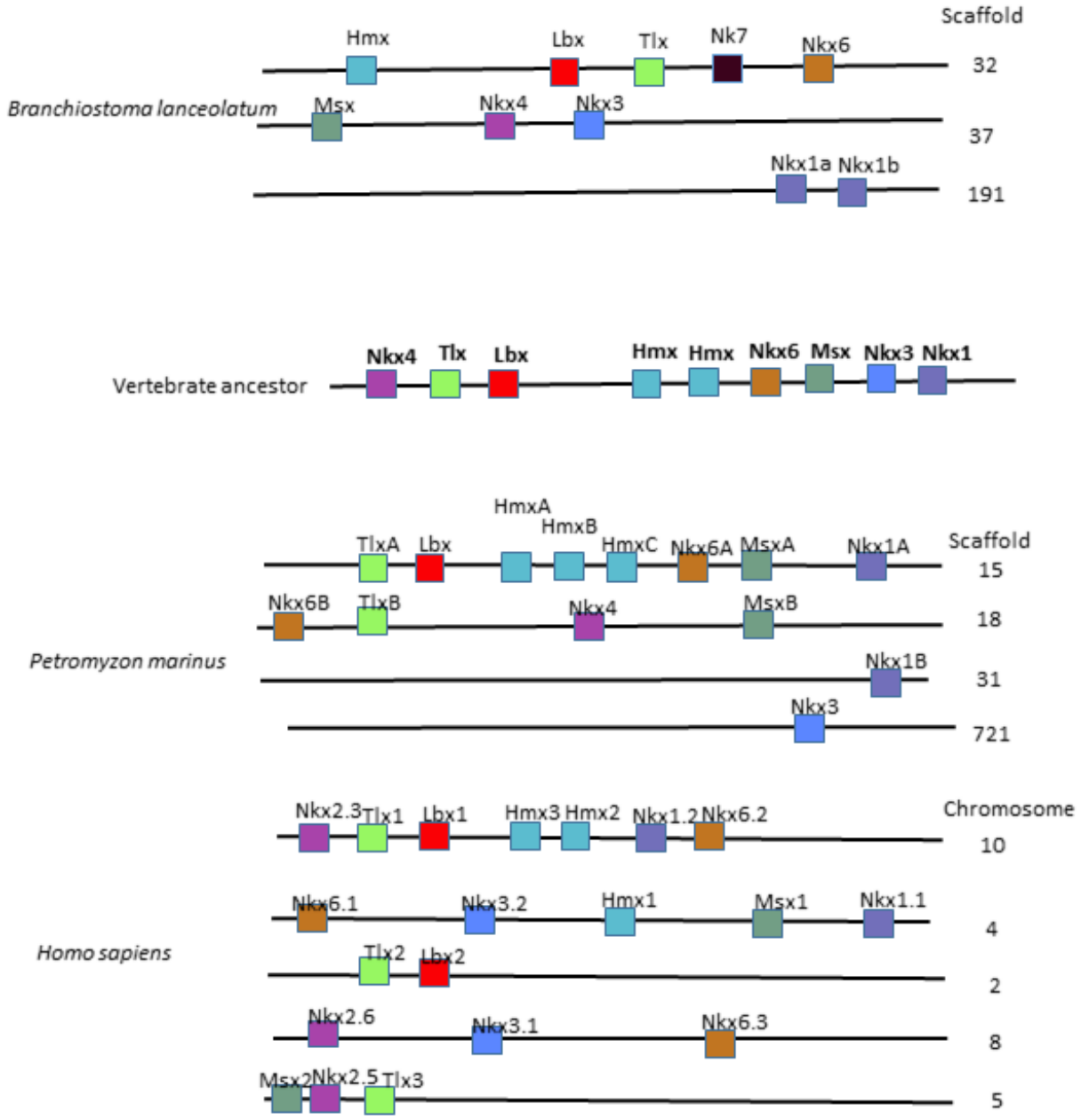


Figure 4.2. NK cluster synteny. Clustered NK genes in human, lamprey, amphioxus and predicted vertebrate ancestral state of nine genes. Orthologues and paralogues are depicted in matching colours. Chromosome or scaffold number is shown on the right. Genomes used to extract these syntenic data: Amphioxus ((Marletaz et al., 2018); http://amphiencode.github.io/Data/), Lamprey ((Smith et al., 2018); https://genomes.stowers.org/sealamprey), Human (UCSC genome browser GRCh38/hg38).

The NK cluster has been identified in hemichordates *Saccoglossus kowalevskii* and *Ptychodera flava* by Simakov et al. (2015) as *nkx5-msx-nkx3.2-nkx4-lbx-hex, nkx1a-vent1-vent2.1-vent2.2, voxvent-nk7like-nk7like2* in three different scaffolds. In the

cephalochordate *Branchiostoma lanceolatum* using the more recent Marletaz et al. (2018) genome I have updated the work of Wotton et al. (2009) and found that the NK cluster is dispersed as *hmx-lbx-tlx-nk7-nk6*, *msx-nk4-nk3*, and *nk1a-nk1b* in three separate scaffolds. In *Petromyzon marinus* I identified *TlxA-Lbx1-HmxA-HmxB-HmxC-Nkx6.2-MsxA-Nkx3.2-Nkx1A, Nkx6.1-TlxB-Nkx2.3(4)-MsxB, Nkx1B, Nkx3.1* in four different scaffolds (Figure 4.2). The tight linkage between *Nkx4-Nkx3* cannot be seen in the *Petromyzon marinus* 2018 genome assembly. This may be due to genome assembly artefacts because such linkage is conserved in hemichordate, cephalochordate and human, though could also reflect evolutionary changes in lamprey. Currently, there are no other high quality cyclostome genome assemblies to confirm or deny this. The *Tlx-Lbx* linkage is likely to have been lost independently in the hemichordate lineage, after the deuterostome split because such linkage is conserved in *Drosophila melanogaster* and *Tribolium castaneum* (Wotton et al., 2009). When comparing NK cluster synteny between lamprey and human, vertebrate evolution is consistent with Holland's hypothesis (Holland, 2015) that NK cluster genes duplicated and then broke apart into separate chromosomes by differential paralogue loss.

4.3.2 Chromatin landscape around lamprey genes

In an attempt to identify novel lamprey specific cis-regulatory elements, ATAC-seq was performed on dissected lamprey spinal cord tissue at a younger (T25, three replicates) stage and an older (T27, three replicates) stage. Hockman et al. (2019) also analysed lamprey cis-regulatory elements. The authors found conservation of functional non-coding regions between *P. marinus* and *L. planeri* which facilitated the hunting of cis-regulatory elements when mapping *L. planeri* ATAC-seq reads onto the *P. marinus* genome. My analysis followed the same approach. In theory, the experiments should provide a snapshot of the openness of the chromatin state. I was expecting to see consistency between test groups for genes which are expressed throughout the stages and differences between the younger and older test groups for genes which cease to express in the older stage. For example, all three T25 replicates should have the same open chromatin areas hinting that the potential enhancers or repressors are controlling the transcription, or at least for genes that I know are expressed in the dissected area at that stage should have their transcription start site open. Whereas in the older test group, genes which have stopped expressing in the dissected area of spinal cord should have a closed transcription start site with different open chromatin areas from T25. However, the locations of transcription start sites can sometimes be difficult to tell because the

gene models are truncated and shorter than their gnathostomes orthologues. For example, the *P. marinus OligB* gene model has a predicted transcript of 307bp, even though the HLH amino acid sequence alone should have a longer coding region than that, while *Olig* orthologues in human, chick, and mouse are 2kb to 2.5kb long. Accordingly, I have estimated the lamprey *Olig* genes to be around 2.5kb when depicting their ATAC-seq data (Figure 4.3, 4.4). I define relevant open chromatin areas to be from 5kb upstream to 5kb downstream of the genes including introns, unless there is another gene abutting the gene of interest, in which case the relevant open chromatin area would only extend to the beginning of the abutting gene (Figure 4.12). The viability of this ATAC-seq experiment is not good considering the percentage of successfully mapped reads for each replicate are low (Table 4.2): I would expect this to be lower than normal for an ATAC-seq experiment because of the cross-species mapping, but not as low as that shown in Table 4.1. Also, the insert size distributions do not show the expected nucleosome occupancy of 150bp periodicity (data not shown), so I would judge that the quality of the ATAC-seq libraries were low. However, I am not doing genome wide analysis with these data, so despite the low quality, they could still potentially be useful for finding cis-regulatory elements.

Table 4.2. ATAC-seq number of reads in each replicate

Stages and replicates	Total sequences	Mapped successfully
T25 R1	16,621,492	4.88%
T25 R2	9,411,000	2.82%
T25 R3	6,687,072	1.85%
T27 R1	19,053,675	3.55%
T27 R2	2,884,306	12.7%
T27 R3	12,947,993	4.08%

I would expect *OligA, OligB, Nkx2.2, PtchB,* and *AxinA* to have different open chromatin areas when comparing between T25 and T27 because those genes are expressed in T25 but ceased expression in T27 at the anterior spinal cord (Figures 2.1, 2.2, 2.7, 2.18, 2.19). Contrary to my predictions, *OligA, Nkx2.2* and *AxinA* do not have consistent open chromatin sites within the three replicates of each test group (Figures 4.3, 4.6, 4.12). There is a consistent open chromatin site across all three replicates of T27 downstream of *OligB* (Figure 4.4). The function of that sequence could be tested in future work. I would hypothesise that the open chromatin site could potentially function as a repressor of *OligB* because *OligB* has ceased

expression in anterior spinal cord at T27 (Figure 2.2). It is inconclusive as to whether my ATAC-seq data are accurate depictions of cis-regulatory elements. Also, contrary to my predictions, there are consistent open chromatin sites across all the replicates for *PtchB* (Figure 4.11). Some of these apparent open chromatin sites are read mapping artifacts. When these segments of non-coding sequences are filled with short repeated motifs, reads including similar repeats were also mapped to these areas. Other open chromatin sites can be potential lamprey-specific cis-regulatory elements because when I searched the sequences against other species, there were no matches. A reason could be that the GC-rich lamprey sequences are rarely found in other chordates (Figures 4.4, 4.8, 4.11, 4.12 legend). A potential explanation for having consistent open chromatin sites across two test groups could be that there were still low levels of transcription of *PtchB* at T27 in the anterior spinal cord which were not picked up in the in situ hybridisation data, or during dissection some of the hindbrain tissue was included. *PtchB* had fading expression in anterior spinal cord as the embryos aged while maintaining hindbrain expression.

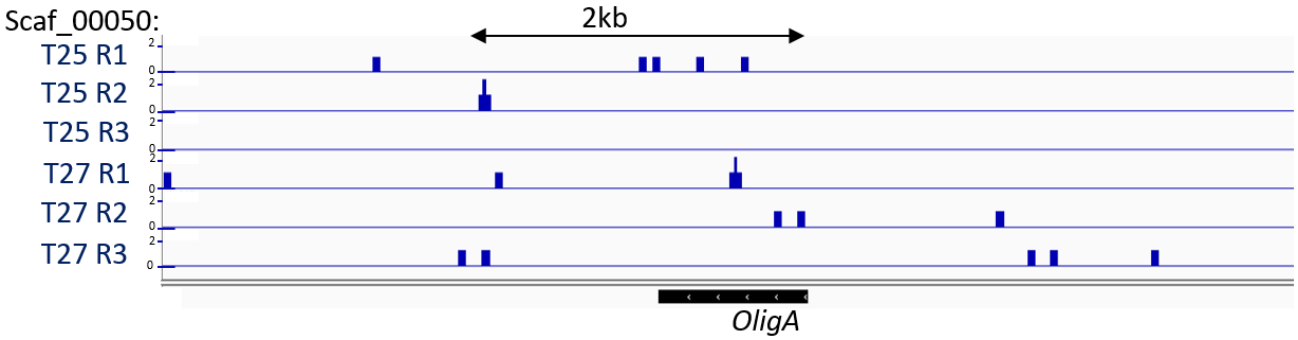


Figure 4.3. Open chromatin state related to lamprey *OligA*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. Only one exon is shown for *OligA*. The gene model is likely to be truncated or incomplete because *Olig* orthologues in other species are usually around 2-2.5kb with multiple exons.

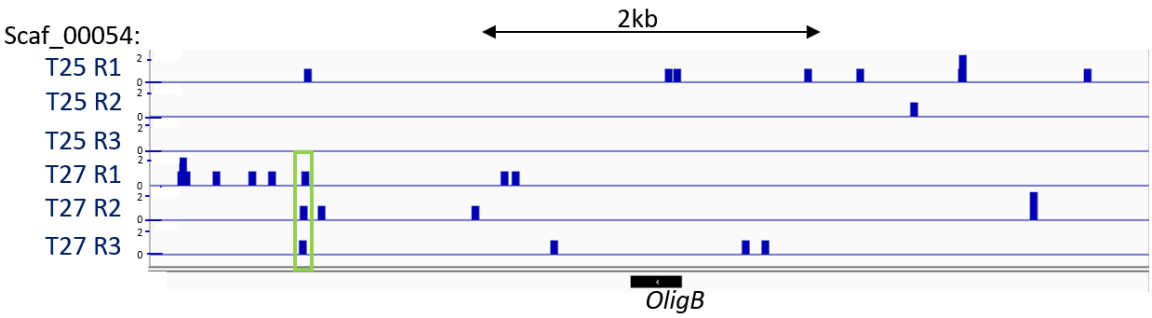


Figure 4.4. Open chromatin state related to lamprey *OligB.* Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. Only one exon is shown for *OligB*. The gene model is likely to be truncated or incomplete because other *Olig* orthologues in other species are usually around 2-2.5kb with multiple exons. The green box highlighting common open chromatin site within all three T27 replicates and its sequence is "ACCACCCACCCCCCTCGTGTGCCGTGCCGGCCTTCATAGGTACTACTCGCT".

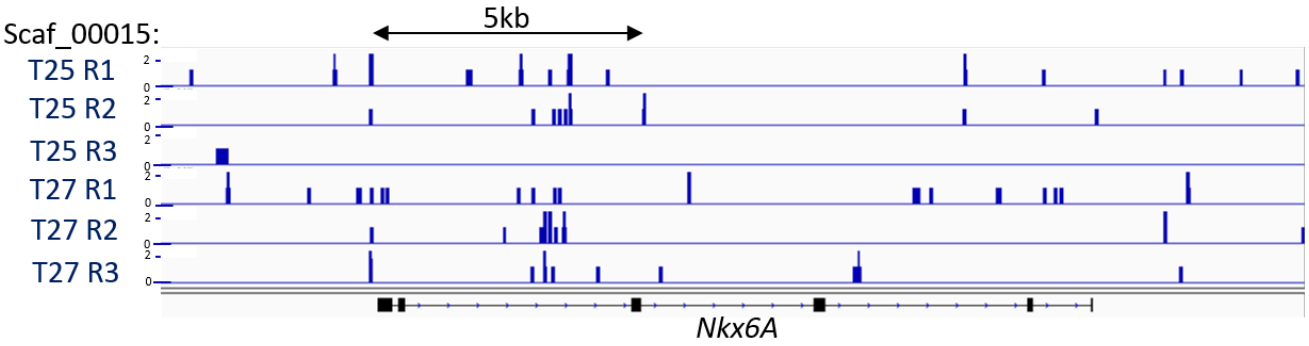


Figure 4.5. Open chromatin state related to lamprey *Nkx6A.* Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns.

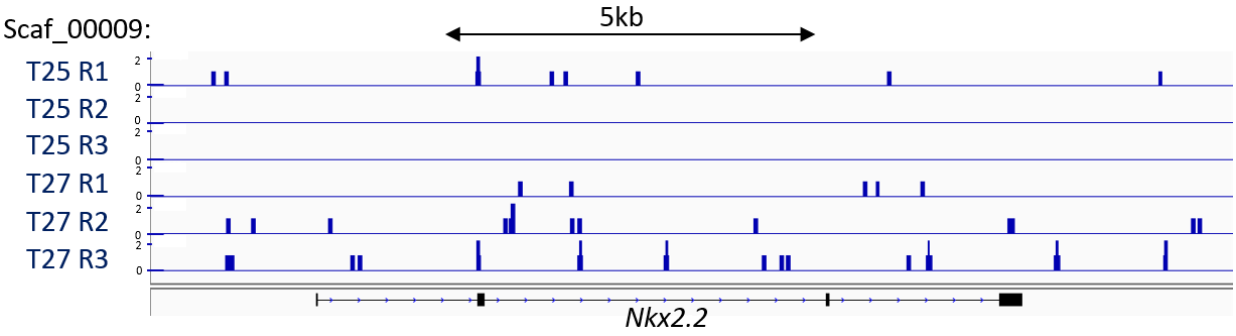


Figure 4.6. Open chromatin state related to lamprey *Nkx2.2*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns.

As for the other genes which are expressed (or not expressed, i.e. *MsxA*) throughout T25 to T27, I would expect their chromatin state to be similar across the two test groups, which is true in some cases but not in others. For *Pax6β,* there is a consistent open chromatin site which could be a potential lamprey-specific cis-regulatory element (Figure 4.8). For *Dbx, Nkx6, MsxA* and *Pax7,* no consistent open chromatin sites were observed between the two test groups (Figure 4.5, 4.7, 4.9, 4.10). I think this is as far as this line of investigation can be pushed because the next step would be testing the function of these open chromatin sequences in vivo in lamprey embryos or in vitro with electromobility shift assay, which I tried unsuccessfully (see below). Alternatively, ATAC-seq or chromosome conformation capture sequencing (4C-seq) data can be compared between drug treated embryos and controls to observe their change in chromatin state and yield insight into gene regulation. There is also now a genome assembly of a different lamprey species, *Lampetra richardsoni,* a western brook lamprey. This species of lamprey is found in the United States and Canada, and is different from the brook lamprey, *Lampetra planeri*, used in this thesis (Vladykov and Follett, 1965). I do not think mapping my ATAC-seq to the more recent *L. richardsoni* genome assembly would significantly help with the percentage of successful mapped reads because the quality of the libraries are not good to begin with. However, it might be interesting to compare non-coding regions between different lamprey species to see if the phylogenetic "goldilocks" distance exists within lampreys like it does with amphioxus *A. lucayanum* and *B. floridae* (Yue et al., 2016).

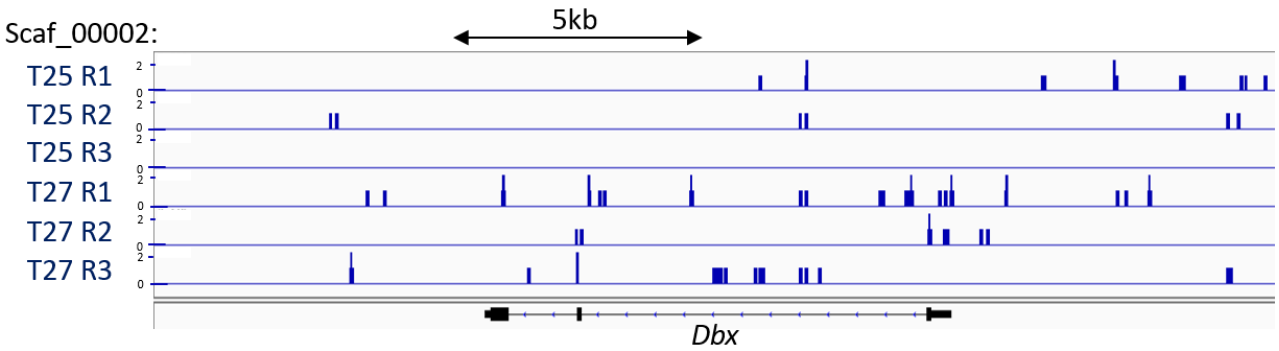


Figure 4.7. Open chromatin state related to lamprey *Dbx*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns.

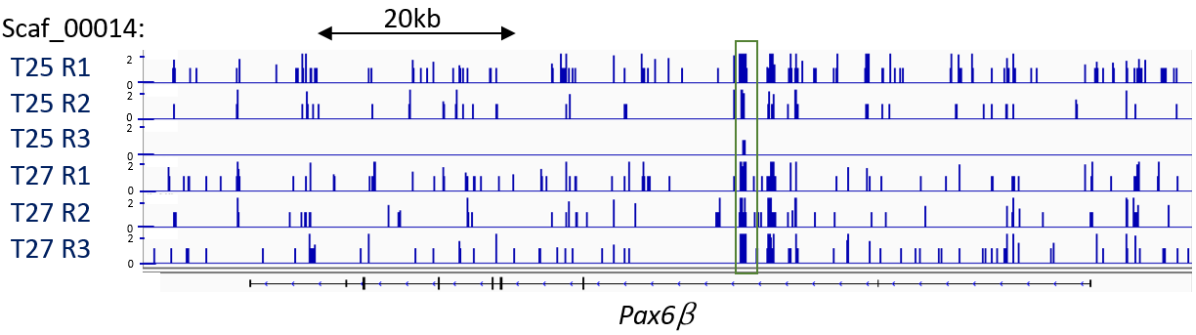


Figure 4.8. Open chromatin state related to lamprey *Pax6β*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns. Green box highlighting consistent open chromatin region, but upon closer examination it was apparent that the sequence is short repeated motifs.

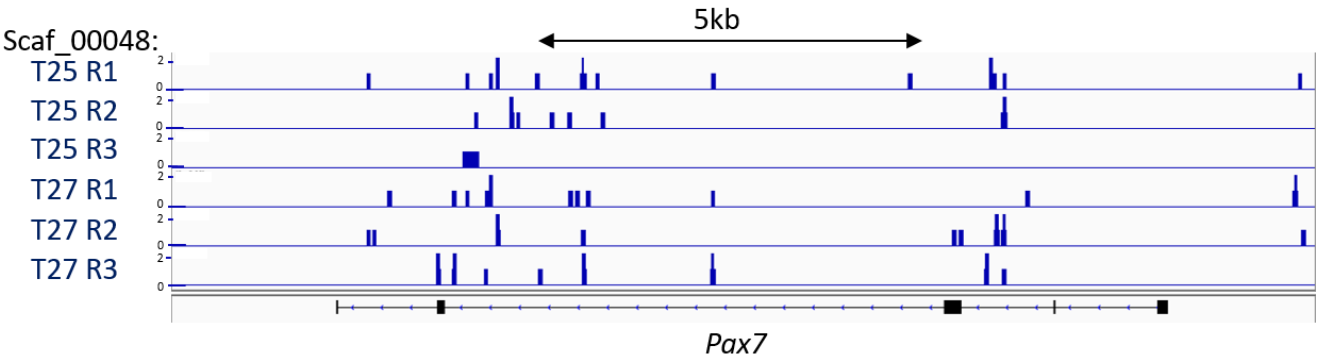


Figure 4.9. Open chromatin state related to lamprey *Pax7*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns.

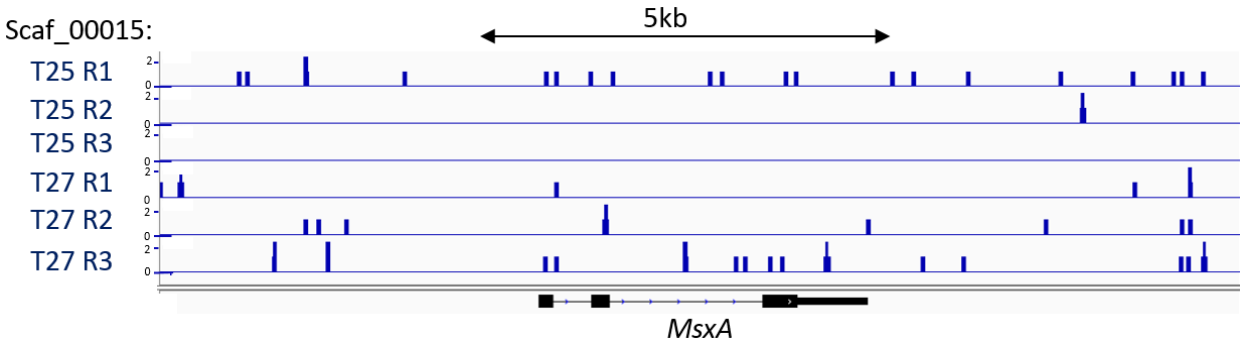


Figure 4.10. Open chromatin state related to lamprey *MsxA*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns.

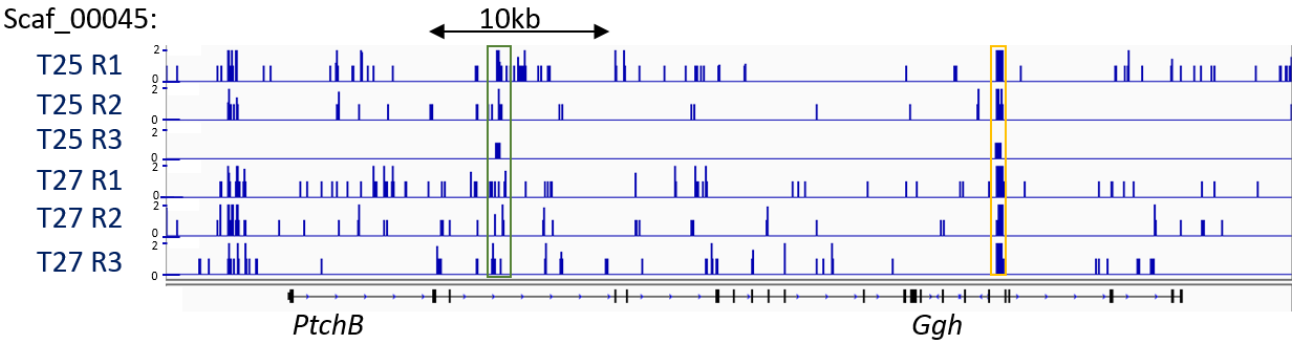


Figure 4.11. Open chromatin state related to lamprey *PtchB*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns. The *Ggh* gene model is about 5kb long and overlaps with *PtchB* 13th intron. Green box highlighting consistent open chromatin region, but upon closer examination it was apparent that the sequence is short repeated motifs. Orange box highlighting consistent open chromatin region with the sequence "CTCCATCCTCTCACGCACCTCCTCTCATCTCACTCTACTCTCCCCTCTCCATCCTCTCACACACCACTCATCTCACTCT ACTCTCACCT".

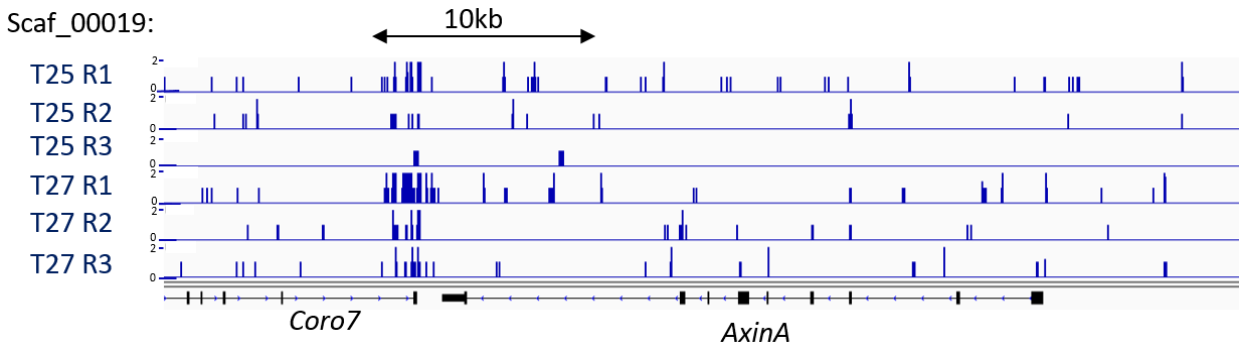


Figure 4.12. Open chromatin state related to lamprey *AxinA*. Top three rows are showing the ATAC-seq data for the three replicates of T25 lamprey spinal cord while the bottom three rows are showing the three replicates of T27 lamprey spinal cord. The blue peaks represent open chromatin sites identified by MACS2. The gene model is represented at the bottom with the arrows within showing the direction of the transcript. The black bars represent exons while the lines in between represent introns.

4.3.3 Phylogenetic foot-printing in the lamprey genome

To look for slow evolving conserved non-coding sequences which could potentially be functioning as cis-regulatory elements, I used the UCSC human genome browser which has a track showing the comparison of a hundred vertebrate genomes with more conserved sites shown as upward blue peaks and less conserved sites as downward red peaks. I then extracted the conserved sequences and searched them against the lamprey and amphioxus genomes. Out of all my lamprey genes of interest, only *Nkx6, Pax7, Pax6,* and *Nkx2.2* had related conserved non-coding sequences (Figure 4.13, 4.15, 4.17, 4.19), while none of them were found in the amphioxus genome. I aligned the homologous sequences from six other vertebrates and looked for conserved motifs. I also looked for other conserved non-coding sequences related to *Dbx1, Dbx2, Msx1, Msx2, Nkx2.1,* and *Nkx2.4,* which were not found in lamprey nor amphioxus genomes (Supp. Figure 4.1-4.3). All the conserved motifs from the conserved sequence alignments were searched through the JASPAR database but no experimentally validated results were found.

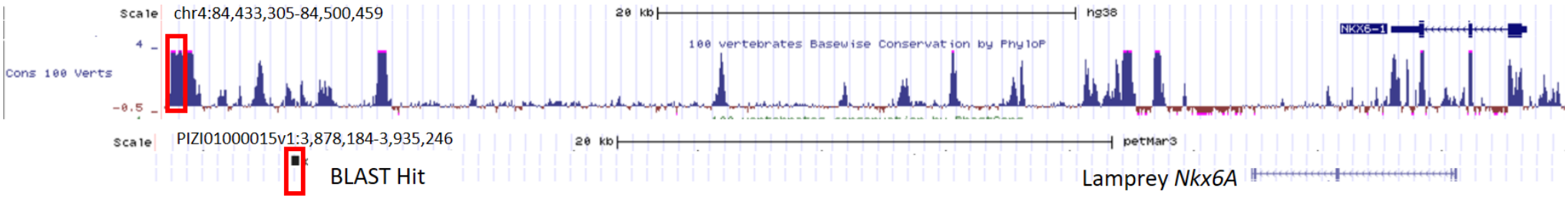


Figure 4.13. *Nkx6.1* conserved non-coding sequence location in human and lamprey. The panel above is adapted from the UCSC human genome browser. The row of peaks below the human *Nkx6.1* gene model is the track of sequences conserved in 100 vertebrate genomes. The panel below is adapted from the UCSC lamprey genome browser. The red boxes highlight regions of the genomes where the sequence is conserved in lamprey. The human and lamprey genome browser displays are not to the same scale and not aligned in the figure.

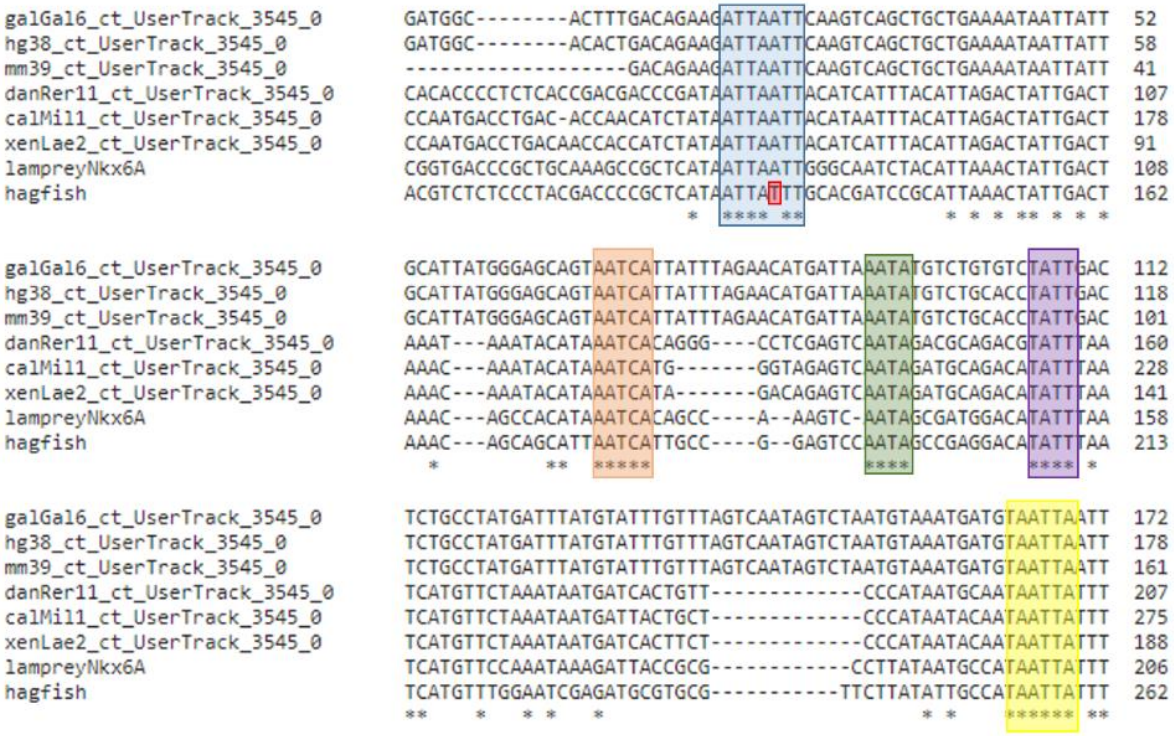


Figure 4.14. *Nkx6.1* conserved non-coding sequence alignment. Conserved motifs across chick, human, mouse, zebrafish, elephant shark, frog, lamprey, and hagfish are highlighted in coloured boxes with a small red box highlighting single nucleotide difference in hagfish (T instead of an A). Sequences used for the alignment are in the supplementary data.

The *Nkx6.1* conserved non-coding sequence is located around 32kb downstream of the gene in both human and lamprey (Figure 4.13). It has five conserved motifs with the first one having a single nucleotide difference in hagfish (Figure 4.14).

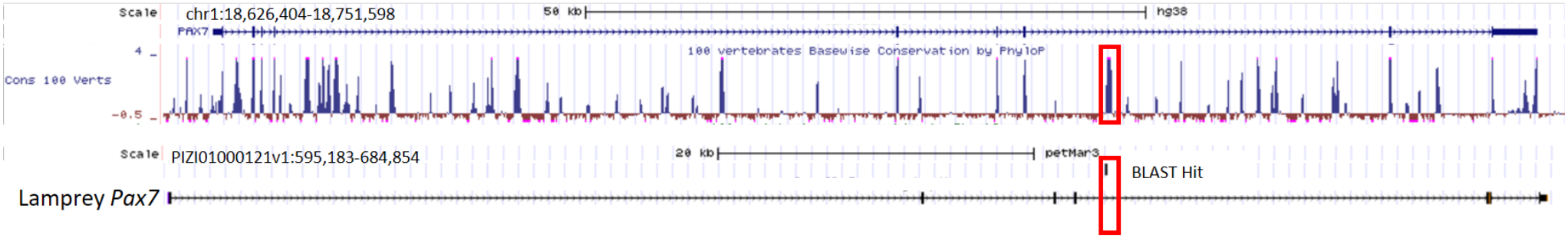


Figure 4.15. *Pax7* conserved non-coding sequence location in human and lamprey. The panel above is adapted from the UCSC human genome browser. The row of peaks below the human *Pax7* gene model is the track of sequences conserved in 100 vertebrate genomes. The panel below is adapted from the UCSC lamprey genome browser. The red boxes highlight regions of the genomes where the sequence is conserved in lamprey. The human and lamprey genome browser displays are not to the same scale and not aligned in the figure.

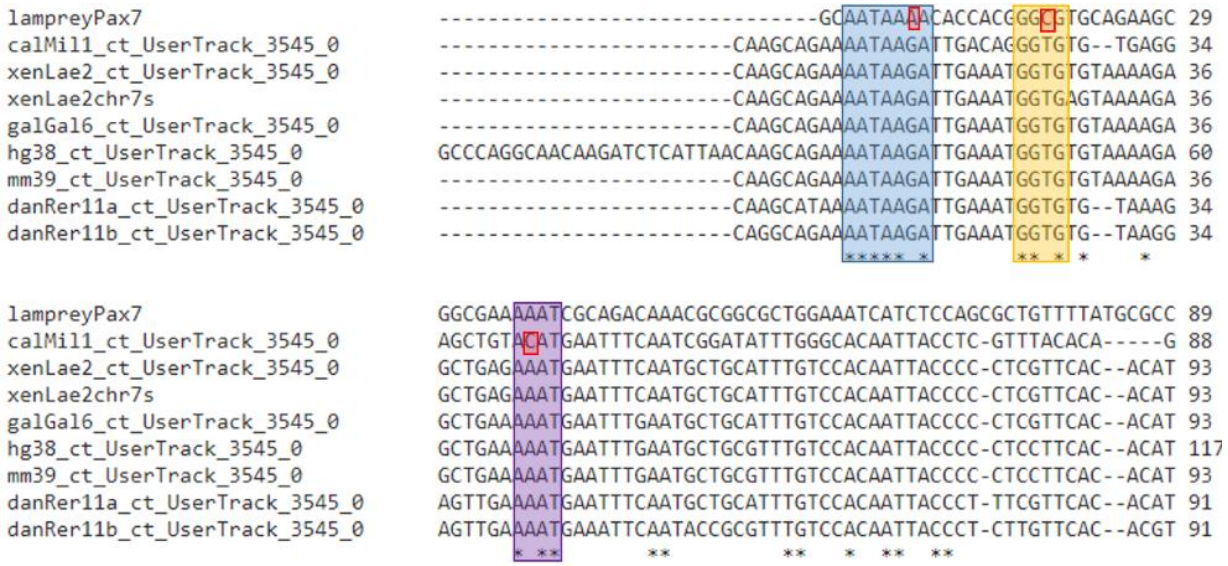


Figure 4.16. *Pax7* conserved non-coding sequence alignment. Conserved motif across lamprey, elephant shark, frog, chick, human, mouse, and zebrafish are highlighted in colour boxes with single nucleotide variants being highlighted in red boxes. Sequences used for alignment are in the supplementary data. I could not find an orthologous sequence in the hagfish genome.

The *Pax7* conserved non-coding sequence is located in the seventh intron in human and fourth intron in lamprey (Figure 4.15). These are not homologous introns, I have examined the exons on either side and they encode different parts of the protein. The sequence has one conserved motif including two paralogous sequences in zebrafish while an orthologous sequence was not found in the hagfish genome (Figure 4.16).

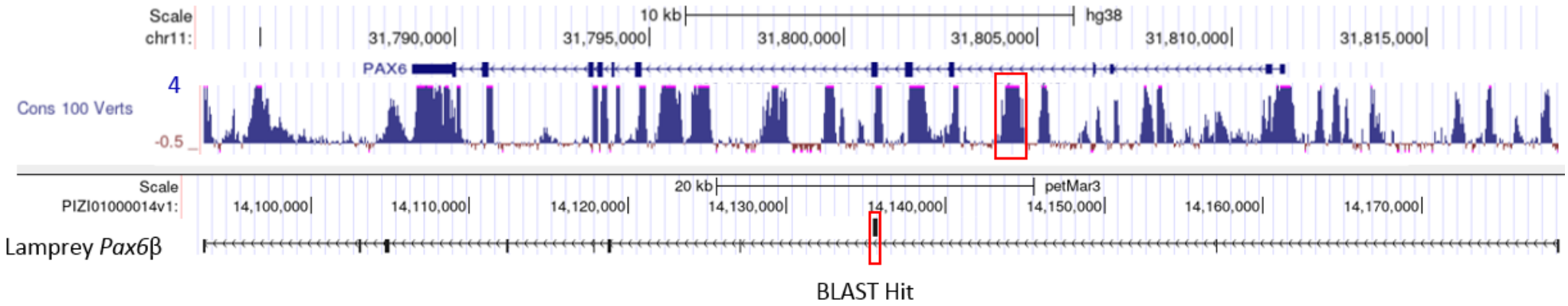


Figure 4.17. *Pax6* conserved non-coding sequence location in human and lamprey. The panel above is adapted from the UCSC human genome browser. The row of peaks below the human *Pax6* gene model is the track of sequences conserved in 100 vertebrate genomes. The panel below is adapted from the UCSC lamprey genome browser. The red boxes highlight regions of the genomes where the sequence is conserved in lamprey. The human and lamprey genome browser displays are not to the same scale and not aligned in the figure.



Figure 4.18. *Pax6* conserved non-coding sequence alignment. Conserved motif across elephant shark, zebrafish, human, mouse, chick, frog, hagfish, and lamprey are highlighted in coloured boxes with the hagfish single nucleotide variants being highlighted in red boxes. Sequences used for alignment are in the supplementary data.

The *Pax6* conserved non-coding sequence is located in the fourth intron in human and the second intron in lamprey (Figure 4.17). These are not homologous introns, I have examined the exons on either side and they encode different parts of the protein. It has one conserved motif with four single nucleotide variants in the hagfish sequence (Figure 4.18).

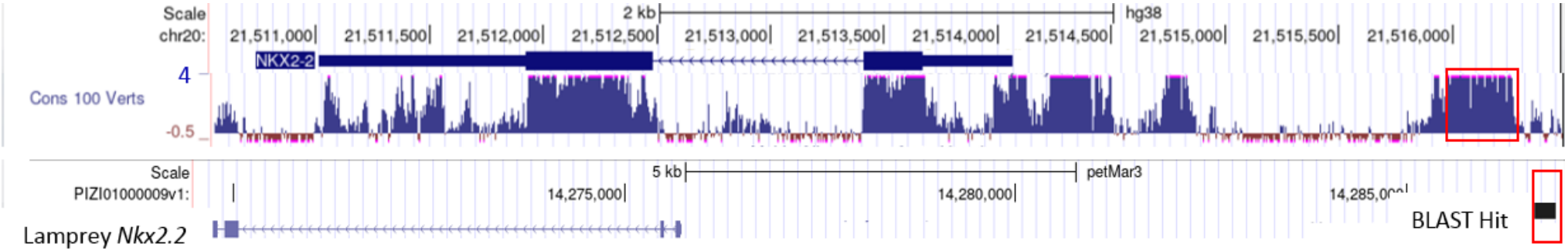


Figure 4.19. *Nkx2.2* conserved non-coding sequence location in human and lamprey. The panel above is adapted from the UCSC human genome browser. The row of peaks below the human *Nkx2.2* gene model is the track of sequences conserved in 100 vertebrate genomes. The panel below is adapted from the UCSC lamprey genome browser. The red boxes highlight regions of the genomes where the sequence is conserved in lamprey. The human and lamprey genome browser displays are not to the same scale and not aligned in the figure.

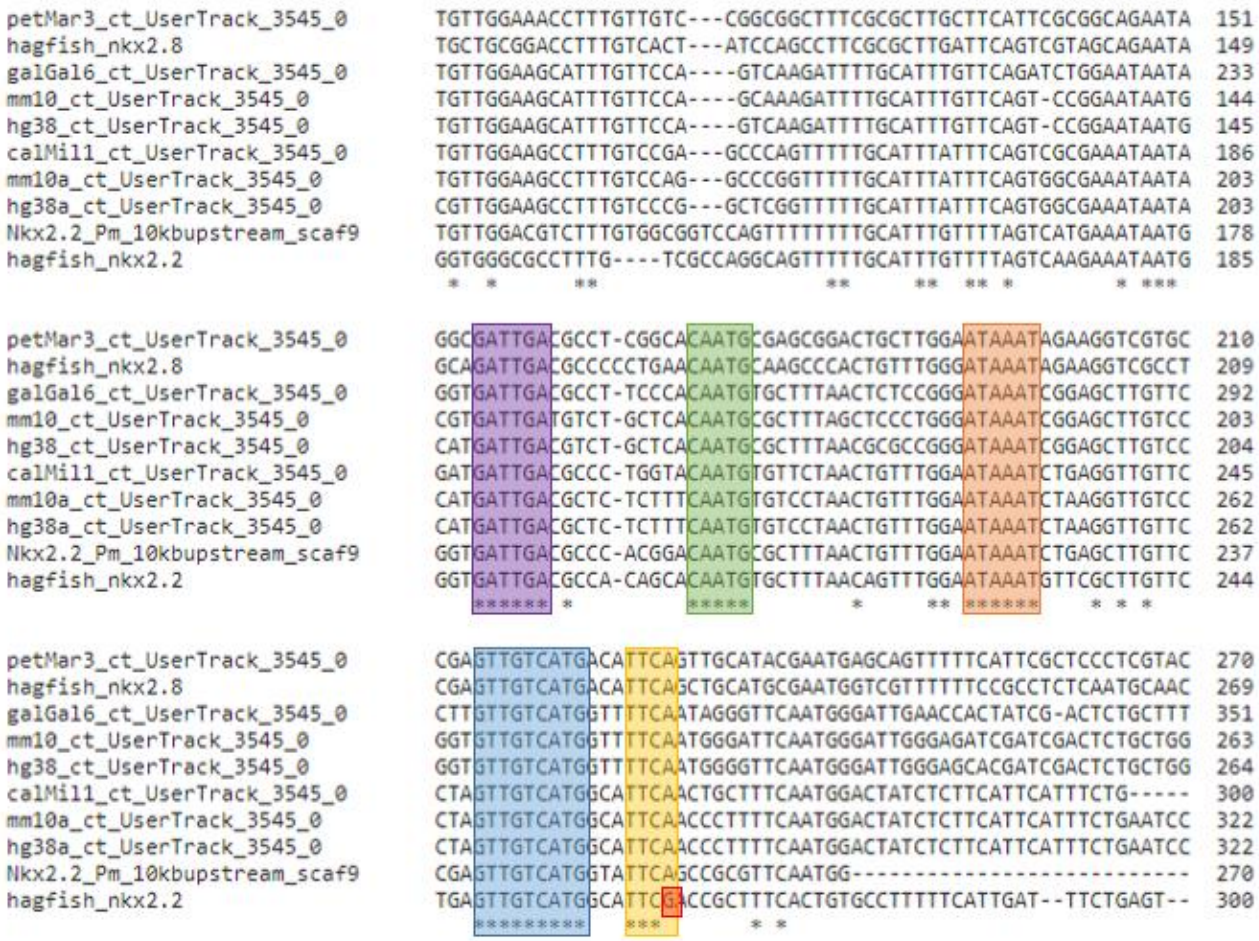


Figure 4.20. *Nkx2.2/Nkx2.8* conserved non-coding sequence alignment. Conserved motifs across lamprey, hagfish, chick, mouse, human, elephant shark are highlighted in coloured boxes with a small red box highlighting a single nucleotide variant (G instead of A) in hagfish. Sequences used for alignment are in the supplementary data.

The *Nkx2.2* conserved non-coding sequence is located around 2kb upstream in human and 11kb upstream in lamprey (Figure 4.19). It has five conserved motifs with the last one having a single nucleotide variant in the hagfish sequence. *Nkx2.2* and *Nkx2.8* are paralogues in vertebrates and their upstream sequences seem to also be paralogous in some vertebrates, so two hagfish, two mouse, two lamprey, and two human sequences were aligned (Figure 4.20). Orthologous sequences seem to have diverged in frog and zebrafish.

4.3.4 Validating potential cis-regulatory elements

To validate that the above lamprey *Nkx6A* (n=12)*, Pax7* (n=11)*, Pax6β* (n=13)*,* and *Nkx2.2* (n=14) conserved non-coding sequences are functional enhancers in spinal cord, an enhancer reporter assay was performed in chick. Out of the four, only *Nkx2.2* was a functional enhancer in chick neural tube (Figure4.21). The other three non-coding sequences gave no signal at all.

To try and test if the *Nkx2.2* enhancer function is conserved in lamprey, an enhancer reporter assay was performed in lamprey embryos. Due to technical difficulty with adult brook lamprey husbandry, in vitro fertilisation techniques are not developed for the transgenesis used in *P. marinus*. Hence, I attempted to develop a novel method of whole mount lamprey electroporation. In transgenic lamprey embryos green fluorescent protein expression in the head shows enhancer activity in the ventral CNS that seems to mimic *Nkx2.2* transcript location in the anterioventral CNS (Figure 4.22). However, enhancer activity in the lamprey spinal cord is less convincing because there are three stripes of green fluorescent tissue down its trunk and tail instead of just one adjacent to the floorplate (Figure 4.22 B white arrowheads). It could potentially be labelling the nephric ducts in addition to ventral spinal cord. Out of the four hundred embryos electroporated, only 7% have GFP signal activation and their expressions are heterogeneous. Some of the embryos have enhancer activated in somites, branchial arches, tailbud or ubiquitously along the trunk and tail tissue. With such low reproducibility, this lamprey transgenic method is probably not worth developing further. I also performed a positive control experiment to test the validity of this novel electroporation method using lamprey *Hmx* enhancers which had already been shown to be active in cranial ganglia and spinal cord by the I-SceI meganuclease method (Vasileios Papadogiannis, pers. comm.). My method only replicated their epibranchial ganglion expression but not the spinal cord expression (data not shown), suggesting penetration of construct into the interior of the embryos was poor. With such low reproducibility, the original plan to test lamprey specific open chromatin sequences derived from ATAC-seq for potential enhancers was foiled. Perhaps with more time and optimization of the protocol, more could have been done, but this was not possible with the loss of the 2020 experimental season. In the future, the orthologous *Nkx2.2* enhancer sequence in chick, zebrafish, and mouse could be introduced with a reporter into their respective neural tubes to test if the enhancer has the same function across vertebrates.

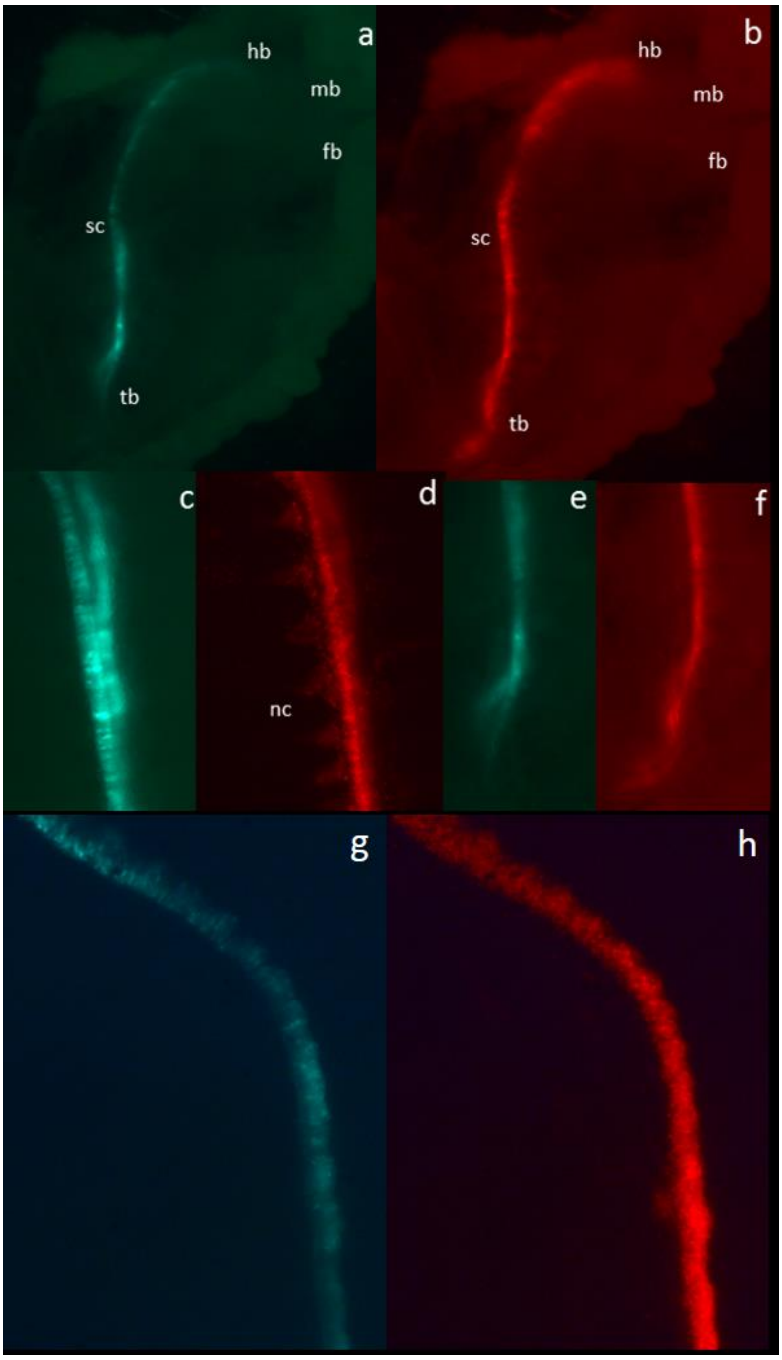


Figure 4.21. Chick neural tube enhancer reporter electroporation. The enhancer reporter GFP (TAG3) is being expressed in green (a,c,e), the red RFP is a nuclear marking positive control (pCI-H2B-RFP) (b,d,f). Whole embryo is shown in a and b, trunk spinal cord in c and d, caudal spinal cord and tailbud in e and f. Panel c appears to have two sides of the neual tube expressing GFP because both sides of the spinal cord was being electroporated. Panel g and h is a different embryo where only one side of the spinal cord is electroporated. fb: forebrain; mb: midbrain; hd: hind-brain; sc: spinal cord; tb: tailbud; nc: neural crest cells migrating laterally. (n=14) The green signal is not bleed-through from the red channel because the red neural crest cells migrating laterally are not seen in the green channel.

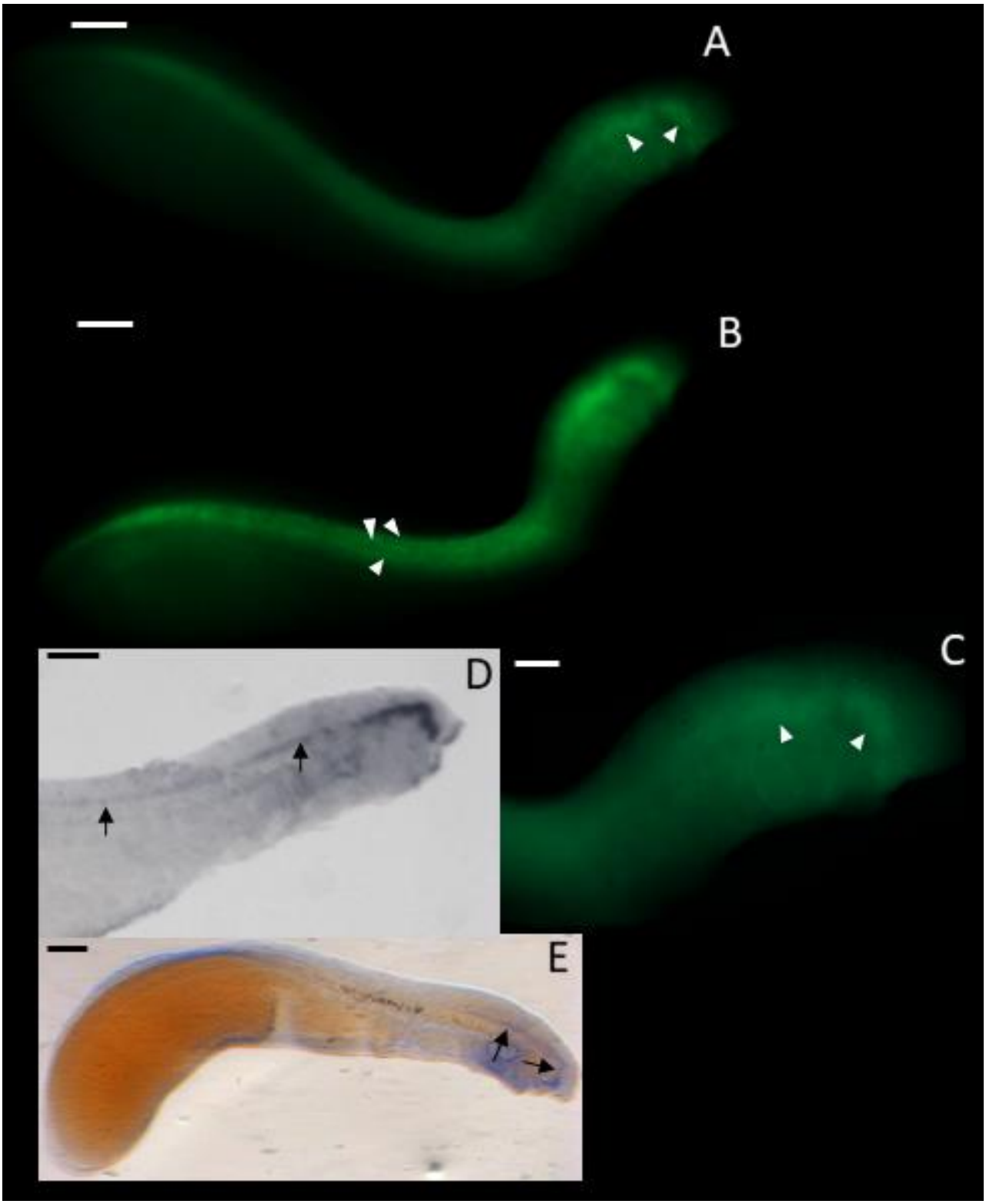


Figure 4.22. Lamprey whole mount enhancer reporter electroporation. Green florescence showing enhancer activation (n=1 out of 400; A-C). White arrowheads pointing at GFP expression. Bright field showing *Nkx2.2* whole mount in situ hybridisation results in the same developmental stage (D, E). Black arrows pointing at *Nkx2.2* expression. Scale bars: 100um.

After identifying enhancers through phylogenetic footprinting and genetic methods, I planned to investigate enhancer function biochemically by performing electromobility shift assays followed by SDS-PAGE then mass spectrometry to identify the nature of the trans-regulatory elements. However, due to technical difficulties then facility shutdown, no results were obtained.

If there are conserved non-coding elements near orthologues then we can provisionally assume that at least some of their transcriptional regulation is conserved. If there are no conserved non-coding elements found then open chromatin sites would suggest differences in regulation between species. There are also possibilities that the trans-regulator is conserved but the cis-regulatory primary sequence has diverged to a level such that it is not detected in cross-genome comparisons. This is especially possible when most of the conserved non-coding sequences are a few hundred base-pairs while trans-regulator binding sites are usually only a few (4-10) base-pairs long. Conversely, if short binding motifs are used for cross-genome comparisons, it generates a lot of false positives. However, more recent data seems to contradict such a hypothesis. Through testing a library of random mutations within an enhancer, Fuqua et al. (2020) found that most mutations alter gene expression and sequence conservation is not an accurate measure for predicting its role in the enhancer.

4.4 Summary, conclusion, and future directions

The first transgenic lamprey was described by Kusakabe et al. (2003) using GFP constructs with a virus promotor or medaka actin promotor. Their microinjection at the fertilisation stage caused a lot of deformity and mortality. The medaka actin promotor construct embryos showed highly mosaic results regarding somite expression, which lasted until the ammocoete larval stage. The next described attempt at lamprey transgenesis was by Parker et al. (2011), adapting a method originally developed in *Xenopus,* the I-SceI meganuclease method mentioned above, and using a mouse cFos minimal promoter. The survival rate of injected lamprey embryos were also low with highly mosaic background expression. The method was later repeated and optimized (Parker et al., 2014b). Other authors have now successfully repeated the same enhancer-reporter assays in lamprey, identifying enhancers of *SoxE1* and *Tfap2B* active in lamprey neural crest cells (Hockman et al., 2019). Neural crest cells are derived from neural plate border cells and have a relatively homogenous progenitor population, while spinal cord progenitor cells are more heterogenous, potentially with a more complex gene regulatory network in

comparison. Since the established I-SceI meganuclease method requires injection of a fertilised egg during or before the first cell division, there is a tight time constraint that could only be feasible when in vitro fertilisation of lamprey is well established. When I was collecting lamprey embryos in the wild, I had no control over when the eggs are fertilised. Instead, I tried to develop my own method of transgenesis in lamprey. I electroporated lamprey embryos at my stages of interest with an enhancer construct with a minimal pTK promoter originally developed in chick (Chen and Streit, 2015). The survival rate is very low and the heterogeneity of expression made the method very difficult to further optimise. The ATAC-seq experiments were done to look for lamprey spinal cord specific cis-regulatory elements. Since my lamprey electroporation method was unsuccessful and the I-SceI meganuclease method was not available, there were no ways to validate potential enhancers. There was also no feasible way to test the potential lamprey specific repressor in vivo associated with *OligB* suggested by my ATAC-seq data. Future experiments are required to accompany my ATAC-seq results to put it into biological context. The potential *OligB* repressor could be binding to Nkx2.2 as Olig2 and Nkx2.2 cross-repress each other in gnathostome (Sun et al., 2003).

The phylogenetic footprinting method has provided some headway into deciphering the gene regulatory network by identifying conserved non-coding sequences across vertebrates. Such a method has been used to identify cis-regulatory elements in lamprey before (Parker et al., 2011, Parker et al., 2014b, Parker et al., 2019b, McEwen et al., 2009). Even though I did not manage find out whether Gli is directly regulating the transcription of *Nkx2.2* and *Pax6* as suggested by chick tissue culture experiments (Soula et al., 2001, Ericson et al., 1997), I did manage to validate an *Nkx2.2* enhancer with highly conserved sequence in vertebrates. The regulation of *Nkx2.2* is well conserved in vertebrates as shown by conserved motifs in the sequence alignment and that this lamprey sequence works as an enhancer in chick embryos, which might explain why *Nkx2.2* is expressed in neural tube p3 cells in different classes of vertebrates. The non-coding sequence is not found is amphioxus, in which *Nkx2.2* is only expressed in cerebral vesicle. *Nkx2.2* has probably adopted new regulation and function in the vertebrate ancestor. I also found other conserved non-coding sequence associated with *Pax6, Pax7,* and *Nkx6.1*. The *Pax6* conserved non-coding sequence could be a potential repressor for *Nkx2.2* to bind, forming the p3/pMN cell type boundary in gnathostomes. The *Nkx6.1* conserved non-coding sequence could be a potential repressor for *Dbx2* to bind, forming the p2/p1 cell type

boundary in gnathostomes. Further experiments are needed to validate their repressor activities.

With more resource, ATAC-seq could have been done on other lamprey tissues to help identify non-spinal-cord-specific cis-regulatory elements, so elements which are actively involved in spinal cord development would stand out. This would make it easier to find potential enhancers to test in vivo. Also, the ATAC-seq experiments could be optimised to improve library quality. Fluorescence-activated cell sorting (FACS) could also be used to separate antibody labelled cell types and obtain cell type resolution of ATAC-seq data instead of the current heterogeneous cell type results. However, with new technology maturing, single cell next generation sequencing would yield more and higher resolution data.

Recent technological advancements have led to high-throughput data-driven functional testing of enhancer sequences. Hoellerer et al. (2020) created a system that can test 2.7 million sequence-function pairs in an experiment making use of a DNA recombinase that can recognise if a transcription factor is binding to the cis-regulatory element/motif. It would then flip the *mCherry* coding sequence around so the *Escherichia coli* would glow fluorescent red. When more than one *E. coli* are containing the same motif sequence and transcription factor then the red fluorescence can provide a quantitative result for how likely the pair would bind. Conversely, if only one *E. coli* for each sequence pair is used then it would give a binary result. The *E. coli* library is then sent for bulk sequencing. The orientation of the *mCherry* and the cis-regulatory element read would show if the sequence is an active binding site. This offers a potential route for screening the large number of potential sequences and transcription factors involved in patterning a complex structure like the spinal cord. With such technology, a higher number of shorter non-coding sequences can be tested. For example, with my current *Nkx2.2* enhancer which is functional in chick, the enhancer sequence can be divided into shorter sequences of around ten base-pairs to test for their functions with all known transcription factors expressing in the spinal cord at that stage, and identify what transcription factor is binding to which lamprey short non-coding sequence. Such a method would provide a much higher accuracy than a computational predictive modelling which is known to have false positives. This method would also be more precise than my original plan of electromobility shift assay followed by SDS-PAGE and mass spectrometry to identify the transcription factors, since the transcription factor and binding sequence are already known when designing the *E. coli* construct.

Chapter5: General Discussion

In theory, the spinal cord is functionally the most conserved part of the central nervous system of the chordates because somatosensory, rhythmic, and motor neurons involved in locomotion and reflex circuits are important for survival, to hunt for food and avoid danger. Chordates come in varying shapes and sizes, some with limbs, some with fins, cephalochordates and agnathans are without lateral fins nor limbs. So how does this anatomical structure have such conserved functions, but it has evolved with such diverse shapes and sizes? To answer this question, I tried comparing adult cell type neuroanatomy by synthesizing published data (Figure 1.4), the results of which I published as part of a review article in 2019. This analysis was useful in framing the distrubutions of cell types in different chordate spinal cords, but was of limited help in understanding their evolution because of the very different spinal cord morphology in different species, and due to insufficient published data for some lineages and cell type markers like neurotransmitters and neuropeptides. So, I sought inspiration from Haeckel's (flawed but still provocative) classical hypothesis "ontogeny recapitulates phylogeny" and proceeded to compare the development of the embryonic spinal cord in different chordates.

A considerable number of studies on gnathostome spinal cords have been published, but far fewer on other chordates. Accordingly, I focused on the less well studied lamprey, a basally divergent vertebrate, and amphioxus, a basally divergent chordate. The phylogenetic positions of these two species allow me to take a glimpse into the evolutionary time when vertebrates began to explode in diversity and compare it to the outgroup, cephalochordates, to gain a perspective on when specific developmental features have evolved. The explosion of diversity time period approximately coincides with genome duplications followed by subsequent gene losses. These genome level evolutionary events could have provided the genetic material used to evolve different gene functions (Table 1.2) and gene regulatory networks, in turn, leading to variations in form and function. Mapping gene duplications, gene regulation and gene function across phylogeny can help identify specific genetic changes involved in morphological evolution.

Considering the spinal cord as a three-dimensional structure, surrounded by other tissues and receiving signals from them (Figure 1.2), I categorized the development of the spinal cord across the AP, DV, and ML axes, while recognising that patterning of these three axes influence one another. I chose to focus principally on the DV axis

patterning as this is where understanding of gene regulatory networks controlling development is most advanced. Amphioxus does not seem to pattern its spinal cord across the DV axis in the same way as vertebrates. While some genes are expressed with DV restrictions showing cells in different DV locations do have different specification, it is difficult to imagine a morphogen gradient being deployed when there are only a few cells across the entire DV axis in the larval stages. Indeed, work from a collaborators group in which I was involved (and was a co-author) showed its neural tube cells seem to differentiate on an individual neuron level rather than as populations of cells with the same specification (Ren et al., 2020). This difference made any kind of morphogen gradient for DV spinal cord patterning a possible vertebrate novelty. More experiments regarding signal concentration variation would be needed to concretely prove this. My study focused on a detailed examination of transcription factor gene expression in the spinal cord, and on manipulation of signalling pathways with small pharmacological molecules as discussed more below.

Quite a number of previous studies have touched on lamprey neural development but their usefulness for understanding spinal cord development is limited by two factors: many focus on brain development or other organs and report little or not at all on spinal cord, and several different species of lamprey have been used. I focused my work on one species *L. planeri,* and adapted gene expression protocols to ensure the spinal cord was well visible and clear. Comparing my data to published work showed that gene expression is generally conserved across different species of lamprey. My amphioxus data also recapitulates published records which also stem from multiple amphioxus species (Table 2.1). This might have been expected based on previous studies and the very similar morphology of the species concerned, but is important to confirm.

I also characterized some undocumented genes, and for many genes examined a wider range of temporal stages in both species. This is important to start to identify when genes are definitively absent in spinal expression in a species, as oppose to what might be differences in the timing of expression. This gene expression characterization (described in detail in Chapter two) showed that *Pax4/6, Pax3/7,* and *Dbx* spinal cord expression are likely to be vertebrate novelties because they are either expressed in gnathostome and lamprey spinal cord but only in the cerebral vesicle in amphioxus (*Pax4/6* and *Pax3/7*) or just in the tailbud (*Dbx*) in amphioxus. These expression patterns do not preclude roles in neural development more broadly: tailbud expression could potentially be involved in AP patterning of neural

tube and cerebral vesicle expression in brain patterning. However, their roles in spinal cord neuron specification are disproved (with the proviso that later larval stages in amphioxus have not been studied: more below). I also found a difference in *Msx* spinal cord expression, which is likely to be ancestral and conserved in chordates but diverged in lamprey, because of its lack of expression in lamprey spinal cord and presence in amphioxus, ascidian and gnathostome spinal cords. *Zic, Irx, SoxB1, Pax2/5/8, Olig, Nkx6, Mnx/Hb9* and *Nkx2* have at least one paralogue expressing in the spinal cord in amphioxus, lamprey, and gnathostomes. These genes are conserved across the three lineages at least to the extent that they are expressed in similar regions of the nervous system.

Most of the amphioxus genes I studied are marking the anterior spinal cord. As considered more extensively in the Introduction above, this region may be better described as the "rhombospinal" region, as it covers both hindbrain and spinal cord. Irrespective of these considerations of homology, the restriction of gene expression to the anterior means patterning of a stretch of posterior amphioxus spinal cord cell types, especially in the larval stage, is still mostly unknown. Studying this region is more technically challenging as much of it develops late in larval life or after metamorphosis. While amphioxus culture is advanced enough to get animals to these stages, numbers can be quite low and gene expression protocols are less well developed. Transgenic and single cell-based methods may be a better approach for studying this, and are considered a bit more below.

Judging by the expression of orthologues across different species, lamprey is more similar to gnathostomes than amphioxus is to vertebrates. These results match my initial predictions, made fundamentally because, anatomically, lamprey developing spinal cord is similar to gnathostomes with a medial ventricular progenitor zone and a lateral mantle zone, while amphioxus developing spinal cord is one-cell thick in the larva. I hypothesized that similar morphology would have similar spinal cord patterning. In addition to exploring how expression and function of these transcription factor genes had evolved in playing a role in spinal cord development, the gene characterization also provided a foundation for choosing which gene to use for signal manipulation experiments in Chapter three. I also examined the evolution of many of these gene families through molecular phylogenetics. The purpose of these analyses was primarily to be certain genes within the same family were being compared and the limited number I have analysed is not sufficient to give insight into genome duplication history. The exception to this may be my analysis of NK cluster

synteny in lamprey. The outcome of this analysis is consistent with the evolutionary model proposed by Simakov et al. (2020), autotetraploidy in the vertebrate ancestor followed by unresolved gene/genome duplications and gene losses in the agnathan lineage, but it still just represents one collection of loci. Genome wide analysis across many species would give a more representative evaluation on the genome duplication and gene loss history.

In Chapter three, I reported manipulation of signalling pathways and observations on how this effected gene expression with a view to understanding how pattern was being regulated (Table 3.2). With the same rationale as applied in my gene expression study described in Chapter two, I also hypothesized that lamprey patterning would be similar to gnathostomes while amphioxus patterning would be different from vertebrates. My bullet points in Section 3.4 summarised the comparative observations. They can be generalized into some similarities and some differences between lamprey and gnathostome DV spinal cord patterning, but the most provocative point is that my results suggest lamprey patterning is different from gnathostomes and that these divergent vertebrate lineages have their own specific facets of gene regulation. This is despite having similarities in gene expression and embryonic anatomy. A potential explanation could be parallel evolution where different vertebrate species maintain similar developmental anatomy and gene expression from the vertebrate ancestor because these spinal cord features are under more evolutionary pressure to produce certain neural cell types and circuitries while gene regulatory network and signal patterning diverged because as long as a functional end product is developed, the methods/mechanisms do not matter that much. Such a function-driven evolutionary hypothesis would also explain why varying shapes and sizes of adult vertebrate spinal cords, i.e. differences in grey matter, white matter, number of cells, locations and sizes of axons running along the AP axis, perform the same function of locomotion and withdrawal reflexes.

Amphioxus results fit my hypothesis for being different from vertebrate patterning, but there is also some conservation, for example Notch is also involved in progenitor marker maintenance and prevention of premature differentiation marker expression. Also, RA is involved in motor neuron specification throughout chordates. In the vertebrate ancestor, Hh was involved in *Nkx2.2* induction and repression of *Pax3* and *Pax7.* RA is required for *Pax3* and *Pax7* expression and RA represses *Nkx2.2.* Figure 5.1 puts these spinal cord patterning features on an evolution timeline. A lot remains unknown as listed in Section 3.4 summary. In the future, signal manipulation

followed by single cell RNA-seq would give a more comprehensive picture on how gene expression and neuron differentiation are being affected.

It is important to recognize that comparing my lamprey data to other vertebrate spinal cord signal manipulation data does have limitations. These are due to difficulty in comparing the developmental timeline across species and because studies in different species have used different methods of manipulating signals. This could well have different impacts on results obtained. There are also differences between different species of gnathostomes, so there are also limitations to lumping all gnathostome data into a group to compare to lamprey. Regardless, I still try to put my lamprey data into biological context.

The conserved aspects of spinal cord patterning could be involved in the development of conserved aspects of spinal neuroanatomy across vertebrates. For example, motor neurons, rhythmic neurons, and sensory neurons, which govern some of the core functions of the spinal cord such as central pattern control and reflex responses. The diverged aspects of spinal cord patterning could be involved in the development of varied form and function of the spinal cord across different species. For example, different vertebrate species vary a lot in size and have different neural circuits adapted to innervate different structures like limbs versus fins, while lamprey has neither, so might be expected not to have these circuits and perhaps some of their cell types. So far, there are very few studies bridging these levels by linking early spinal cord patterning with neural circuit formation and anatomical functions. Two review papers that have tried to bridge the gap between developmental and functional studies by discussing current data and presenting hypotheses (Osseward II and Pfaff, 2019, Lu et al., 2015). However, direct tests of these, or generation of functional data linking development in the embryo with circuit formation and behavioral output are lacking, so I remain cautious with my interpretation of my Chapter three data.

Some inferences, however, can be made. For example, a study showed that *Nkx2.2*$^{+/-}$*; Nkx2.9*$^{-/-}$ compound-mutant mice have defects in rhythmicity and gait of locomotion (Holz et al., 2010). This could explain why the induction and regulation of *Nkx2.2* is conserved in lamprey as locomotion is important for an animal to behave as an active hunter rather than the filter feeder they evolved from. Similarly, lineage tracing experiments showed that *Dbx1* labelled progenitor neurons differentiate into glia cells and commissural interneurons, with a glycinergic/GABAergic character. In the

Dbx1 knockout mouse, there is a disruption in left-right coordination of locomotion (Lanuza et al., 2004). The expression of *Dbx* is affected by different signals in different vertebrate species. Lamprey lost its *Dbx* expression when RA or Notch is inhibited, but not when signals are manipulated in other ways (Figure 3.9 G, H; 3.11 E-H; Table 3.2). RA inhibition does not affect *Dbx* expression in zebrafish (Gribble et al., 2007). RA and Shh induce *Dbx1* and *Dbx2* in chick (Pierani et al., 1999). Since left-right coordination is also an important and conserved part of locomotion, this could mean that the specification of Dbx^{+} progenitor cells, which will become glycinergic / GABAergic neurons, have diverged across different species. Amphioxus is a filter feeder so spinal cord circuitry and locomotion might be different from vertebrates, with a lower demand for controlled swimming: although amphioxus adults can swim very rapidly as an escape response, they do not swim for long periods but rather swim quickly away before burying themselves in sediment. As stated in previous chapters, the genes expressed in the amphioxus spinal cord show differences to vertebrates and there are very few studies linking progenitor or differentiation markers to neuron functions. The development of motor neurons is the most well studied neural type so far (Ren et al., 2020), with sensory and interneurons development unstudied. There are also no studies which explore whether variation in spinal cord size and number of neurons across different vertebrate species leads to different functions. So unfortunately, it is difficult to explore the biological significance of some of my data in terms of neuronal circuit functionality. However, I think my results provide a foundation for future experiments.

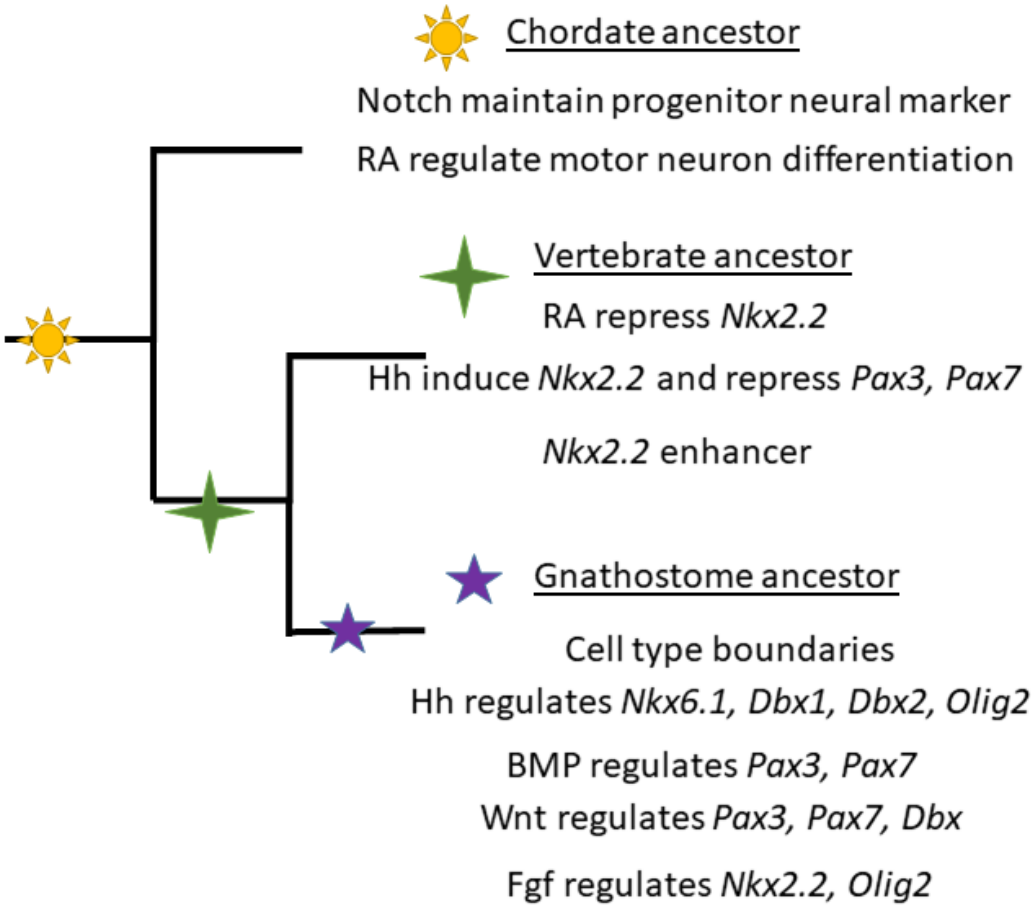


Figure 5.1 Summary of spinal cord patterning evolution. *Nkx2.2* enhancer is referring to the results from Chapter four. Cell type boundaries is referring to the transcription factor cross-repressions that are observed in gnathostomes but not in lamprey, as the ectopic ventralisation of lamprey *Nkx2.2* spatially overlaps with other transcription factor expression.

In addition to studying gene expression and signal patterning, I also attempted to dissect the gene regulation network of spinal cord development at a molecular interaction level, especially when there is evidence in gnathostomes showing that transcription factors cross-repress each other to maintain cell type specification boundaries, for example *Olig* cross repressing *Nkx2.2* to maintain the pMN/p3 boundary (Sun et al., 2003). I performed lamprey ATAC-seq and phylogenetic footprinting to look for potential cis-regulatory elements and find out if transcription factors interact with each other in the way they do in gnathostomes. I identified a potential lamprey *OligB* repressor from ATAC-seq data which could be tested in future experiments. An enhancer associated with *Nkx2.2* was discovered with phylogenetic footprinting and the lamprey sequence was tested and shown to be functional in chick neural tube. This shows this route can identify enhancers, though

perhaps only the most highly conserved ones. The sequence is not found in the amphioxus genome and this suggested that the *Nkx2.2* enhancer is a vertebrate novelty. Future work could validate the enhancer sequence in lamprey embryos when there is access to in vitro fertilization and I-SceI meganuclease mediated transgenesis. It could also be used to identify the nature of transcription factors binding to the enhancer sequence and to establish if it were a route to repression by factors like Olig.

In addition to ATAC-seq and phylogenetic footprinting, a study had revealed that enhancer functions are conserved as long as they have conserved transcription factor binding motifs even when the enhancer sequence have diverged. The authors identified the orthologous region in sponge, *Amphimedon queenslandica,* via histone marks and microsynteny conservation. They found orthologous enhancers in sponge, zebrafish, human, and mouse which are active in the same neural tissues in zebrafish (Wong et al., 2020). Such study is remarkable considering that sponges do not have neural tissue. They also found other sponge enhancers outside of microsyntenic regions which also drive specific zebrafish expression. So perhaps even when conserved non-coding sequence in gnathostomes rarely extends into the agnathan lineage and even fewer into the cephalochordate lineage, less conserved non-coding lamprey or amphioxus sequences are also worth testing their cis-regulatory functions. Higher quality ATAC-seq and high-throughput chromosome conformation capture (Hi-C) data across different chordate species might help with deciding which non-coding sequence to test. From a technical point of view, I think the high GC percentage of the lamprey genome can be used to eliminate potential cis-regulatory elements for testing because high GC sequences are difficult to amplify in the lab (anecdotally guessed to be caused by secondary stuctures formation) and it is difficult to imagine such conserved transcription factor homeodomains would bind to such divergent motifs.

For more detailed dissection of development, some new technologies have been developed. Fluorescent tagged amphioxus lines have been made to trace gene expression in real time (Hua et al., 2018). Such a method is unlikely to be developed in lamprey because of its seven-year-long life cycle and because males and females die soon after spawning. Single cell RNA-seq on amphioxus and lamprey can be done to compare cell types and expression profiles. I propose that short-range relay of sequential signalling events could be a hypothesis for how neurons are patterned in amphioxus, as discovered in the *Ciona intestinalis* tadpole by Stolfi et al. (2011).

Since lamprey and amphioxus are non-model organisms, many of the experimental tools which are used to look for regulatory elements in model organisms are not applicable. For example, a lot of primary antibodies raised against chick/mouse/fish proteins do not work in amphioxus and lamprey, which prevents the easy use of chromatin immunoprecipitation sequencing (ChIP-seq) to identify where specific transcription factors bind. Even when antibodies have been raised specifically against amphioxus proteins, they often still do not work well for these applications (Cedric Patthey and Guang Li, pers. comm.). There are also no established cell cultures for lamprey or amphioxus neural cells, so ATAC-seq results are bound to be heterogeneous unless single cell ATAC-seq can be applied. The lack of working primary antibodies also presents a challenge here. The heterogeneous nature of neural tissue/whole embryos cannot be resolved by the usual method of FACS to separate specific cell types. The lack of transgenic fluorescent protein tagged lines in lamprey also contributes the unusability of FACS.

If I were to plan a study with the most up-to-date technology and abundant resources to answer the same question, I would do single cell RNA-seq on the posterior half of the amphioxus larvae to figure out what the posterior spinal cord is expressing, followed by in situ hybridization to confirm that those transcripts are present in the neural tube. I would take advantage of the rapid developmental time of amphioxus embryos and the new flourescent protein tagging technology in amphioxus to do real-time confocal imaging for lineage tracing experiments. This would be a good replacement for the classical plasma membrane soluble dye, DiI/DiO and would overcome the small embryo size problem of not being able to introduce dye to only a few cells. For example, it would be interesting to study AP spinal cord patterning by observing how tailbud cells contribute to the neural tube and to other germ layers. Tailbud expressing genes like *dbx, axin, wnt,* etc. could be used for this purpose. An historical problem of fluorescent proteins not degrading at the same rate as the endogenous protein, and therefore not being a good representation of protein dynamics, would become an advantage in the lineage tracing case.

Lamprey spinal cord single cell RNA-seq could be done to compare with the published mouse spinal cord single cell RNA-seq data (Delile et al., 2019). More transcripts across different developmental time points could be compared between the two species. I would expect orthologue expression between lamprey and mouse

to be similar, like my Chapter two data. From my experience of dissecting lamprey spinal cord, I would suggest that dissecting fixed embryos would be technically less challenging and biologically more representative because the cells would not be under stress when analysed. Alles et al. (2017) would be a good starting point for developing this method. Since the amphioxus spinal cord is too small to dissect, whole embryo single cell RNA-seq could be done instead. With sufficient depth, this would show all the transcripts within a cell which saves a lot of double in situ hybridization experiments to confirm whether the gene expression coincides within the same neurons. However, double in situ hybridization would still be needed to show spatial information, i.e. where the cells of interest are in the embryo.

To further study the evolution of signalling effects on spinal cord pattern, qPCR could be performed on drug treated amphioxus and lamprey embryos as a complement to in situ data as described in Chapter three. For a larger scale comparison, lamprey spinal cord and whole mount amphioxus single cell RNA-seq could be performed on signal manipulated embryos to see what transcripts are affected and how cell types have changed under signal manipulations. Candidate gene studies like those presented in Chapter three can only observe the genes that have been initially selected, but with single cell RNA-seq, it would be theoretically possible to see all transcripts that have changed their levels. Ideally, double in situ hybridization would follow to provide spatial information and qPCR to confirm increase or decrease of cells expressing certain genes. Experiments could also make use of the knockout amphioxus lines and expand upon the work of Ren et al. (2020) to decipher the gene regulatory networks of other specialised neurons like glycinergic, GABAergic, glutamatergic and other subsets of motor neurons.

Other new conditional knockout or knockin technologies are worth developing in amphioxus as well. Something similar to the UAS-Gal4 system in *Drosophila* or the Cre/loxP system in mice for spatial specific conditional knockout, and heat-shock protein recombinants for temporal specific conditional knockout would be helpful. For example, having a spinal cord conditional signalling pathway gene knockout during the mid-neurula to early larval stage would more accurately study the signal function on neuron differentiation when compared to knockout lines where the whole embryo and developmental time is deprived of that signal. This would also allow the study of leathal knockout genes. The problem of heterogeneity when testing enhancers in amphoxus embryos could potentially be replaced by luciferase

reporter assay in cell culture, but the spatial information within an embryo would be lost.

CRISPR/Cas9 has been done in Northeast Chinese lamprey and the sea lamprey, *Petromyzon marinus,* and both require microinjection at the one-cell stage (Square et al., 2015; Zu et al., 2016). If rearing lamprey adults and in vitro fertilization could be achieved for *Lampetra planeri* then more of these gene editing and transgenic technologies could be applied. Signalling and transcription factor genes could be knocked out to observe their effects on spinal cord patterning. Potential cis-regulatory elements can also be tested in lamprey if zygotes can be accessed. On a genomic level, more and more lamprey species are being sequenced with chromosomal resolution. This could potentially provide insights on evolution within the agnathan lineage and comparing distantly related lamprey species and perhaps hagfish also could be helpful in finding cis-regulatory elements with phylogenetic footprinting. Coupled with better ATAC-seq, this might sidestep the problem of the phylogenetic distance between agnathans and gnathostomes, rendering such enhancers difficult to find.

Acknowledgement
Dorit Hockman helped with lamprey spinal cord ATAC-seq experiments
Irving Ling demonstrated in ovo chick neural tube electroporation experiment.

Guang Li hosted me for two weeks in his lab in Xiamen, China.
Qiongqiong Ren transferred the skills and materials for amphioxus in situ hybridisation/drug treatment to me.

Linda Holland gifted me *Branchiostoma floridae SoxB1a* plasmid.
Sky Yu gifted me *Branchiostoma floridae SoxB1b* and *SoxB1c* plasmid.

Peter Holland gifted me *Branchiostoma floridae Gbx* plasmid and helpful feedbacks along the way.
Sebastian Shimeld has been a very supportive supervisor.
Serena Yichen Dai has been helpful with technical and scientific advice.

References

ACEMEL, R. D., TENA, J. J., IRASTORZA-AZCARATE, I., MARLETAZ, F., GOMEZ-MARIN, C., DE LA CALLE-MUSTIENES, E., BERTRAND, S., DIAZ, S. G., ALDEA, D., AURY, J. M., MANGENOT, S., HOLLAND, P. W., DEVOS, D. P., MAESO, I., ESCRIVA, H. & GOMEZ-SKARMETA, J. L. 2016. A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat Genet,* 48**,** 336-41.

AKAI, J., HALLEY, P. A. & STOREY, K. G. 2005. FGF-dependent Notch signaling maintains the spinal cord stem zone. *Genes Dev,* 19**,** 2877-87.

AL OUSTAH, A., DANESIN, C., KHOURI-FARAH, N., FARRENY, M.-A., ESCALAS, N., COCHARD, P., GLISE, B. & SOULA, C. 2014. Dynamics of sonic hedgehog signaling in the ventral spinal cord are controlled by intrinsic changes in source cells requiring sulfatase 1. *Development,* 141**,** 1392-1403.

ALAYNICK, W. A., JESSELL, T. M. & PFAFF, S. L. 2011. SnapShot: spinal cord development. *Cell,* 146**,** 178-178 e1.

ALBUIXECH-CRESPO, B., LOPEZ-BLANCH, L., BURGUERA, D., MAESO, I., SANCHEZ-ARRONES, L., MORENO-BRAVO, J. A., SOMORJAI, I., PASCUAL-ANAYA, J., PUELLES, E., BOVOLENTA, P., GARCIA-FERNANDEZ, J., PUELLES, L., IRIMIA, M. & FERRAN, J. L. 2017. Molecular regionalization of the developing amphioxus neural tube challenges major partitions of the vertebrate brain. *PLoS Biol,* 15**,** e2001573.

ALLES, J., KARAISKOS, N., PRAKTIKNJO, S. D., GROSSWENDT, S., WAHLE, P., RUFFAULT, P.-L., AYOUB, S., SCHREYER, L., BOLTENGAGEN, A. & BIRCHMEIER, C. J. B. B. 2017. Cell fixation and preservation for droplet-based single-cell transcriptomics. 15**,** 1-14.

ALREFAEI, A. F., MÜNSTERBERG, A. E. & WHEELER, G. N. 2020. FZD10 regulates cell proliferation and mediates Wnt1 induced neurogenesis in the developing spinal cord. *Plos one,* 15**,** e0219721.

ALVAREZ-MEDINA, R., CAYUSO, J., OKUBO, T., TAKADA, S. & MARTI, E. 2008. Wnt canonical pathway restricts graded Shh/Gli patterning activity through the regulation of Gli3 expression. *Development,* 135**,** 237-47.

ANADON, R., ADRIO, F. & RODRIGUEZ-MOLDES, I. 1998. Distribution of GABA immunoreactivity in the central and peripheral nervous system of amphioxus (Branchiostoma lanceolatum Pallas). *Journal of Comparative Neurology,* 401**,** 293-307.

ANADÓN, R., ADRIO, F. & RODRÍGUEZ‐MOLDES, I. 1998. Distribution of GABA immunoreactivity in the central and peripheral nervous system of amphioxus (Branchiostoma lanceolatum Pallas). *Journal of Comparative Neurology,* 401**,** 293-307.

ANDERSON, R. M., LAWRENCE, A. R., STOTTMANN, R. W., BACHILLER, D. & KLINGENSMITH, J. 2002. Chordin and noggin promote organizing centers of forebrain development in the mouse. *Development,* 129**,** 4975-4987.

ANDREWS, M. G., DEL CASTILLO, L. M., OCHOA-BOLTON, E., YAMAUCHI, K., SMOGORZEWSKI, J. & BUTLER, S. J. 2017. BMPs direct sensory interneuron identity in the developing spinal cord using signal-specific not morphogenic activities. *Elife,* 6.

APPEL, B., GIVAN, L. A. & EISEN, J. S. 2001. Delta-Notch signaling and lateral inhibition in zebrafish spinal cord development. *BMC Dev Biol,* 1**,** 13.

ARBEILLE, E., REYNAUD, F., SANYAS, I., BOZON, M., KINDBEITER, K., CAUSERET, F., PIERANI, A., FALK, J., MORET, F. & CASTELLANI, V. 2015. Cerebrospinal fluid-

derived Semaphorin3B orients neuroepithelial cell divisions in the apicobasal axis. *Nat Commun,* 6**,** 6366.

ARENDT, D., TOSCHES, M. A. & MARLOW, H. 2016. From nerve net to nerve ring, nerve cord and brain--evolution of the nervous system. *Nat Rev Neurosci,* 17**,** 61-72.

AUGSBURGER, A., SCHUCHARDT, A., HOSKINS, S., DODD, J. & BUTLER, S. 1999. BMPs as mediators of roof plate repulsion of commissural neurons. *Neuron,* 24**,** 127-141.

AZPIAZU, N., FRASCH, M. J. G. & DEVELOPMENT 1993. tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of Drosophila. 7**,** 1325-1340.

BAI, C. B., AUERBACH, W., LEE, J. S., STEPHEN, D. & JOYNER, A. L. 2002. Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. *Development,* 129**,** 4753-4761.

BARREIRO‐IGLESIAS, A., CORNIDE‐PETRONIO, M. E., ANADÓN, R. & RODICIO, M. C. 2009a. Serotonin and GABA are colocalized in restricted groups of neurons in the larval sea lamprey brain: insights into the early evolution of neurotransmitter colocalization in vertebrates. *Journal of anatomy,* 215**,** 435-443.

BARREIRO‐IGLESIAS, A., VILLAR‐CERVIÑO, V., ANADÓN, R. & RODICIO, M. C. 2009b. Dopamine and γ‐aminobutyric acid are colocalized in restricted groups of neurons in the sea lamprey brain: insights into the early evolution of neurotransmitter colocalization in vertebrates. *Journal of Anatomy,* 215**,** 601-610.

BARTON-OWEN, T. B., FERRIER, D. E. & SOMORJAI, I. M. 2018. Pax3/7 duplicated and diverged independently in amphioxus, the basal chordate lineage. *Scientific reports,* 8**,** 1-11.

BATISTA, M. F. & LEWIS, K. E. 2008. Pax2/8 act redundantly to specify glycinergic and GABAergic fates of multiple spinal interneurons. *Developmental biology,* 323**,** 88-97.

BEASTER-JONES, L., KALTENBACH, S. L., KOOP, D., YUAN, S., CHASTAIN, R. & HOLLAND, L. Z. 2008. Expression of somite segmentation genes in amphioxus: a clock without a wavefront? *Development genes and evolution,* 218**,** 599-611.

BEL-VIALAR, S., ITASAKI, N. & KRUMLAUF, R. 2002. Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. *Development,* 129**,** 5103-15.

BEL-VIALAR, S., MEDEVIELLE, F. & PITUELLO, F. 2007. The on/off of Pax6 controls the tempo of neuronal differentiation in the developing spinal cord. *Developmental biology,* 305**,** 659-673.

BELLO, S. M., MILLO, H., RAJEBHOSALE, M. & PRICE, S. R. 2012. Catenin-dependent cadherin function drives divisional segregation of spinal motor neurons. *Journal of Neuroscience,* 32**,** 490-505.

BERTRAND, N., MÉDEVIELLE, F. & PITUELLO, F. 2000. FGF signalling controls the timing of Pax6 activation in the neural tube. *Development,* 127**,** 4837-4843.

BERTRAND, S., ALDEA, D., OULION, S., SUBIRANA, L., DE LERA, A. R., SOMORJAI, I. & ESCRIVA, H. 2015. Evolution of the Role of RA and FGF Signals in the Control of Somitogenesis in Chordates. *Plos One,* 10.

BERTRAND, S., CAMASSES, A., SOMORJAI, I., BELGACEM, M. R., CHABROL, O., ESCANDE, M.-L., PONTAROTTI, P. & ESCRIVA, H. 2011. Amphioxus FGF signaling predicts the acquisition of vertebrate morphological traits. *Proceedings of the National Academy of Sciences,* 108**,** 9160-9165.

BERTRAND, S. & ESCRIVA, H. 2011. Evolutionary crossroads in developmental biology: amphioxus. *Development,* 138**,** 4819-4830.

BOCINA, I. & SARAGA-BABIC, M. 2006. The notochordal sheath in amphioxus - An ultrastructural and histochemical study. *Collegium Antropologicum,* 30**,** 361-367.

BONE, Q. 1959. The Central Nervous System in Larval Acraniates. *Quarterly Journal of Microscopical Science,* 100**,** 509-527.

BONE, Q. 1960. The Central Nervous System in Amphioxus. *J Comp Neurol,* 115**,** 27-64.

BONNER, J., GRIBBLE, S. L., VEIEN, E. S., NIKOLAUS, O. B., WEIDINGER, G. & DORSKY, R. I. 2008. Proliferation and patterning are mediated independently in the dorsal spinal cord downstream of canonical Wnt signaling. *Developmental biology,* 313**,** 398-407.

BORROMEO, M. D., MEREDITH, D. M., CASTRO, D. S., CHANG, J. C., TUNG, K.-C., GUILLEMOT, F. & JOHNSON, J. E. 2014. A transcription factor network specifying inhibitory versus excitatory neurons in the dorsal spinal cord. *Development,* 141**,** 2803-2812.

BRISCOE, J., PIERANI, A., JESSELL, T. M. & ERICSON, J. 2000. A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell,* 101**,** 435-445.

BRISCOE, J. & SMALL, S. 2015. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development,* 142**,** 3996-4009.

BUCKLEY, D. M., BURROUGHS-GARCIA, J., KRIKS, S., LEWANDOSKI, M. & WATERS, S. T. J. J. O. D. B. 2020. Gbx1 and Gbx2 Are Essential for Normal Patterning and Development of Interneurons and Motor Neurons in the Embryonic Spinal Cord. 8**,** 9.

BULFONE, A., CACCIOPPOLI, C., PARDINI, C., FAEDO, A., MARTINEZ, S. & BANFI, S. 2004. Pcp4l1, a novel gene encoding a Pcp4-like polypeptide, is expressed in specific domains of the developing brain. *Gene expression patterns,* 4**,** 297-301.

BULFONE, A., MENGUZZATO, E., BROCCOLI, V., MARCHITIELLO, A., GATTUSO, C., MARIANI, M., CONSALEZ, G. G., MARTINEZ, S., BALLABIO, A. & BANFI, S. 2000. Barhl1, a gene belonging to a new subfamily of mammalian homeobox genes, is expressed in migrating neurons of the CNS. *Human Molecular Genetics,* 9**,** 1443-1452.

BURRILL, J. D., MORAN, L., GOULDING, M. D. & SAUERESSIG, H. 1997. PAX2 is expressed in multiple spinal cord interneurons, including a population of EN1+ interneurons that require PAX6 for their development. *Development,* 124**,** 4493-4503.

BUTLER, S. J. & DODD, J. 2003. A role for BMP heterodimers in roof plate-mediated repulsion of commissural axons. *Neuron,* 38**,** 389-401.

CAI, J., XU, X., YIN, H., WU, R., MODDERMAN, G., CHEN, Y., JENSEN, J., HUI, C. C. & QIU, M. 2000. Evidence for the differential regulation of Nkx‐6.1 expression in

the ventral spinal cord and foregut by Shh‐dependent and‐independent mechanisms. *genesis,* 27**,** 6-11.

CANDE, J. D., CHOPRA, V. S. & LEVINE, M. J. D. 2009. Evolving enhancer-promoter interactions within the tinman complex of the flour beetle, Tribolium castaneum. 136**,** 3153-3160.

CANDIANI, S., AUGELLO, A., OLIVERI, D., PASSALACQUA, M., PENNATI, R., DE BERNARDI, F. & PESTARINO, M. 2001. Immunocytochemical localization of serotonin in embryos, larvae and adults of the lancelet, Branchiostoma floridae. *The Histochemical Journal,* 33**,** 413-420.

CANDIANI, S., LACALLI, T. C., PARODI, M., OLIVERI, D. & PESTARINO, M. J. D. D. A. O. P. O. T. A. A. O. A. 2008. The cholinergic gene locus in amphioxus: molecular characterization and developmental expression patterns. 237**,** 1399-1411.

CANDIANI, S., MORONTI, L., RAMOINO, P., SCHUBERT, M. & PESTARINO, M. 2012. A neurochemical map of the developing amphioxus nervous system. *Bmc Neuroscience,* 13.

CAÑESTRO, C., HJELMQVIST, L., ALBALAT, R., GARCIA‐FERNÀNDEZ, J., GONZÀLEZ‐DUARTE, R. & JÖRNVALL, H. J. E. J. O. B. 2000. Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural conservation but with unique developmental expression. 267**,** 6511-6518.

CARMENA, A., GISSELBRECHT, S., HARRISON, J., JIMÉNEZ, F., MICHELSON, A. M. J. G. & DEVELOPMENT 1998. Combinatorial signaling codes for the progressive determination of cell fates in the Drosophila embryonic mesoderm. 12**,** 3910-3922.

CARVALHO, J. E., LAHAYE, F., CROCE, J. C. & SCHUBERT, M. 2017. CYP26 function is required for the tissue-specific modulation of retinoic acid signaling during amphioxus development. *The International journal of developmental biology,* 61**,** 733-747.

CASTILLO, H. A., CRAVO, R. M., AZAMBUJA, A. P., SIMÕES-COSTA, M. S., SURA-TRUEBA, S., GONZALEZ, J., SLONIMSKY, E., ALMEIDA, K., ABREU, J. G. & DE ALMEIDA, M. A. A. 2010. Insights into the organization of dorsal spinal cord pathways from an evolutionarily conserved raldh2 intronic enhancer. *Development,* 137**,** 507-518.

CASTRO, A., BECERRA, M., MANSO, M. J. & ANADON, R. 2004. Somatomotor system of the adult amphioxus (Branchiostoma lanceolatum) revealed by an anticalretinin antiserum: An immunocytochemical study. *Journal of Comparative Neurology,* 477**,** 161-171.

CASTRO, A., BECERRA, M., MANSO, M. J. & ANADON, R. 2015. Neuronal organization of the brain in the adult amphioxus (Branchiostoma lanceolatum): A study with acetylated tubulin immunohistochemistry. *Journal of Comparative Neurology,* 523**,** 2211-2232.

CASTRO, A., MANSO, M. J. & ANADON, R. 2003. Distribution of neuropeptide Y immunoreactivity in the central and peripheral nervous systems of amphioxus (Branchiostoma lanceolatum Pallas). *Journal of Comparative Neurology,* 461**,** 350-361.

CASTRO, L. F. C., RASMUSSEN, S. L., HOLLAND, P. W., HOLLAND, N. D. & HOLLAND, L. Z. J. D. B. 2006. A Gbx homeobox gene in amphioxus: insights into ancestry of the ANTP class and evolution of the midbrain/hindbrain boundary. 295**,** 40-51.

CATTELL, M. V., GARNETT, A. T., KLYMKOWSKY, M. W. & MEDEIROS, D. M. 2012. A maternally established S ox B 1/S ox F axis is a conserved feature of chordate germ layer patterning. *Evolution & development,* 14**,** 104-115.

CERNY, R., CATTELL, M., SAUKA-SPENGLER, T., BRONNER-FRASER, M., YU, F. & MEDEIROS, D. M. 2010. Evidence for the prepattern/cooption model of vertebrate jaw evolution. *Proc Natl Acad Sci U S A,* 107**,** 17262-7.

CHANDRAN, S., KATO, H., GERRELI, D., COMPSTON, A., SVENDSEN, C. N. & ALLEN, N. D. 2003. FGF-dependent generation of oligodendrocytes by a hedgehog-independent pathway.

CHEN, J. & STREIT, A. 2015. A medium‐scale assay for enhancer validation in amniotes. *Developmental Dynamics,* 244**,** 1291-1299.

CHIANG, C., LITINGTUNG, Y., LEE, E., YOUNG, K. E., CORDEN, J. L., WESTPHAL, H. & BEACHY, P. A. 1996. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature,* 383**,** 407-13.

CHIZHIKOV, V. V. & MILLEN, K. J. 2004. Control of roof plate formation by Lmx1a in the developing spinal cord. *Development,* 131**,** 2693-2705.

CLOVIS, Y. M., SEO, S. Y., KWON, J.-S., RHEE, J. C., YEO, S., LEE, J. W., LEE, S. & LEE, S.-K. 2016. Chx10 consolidates V2a interneuron identity through two distinct gene repression modes. *Cell reports,* 16**,** 1642-1652.

COHN, M. J. J. N. 2002. Lamprey Hox genes and the origin of jaws. 416**,** 386-387.

COPPOLA, U., CACCAVALE, F., SCELZO, M., HOLLAND, N. D., RISTORATORE, F. & D'ANIELLO, S. 2018. Ran GTPase, an eukaryotic gene novelty, is involved in amphioxus mitosis. *PloS one,* 13**,** e0196930.

CORNIDE-PETRONIO, M. E., FERNÁNDEZ-LÓPEZ, B., BARREIRO-IGLESIAS, A. & RODICIO, M. C. 2014. Traumatic injury induces changes in the expression of

the serotonin 1A receptor in the spinal cord of lampreys. *Neuropharmacology,* 77**,** 369-378.

COX, V. T. & BAYLIES, M. K. 2005. Specification of individual Slouch muscle progenitors in Drosophila requires sequential Wingless signaling.

CROTEAU, L.-P. & KANIA, A. J. J. O. N. M. 2011. Optimisation of in ovo electroporation of the chick neural tube. 201**,** 381-384.

CUI, M., SIRIWON, N., LI, E., DAVIDSON, E. H. & PETER, I. S. 2014. Specific functions of the Wnt signaling system in gene regulatory networks throughout the early sea urchin embryo. *Proceedings of the National Academy of Sciences,* 111**,** E5029-E5038.

DAILEY, S. C., KOZMIKOVA, I. & SOMORJAI, I. M. L. 2017. Amphioxus Sp5 is a member of a conserved Specificity Protein complement and is modulated by Wnt/beta-catenin signalling. *Int J Dev Biol,* 61**,** 723-732.

DANESIN, C., AGIUS, E., ESCALAS, N., AI, X., EMERSON, C., COCHARD, P. & SOULA, C. 2006. Ventral neural progenitors switch toward an oligodendroglial fate in response to increased Sonic hedgehog (Shh) activity: involvement of Sulfatase 1 in modulating Shh signaling in the ventral spinal cord. *Journal of Neuroscience,* 26**,** 5037-5048.

DARRAS, S., FRITZENWANKER, J. H., UHLINGER, K. R., FARRELLY, E., PANI, A. M., HURLEY, I. A., NORRIS, R. P., OSOVITZ, M., TERASAKI, M. & WU, M. 2018. Anteroposterior axis patterning by early canonical Wnt signaling during hemichordate development. *PLoS biology,* 16**,** e2003698.

DASEN, J. S., LIU, J. P. & JESSELL, T. M. 2003. Motor neuron columnar fate imposed by sequential phases of Hox-c activity. *Nature,* 425**,** 926-33.

DASEN, J. S. J. T. I. N. 2018. Evolution of locomotor rhythms. 41**,** 648-651.

DAVIDSON, A. J., POSTLETHWAIT, J. H., YAN, Y.-L., BEIER, D. R., VAN DOREN, C., FOERNZLER, D., CELESTE, A. J., CROSIER, K. E. & CROSIER, P. S. 1999. Isolation of Zebrafish GDF7 and comparative genetic mapping of genes belonging to the growth/differentiation factor 5, 6, 7 subgroup of the TGF-β superfamily. *Genome research,* 9**,** 121-129.

DELILE, J., RAYON, T., MELCHIONDA, M., EDWARDS, A., BRISCOE, J. & SAGNER, A. 2019. Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development*.

DEROBERT, Y., BARATTE, B., LEPAGE, M. & MAZAN, S. J. B. R. B. 2002. Pax6 expression patterns in Lampetra fluviatilis and Scyliorhinus canicula embryos suggest highly conserved roles in the early regionalization of the vertebrate brain. 57**,** 277-280.

DESCHAMPS, J. & DUBOULE, D. 2017. Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev,* 31**,** 1406-1416.

DICK, A., MEIER, A. & HAMMERSCHMIDT, M. 1999. Smad1 and Smad5 have distinct roles during dorsoventral patterning of the zebrafish embryo. *Developmental dynamics,* 216**,** 285-298.

DICKSON, B. J. 2005. Wnts send axons up and down the spinal cord. *Nature neuroscience,* 8**,** 1130-1132.

DIEZ DEL CORRAL, R. & MORALES, A. V. 2017. The Multiple Roles of FGF Signaling in the Developing Spinal Cord. *Front Cell Dev Biol,* 5**,** 58.

DIEZ DEL CORRAL, R., OLIVERA-MARTINEZ, I., GORIELY, A., GALE, E., MADEN, M. & STOREY, K. 2003. Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron,* 40**,** 65-79.

DING, Q., JOSHI, P. S., XIE, Z.-H., XIANG, M. & GAN, L. 2012. BARHL2 transcription factor regulates the ipsilateral/contralateral subtype divergence in postmitotic dI1 neurons of the developing spinal cord. *Proceedings of the National Academy of Sciences,* 109**,** 1566-1571.

DJIOTSA, J., VERBRUGGEN, V., GIACOMOTTO, J., ISHIBASHI, M., MANNING, E., RINKWITZ, S., MANFROID, I., VOZ, M. L. & PEERS, B. 2012. Pax4 is not essential for beta-cell differentiation in zebrafish embryos but modulates alpha-cell generation by repressing arx gene expression. *BMC developmental biology,* 12**,** 37.

DUNCAN, R. N., PANAHI, S., PIOTROWSKI, T. & DORSKY, R. I. 2015. Identification of Wnt genes expressed in neural progenitor zones during zebrafish brain development. *PloS one,* 10**,** e0145810.

DUVAL, N., DAUBAS, P., DE CARBON, C. B., ST CLOMENT, C., TINEVEZ, J.-Y., LOPES, M., RIBES, V. & ROBERT, B. 2014. Msx1 and Msx2 act as essential activators of Atoh1 expression in the murine spinal cord. *Development,* 141**,** 1726-1736.

ECKEI, G., BÖING, M., BRAND-SABERI, B. & MOROSAN-PUOPOLO, G. 2016. Expression pattern of Axin2 during chicken development. *PloS one,* 11**,** e0163610.

ERB, M., LEE, B., SEO, S. Y., LEE, J. W., LEE, S. & LEE, S.-K. 2017. The Isl1-Lhx3 complex promotes motor neuron specification by activating transcriptional pathways that enhance its own expression and formation. *eneuro*.

ERICSON, J., RASHBASS, P., SCHEDL, A., BRENNER-MORTON, S., KAWAKAMI, A., VAN HEYNINGEN, V., JESSELL, T. & BRISCOE, J. J. C. 1997. Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. 90**,** 169-180.

ESCRIVA, H., HOLLAND, N. D., GRONEMEYER, H., LAUDET, V. & HOLLAND, L. Z. 2002a. The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. *Development,* 129**,** 2905-2916.

ESCRIVA, H., MANZON, L., YOUSON, J. & LAUDET, V. 2002b. Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. *Molecular biology and evolution,* 19**,** 1440-1450.

FABIAN, P., KOZMIKOVA, I., KOZMIK, Z. & PANTZARTZI, C. N. 2015. Pax2/5/8 and Pax6 alternative splicing events in basal chordates and vertebrates: a focus on paired box domain. *Frontiers in genetics,* 6**,** 228.

FARRENY, M.-A., AGIUS, E., BEL-VIALAR, S., ESCALAS, N., KHOURI-FARAH, N., SOUKKARIEH, C., DANESIN, C., PITUELLO, F., COCHARD, P. & SOULA, C. J. N. D. 2018. FGF signaling controls Shh-dependent oligodendroglial fate specification in the ventral spinal cord. 13**,** 1-19.

FEINBERG, S., ROURE, A., PIRON, J. & DARRAS, S. 2019. Antero-posterior ectoderm patterning by canonical Wnt signaling during ascidian development. *PLoS genetics,* 15**,** e1008054.

FENG, J., LI, G., LIU, X., WANG, J. & WANG, Y. Q. 2014. Functional analysis of the promoter region of amphioxus beta-actin gene: a useful tool for driving gene expression in vivo. *Mol Biol Rep,* 41**,** 6817-26.

FERNÁNDEZ-LÓPEZ, B., ROMAUS-SANJURJO, D., CORNIDE-PETRONIO, M. E., GÓMEZ-FERNÁNDEZ, S., BARREIRO-IGLESIAS, A. & RODICIO, M. C. 2015. Full anatomical recovery of the dopaminergic system after a complete spinal cord injury in lampreys. *Neural plasticity,* 2015.

FERNÁNDEZ-LÓPEZ, B., VILLAR-CERVIÑO, V., VALLE-MAROTO, S. M., BARREIRO-IGLESIAS, A., ANADÓN, R. & RODICIO, M. C. 2012. The glutamatergic neurons in the spinal cord of the sea lamprey: an in situ hybridization and immunohistochemical study. *PLoS One,* 7**,** e47898.

FERNÁNDEZ‐LÓPEZ, B., SOBRIDO‐CAMEÁN, D., ANADÓN, R., RODICIO, M. & BARREIRO‐IGLESIAS, A. 2017. Restricted co‐localization of glutamate and dopamine in neurons of the adult sea lamprey brain. *Journal of anatomy,* 231**,** 776-784.

FERRIER, D. E. 2016. Evolution of homeobox gene clusters in animals: the giga-cluster and primary vs. secondary clustering. *Frontiers in Ecology and Evolution,* 4**,** 36.

FERRIER, D. E., BROOKE, N. M., PANOPOULOU, G. & HOLLAND, P. W. 2001a. The Mnx homeobox gene class defined by HB9, MNR2 and amphioxus AmphiMnx. *Development Genes & Evolution,* 211.

FERRIER, D. E., MINGUILLON, C., CEBRIÁN, C. & GARCIA-FERNANDEZ, J. 2001b. Amphioxus Evx genes: implications for the evolution of the midbrain–hindbrain boundary and the chordate tailbud. *Developmental biology,* 237**,** 270-281.

FIOR, R., MAXWELL, A. A., MA, T. P., VEZZARO, A., MOENS, C. B., AMACHER, S. L. & LEWIS, J. 2012. The differentiation and movement of presomitic mesoderm progenitor cells are controlled by Mesogenin 1. *Development,* 139**,** 4656-4665.

FOGARTY, M., RICHARDSON, W. D. & KESSARIS, N. 2005. A subset of oligodendrocytes generated from radial glia in the dorsal spinal cord. *Development,* 132**,** 1951-1959.

FORNES, O., CASTRO-MONDRAGON, J. A., KHAN, A., VAN DER LEE, R., ZHANG, X., RICHMOND, P. A., MODI, B. P., CORREARD, S., GHEORGHE, M. & BARANAŠIĆ, D. 2020. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic acids research,* 48**,** D87-D92.

FRIESE, A., KALTSCHMIDT, J. A., LADLE, D. R., SIGRIST, M., JESSELL, T. M. & ARBER, S. 2009. Gamma and alpha motor neurons distinguished by expression of transcription factor Err3. *Proceedings of the National Academy of Sciences,* 106**,** 13588-13593.

FU, H., QI, Y., TAN, M., CAI, J., TAKEBAYASHI, H., NAKAFUKU, M., RICHARDSON, W. & QIU, M. 2002. Dual origin of spinal oligodendrocyte progenitors and evidence for the cooperative role of Olig2 and Nkx2. 2 in the control of oligodendrocyte differentiation. *Development,* 129**,** 681-693.

FUQUA, T., JORDAN, J., VAN BREUGEL, M. E., HALAVATYI, A., TISCHER, C., POLIDORO, P., ABE, N., TSAI, A., MANN, R. S. & STERN, D. L. J. N. 2020. Dense and pleiotropic regulatory information in a developmental enhancer. 587**,** 235-239.

FURLONG, R. F. & HOLLAND, P. W. J. P. T. O. T. R. S. O. L. S. B. B. S. 2002. Were vertebrates octoploid? 357**,** 531-544.

GANS, C. & NORTHCUTT, R. G. 1983. Neural Crest and the Origin of Vertebrates - a New Head. *Science,* 220**,** 268-273.

GARCIA-FERNANDEZ, J. & HOLLAND, P. W. 1994. Archetypal organization of the amphioxus Hox gene cluster. *Nature,* 370**,** 563-6.

GARD, C., CURTO, G. G., FRARMA, Y. E.-M., CHOLLET, E., DUVAL, N., AUZIÉ, V., AURADÉ, F., VIGIER, L., RELAIX, F. & PIERANI, A. 2017. Pax3-and Pax7-mediated Dbx1 regulation orchestrates the patterning of intermediate spinal interneurons. *Developmental biology,* 432**,** 24-33.

GAUNT, S. J. 2017. Gdf11/Smad signalling and Cdx proteins cooperate to activate the Hoxc8 early enhancer in HepG2 cells. *Int J Dev Biol,* 61**,** 427-432.

GIL-GALVEZ, A., JIMENEZ-GANCEDO, S., ACEMEL, R. D., BERTRAND, S., SCHUBERT, M., ESCRIVA, H., TENA, J. J. & GOMEZ-SKARMETA, J. L. 2020. Gain of gene regulatory network interconnectivity at the origin of vertebrates. *bioRxiv*.

GILLIS, J. A. & HALL, B. K. 2016. A shared role for sonic hedgehog signalling in patterning chondrichthyan gill arch appendages and tetrapod limbs. *Development,* 143**,** 1313-1317.

GLARDON, S., HOLLAND, L. Z., GEHRING, W. J. & HOLLAND, N. D. 1998. Isolation and developmental expression of the amphioxus Pax-6 gene (AmphiPax-6): insights into eye and photoreceptor evolution. *Development,* 125**,** 2701-2710.

GOODRICH, L. V., JOHNSON, R. L., MILENKOVIC, L., MCMAHON, J. A. & SCOTT, M. P. 1996. Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. *Genes & development,* 10**,** 301-312.

GOSTLING, N. J. & SHIMELD, S. M. 2003. Protochordate Zic genes define primitive somite compartments and highlight molecular changes underlying neural crest evolution. *Evolution & development,* 5**,** 136-144.

GOTO, H., KIMMEY, S. C., ROW, R. H., MATUS, D. Q. & MARTIN, B. L. J. D. 2017. FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from

tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition. 144**,** 1412-1424.

GOULDING, M. 2009. Circuits controlling vertebrate locomotion: moving in a new direction. *Nat Rev Neurosci,* 10**,** 507-18.

GOUTI, M., METZIS, V. & BRISCOE, J. 2015. The route to spinal cord cell types: a tale of signals and switches. *Trends Genet,* 31**,** 282-9.

GRAHAM, V., KHUDYAKOV, J., ELLIS, P. & PEVNY, L. 2003. SOX2 functions to maintain neural progenitor identity. *Neuron,* 39**,** 749-65.

GRANDEL, H. & BRAND, M. 2013. Comparative aspects of adult neural stem cell activity in vertebrates. *Dev Genes Evol,* 223**,** 131-47.

GREEN, S. A., SIMOES-COSTA, M. & BRONNER, M. E. 2015. Evolution of vertebrates as viewed from the crest. *Nature,* 520**,** 474.

GRIBBLE, S. L., NIKOLAUS, O. B. & DORSKY, R. I. 2007. Regulation and function of Dbx genes in the zebrafish spinal cord. *Developmental dynamics: an official publication of the American Association of Anatomists,* 236**,** 3472-3483.

GRUMOLATO, L., LIU, G., MONG, P., MUDBHARY, R., BISWAS, R., ARROYAVE, R., VIJAYAKUMAR, S., ECONOMIDES, A. N. & AARONSON, S. A. 2010. Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. *Genes & development,* 24**,** 2517-2530.

GUERIN, A., D'AUBENTON-CARAFA, Y., MARRAKCHI, E., DA SILVA, C., WINCKER, P., MAZAN, S. & RETAUX, S. 2009. Neurodevelopment genes in lampreys reveal trends for forebrain evolution in craniates. *PLoS One,* 4**,** e5374.

HÄMMERLE, B. & TEJEDOR, F. J. 2007. A novel function of DELTA-NOTCH signalling mediates the transition from proliferation to neurogenesis in neural progenitor cells. *PloS one,* 2**,** e1169.

HAMMOND, K. L., BAXENDALE, S., MCCAULEY, D. W., INGHAM, P. W. & WHITFIELD, T. T. 2009. Expression of patched, prdm1 and engrailed in the lamprey somite reveals conserved responses to Hedgehog signaling. *Evol Dev,* 11**,** 27-40.

HERMAN, P. E., PAPATHEODOROU, A., BRYANT, S. A., WATERBURY, C. K. M., HERDY, J. R., ARCESE, A. A., BUXBAUM, J. D., SMITH, J. J., MORGAN, J. R. & BLOOM, O. 2018. Highly conserved molecular pathways, including Wnt signaling, promote functional recovery from spinal cord injury in lampreys. *Sci Rep,* 8**,** 742.

HOCKMAN, D., CHONG-MORRISON, V., GREEN, S. A., GAVRIOUCHKINA, D., CANDIDO-FERREIRA, I., LING, I. T., WILLIAMS, R. M., AMEMIYA, C. T., SMITH, J. J. & BRONNER, M. E. J. N. C. 2019. A genome-wide assessment of the ancestral neural crest gene regulatory network. 10**,** 1-15.

HOELLERER, S., PAPAXANTHOS, L., GUMPINGER, A. C., FISCHER, K., BEISEL, C., BORGWARDT, K., BENENSON, Y. & JESCHEK, M. J. B. 2020. Large-scale DNA-based phenotypic recording and deep learning enable highly accurate sequence-function mapping.

HOLLAND, L. Z., CARVALHO, J. E., ESCRIVA, H., LAUDET, V., SCHUBERT, M., SHIMELD, S. M. & YU, J.-K. 2013. Evolution of bilaterian central nervous systems: a single origin? *EvoDevo,* 4**,** 27.

HOLLAND, L. Z. & HOLLAND, N. D. J. C. O. I. N. 1999. Chordate origins of the vertebrate central nervous system. 9**,** 596-602.

HOLLAND, L. Z., RACHED, L. A., TAMME, R., HOLLAND, N. D., INOKO, H., SHIINA, T., BURGTORF, C. & LARDELLI, M. 2001a. Characterization and developmental expression of the amphioxus homolog of Notch (AmphiNotch): evolutionary

conservation of multiple expression domains in amphioxus and vertebrates. *Developmental biology,* 232**,** 493-507.

HOLLAND, L. Z., RACHED, L. A., TAMME, R., HOLLAND, N. D., KORTSCHAK, D., INOKO, H., SHIINA, T., BURGTORF, C. & LARDELLI, M. 2001b. Characterization and developmental expression of the amphioxus homolog of Notch (AmphiNotch): evolutionary conservation of multiple expression domains in amphioxus and vertebrates. *Dev Biol,* 232**,** 493-507.

HOLLAND, L. Z., SCHUBERT, M., HOLLAND, N. & NEUMAN, T. 2000. Evolutionary conservation of the presumptive neural plate markers AmphiSox1/2/3 and AmphiNeurogenin in the invertebrate chordate amphioxus. *Developmental biology,* 226**,** 18-33.

HOLLAND, L. Z., SCHUBERT, M., KOZMIK, Z. & HOLLAND, N. D. 1999. AmphiPax3/7, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and the possible evolutionary precursor of definitive vertebrate neural crest. *Evolution & development,* 1**,** 153-165.

HOLLAND, P. W. 2015. Did homeobox gene duplications contribute to the Cambrian explosion? *Zoological letters,* 1**,** 1-8.

HOLTZ, A. M., PETERSON, K. A., NISHI, Y., MORIN, S., SONG, J. Y., CHARRON, F., MCMAHON, A. P. & ALLEN, B. L. 2013. Essential role for ligand-dependent feedback antagonism of vertebrate hedgehog signaling by PTCH1, PTCH2 and HHIP1 during neural patterning. *Development,* 140**,** 3423-3434.

HOLZ, A., KOLLMUS, H., RYGE, J., NIEDERKOFLER, V., DIAS, J., ERICSON, J., STOECKLI, E. T., KIEHN, O. & ARNOLD, H.-H. J. D. 2010. The transcription factors Nkx2. 2 and Nkx2. 9 play a novel role in floor plate development and commissural axon guidance. 137**,** 4249-4260.

HOOVER, A. N., WYNKOOP, A., ZENG, H., JIA, J., NISWANDER, L. A. & LIU, A. 2008. C2cd3 is required for cilia formation and Hedgehog signaling in mouse. *Development,* 135**,** 4049-4058.

HOUWELING, A. C., DILDROP, R., PETERS, T., MUMMENHOFF, J., MOORMAN, A. F., RÜTHER, U. & CHRISTOFFELS, V. M. 2001. Gene and cluster-specific expression of the Iroquois family members during mouse development. *Mechanisms of development,* 107**,** 169-174.

HU, G., LI, G., WANG, H. & WANG, Y. 2017. Hedgehog participates in the establishment of left-right asymmetry during amphioxus development by controlling Cerberus expression. *Development,* 144**,** 4694-4703.

HU, Q., UENO, N. & BEHRINGER, R. R. 2004. Restriction of BMP4 activity domains in the developing neural tube of the mouse embryo. *EMBO reports,* 5**,** 734-739.

HUA, J., LI, G. & WANG, Y. 2018. Construction of a 2A Peptide-linked Multicistronic Expression Vector for Amphioxus. *J Xiamen,* 57.

HUANG, M., HUANG, T., XIANG, Y., XIE, Z., CHEN, Y., YAN, R., XU, J. & CHENG, L. 2008. Ptf1a, Lbx1 and Pax2 coordinate glycinergic and peptidergic transmitter phenotypes in dorsal spinal inhibitory neurons. *Developmental biology,* 322**,** 394-405.

HUI, J. H., MCDOUGALL, C., MONTEIRO, A. S., HOLLAND, P. W., ARENDT, D., BALAVOINE, G., FERRIER, D. E. J. M. B. & EVOLUTION 2012. Extensive chordate and annelid macrosynteny reveals ancestral homeobox gene organization. 29**,** 157-165.

HUTCHINSON, S. A. & EISEN, J. S. 2006. Islet1 and Islet2 have equivalent abilities to promote motoneuron formation and to specify motoneuron subtype identity. *Development,* 133**,** 2137-2147.

IKEYA, M., LEE, S. M., JOHNSON, J. E., MCMAHON, A. P. & TAKADA, S. J. N. 1997. Wnt signalling required for expansion of neural crest and CNS progenitors. 389**,** 966-970.

IKUTA, T., YOSHIDA, N., SATOH, N. & SAIGA, H. 2004. Ciona intestinalis Hox gene cluster: Its dispersed structure and residual colinear expression in development. *Proc Natl Acad Sci U S A,* 101**,** 15118-23.

ISOMURA, A., OGUSHI, F., KORI, H. & KAGEYAMA, R. 2017. Optogenetic perturbation and bioluminescence imaging to analyze cell-to-cell transfer of oscillatory information. *Genes Dev,* 31**,** 524-535.

JACKMAN, W. R. & KIMMEL, C. B. 2002. Coincident iterated gene expression in the amphioxus neural tube. *Evolution & development,* 4**,** 366-374.

JACOBS, C. T. & HUANG, P. 2019. Notch signalling maintains Hedgehog responsiveness via a Gli-dependent mechanism during spinal cord patterning in zebrafish. *Elife,* 8**,** e49252.

JAGLA, K., BELLARD, M. & FRASCH, M. J. B. 2001. A cluster of Drosophila homeobox genes involved in mesoderm differentiation programs. 23**,** 125-133.

JANDZIK, D., HAWKINS, M. B., CATTELL, M. V., CERNY, R., SQUARE, T. A. & MEDEIROS, D. M. 2014. Roles for FGF in lamprey pharyngeal pouch formation and skeletogenesis highlight ancestral functions in the vertebrate head. *Development,* 141**,** 629-638.

JING, W., GUANG, L., GUANG-HUI, Q., JUN-HAO, H. & YI-QUAN, W. 2016. Expression analysis of eight amphioxus genes involved in the Wnt/β-catenin signaling pathway. *Zoological research,* 37**,** 136.

JUÁREZ-MORALES, J. L., SCHULTE, C. J., PEZOA, S. A., VALLEJO, G. K., HILINSKI, W. C., ENGLAND, S. J., DE JAGER, S. & LEWIS, K. E. 2016. Evx1 and Evx2 specify

excitatory neurotransmitter fates and suppress inhibitory fates through a Pax2-independent mechanism. *Neural development,* 11**,** 5.

KALTENBACH, S. L., HOLLAND, L. Z., HOLLAND, N. D. & KOOP, D. 2009a. Developmental expression of the three iroquois genes of amphioxus (BfIrxA, BfIrxB, and BfIrxC) with special attention to the gastrula organizer and anteroposterior boundaries in the central nervous system. *Gene Expression Patterns,* 9**,** 329-334.

KALTENBACH, S. L., YU, J. K. & HOLLAND, N. D. 2009b. The origin and migration of the earliest‐developing sensory neurons in the peripheral nervous system of amphioxus. *Evolution & development,* 11**,** 142-151.

KANO, S., XIAO, J. H., OSORIO, J., EKKER, M., HADZHIEV, Y., MULLER, F., CASANE, D., MAGDELENAT, G. & RETAUX, S. 2010. Two Lamprey Hedgehog Genes Share Non-Coding Regulatory Sequences and Expression Patterns with Gnathostome Hedgehogs. *Plos One,* 5.

KATSUYAMA, Y., WADA, S., YASUGI, S. & SAIGA, H. 1995. Expression of the labial group Hox gene HrHox-1 and its alteration induced by retinoic acid in development of the ascidian Halocynthia roretzi. *Development,* 121**,** 3197-205.

KENNEDY, T. E., WANG, H., MARSHALL, W. & TESSIER-LAVIGNE, M. 2006. Axon guidance by diffusible chemoattractants: a gradient of netrin protein in the developing spinal cord. *Journal of Neuroscience,* 26**,** 8866-8874.

KERA, S. A., AGERWALA, S. M. & HORNE, J. H. J. Z. 2010. The temporal resolution of in vivo electroporation in zebrafish: a method for time-resolved loss of function. 7**,** 97-108.

KHADKA, A., MARTÍNEZ-BARTOLOMÉ, M., BURR, S. D. & RANGE, R. C. 2018. A novel gene's role in an ancient mechanism: secreted Frizzled-related protein 1 is a critical component in the anterior–posterior Wnt signaling network that governs the establishment of the anterior neuroectoderm in sea urchin embryos. *EvoDevo,* 9**,** 1-15.

KIEHN, O. 2016. Decoding the organization of spinal circuits that control locomotion. *Nat Rev Neurosci,* 17**,** 224-38.

KIM, H., SHIN, J., KIM, S., POLING, J., PARK, H. C. & APPEL, B. 2008. Notch-regulated oligodendrocyte specification from radial glia in the spinal cord of zebrafish embryos. *Developmental dynamics: an official publication of the American Association of Anatomists,* 237**,** 2081-2089.

KIM, N., PARK, C., JEONG, Y. & SONG, M.-R. 2015. Functional diversification of motor neuron-specific Isl1 enhancers during evolution. *PLoS Genet,* 11**,** e1005560.

KOIDE, T., HAYATA, T. & CHO, K. W. 2006. Negative regulation of Hedgehog signaling by the cholesterogenic enzyme 7-dehydrocholesterol reductase. *Development,* 133**,** 2395-2405.

KONG, J. H., YANG, L., DESSAUD, E., CHUANG, K., MOORE, D. M., ROHATGI, R., BRISCOE, J. & NOVITCH, B. G. 2015. Notch activity modulates the responsiveness of neural progenitors to sonic hedgehog signaling. *Dev Cell,* 33**,** 373-87.

KOOP, D., CHEN, J., THEODOSIOU, M., CARVALHO, J. E., ALVAREZ, S., DE LERA, A. R., HOLLAND, L. Z. & SCHUBERT, M. 2014. Roles of retinoic acid and Tbx1/10 in pharyngeal segmentation: amphioxus and the ancestral chordate condition. *EvoDevo,* 5**,** 1-16.

KOZMIK, Z., HOLLAND, N. D., KALOUSOVA, A., PACES, J., SCHUBERT, M. & HOLLAND, L. Z. 1999. Characterization of an amphioxus paired box gene, AmphiPax2/5/8: developmental expression patterns in optic support cells, nephridium, thyroid-like structures and pharyngeal gill slits, but not in the midbrain-hindbrain boundary region. *Development,* 126**,** 1295-1304.

KOZMIKOVA, I., CANDIANI, S., FABIAN, P., GURSKA, D. & KOZMIK, Z. 2013. Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates. *Developmental biology,* 382**,** 538-554.

KRIKS, S., LANUZA, G. M., MIZUGUCHI, R., NAKAFUKU, M. & GOULDING, M. 2005. Gsh2 is required for the repression of Ngn1 and specification of dorsal interneuron fate in the spinal cord. *Development,* 132**,** 2991-3002.

KRUMLAUF, R. 2016. Hox Genes and the Hindbrain: A Study in Segments. *Curr Top Dev Biol,* 116**,** 581-96.

KUMAR, S., STECHER, G. & TAMURA, K. 2015. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution (submitted).

KURAKU, S., MEYER, A. & KURATANI, S. 2009. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol Biol Evol,* 26**,** 47-59.

KURATANI, S., UEKI, T., HIRANO, S. & AIZAWA, S. 1998. Rostral truncation of a cyclostome, Lampetra japonica, induced by all‐trans retinoic acid defines the head/trunk interface of the vertebrate body. *Developmental dynamics: an official publication of the American Association of Anatomists,* 211**,** 35-51.

KUSAKABE, R., KURAKU, S. & KURATANI, S. 2011. Expression and interaction of muscle-related genes in the lamprey imply the evolutionary scenario for

vertebrate skeletal muscle, in association with the acquisition of the neck and fins. *Developmental biology,* 350**,** 217-227.

KUSAKABE, R. & KURATANI, S. 2005. The origin of epaxial/hypaxial distinction of trunk skeletal muscle: Implications from the gene expression patterns in the lamprey embryos. *Zoological Science,* 22**,** 1402-1402.

KUSAKABE, R., TOCHINAI, S., KURATANI, S. J. J. O. E. Z. P. B. M. & EVOLUTION, D. 2003. Expression of foreign genes in lamprey embryos: an approach to study evolutionary changes in gene regulation. 296**,** 87-97.

KUTEJOVA, E., SASAI, N., SHAH, A., GOUTI, M. & BRISCOE, J. 2016. Neural Progenitors Adopt Specific Identities by Directly Repressing All Alternative Progenitor Transcriptional Programs. *Dev Cell,* 36**,** 639-53.

LAI, H. C., SEAL, R. P. & JOHNSON, J. E. 2016. Making sense out of spinal cord somatosensory development. *Development,* 143**,** 3434-3448.

LANUZA, G. M., GOSGNACH, S., PIERANI, A., JESSELL, T. M. & GOULDING, M. J. N. 2004. Genetic identification of spinal interneurons that coordinate left-right locomotor activity necessary for walking movements. 42**,** 375-386.

LARA-RAMIREZ, R., PEREZ-GONZALEZ, C., ANSELMI, C., PATTHEY, C. & SHIMELD, S. M. 2019. A Notch-regulated proliferative stem cell zone in the developing spinal cord is an ancestral vertebrate trait. *Development,* 146.

LARA-RAMÍREZ, R., PONCELET, G., PATTHEY, C., SHIMELD, S. M. J. D. G. & EVOLUTION 2017. The structure, splicing, synteny and expression of lamprey COE genes and the evolution of the COE gene family in chordates. 227**,** 319-338.

LARA‐RAMÍREZ, R., PATTHEY, C. & SHIMELD, S. M. 2015. Characterization of two neurogenin genes from the brook lamprey lampetra planeri and their

expression in the lamprey nervous system. *Developmental Dynamics,* 244**,** 1096-1108.

LARSSON, M. 2017. Pax2 is persistently expressed by GABAergic neurons throughout the adult rat dorsal horn. *Neuroscience letters,* 638**,** 96-101.

LAUSSU, J., AUDOUARD, C., KISCHEL, A., ASSIS-NASCIMENTO, P., ESCALAS, N., LIEBL, D. J., SOULA, C. & DAVY, A. 2017. Eph/Ephrin Signaling Controls Progenitor Identities In The Ventral Spinal Cord. *Neural Dev,* 12**,** 10.

LE DREAU, G., SAADE, M., GUTIERREZ-VALLEJO, I. & MARTI, E. 2014. The strength of SMAD1/5 activity determines the mode of stem cell division in the developing spinal cord. *J Cell Biol,* 204**,** 591-605.

LEE, J., WU, Y., QI, Y., XUE, H., LIU, Y., SCHEEL, D., GERMAN, M., QIU, M., GUILLEMOT, F. & RAO, M. 2003. Neurogenin3 participates in gliogenesis in the developing vertebrate spinal cord. *Developmental biology,* 253**,** 84-98.

LEE, S.-K., LEE, B., RUIZ, E. C. & PFAFF, S. L. 2005. Olig2 and Ngn2 function in opposition to modulate gene expression in motor neuron progenitor cells. *Genes & development,* 19**,** 282-294.

LEI, Q., JEONG, Y., MISRA, K., LI, S., ZELMAN, A. K., EPSTEIN, D. J. & MATISE, M. P. J. D. C. 2006. Wnt signaling inhibitors regulate the transcriptional response to morphogenetic Shh-Gli signaling in the neural tube. 11**,** 325-337.

LEWIS, K. E., BATES, J. & EISEN, J. S. 2005. Regulation of iro3 expression in the zebrafish spinal cord. *Developmental dynamics: an official publication of the American Association of Anatomists,* 232**,** 140-148.

LI, G., FENG, J., LEI, Y., WANG, J., WANG, H., SHANG, L. K., LIU, D. T., ZHAO, H., ZHU, Y. & WANG, Y. Q. 2014a. Mutagenesis at specific genomic loci of amphioxus Branchiostoma belcheri using TALEN method. *J Genet Genomics,* 41**,** 215-9.

LI, G., SHU, Z. & WANG, Y. 2013. Year-round reproduction and induced spawning of Chinese amphioxus, Branchiostoma belcheri, in laboratory. *PLoS One,* 8**,** e75461.

LI, G., WANG, J., YUAN, L., WANG, H. & WANG, Y. Q. 2015. A simple method for selecting spawning‐ready individuals out from laboratorial cultured amphioxus population. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution,* 324**,** 629-635.

LI, G., YANG, X., SHU, Z., CHEN, X. & WANG, Y. 2012. Consecutive spawnings of Chinese amphioxus, Branchiostoma belcheri, in captivity. *PLoS One,* 7**,** e50838.

LI, X., LIU, Z., QIU, M. & YANG, Z. 2014b. Sp8 plays a supplementary role to Pax6 in establishing the pMN/p3 domain boundary in the spinal cord. *Development,* 141**,** 2875-2884.

LI, Y., NONG, W., BARIL, T., YIP, H. Y., SWALE, T., HAYWARD, A., FERRIER, D. E. & HUI, J. H. J. B. G. 2020. Reconstruction of ancient homeobox gene linkages inferred from a new high-quality assembly of the Hong Kong oyster (Magallana hongkongensis) genome. 21**,** 1-17.

LI, Y., TZATZALOS, E., KWAN, K. Y., GRUMET, M. & CAI, L. 2016. Transcriptional regulation of Notch1 expression by Nkx6. 1 in Neural Stem/Progenitor cells during ventral spinal cord development. *Scientific reports,* 6**,** 38665.

LIN, G. & SLACK, J. M. 2008. Requirement for Wnt and FGF signaling in Xenopus tadpole tail regeneration. *Developmental biology,* 316**,** 323-335.

LIN, J., WANG, C., YANG, C., FU, S. & REDIES, C. 2016. Pax3 and Pax7 interact reciprocally and regulate the expression of cadherin‐7 through inducing

neuron differentiation in the developing chicken spinal cord. *Journal of Comparative Neurology,* 524**,** 940-962.

LIN, Y., CAI, Z., HUANG, S., YANG, L., WANG, C., LIU, Z., CAO, J., AN, Y., ZHANG, H. J. E. & DEVELOPMENT 2009. Ptc, Smo, Sufu, and the Hedgehog signaling pathway in amphioxus. 11**,** 710-718.

LIU, J. P., LAUFER, E. & JESSELL, T. M. 2001. Assigning the positional identity of spinal motor neurons: rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. *Neuron,* 32**,** 997-1012.

LIU, X., LI, G., FENG, J., YANG, X. & WANG, Y. Q. 2013. An efficient microinjection method for unfertilized eggs of Asian amphioxus Branchiostoma belcheri. *Dev Genes Evol,* 223**,** 269-78.

LIU, Z., HU, X., HUANG, C., ZHENG, K., TAKEBAYASHI, H., CAO, C. & QIU, M. 2014. Olig3 is not involved in the ventral patterning of spinal cord. *PloS one,* 9**,** e111076.

LOGAN, C., WINGATE, R. J., MCKAY, I. J. & LUMSDEN, A. 1998. Tlx-1 and Tlx-3 homeobox gene expression in cranial sensory ganglia and hindbrain of the chick embryo: markers of patterned connectivity. *Journal of Neuroscience,* 18**,** 5389-5402.

LOWE, C. J., WU, M., SALIC, A., EVANS, L., LANDER, E., STANGE-THOMANN, N., GRUBER, C. E., GERHART, J. & KIRSCHNER, M. 2003. Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. *Cell,* 113**,** 853-865.

LU, D. C., NIU, T. & ALAYNICK, W. A. J. F. I. M. N. 2015. Molecular and cellular development of spinal cord locomotor circuitry. 8**,** 25.

LU, Q. R., SUN, T., ZHU, Z., MA, N., GARCIA, M., STILES, C. D. & ROWITCH, D. H. 2002. Common developmental requirement for Olig function indicates a motor neuron/oligodendrocyte connection. *Cell,* 109**,** 75-86.

LU, T. M., LUO, Y. J. & YU, J. K. 2012. BMP and Delta/Notch signaling control the development of amphioxus epidermal sensory neurons: insights into the evolution of the peripheral sensory system. *Development,* 139**,** 2020-30.

MACHADO, C. B., KANNING, K. C., KREIS, P., STEVENSON, D., CROSSLEY, M., NOWAK, M., IACOVINO, M., KYBA, M., CHAMBERS, D. & BLANC, E. 2014. Reconstruction of phrenic neuron identity in embryonic stem cell-derived motor neurons. *Development,* 141**,** 784-794.

MARLETAZ, F., FIRBAS, P. N., MAESO, I., TENA, J. J., BOGDANOVIC, O., PERRY, M., WYATT, C. D. R., DE LA CALLE-MUSTIENES, E., BERTRAND, S., BURGUERA, D., ACEMEL, R. D., VAN HEERINGEN, S. J., NARANJO, S., HERRERA-UBEDA, C., SKVORTSOVA, K., JIMENEZ-GANCEDO, S., ALDEA, D., MARQUEZ, Y., BUONO, L., KOZMIKOVA, I., PERMANYER, J., LOUIS, A., ALBUIXECH-CRESPO, B., LE PETILLON, Y., LEON, A., SUBIRANA, L., BALWIERZ, P. J., DUCKETT, P. E., FARAHANI, E., AURY, J. M., MANGENOT, S., WINCKER, P., ALBALAT, R., BENITO-GUTIERREZ, E., CANESTRO, C., CASTRO, F., D'ANIELLO, S., FERRIER, D. E. K., HUANG, S., LAUDET, V., MARAIS, G. A. B., PONTAROTTI, P., SCHUBERT, M., SEITZ, H., SOMORJAI, I., TAKAHASHI, T., MIRABEAU, O., XU, A., YU, J. K., CARNINCI, P., MARTINEZ-MORALES, J. R., CROLLIUS, H. R., KOZMIK, Z., WEIRAUCH, M. T., GARCIA-FERNANDEZ, J., LISTER, R., LENHARD, B., HOLLAND, P. W. H., ESCRIVA, H., GOMEZ-SKARMETA, J. L. & IRIMIA, M. 2018. Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature,* 564**,** 64-70.

MARLÉTAZ, F., HOLLAND, L. Z., LAUDET, V. & SCHUBERT, M. 2006. Retinoic acid signaling and the evolution of chordates. *International journal of biological sciences,* 2**,** 38.

MARTÍN-DURÁN, J. M., PANG, K., BØRVE, A., LÊ, H. S., FURU, A., CANNON, J. T., JONDELIUS, U. & HEJNOL, A. 2018. Convergent evolution of bilaterian nerve cords. *Nature,* 553**,** 45-50.

MATHIS, L., KULESA, P. M. & FRASER, S. E. 2001. FGF receptor signalling is required to maintain neural progenitors during Hensen's node progression. *Nat Cell Biol,* 3**,** 559-66.

MATSUMOTO, Y., IRIE, F., INATANI, M., TESSIER-LAVIGNE, M. & YAMAGUCHI, Y. 2007. Netrin-1/DCC signaling in commissural axon guidance requires cell-autonomous expression of heparan sulfate. *Journal of Neuroscience,* 27**,** 4342-4350.

MAZZONI, E. O., MAHONY, S., PELJTO, M., PATEL, T., THORNTON, S. R., MCCUINE, S., REEDER, C., BOYER, L. A., YOUNG, R. A. & GIFFORD, D. K. J. N. N. 2013. Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. 16**,** 1191-1198.

MCCAULEY, D. W. & BRONNER-FRASER, M. 2002. Conservation of Pax gene expression in ectodermal placodes of the lamprey. *Gene,* 287**,** 129-139.

MCCAULEY, D. W. & BRONNER‐FRASER, M. 2004. Conservation and divergence of BMP2/4 genes in the lamprey: expression and phylogenetic analysis suggest a single ancestral vertebrate gene. *Evolution & development,* 6**,** 411-422.

MCEWEN, G. K., GOODE, D. K., PARKER, H. J., WOOLFE, A., CALLAWAY, H. & ELGAR, G. J. P. G. 2009. Early evolution of conserved regulatory sequences associated with development in vertebrates. 5**,** e1000762.

MEHTA, T. K., RAVI, V., YAMASAKI, S., LEE, A. P., LIAN, M. M., TAY, B. H., TOHARI, S., YANAI, S., TAY, A., BRENNER, S. & VENKATESH, B. 2013. Evidence for at least six Hox clusters in the Japanese lamprey (Lethenteron japonicum). *Proc Natl Acad Sci U S A,* 110**,** 16044-9.

MERKES, C., TURKALO, T. K., WILDER, N., PARK, H., WENGER, L. W., LEWIN, S. J. & AZUMA, M. 2015. Ewing sarcoma ewsa protein regulates chondrogenesis of Meckel's cartilage through modulation of Sox9 in zebrafish. *PloS one,* 10**,** e0116627.

MERZDORF, C. S. 2007. Emerging roles for zic genes in early development. *Developmental dynamics: an official publication of the American Association of Anatomists,* 236**,** 922-940.

METZIS, V., STEINHAUSER, S., PAKANAVICIUS, E., GOUTI, M., STAMATAKI, D., IVANOVITCH, K., WATSON, T., RAYON, T., MOUSAVY GHARAVY, S. N., LOVELL-BADGE, R., LUSCOMBE, N. M. & BRISCOE, J. 2018. Nervous System Regionalization Entails Axial Allocation before Neural Differentiation. *Cell,* 175**,** 1105-1118 e17.

MEULEMANS, D. & BRONNER-FRASER, M. 2007. The amphioxus SoxB family: implications for the evolution of vertebrate placodes. *International journal of biological sciences,* 3**,** 356.

MISRA, K., LUO, H., LI, S., MATISE, M. & XIANG, M. 2014. Asymmetric activation of Dll4-Notch signaling by Foxn4 and proneural factors activates BMP/TGFβ signaling to specify V2b interneurons in the spinal cord. *Development,* 141**,** 187-198.

MODRELL, M. S., HOCKMAN, D., UY, B., BUCKLEY, D., SAUKA-SPENGLER, T., BRONNER, M. E. & BAKER, C. V. J. D. B. 2014. A fate-map for cranial sensory ganglia in the sea lamprey. 385**,** 405-416.

MOLINA, A. & PITUELLO, F. 2017. Playing with the cell cycle to build the spinal cord. *Dev Biol,* 432**,** 14-23.

MOORE, S., RIBES, V., TERRIENTE, J., WILKINSON, D., RELAIX, F. & BRISCOE, J. 2013. Distinct regulatory mechanisms act to establish and maintain Pax3 expression in the developing neural tube. *PLoS genetics,* 9**,** e1003811.

MORALES, A. V., ESPESO‐GIL, S., OCAÑA, I., NIETO‐LOPEZ, F., CALLEJA, E., BOVOLENTA, P., LEWANDOSKI, M. & DIEZ DEL CORRAL, R. 2016. FGF signaling enhances a sonic hedgehog negative feedback loop at the initiation of spinal cord ventral patterning. *Developmental Neurobiology,* 76**,** 956-971.

MORAN-RIVARD, L., KAGAWA, T., SAUERESSIG, H., GROSS, M. K., BURRILL, J. & GOULDING, M. 2001. Evx1 is a postmitotic determinant of v0 interneuron identity in the spinal cord. *Neuron,* 29**,** 385-399.

MORET, F., GUILLAND, J. C., COUDOUEL, S., ROCHETTE, L. & VERNIER, P. 2004. Distribution of tyrosine hydroxylase, dopamine, and serotonin in the central nervous system of amphioxus (Branchiostoma lanceolatum): implications for the evolution of catecholamine systems in vertebrates. *Journal of Comparative Neurology,* 468**,** 135-150.

MÜLLER, T., ANLAG, K., WILDNER, H., BRITSCH, S., TREIER, M. & BIRCHMEIER, C. 2005. The bHLH factor Olig3 coordinates the specification of dorsal neurons in the spinal cord. *Genes & development,* 19**,** 733-743.

MURAKAMI, Y., OGASAWARA, M., SATOH, N., SUGAHARA, F., MYOJIN, M., HIRANO, S. & KURATANI, S. 2002. Compartments in the lamprey embryonic brain as

revealed by regulatory gene expression and the distribution of reticulospinal neurons. *Brain research bulletin,* 57**,** 271-275.

MURAKAMI, Y., OGASAWARA, M., SUGAHARA, F., HIRANO, S., SATOH, N. & KURATANI, S. 2001. Identification and expression of the lamprey Pax6 gene: evolutionary origin of the segmented brain of vertebrates. *Development,* 128**,** 3521-3531.

MUROYAMA, Y., FUJIHARA, M., IKEYA, M., KONDOH, H. & TAKADA, S. 2002. Wnt signaling plays an essential role in neuronal specification of the dorsal spinal cord. *Genes Dev,* 16**,** 548-53.

NAMM, A., AREND, A. & AUNAPUU, M. 2015. BMP-2 and BMP-4 signalling in the developing spinal cord of human and rat embryos. *Folia morphologica,* 74**,** 359-364.

NARDELLI, J., THIESSON, D., FUJIWARA, Y., TSAI, F.-Y. & ORKIN, S. H. 1999. Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. *Developmental biology,* 210**,** 305-321.

NEIJTS, R., AMIN, S., VAN ROOIJEN, C. & DESCHAMPS, J. 2017. Cdx is crucial for the timing mechanism driving colinear Hox activation and defines a trunk segment in the Hox cluster topology. *Dev Biol,* 422**,** 146-154.

NIEHRS, C. 2010. On growth and form: a Cartesian coordinate system of Wnt and BMP signaling specifies bilaterian body axes. *Development,* 137**,** 845-57.

NIKITINA, N., BRONNER-FRASER, M. & SAUKA-SPENGLER, T. 2009. Microinjection of RNA and morpholino oligos into lamprey embryos. *Cold Spring Harb Protoc,* 2009**,** pdb prot5123.

NITZAN, E., AVRAHAM, O., KAHANE, N., OFEK, S., KUMAR, D. & KALCHEIM, C. 2016. Dynamics of BMP and Hes1/Hairy1 signaling in the dorsal neural tube

underlies the transition from neural crest to definitive roof plate. *BMC biology,* 14**,** 23.

NOLTE, C. & KRUMLAUF, R. 2007. Expression of Hox genes in the nervous system of vertebrates. *HOX gene expression.* Springer.

NOVITCH, B. G., CHEN, A. I. & JESSELL, T. M. 2001. Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron,* 31**,** 773-789.

NOVITCH, B. G., WICHTERLE, H., JESSELL, T. M. & SOCKANATHAN, S. 2003. A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification. *Neuron,* 40**,** 81-95.

OHYAMA, T., MOHAMED, O. A., TAKETO, M. M., DUFORT, D. & GROVES, A. K. 2006. Wnt signals mediate a fate decision between otic placode and epidermis. *Development,* 133**,** 865-875.

OKADA, Y., SHIMAZAKI, T., SOBUE, G. & OKANO, H. 2004. Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse embryonic stem cells. *Developmental biology,* 275**,** 124-142.

ONAI, T., LIN, H.-C., SCHUBERT, M., KOOP, D., OSBORNE, P. W., ALVAREZ, S., ALVAREZ, R., HOLLAND, N. D. & HOLLAND, L. Z. 2009. Retinoic acid and Wnt/β-catenin have complementary roles in anterior/posterior patterning embryos of the basal chordate amphioxus. *Developmental biology,* 332**,** 223-233.

ONAI, T., TAKAI, A., SETIAMARGA, D. H. & HOLLAND, L. Z. 2012. Essential role of Dkk3 for head formation by inhibiting Wnt/β-catenin and Nodal/Vg1 signaling pathways in the basal chordate amphioxus. *Evolution & development,* 14**,** 338-350.

ONIMARU, K., MARCON, L., MUSY, M., TANAKA, M. & SHARPE, J. 2016. The fin-to-limb transition as the re-organization of a Turing pattern. *Nature communications,* 7**,** 1-9.

ONO, H., KOOP, D. & HOLLAND, L. Z. 2018. Nodal and Hedgehog synergize in gill slit formation during development of the cephalochordate Branchiostoma floridae. *Development,* 145.

ORQUERA, D. P., TAVELLA, M. B., DE SOUZA, F. S., NASIF, S., LOW, M. J. & RUBINSTEIN, M. 2019. The homeodomain transcription factor NKX2. 1 is essential for the early specification of melanocortin neuron identity and activates Pomc expression in the developing hypothalamus. *Journal of Neuroscience,* 39**,** 4023-4035.

OSBORNE, P. W., BENOIT, G., LAUDET, V., SCHUBERT, M. & FERRIER, D. E. 2009. Differential regulation of ParaHox genes by retinoic acid in the invertebrate chordate amphioxus (Branchiostoma floridae). *Developmental biology,* 327**,** 252-262.

OSSEWARD II, P. J. & PFAFF, S. L. J. C. O. I. N. 2019. Cell type and circuit modules in the spinal cord. 56**,** 175-184.

PABST, O., HERBRAND, H., TAKUMA, N. & ARNOLD, H.-H. 2000. NKX2 gene expression in neuroectoderm but not in mesendodermally derived structures depends on sonic hedgehog in mouse embryos. *Development genes and evolution,* 210**,** 47-50.

PANOPOULOU, G. D., CLARK, M. D., HOLLAND, L. Z., LEHRACH, H. & HOLLAND, N. D. 1998. AmphiBMP2/4, an amphioxus bone morphogenetic protein closely related to *Drosophila* decapentaplegic and vertebrate BMP2 and BMP4:

Insights into evolution of dorsoventral axis specification. *Dev Dynam,* 213**,** 130-139.

PAPS, J., HOLLAND, P. W. & SHIMELD, S. M. 2012. A genome-wide view of transcription factor gene diversity in chordate evolution: less gene loss in amphioxus? *Brief Funct Genomics,* 11**,** 177-86.

PARKER, H. J., BRONNER, M. E. & KRUMLAUF, R. 2014a. A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. *Nature,* 514**,** 490-3.

PARKER, H. J., BRONNER, M. E. & KRUMLAUF, R. 2016. The vertebrate Hox gene regulatory network for hindbrain segmentation: Evolution and diversification: Coupling of a Hox gene regulatory network to hindbrain segmentation is an ancient trait originating at the base of vertebrates. *Bioessays,* 38**,** 526-38.

PARKER, H. J., BRONNER, M. E. & KRUMLAUF, R. J. D. B. 2019a. An atlas of anterior hox gene expression in the embryonic sea lamprey head: hox-code evolution in vertebrates. 453**,** 19-33.

PARKER, H. J., DE KUMAR, B., GREEN, S. A., PRUMMEL, K. D., HESS, C., KAUFMAN, C. K., MOSIMANN, C., WIEDEMANN, L. M., BRONNER, M. E. & KRUMLAUF, R. 2019b. A Hox-TALE regulatory circuit for neural crest patterning is conserved across vertebrates. *Nat Commun,* 10**,** 1189.

PARKER, H. J., PICCINELLI, P., SAUKA-SPENGLER, T., BRONNER, M. & ELGAR, G. J. B. G. 2011. Ancient Pbx-Hox signatures define hundreds of vertebrate developmental enhancers. 12**,** 1-14.

PARKER, H. J., SAUKA-SPENGLER, T., BRONNER, M. & ELGAR, G. J. P. O. 2014b. A reporter assay in lamprey embryos reveals both functional conservation and elaboration of vertebrate enhancers. 9**,** e85492.

PARR, B. A., SHEA, M. J., VASSILEVA, G. & MCMAHON, A. P. 1993. Mouse Wnt genes exhibit discrete domains of expression in the early embryonic CNS and limb buds. *Development,* 119**,** 247-261.

PASCUAL-ANAYA, J., ADACHI, N., ALVAREZ, S., KURATANI, S., D'ANIELLO, S. & GARCIA-FERNANDEZ, J. 2012. Broken colinearity of the amphioxus Hox cluster. *Evodevo,* 3.

PASCUAL-ANAYA, J., SATO, I., SUGAHARA, F., HIGUCHI, S., PAPS, J., REN, Y., TAKAGI, W., RUIZ-VILLALBA, A., OTA, K. G., WANG, W. & KURATANI, S. 2018. Hagfish and lamprey Hox genes reveal conservation of temporal colinearity in vertebrates. *Nat Ecol Evol,* 2**,** 859-866.

PASINI, A., MANENTI, R., ROTHBACHER, U. & LEMAIRE, P. 2012. Antagonizing retinoic acid and FGF/MAPK pathways control posterior body patterning in the invertebrate chordate Ciona intestinalis. *PLoS One,* 7**,** e46193.

PATTHEY, C., SCHLOSSER, G. & SHIMELD, S. M. 2014. The evolutionary history of vertebrate cranial placodes--I: cell type evolution. *Dev Biol,* 389**,** 82-97.

PAUL, B. Y., HONG, C. C., SACHIDANANDAN, C., BABITT, J. L., DENG, D. Y., HOYNG, S. A., LIN, H. Y., BLOCH, K. D. & PETERSON, R. T. 2008. Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. *Nature chemical biology,* 4**,** 33-41.

PEDRONI, A. & AMPATZIS, K. 2019. Large scale analysis of the diversity and complexity of the adult spinal cord neurotransmitter typology. *bioRxiv***,** 518845.

PESTARINO, M. & LUCARONI, B. 1996. FMRFamide-like immunoreactivity in the central nervous system of the lancelet Branchiostoma lanceolatum. *Israel Journal of Zoology,* 42**,** S227-S234.

PFEFFER, P. L., BOUCHARD, M. & BUSSLINGER, M. 2000. Pax2 and homeodomain proteins cooperatively regulate a 435 bp enhancer of the mouse Pax5 gene at the midbrain-hindbrain boundary. *Development,* 127**,** 1017-1028.

PFEFFER, P. L., PAYER, B., REIM, G., DI MAGLIANO, M. P. & BUSSLINGER, M. 2002. The activation and maintenance of Pax2 expression at the mid-hindbrain boundary is controlled by separate enhancers. *Development,* 129**,** 307-318.

PIERANI, A., BRENNER-MORTON, S., CHIANG, C. & JESSELL, T. M. 1999. A sonic hedgehog–independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell,* 97**,** 903-915.

PILLAI, A., MANSOURI, A., BEHRINGER, R., WESTPHAL, H. & GOULDING, M. 2007. Lhx1 and Lhx5 maintain the inhibitory-neurotransmitter status of interneurons in the dorsal spinal cord. *Development,* 134**,** 357-366.

POSNIEN, N., KONISZEWSKI, N. D. B., HEIN, H. J. & BUCHER, G. 2011. Candidate gene screen in the red flour beetle Tribolium reveals six3 as ancient regulator of anterior median head and central complex development. *PLoS Genet,* 7**,** e1002416.

PUTNAM, N. H., BUTTS, T., FERRIER, D. E., FURLONG, R. F., HELLSTEN, U., KAWASHIMA, T., ROBINSON-RECHAVI, M., SHOGUCHI, E., TERRY, A., YU, J. K., BENITO-GUTIERREZ, E. L., DUBCHAK, I., GARCIA-FERNANDEZ, J., GIBSON-BROWN, J. J., GRIGORIEV, I. V., HORTON, A. C., DE JONG, P. J., JURKA, J., KAPITONOV, V. V., KOHARA, Y., KUROKI, Y., LINDQUIST, E., LUCAS, S., OSOEGAWA, K., PENNACCHIO, L. A., SALAMOV, A. A., SATOU, Y., SAUKA-SPENGLER, T., SCHMUTZ, J., SHIN, I. T., TOYODA, A., BRONNER-FRASER, M., FUJIYAMA, A., HOLLAND, L. Z., HOLLAND, P. W., SATOH, N. & ROKHSAR, D. S.

2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature,* 453**,** 1064-71.

QI, Y., TAN, M., HUI, C.-C. & QIU, M. 2003. Gli2 is required for normal Shh signaling and oligodendrocyte development in the spinal cord. *Molecular and Cellular Neuroscience,* 23**,** 440-450.

QIAN, Y., SHIRASAWA, S., CHEN, C.-L., CHENG, L. & MA, Q. 2002. Proper development of relay somatic sensory neurons and D2/D4 interneurons requires homeobox genes Rnx/Tlx-3 and Tlx-1. *Genes & Development,* 16**,** 1220-1233.

QUINLAN, K. A. & BUCHANAN, J. T. 2008. Cellular and synaptic actions of acetylcholine in the lamprey spinal cord. *Journal of neurophysiology,* 100**,** 1020-1031.

RASMUSSEN, S. L., HOLLAND, L. Z., SCHUBERT, M., BEASTER-JONES, L. & HOLLAND, N. D. 2007. Amphioxus AmphiDelta: evolution of Delta protein structure, segmentation, and neurogenesis. *Genesis,* 45**,** 113-22.

RAVI, V., BHATIA, S., SHINGATE, P., TAY, B.-H., VENKATESH, B. & KLEINJAN, D. A. 2019. Lampreys, the jawless vertebrates, contain three Pax6 genes with distinct expression in eye, brain and pancreas. *Scientific reports,* 9**,** 1-12.

REN, Q., ZHONG, Y., HUANG, X., LEUNG, B., XING, C., WANG, H., HU, G., WANG, Y., SHIMELD, S. M. & LI, G. 2020. Step-wise evolution of neural patterning by Hedgehog signalling in chordates. *Nature Ecology & Evolution***,** 1-9.

RIBES, V., LE ROUX, I., RHINN, M., SCHUHBAUR, B. & DOLLÉ, P. 2009. Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. *Development,* 136**,** 665-676.

ROUSSO, D. L., PEARSON, C. A., GABER, Z. B., MIQUELAJAUREGUI, A., LI, S., PORTERA-CAILLIAU, C., MORRISEY, E. E. & NOVITCH, B. G. 2012. Foxp-mediated

suppression of N-cadherin regulates neuroepithelial character and progenitor maintenance in the CNS. *Neuron,* 74**,** 314-30.

SAADE, M., BLANCO-AMEIJEIRAS, J., GONZALEZ-GOBARTT, E. & MARTI, E. 2018. A centrosomal view of CNS growth. *Development,* 145.

SACERDOT, C., LOUIS, A., BON, C., BERTHELOT, C. & ROEST CROLLIUS, H. 2018. Chromosome evolution at the origin of the ancestral vertebrate genome. *Genome Biol,* 19**,** 166.

SAITO, S., SUZUKI, T. J. D., GROWTH & DIFFERENTIATION 2020. How do signaling and transcription factors regulate both axis elongation and Hox gene expression along the anteroposterior axis?

SASAI, N., KUTEJOVA, E. & BRISCOE, J. 2014. Integration of signals along orthogonal axes of the vertebrate neural tube controls progenitor competence and increases cell diversity. *PLoS Biol,* 12**,** e1001907.

SASAI, Y., LU, B., STEINBEISSER, H., GEISSERT, D., GONT, L. K. & DE ROBERTIS, E. M. 1994. Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell,* 79**,** 779-90.

SAUDEMONT, A., DRAY, N., HUDRY, B., LE GOUAR, M., VERVOORT, M. & BALAVOINE, G. J. D. B. 2008. Complementary striped expression patterns of NK homeobox genes during segment formation in the annelid Platynereis. 317**,** 430-443.

SCHÄFER, M., KINZEL, D., NEUNER, C., SCHARTL, M., VOLFF, J.-N. & WINKLER, C. 2005. Hedgehog and retinoid signalling confines nkx2. 2b expression to the lateral floor plate of the zebrafish trunk. *Mechanisms of development,* 122**,** 43-56.

SCHIERWATER, B., KAMM, K., SRIVASTAVA, M., ROKHSAR, D., ROSENGARTEN, R. D. & DELLAPORTA, S. L. J. P. O. 2008. The early ANTP gene repertoire: insights from the placozoan genome. 3**,** e2457.

SCHMID, B., FURTHAUER, M., CONNORS, S. A., TROUT, J., THISSE, B., THISSE, C. & MULLINS, M. C. 2000. Equivalent genetic roles for bmp7/snailhouse and bmp2b/swirl in dorsoventral pattern formation. *Development,* 127**,** 957-967.

SCHUBERT, F. R., MOOTOOSAMY, R. C., WALTERS, E. H., GRAHAM, A., TUMIOTTO, L., MÜNSTERBERG, A. E., LUMSDEN, A. & DIETRICH, S. 2002. Wnt6 marks sites of epithelial transformations in the chick embryo. *Mechanisms of development,* 114**,** 143-148.

SCHUBERT, M., HOLLAND, L. Z., STOKES, M. D. & HOLLAND, N. D. 2001. Three amphioxus Wnt genes (AmphiWnt3, AmphiWnt5, and AmphiWnt6) associated with the tail bud: the evolution of somitogenesis in chordates. *Dev Biol,* 240**,** 262-73.

SCHUBERT, M., HOLLAND, N. D., ESCRIVA, H., HOLLAND, L. Z. & LAUDET, V. 2004. Retinoic acid influences anteroposterior positioning of epidermal sensory neurons and their gene expression in a developing chordate (amphioxus). *Proceedings of the National Academy of Sciences,* 101**,** 10320-10325.

SCHUBERT, M., HOLLAND, N. D., LAUDET, V. & HOLLAND, L. Z. 2006. A retinoic acid-Hox hierarchy controls both anterior/posterior patterning and neuronal specification in the developing central nervous system of the cephalochordate amphioxus. *Developmental Biology,* 296**,** 190-202.

SEKIGAMI, Y., KOBAYASHI, T., OMI, A., NISHITSUJI, K., IKUTA, T., FUJIYAMA, A., SATOH, N. & SAIGA, H. 2017. Hox gene cluster of the ascidian, Halocynthia roretzi, reveals multiple ancient steps of cluster disintegration during ascidian evolution. *Zoological Lett,* 3**,** 17.

SEREDICK, S. D., VAN RYSWYK, L., HUTCHINSON, S. A. & EISEN, J. S. 2012. Zebrafish Mnx proteins specify one motoneuron subtype and suppress acquisition of interneuron characteristics. *Neural development,* 7**,** 1-14.

SHARMAN, A., SHIMELD, S. M. & HOLLAND, P. 1999. An amphioxus Msx gene expressed predominantly in the dorsal neural tube. *Development genes and evolution,* 209**,** 260-263.

SHAWI, M. & SERLUCA, F. C. 2008. Identification of a BMP7 homolog in zebrafish expressed in developing organ systems. *Gene Expression Patterns,* 8**,** 369-375.

SHEN, M. C., OZACAR, A. T., OSGOOD, M., BOERAS, C., PINK, J., THOMAS, J., KOHTZ, J. D. & KARLSTROM, R. 2013. Heat‐shock–mediated conditional regulation of hedgehog/gli signaling in zebrafish. *Developmental Dynamics,* 242**,** 539-549.

SHIFAMN, M. I. & SELZER, M. E. 2015. Axon regeneration in the lamprey spinal cord. *In:* SO, K.-F. & XU, X.-M. (eds.) *Neural Regeneration.* Beijing: Academic Press.

SHIGETANI, Y., SUGAHARA, F., KAWAKAMI, Y., MURAKAMI, Y., HIRANO, S. & KURATANI, S. 2002. Heterotopic shift of epithelial-mesenchymal interactions in vertebrate jaw evolution. *Science,* 296**,** 1316-9.

SHIMELD, S. M. 1999. The evolution of the hedgehog gene family in chordates: Insights from amphioxus hedgehog. *Dev. Genes Evol.,* 209**,** 40-47.

SHIMELD, S. M. 2000. An amphioxus netrin gene is expressed in midline structures during embryonic and larval development. 210**,** 337-344.

SHIMELD, S. M. & HOLLAND, P. W. 2000. Vertebrate innovations. *Proc Natl Acad Sci U S A,* 97**,** 4449-52.

SHINOZUKA, T., TAKADA, R., YOSHIDA, S., YONEMURA, S. & TAKADA, S. 2019. Wnt produced by stretched roof-plate cells is required for the promotion of cell

proliferation around the central canal of the spinal cord. *Development,* 146**,** dev159343.

SHORT, S., KOZMIK, Z. & HOLLAND, L. Z. 2012. The function and developmental expression of alternatively spliced isoforms of amphioxus and Xenopus laevis Pax2/5/8 genes: revealing divergence at the invertebrate to vertebrate transition. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution,* 318**,** 555-571.

SIMAKOV, O., KAWASHIMA, T., MARLÉTAZ, F., JENKINS, J., KOYANAGI, R., MITROS, T., HISATA, K., BREDESON, J., SHOGUCHI, E. & GYOJA, F. 2015. Hemichordate genomes and deuterostome origins. *Nature,* 527**,** 459-465.

SIMAKOV, O., MARLÉTAZ, F., YUE, J.-X., O'CONNELL, B., JENKINS, J., BRANDT, A., CALEF, R., TUNG, C.-H., HUANG, T.-K. & SCHMUTZ, J. 2020. Deeply conserved synteny resolves early events in vertebrate evolution. *Nature Ecology & Evolution***,** 1-11.

SMITH, J. J., KURAKU, S., HOLT, C., SAUKA-SPENGLER, T., JIANG, N., CAMPBELL, M. S., YANDELL, M. D., MANOUSAKI, T., MEYER, A., BLOOM, O. E., MORGAN, J. R., BUXBAUM, J. D., SACHIDANANDAM, R., SIMS, C., GARRUSS, A. S., COOK, M., KRUMLAUF, R., WIEDEMANN, L. M., SOWER, S. A., DECATUR, W. A., HALL, J. A., AMEMIYA, C. T., SAHA, N. R., BUCKLEY, K. M., RAST, J. P., DAS, S., HIRANO, M., MCCURLEY, N., GUO, P., ROHNER, N., TABIN, C. J., PICCINELLI, P., ELGAR, G., RUFFIER, M., AKEN, B. L., SEARLE, S. M., MUFFATO, M., PIGNATELLI, M., HERRERO, J., JONES, M., BROWN, C. T., CHUNG-DAVIDSON, Y. W., NANLOHY, K. G., LIBANTS, S. V., YEH, C. Y., MCCAULEY, D. W., LANGELAND, J. A., PANCER, Z., FRITZSCH, B., DE JONG, P. J., ZHU, B., FULTON, L. L., THEISING, B., FLICEK, P., BRONNER, M. E., WARREN, W. C., CLIFTON, S. W., WILSON, R. K. & LI, W. 2013.

Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nat Genet,* 45**,** 415-21, 421e1-2.

SMITH, J. J., TIMOSHEVSKAYA, N., YE, C., HOLT, C., KEINATH, M. C., PARKER, H. J., COOK, M. E., HESS, J. E., NARUM, S. R., LAMANNA, F., KAESSMANN, H., TIMOSHEVSKIY, V. A., WATERBURY, C. K. M., SARACENO, C., WIEDEMANN, L. M., ROBB, S. M. C., BAKER, C., EICHLER, E. E., HOCKMAN, D., SAUKA-SPENGLER, T., YANDELL, M., KRUMLAUF, R., ELGAR, G. & AMEMIYA, C. T. 2018. The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nat Genet,* 50**,** 270-277.

SOCKANATHAN, S., PERLMANN, T. & JESSELL, T. M. 2003. Retinoid receptor signaling in postmitotic motor neurons regulates rostrocaudal positional identity and axonal projection pattern. *Neuron,* 40**,** 97-111.

SOMORJAI, I. M., MARTÍ-SOLANS, J., DIAZ-GRACIA, M., NISHIDA, H., IMAI, K. S., ESCRIVÀ, H., CAÑESTRO, C. & ALBALAT, R. 2018. Wnt evolution and function shuffling in liberal and conservative chordate genomes. *Genome biology,* 19**,** 98.

SONG, M.-R., SUN, Y., BRYSON, A., GILL, G. N., EVANS, S. M. & PFAFF, S. L. 2009. Islet-to-LMO stoichiometries control the function of transcription complexes that specify motor neuron and V2a interneuron identity. *Development,* 136**,** 2923-2932.

SOSHNIKOVA, N. & DUBOULE, D. 2009. Epigenetic temporal control of mouse Hox genes in vivo. *Science,* 324**,** 1320-3.

SOULA, C., DANESIN, C., KAN, P., GROB, M., PONCET, C. & COCHARD, P. 2001. Distinct sites of origin of oligodendrocytes and somatic motoneurons in the chick

spinal cord: oligodendrocytes arise from Nkx2. 2-expressing progenitors by a Shh-dependent mechanism. *Development,* 128**,** 1369-1379.

SQUARE, T., ROMASEK, M., JANDZIK, D., CATTELL, M. V., KLYMKOWSKY, M. & MEDEIROS, D. M. 2015. CRISPR/Cas9-mediated mutagenesis in the sea lamprey Petromyzon marinus: a powerful tool for understanding ancestral gene functions in vertebrates. *Development,* 142**,** 4180-7.

STAFFORD, D. & PRINCE, V. E. 2002. Retinoic acid signaling is required for a critical early step in zebrafish pancreatic development. *Current Biology,* 12**,** 1215-1220.

STASIULEWICZ, M., GRAY, S. D., MASTROMINA, I., SILVA, J. C., BJÖRKLUND, M., SEYMOUR, P. A., BOOTH, D., THOMPSON, C., GREEN, R. J. & HALL, E. A. 2015. A conserved role for Notch signaling in priming the cellular response to Shh through ciliary localisation of the key Shh transducer Smo. *Development,* 142**,** 2291-2303.

STEINBUCK, M. P. & WINANDY, S. 2018. A review of notch processing with new insights into ligand-independent notch signaling in T-cells. *Frontiers in immunology,* 9**,** 1230.

STOLFI, A., WAGNER, E., TALIAFERRO, J. M., CHOU, S. & LEVINE, M. 2011. Neural tube patterning by Ephrin, FGF and Notch signaling relays. *Development,* 138**,** 5429-5439.

STULBERG, M. J., LIN, A., ZHAO, H. & HOLLEY, S. A. 2012. Crosstalk between Fgf and Wnt signaling in the zebrafish tailbud. *Developmental biology,* 369**,** 298-307.

SUGAHARA, F., AOTA, S., KURAKU, S., MURAKAMI, Y., TAKIO-OGAWA, Y., HIRANO, S. & KURATANI, S. 2011. Involvement of Hedgehog and FGF signalling in the

lamprey telencephalon: evolution of regionalization and dorsoventral patterning of the vertebrate forebrain. *Development,* 138**,** 1217-26.

SUGAHARA, F., PASCUAL-ANAYA, J., OISI, Y., KURAKU, S., AOTA, S.-I., ADACHI, N., TAKAGI, W., HIRAI, T., SATO, N. & MURAKAMI, Y. 2016. Evidence from cyclostomes for complex regionalization of the ancestral vertebrate brain. *Nature,* 531**,** 97-100.

SUMMERHURST, K., STARK, M., SHARPE, J., DAVIDSON, D. & MURPHY, P. 2008. 3D representation of Wnt and Frizzled gene expression patterns in the mouse embryo at embryonic day 11.5 (Ts19). *Gene Expression Patterns,* 8**,** 331-348.

SUN, A. X., LONDONO, R., HUDNALL, M. L., TUAN, R. S. & LOZITO, T. P. 2018. Differences in neural stem cell identity and differentiation capacity drive divergent regenerative outcomes in lizards and salamanders. *Proceedings of the National Academy of Sciences,* 115**,** E8256-E8265.

SUN, T., DONG, H., WU, L., KANE, M., ROWITCH, D. H. & STILES, C. D. 2003. Cross-repressive interaction of the Olig2 and Nkx2. 2 transcription factors in developing neural tube associated with formation of a specific physical complex. *Journal of Neuroscience,* 23**,** 9547-9556.

TAKADA, S., STARK, K. L., SHEA, M. J., VASSILEVA, G., MCMAHON, J. A., MCMAHON, A. P. J. G. & DEVELOPMENT 1994. Wnt-3a regulates somite and tailbud formation in the mouse embryo. 8**,** 174-189.

TAKATORI, N., SATOU, Y. & SATOH, N. 2002. Expression of hedgehog genes in Ciona intestinalis embryos. *Mech Dev,* 116**,** 235-8.

TANABE, Y., WILLIAM, C. & JESSELL, T. M. 1998. Specification of motor neuron identity by the MNR2 homeodomain protein. *Cell,* 95**,** 67-80.

TANAKA, H., MORIMURA, R. & OHSHIMA, T. 2012. Dpysl2 (CRMP2) and Dpysl3 (CRMP4) phosphorylation by Cdk5 and DYRK2 is required for proper positioning of Rohon-Beard neurons and neural crest cells during neurulation in zebrafish. *Developmental biology,* 370**,** 223-236.

TANAKA, S., KAMACHI, Y., TANOUCHI, A., HAMADA, H., JING, N. & KONDOH, H. 2004. Interplay of SOX and POU factors in regulation of the Nestin gene in neural primordial cells. *Molecular and cellular biology,* 24**,** 8834-8846.

THÉLIE, A., DESIDERIO, S., HANOTEL, J., QUIGLEY, I., VAN DRIESSCHE, B., RODARI, A., BORROMEO, M. D., KRICHA, S., LAHAYE, F. & CROCE, J. 2015. Prdm12 specifies V1 interneurons through cross-repressive interactions with Dbx1 and Nkx6 genes in Xenopus. *Development,* 142**,** 3416-3428.

THISSE, B. & THISSE, C. 2004. Fast release clones: a high throughput expression analysis. *ZFIN direct data submission,* 2004.

TIMMER, J. R., WANG, C. & NISWANDER, L. 2002. BMP signaling patterns the dorsal and intermediate neural tube via regulation of homeobox and helix-loop-helix transcription factors. *Development,* 129**,** 2459-2472.

TOZER, S., LE DREAU, G., MARTI, E. & BRISCOE, J. 2013. Temporal control of BMP signalling determines neuronal subtype identity in the dorsal neural tube. *Development,* 140**,** 1467-74.

TREFFKORN, S., KAHNKE, L., HERING, L. & MAYER, G. J. E. 2018. Expression of NK cluster genes in the onychophoran Euperipatoides rowelli: implications for the evolution of NK family genes in nephrozoans. 9**,** 1-32.

UEMURA, H., TEZUKA, Y., HASEGAWA, C. & KOBAYASHI, H. 1994. Immunohistochemical Investigation of Neuropeptides in the Central-Nervous-

System of the Amphioxus, Branchiostoma-Belcheri. *Cell and Tissue Research,* 277**,** 279-287.

UYGUR, A., YOUNG, J., HUYCKE, T. R., KOSKA, M., BRISCOE, J. & TABIN, C. J. 2016. Scaling Pattern to Variations in Size during Development of the Vertebrate Neural Tube. *Dev Cell,* 37**,** 127-35.

VAAGA, C. E., BORISOVSKA, M. & WESTBROOK, G. L. 2014. Dual-transmitter neurons: functional implications of co-release and co-transmission. *Curr Opin Neurobiol,* 29**,** 25-32.

VALLSTEDT, A., MUHR, J., PATTYN, A., PIERANI, A., MENDELSOHN, M., SANDER, M., JESSELL, T. M. & ERICSON, J. 2001. Different levels of repressor activity assign redundant and specific roles to Nkx6 genes in motor neuron and interneuron specification. *Neuron,* 31**,** 743-755.

VAN AMERONGEN, R., FUERER, C., MIZUTANI, M. & NUSSE, R. 2012. Wnt5a can both activate and repress Wnt/β-catenin signaling during mouse embryonic development. *Developmental biology,* 369**,** 101-114.

VILLAR-CERVIÑO, V., FERNÁNDEZ-LÓPEZ, B., CELINA RODICIO, M. & ANADÓN, R. 2014. Aspartate-containing neurons of the brainstem and rostral spinal cord of the sea lamprey Petromyzon marinus: Distribution and comparison with γ-aminobutyric acid. *Journal of Comparative Neurology,* 522**,** 1209-1231.

VILLAR-CERVIÑO, V., HOLSTEIN, G. R., MARTINELLI, G. P., ANADÓN, R. & RODICIO, M. C. 2008. Glycine-immunoreactive neurons in the developing spinal cord of the sea lamprey: Comparison with the γ-aminobutyric acidergic system. *Journal of Comparative Neurology,* 508**,** 112-130.

VLADYKOV, V. D. & FOLLETT, W. J. J. O. T. F. B. O. C. 1965. Lampetra richardsoni, a new nonparasitic species of lamprey (Petromyzonidae) from western North America. 22**,** 139-158.

WADA, H., GARCIA-FERNANDEZ, J. & HOLLAND, P. W. H. 1999. Colinear and segmental expression of amphioxus Hox genes. *Developmental Biology,* 213**,** 131-141.

WAHL, M. B., DENG, C., LEWANDOSKI, M. & POURQUIÉ, O. 2007. FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development,* 134**,** 4033-4041.

WANG, H., COPELAND, N. G., GILBERT, D. J., JENKINS, N. A. & TESSIER-LAVIGNE, M. 1999. Netrin-3, a mouse homolog of human NTN2L, is highly expressed in sensory ganglia and shows differential binding to netrin receptors. *Journal of Neuroscience,* 19**,** 4938-4947.

WANG, R., ZHAO, C., LI, J., LI, Y., LIU, Y., DONG, H., WANG, D., ZHAO, B., ZHANG, X. & WANG, S. 2018. Notch1 promotes mouse spinal neural stem and progenitor cells proliferation via p-p38-pax6 induced cyclin D1 activation. *Experimental Cell Research,* 373**,** 80-90.

WATERS, S. T., WILSON, C. P. & LEWANDOSKI, M. 2003. Cloning and embryonic expression analysis of the mouse Gbx1 gene. *Gene expression patterns,* 3**,** 313-317.

WETTSTEIN, D. A., TURNER, D. L. & KINTNER, C. 1997. The Xenopus homolog of Drosophila Suppressor of Hairless mediates Notch signaling during primary neurogenesis. *Development,* 124**,** 693-702.

WIJGERDE, M., MCMAHON, J. A., RULE, M. & MCMAHON, A. P. 2002. A direct requirement for Hedgehog signaling for normal specification of all ventral

progenitor domains in the presumptive mammalian spinal cord. *Genes & development,* 16**,** 2849-2864.

WILSON, L., GALE, E., CHAMBERS, D. & MADEN, M. 2004. Retinoic acid and the control of dorsoventral patterning in the avian spinal cord. *Developmental biology,* 269**,** 433-446.

WILSON, S. I., SHAFER, B., LEE, K. J. & DODD, J. 2008. A molecular program for contralateral trajectory: Rig-1 control by LIM homeodomain transcription factors. *Neuron,* 59**,** 413-424.

WINATA, C. L., KORZH, S., KONDRYCHYN, I., ZHENG, W., KORZH, V. & GONG, Z. 2009. Development of zebrafish swimbladder: The requirement of Hedgehog signaling in specification and organization of the three tissue layers. *Developmental biology,* 331**,** 222-236.

WONG, E. S., ZHENG, D., TAN, S. Z., BOWER, N. L., GARSIDE, V., VANWALLEGHEM, G., GAITI, F., SCOTT, E., HOGAN, B. M. & KIKUCHI, K. J. S. 2020. Deep conservation of the enhancer regulatory code in animals. 370.

WOTTON, K. R. & SHIMELD, S. M. J. G. 2011. Analysis of lamprey clustered Fox genes: insight into Fox gene evolution and expression in vertebrates. 489**,** 30-40.

WOTTON, K. R., WEIERUD, F. K., JUÁREZ-MORALES, J. L., ALVARES, L. E., DIETRICH, S. & LEWIS, K. E. 2009. Conservation of gene linkage in dispersed vertebrate NK homeobox clusters. *Development genes and evolution,* 219**,** 481-496.

YANG, X., TOMITA, T., WINES-SAMUELSON, M., BEGLOPOULOS, V., TANSEY, M. G., KOPAN, R. & SHEN, J. 2006. Notch1 signaling influences v2 interneuron and motor neuron development in the spinal cord. *Developmental neuroscience,* 28**,** 102-117.

YASUOKA, Y. 2020. Enhancer evolution in chordates: Lessons from functional analyses of cephalochordate cis‐regulatory modules. 62**,** 279-300.

YILDIZ, O., DOWNES, G. B. & SAGERSTRÖM, C. G. 2019. Zebrafish prdm12b acts independently of nkx6. 1 repression to promote eng1b expression in the neural tube p1 domain. *Neural development,* 14**,** 1-19.

YIN, Y., SANES, J. R. & MINER, J. H. 2000. Identification and expression of mouse netrin-4. *Mechanisms of development,* 96**,** 115-119.

YIN, Z. & FRASCH, M. J. D. G. 1998. Regulation and function of tinman during dorsal mesoderm induction and heart specification in Drosophila. 22**,** 187-200.

YORK, J. R., ZEHNDER, K., YUAN, T., LAKIZA, O. & MCCAULEY, D. W. 2019. Evolution of Snail-mediated regulation of neural crest and placodes from an ancient role in bilaterian neurogenesis. *Developmental biology,* 453**,** 180-190.

YOSHIKAWA, S., MCKINNON, R. D., KOKEL, M. & THOMAS, J. B. 2003. Wnt-mediated axon guidance via the Drosophila Derailed receptor. *Nature,* 422**,** 583-588.

YU, H.-M. I., LIU, B., COSTANTINI, F. & HSU, W. 2007. Impaired neural development caused by inducible expression of Axin in transgenic mice. *Mechanisms of development,* 124**,** 146-156.

YU, J. S. & HOLLAND, L. Z. 2009. Amphioxus whole-mount in situ hybridization. *Cold Spring Harbor Protocols,* 2009**,** pdb. prot5286.

YUAN, T., YORK, J. R. & MCCAULEY, D. W. 2018. Gliogenesis in lampreys shares gene regulatory interactions with oligodendrocyte development in jawed vertebrates. *Dev Biol,* 441**,** 176-190.

YUE, J.-X., KOZMIKOVA, I., ONO, H., NOSSA, C. W., KOZMIK, Z., PUTNAM, N. H., YU, J.-K., HOLLAND, L. Z. J. G. B. & EVOLUTION 2016. Conserved noncoding

elements in the most distant genera of cephalochordates: the Goldilocks principle. 8**,** 2387-2405.

ZECHNER, D., MÜLLER, T., WENDE, H., WALTHER, I., TAKETO, M. M., CRENSHAW III, E. B., TREIER, M., BIRCHMEIER, W. & BIRCHMEIER, C. 2007. Bmp and Wnt/β-catenin signals control expression of the transcription factor Olig3 and the specification of spinal cord neurons. *Developmental biology,* 303**,** 181-190.

ZHOU, Q., CHOI, G. & ANDERSON, D. J. 2001. The bHLH transcription factor Olig2 promotes oligodendrocyte differentiation in collaboration with Nkx2. 2. *Neuron,* 31**,** 791-807.

ZIEGER, E., CANDIANI, S., GARBARINO, G., CROCE, J. C. & SCHUBERT, M. 2018a. Roles of retinoic acid signaling in shaping the neuronal architecture of the developing amphioxus nervous system. *Molecular neurobiology,* 55**,** 5210-5229.

ZIEGER, E., GARBARINO, G., ROBERT, N. S., YU, J.-K., CROCE, J. C., CANDIANI, S. & SCHUBERT, M. 2018b. Retinoic acid signaling and neurogenic niche regulation in the developing peripheral nervous system of the cephalochordate amphioxus. *Cellular and Molecular Life Sciences,* 75**,** 2407-2429.

ZIKOVA, M., DA PONTE, J.-P., DASTUGUE, B. & JAGLA, K. J. P. O. T. N. A. O. S. 2003. Patterning of the cardiac outflow region in Drosophila. 100**,** 12189-12194.

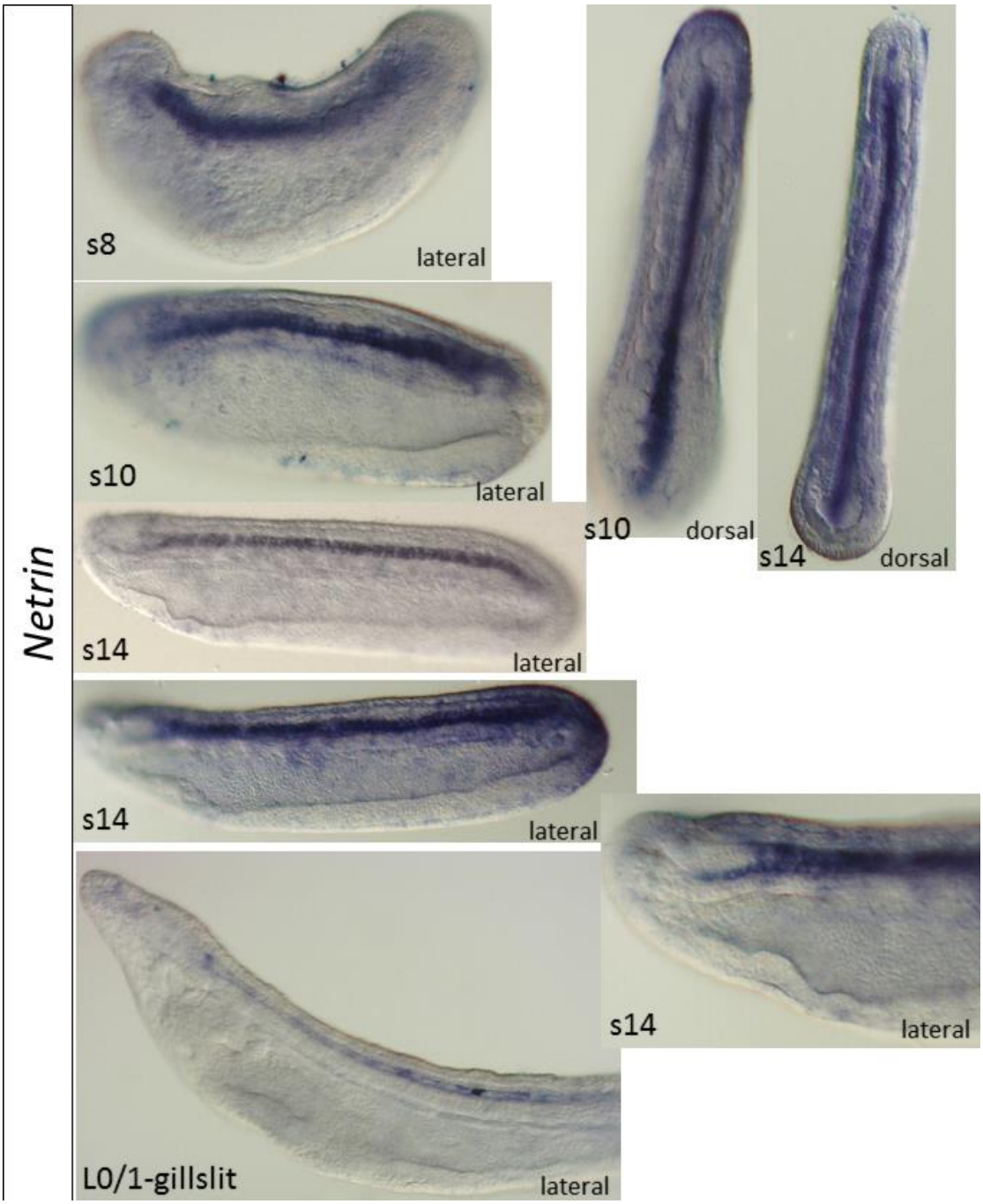
ZU, Y., ZHANG, X., REN, J., DONG, X., ZHU, Z., JIA, L., ZHANG, Q. & LI, W. 2016. Biallelic editing of a lamprey genome using the CRISPR/Cas9 system. *Sci Rep,* 6**,** 23496.

Chapter 2 Supplementary data

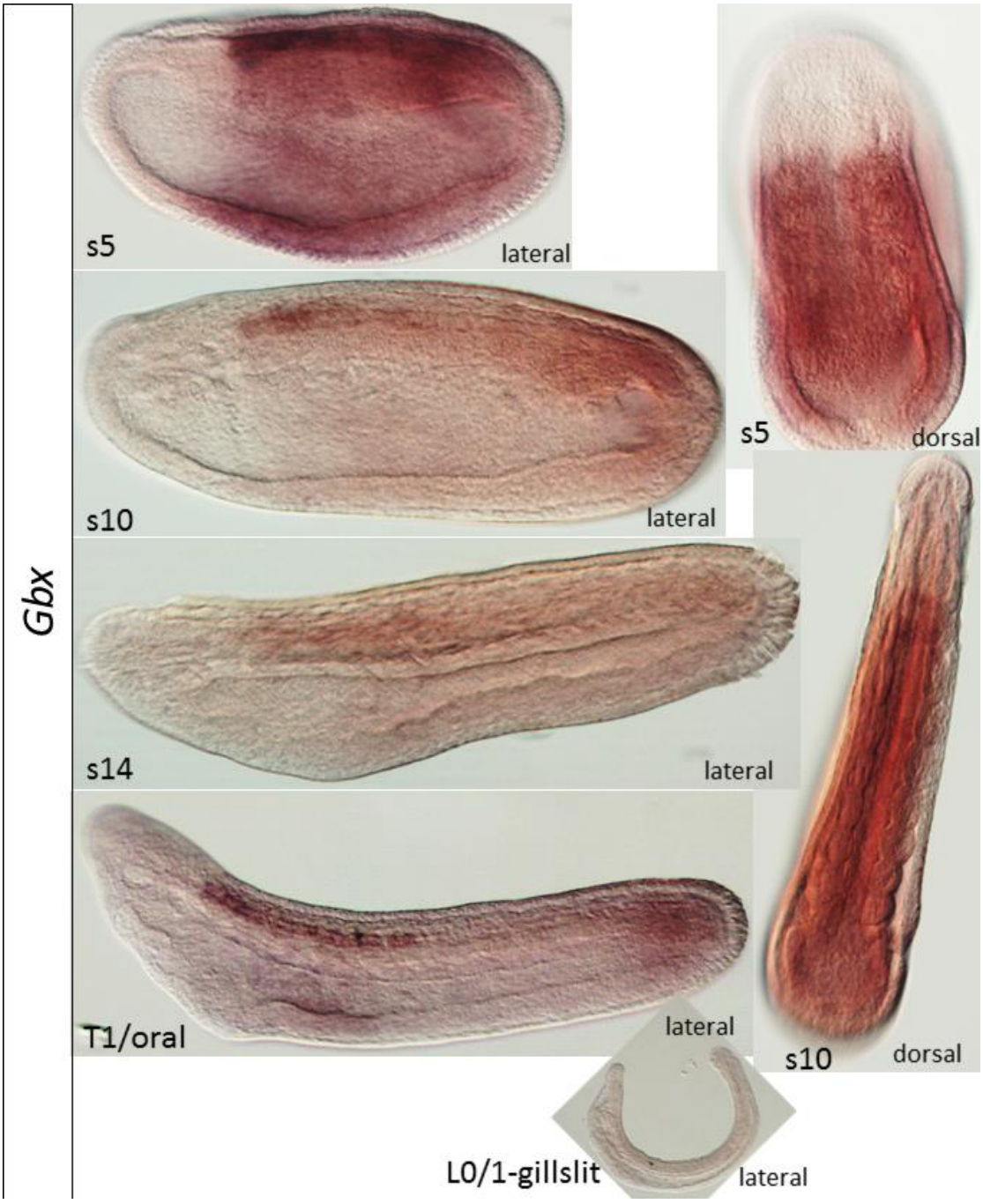
S2.1 Amphioxus *Vacht, Netrin*, and *Gbx* gene expression characterization



Supplementary Figure 2.1. Amphioxus *Vacht* expression characterization. Expressed in a set of neurons in the spinal cord. Embryo stages are at bottom left corner while bottom right shows angle of perspective.



Supplementary Figure 2.2. Amphioxus *Netrin* expression characterization. Expressed in floorplate and notochord. Embryo stages are at bottom left corner while bottom right shows angle of perspective.

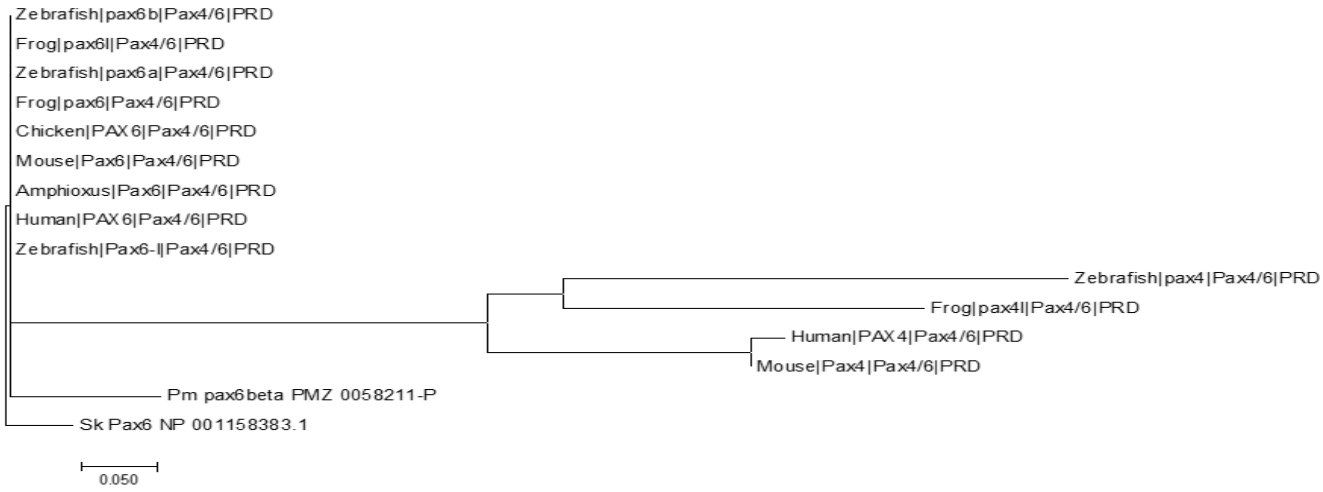


Supplementary Figure 2.3. Amphioxus *Gbx* expression characterization. Expressed in neural plate, a set of neurons at oral stage but those expression are not maintained in 1-gillslit stage. Embryo stages are at bottom left corner while bottom right shows angle of perspective.

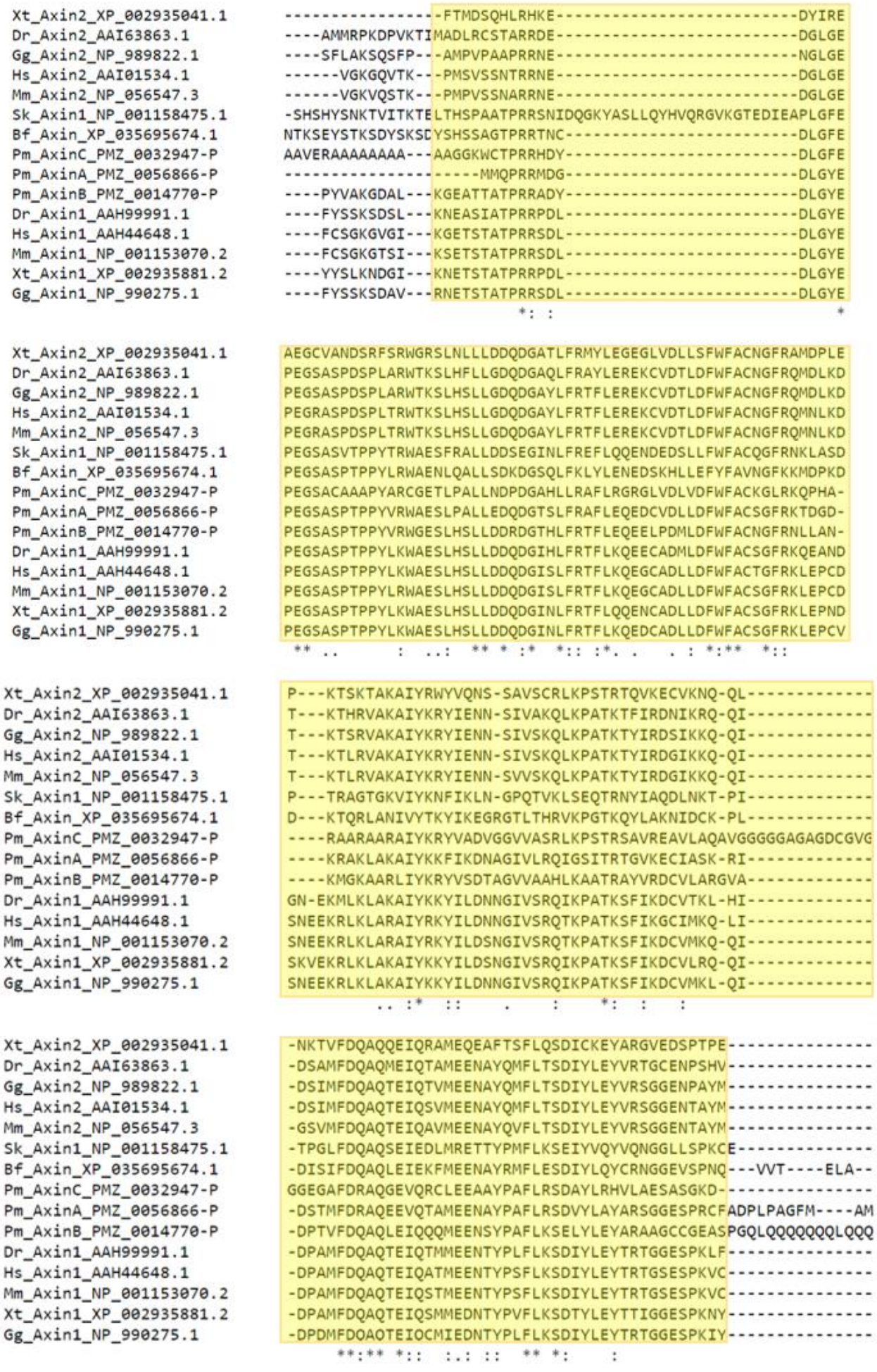
S2.2 Phylogenetic trees and alignments

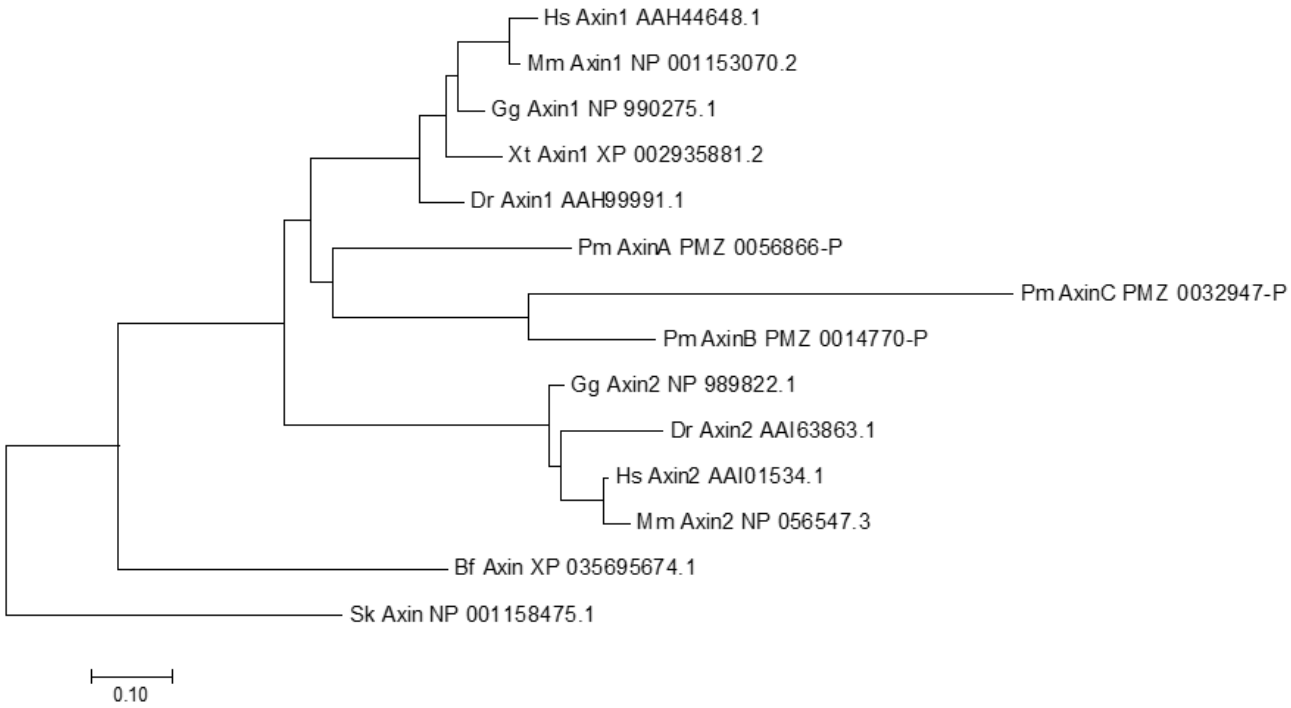
```
Pm_pax6beta_PMZ_0058211-P       ----------PLSA-EFERTHYPDVFARERLASKIDLPEAR
Sk_Pax6_NP_001158383.1          GDDDEAQARLRLKRKLQRNRTSFTQYQIETLEK-EFERTHYPDVFARERLAQKIDLPEAR
Human|PAX6|Pax4/6|PRD           ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Amphioxus|Pax6|Pax4/6|PRD       ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Mouse|Pax6|Pax4/6|PRD           ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Chicken|PAX6|Pax4/6|PRD         ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Frog|pax6|Pax4/6|PRD            ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Zebrafish|pax6a|Pax4/6|PRD      ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Zebrafish|pax6b|Pax4/6|PRD      ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Frog|pax6l|Pax4/6|PRD           ----------LQRNRTSFTQEQIEALEK-EFERTHYPDVFARERLAAKIDLPEAR
Zebrafish|Pax6-l|Pax4/6|PRD     ----------LQRNRTSFTQEQIDALEK-EFERTHYPDVFARERLAAKIDLPEAR
Zebrafish|pax4|Pax4/6|PRD       ----------HRSRTAFTADQSGRLEK-EFTCGLYPDLLTREKLAEETNLSQDT
Frog|pax4l|Pax4/6|PRD           ----------RNRTVFSTAQAEVLEKEEFMRVQYPDISRREKLATDTALSEIT
Human|PAX4|Pax4/6|PRD           ----------GHRNRTIFSPSQAEALEK-EFQRGQYPDSVARGKLATATSLPEDT
Mouse|Pax4|Pax4/6|PRD           ----------SHRNRTIFSPGQAEALEK-EFQRGQYPDSVARGKLAAATSLPEDT
                                          **    ***   * :**    * :

Pm_pax6beta_PMZ_0058211-P       IQVWFSNRRAKWRREEKLRNQRRQGGVGGGGPAGAGGGIPSSSHIPISSNFSAGVYQPLA
Sk_Pax6_NP_001158383.1          IQVWFSNRRAKWRREEKLRNQRRQMSS----------APTSSSHLPINSSFTNSVYQPIP
Human|PAX6|Pax4/6|PRD           IQVWFSNRRAKWRREE----------------------------------------------
Amphioxus|Pax6|Pax4/6|PRD       IQVWFSNRRAKWRREE----------------------------------------------
Mouse|Pax6|Pax4/6|PRD           IQVWFSNRRAKWRREE----------------------------------------------
Chicken|PAX6|Pax4/6|PRD         IQVWFSNRRAKWRREE----------------------------------------------
Frog|pax6|Pax4/6|PRD            IQVWFSNRRAKWRREE----------------------------------------------
Zebrafish|pax6a|Pax4/6|PRD      IQVWFSNRRAKWRREE----------------------------------------------
Zebrafish|pax6b|Pax4/6|PRD      IQVWFSNRRAKWRREE----------------------------------------------
Frog|pax6l|Pax4/6|PRD           IQVWFSNRRAKWRREE----------------------------------------------
Zebrafish|Pax6-l|Pax4/6|PRD     IQVWFSNRRAKWRREE----------------------------------------------
Zebrafish|pax4|Pax4/6|PRD       IKVWFSNRRARMRRE-----------------------------------------------
Frog|pax4l|Pax4/6|PRD           IRVWFSNRRAKWRRE-----------------------------------------------
Human|PAX4|Pax4/6|PRD           VRVWFSNRRAKWRRQE----------------------------------------------
Mouse|Pax4|Pax4/6|PRD           VRVWFSNRRAKWRRQE----------------------------------------------
                                ::********: **:
```

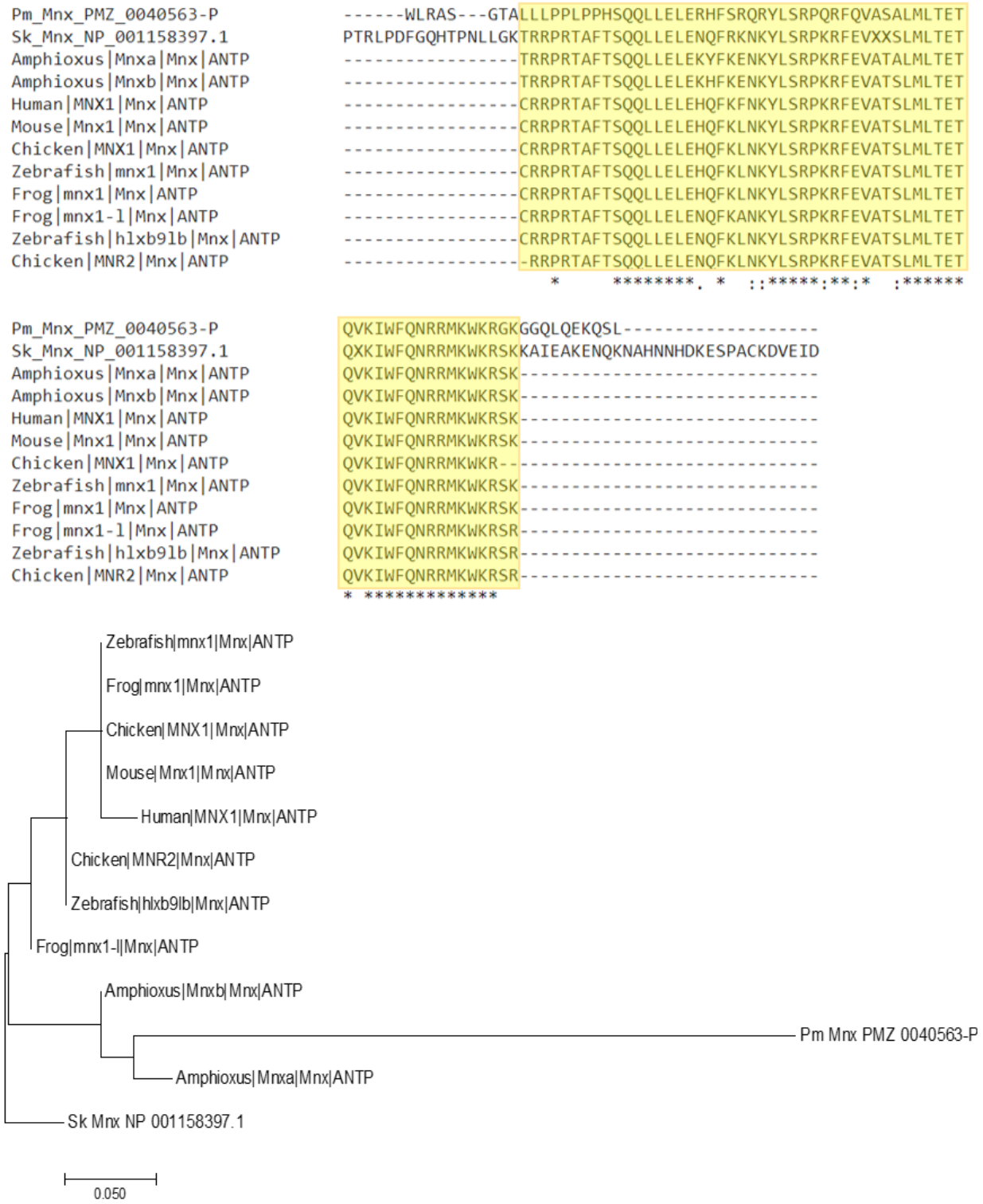


Supplementary Figure 2.4. Molecular phylogenetic analysis of *Pax4/6* by maximum likelihood. The upper panel showing alignment with the highlighted region showing the homeodomain which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Pm: *Petromyzon marinus,* lamprey. Accession number follows homologue name.

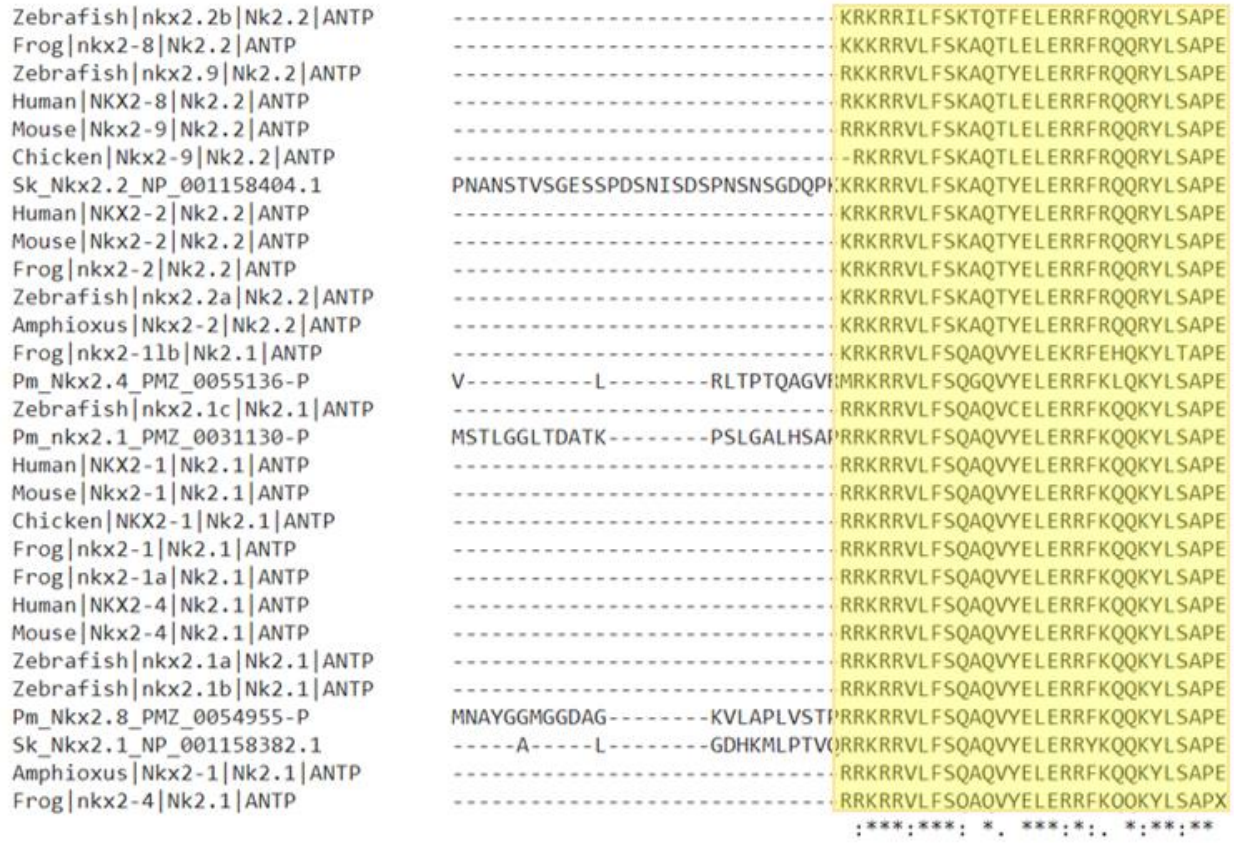

Xt_Axin2_XP_002935041.1 ----------FTMDSQHLRHKE----------DYIRE
Dr_Axin2_AAI63863.1 ----AMMRPKDPVKTIMADLRCSTARRDE----------DGLGE
Gg_Axin2_NP_989822.1 ----SFLAKSQSFP--AMPVPAAPRRNE----------NGLGE
Hs_Axin2_AAI01534.1 ------VGKGQVTK--PMSVSSNTRRNE----------DGLGE
Mm_Axin2_NP_056547.3 ------VGKVQSTK--PMPVSSNARRNE----------DGLGE
Sk_Axin1_NP_001158475.1 -SHSHYSNKTVITKTELTHSPAATPRRSNIDQGKYASLLQYHVQRGVKGTEDIEAPLGFE
Bf_Axin_XP_035695674.1 NTKSEYSTKSDYSKSDYSHSSAGTPRRTNC----------DLGFE
Pm_AxinC_PMZ_0032947-P AAVERAAAAAAAA---AAGGKWCTPRRHDY----------DLGFE
Pm_AxinA_PMZ_0056866-P ----------MMQPRRMDG----------DLGYE
Pm_AxinB_PMZ_0014770-P ----PYVAKGDAL---KGEATTATPRRADY----------DLGYE
Dr_Axin1_AAH99991.1 ----FYSSKSDSL---KNEASIATPRRPDL----------DLGYE
Hs_Axin1_AAH44648.1 ----FCSGKGVGI---KGETSTATPRRSDL----------DLGYE
Mm_Axin1_NP_001153070.2 ----FCSGKGTSI---KSETSTATPRRSDL----------DLGYE
Xt_Axin1_XP_002935881.2 ----YYSLKNDGI---KNETSTATPRRPDL----------DLGYE
Gg_Axin1_NP_990275.1 ----FYSSKSDAV---RNETSTATPRRSDL----------DLGYE
Xt_Axin2_XP_002935041.1 AEGCVANDSRFSRWGRSLNLLLDDQDGATLFRMYLEGEGLVDLLSFWFACNGFRAMDPLE
Dr_Axin2_AAI63863.1 PEGSASPDSPLARWTKSLHFLLGDQDGAQLFRAYLEREKCVDTLDFWFACNGFRQMDLKD
Gg_Axin2_NP_989822.1 PEGSASPDSPLARWTKSLHSLLGDQDGAYLFRTFLEREKCVDTLDFWFACNGFRQMDLKD
Hs_Axin2_AAI01534.1 PEGRASPDSPLTRWTKSLHSLLGDQDGAYLFRTFLEREKCVDTLDFWFACNGFRQMNLKD
Mm_Axin2_NP_056547.3 PEGRASPDSPLTRWTKSLHSLLGDQDGAYLFRTFLEREKCVDTLDFWFACNGFRQMNLKD
Sk_Axin1_NP_001158475.1 PEGSASVTPPYTRWAESFRALLDDSEGINLFREFLQQENDEDSLLFWFACQGFRNKLASD
Bf_Axin_XP_035695674.1 PEGSASPTPPYLRWAENLQALLSDKDGSQLFKLYLENEDSKHLLEFYFAVNGFKKMDPKD
Pm_AxinC_PMZ_0032947-P PEGSACAAAPYARCGETLPALLNDPDGAHLLRAFLRGRGLVDLVDFWFACKGLRKQPHA-
Pm_AxinA_PMZ_0056866-P PEGSASPTPPYVRWAESLPALLEDQDGTSLFRAFLEQEDCVDLLDFWFACSGFRKTDGD-
Pm_AxinB_PMZ_0014770-P PEGSASPTPPYVRWGESLHSLLDDRDGTHLFRTFLEQEELPDMLDFWFACNGFRNLLAN-
Dr_Axin1_AAH99991.1 PEGSASPTPPYLKWAESLHSLLDDQDGIHLFRTFLKQEECADMLDFWFACSGFRKQEAND
Hs_Axin1_AAH44648.1 PEGSASPTPPYLKWAESLHSLLDDQDGISLFRTFLKQEGCADLLDFWFACTGFRKLEPCD
Mm_Axin1_NP_001153070.2 PEGSASPTPPYLRWAESLHSLLDDQDGISLFRTFLKQEGCADLLDFWFACSGFRKLEPCD
Xt_Axin1_XP_002935881.2 PEGSASPTPPYLKWAESLHSLLDDQDGINLFRTFLQQENCADLLDFWFACSGFRKLEPND
Gg_Axin1_NP_990275.1 PEGSASPTPPYLKWAESLHSLLDDQDGINLFRTFLKQEDCADLLDFWFACSGFRKLEPCV
Xt_Axin2_XP_002935041.1 P---KTSKTAKAIYRWYVQNS-SAVSCRLKPSTRTQVKECVKNQ-QL----------
Dr_Axin2_AAI63863.1 T---KTHRVAKAIYKRYIENN-SIVAKQLKPATKTFIRDNIKRQ-QI----------
Gg_Axin2_NP_989822.1 T---KTSRVAKAIYKRYIENN-SIVSKQLKPATKTYIRDSIKKQ-QI----------
Hs_Axin2_AAI01534.1 T---KTLRVAKAIYKRYIENN-SIVSKQLKPATKTYIRDGIKKQ-QI----------
Mm_Axin2_NP_056547.3 T---KTLRVAKAIYKRYIENN-SVVSKQLKPATKTYIRDGIKKQ-QI----------
Sk_Axin1_NP_001158475.1 P---TRAGTGKVIYKNFIKLN-GPQTVKLSEQTRNYIAQDLNKT-PI----------
Bf_Axin_XP_035695674.1 D---KTQRLANIVYTKYIKEGRGTLTHRVKPGTKQYLAKNIDCK-PL----------
Pm_AxinC_PMZ_0032947-P ----RAARAARAIYKRYVADVGGVVASRLKPSTRSAVREAVLAQAVGGGGGAGAGDCGVG
Pm_AxinA_PMZ_0056866-P ----KRAKLAKAIYKKFIKDNAGIVLRQIGSITRTGVKECIASK-RI----------
Pm_AxinB_PMZ_0014770-P ----KMGKAARLIYKRYVSDTAGVVAAHLKAATRAYVRDCVLARGVA----------
Dr_Axin1_AAH99991.1 GN-EKMLKLAKAIYKKYILDNNGIVSRQIKPATKSFIKDCVTKL-HI----------
Hs_Axin1_AAH44648.1 SNEEKRLKLARAIYRKYILDNNGIVSRQTKPATKSFIKGCIMKQ-LI----------
Mm_Axin1_NP_001153070.2 SNEEKRLKLARAIYRKYILDSNGIVSRQTKPATKSFIKDCVMKQ-QI----------
Xt_Axin1_XP_002935881.2 SKVEKRLKLAKAIYKKYILDSNGIVSRQIKPATKSFIKDCVLRQ-QI----------
Gg_Axin1_NP_990275.1 SNEEKRLKLAKAIYKKYILDNNGIVSRQIKPATKSFIKDCVMKL-QI----------
Xt_Axin2_XP_002935041.1 -NKTVFDQAQQEIQRAMEQEAFTSFLQSDICKEYARGVEDSPTPE----------
Dr_Axin2_AAI63863.1 -DSAMFDQAQMEIQTAMEENAYQMFLTSDIYLEYVRTGCENPSHV----------
Gg_Axin2_NP_989822.1 -DSIMFDQAQTEIQTVMEENAYQMFLTSDIYLEYVRSGGENPAYM----------
Hs_Axin2_AAI01534.1 -DSIMFDQAQTEIQSVMEENAYQMFLTSDIYLEYVRSGGENTAYM----------
Mm_Axin2_NP_056547.3 -GSVMFDQAQTEIQAVMEENAYQVFLTSDIYLEYVRSGGENTAYM----------
Sk_Axin1_NP_001158475.1 -TPGLFDQAQSEIEDLMRETTYPMFLKSEIYVQYVQNGGLLSPKCE----------
Bf_Axin_XP_035695674.1 -DISIFDQAQLEIEKFMEENAYRMFLESDIYLQYCRNGGEVSPNQ---VVT----ELA--
Pm_AxinC_PMZ_0032947-P GGEGAFDRAQGEVQRCLEEAAYPAFLRSDAYLRHVLAESASGKD----------
Pm_AxinA_PMZ_0056866-P -DSTMFDRAQEEVQTAMEENAYPAFLRSDVYLAYARSGGESPRCFADPLPAGFM----AM
Pm_AxinB_PMZ_0014770-P -DPTVFDQAQLEIQQQMEENSYPAFLKSELYLEYARAAGCCGEASPGQLQQQQQQQLQQQ
Dr_Axin1_AAH99991.1 -DPAMFDQAQTEIQTMMEENTYPLFLKSDIYLEYTRTGGESPKLF----------
Hs_Axin1_AAH44648.1 -DPAMFDQAQTEIQATMEENTYPSFLKSDIYLEYTRTGSESPKVC----------
Mm_Axin1_NP_001153070.2 -DPAMFDQAQTEIQSTMEENTYPSFLKSDIYLEYTRTGSESPKVC----------
Xt_Axin1_XP_002935881.2 -DPAMFDQAQTEIQSMMEDNTYPVFLKSDTYLEYTTIGGESPKNY----------
Gg_Axin1_NP_990275.1 -DPDMFDQAQTEIQCMIEDNTYPLFLKSDIYLEYTRTGGESPKIY----------

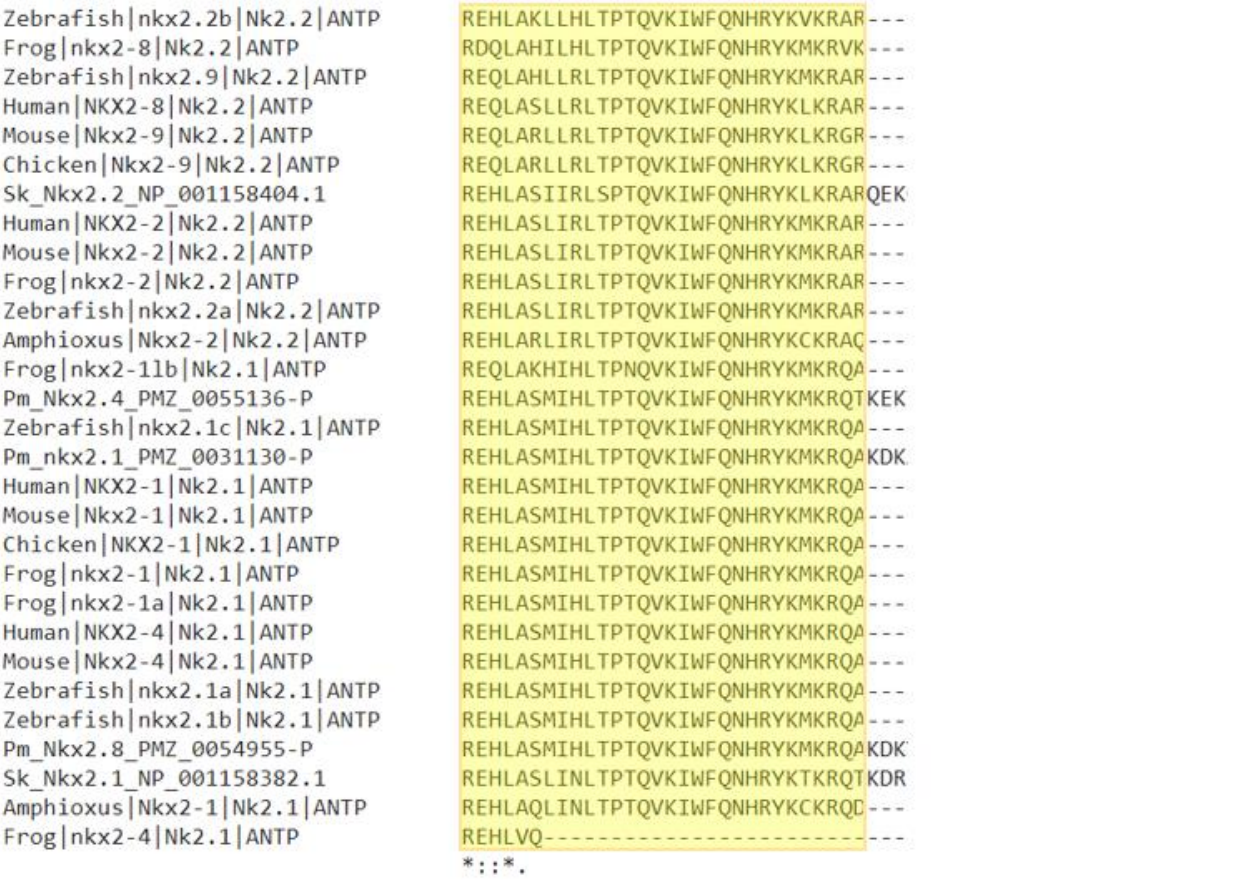


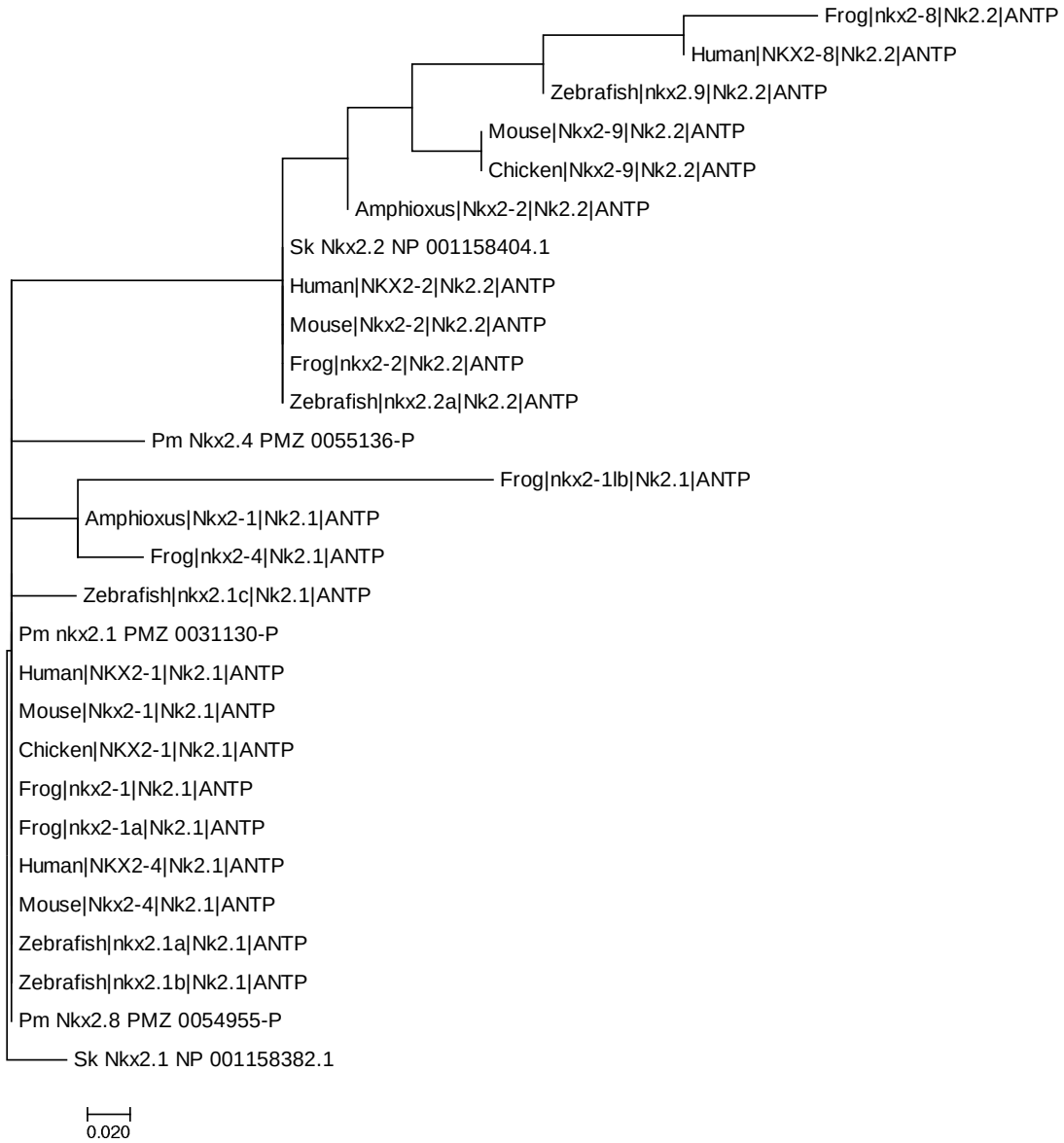
Supplementary Figure 2.5. Molecular phylogenetic analysis of *Axin* by maximum likelihood. The upper panel showing alignment with the highlighted region showing the most conserved region which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Dr: *Danio rerio,* zebrafish; Gg: *Gallus gallus,* chicken; Hs: *Homo sapien, human;* Mm: *Mus musculus,* mouse; Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Bf: *Branchiostoma floridae,* amphioxus; Pm: *Petromyzon marinus,* lamprey; Xt: *Xenopus tropicalis,* frog. Accession number follows each homologue name.



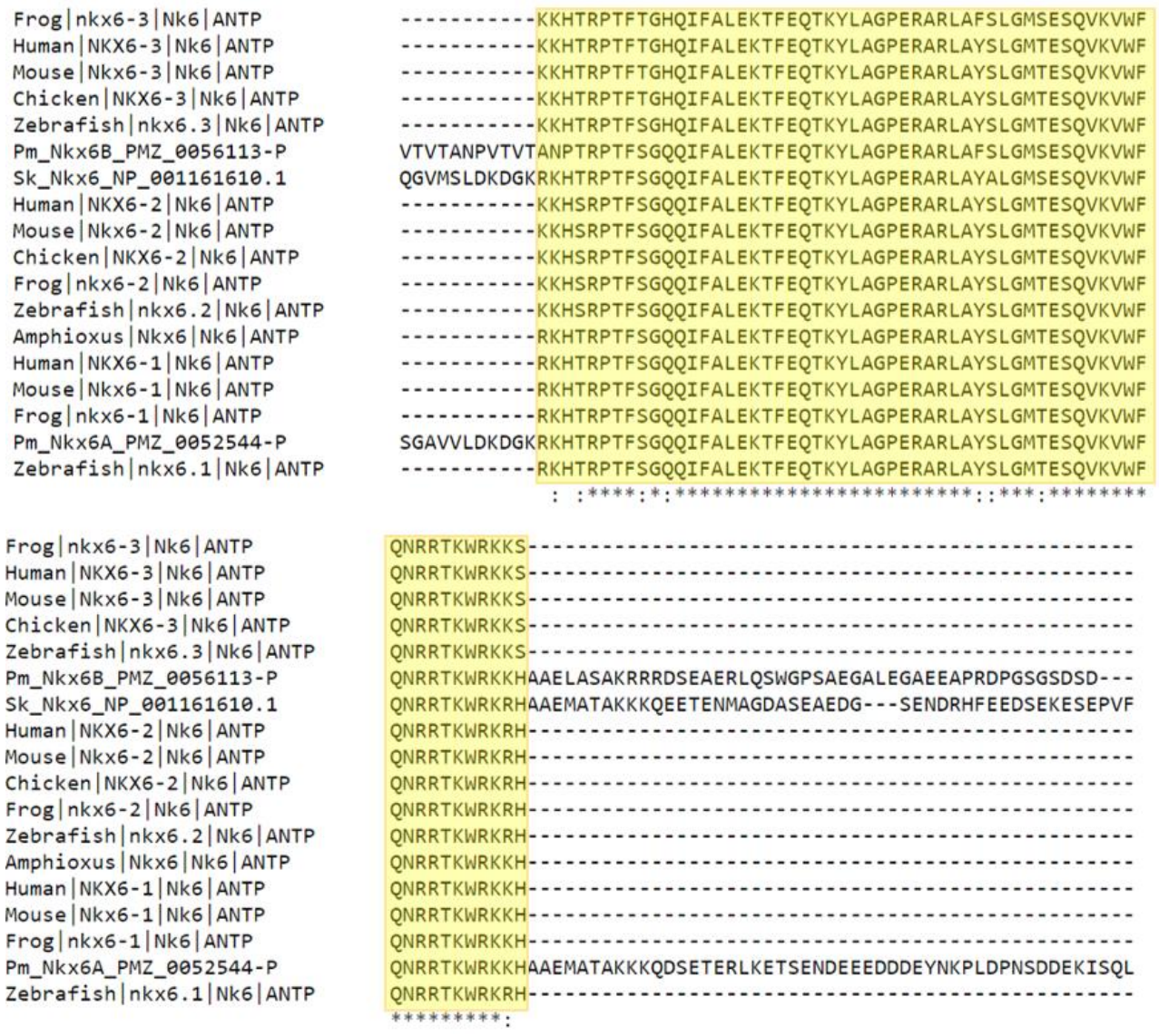
Supplementary Figure 2.6. Molecular phylogenetic analysis of *Mnx/Hb9* by maximum likelihood.
The upper panel showing alignment with the highlighted region showing the homeodomain which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Pm: *Petromyzon marinus,* lamprey. Accession number follows homologue name.

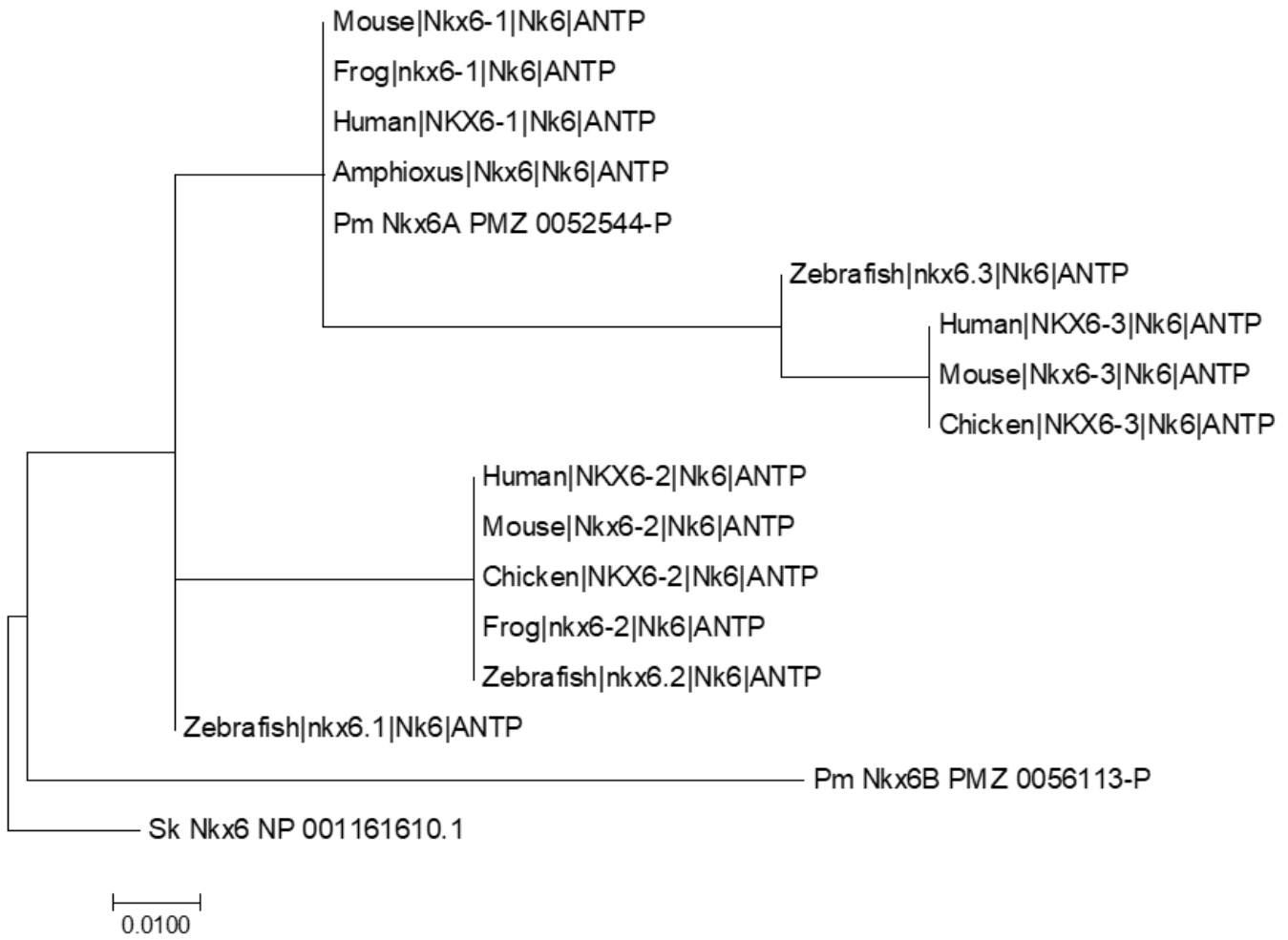

Zebrafish|nkx2.2b|Nk2.2|ANTP
Frog|nkx2-8|Nk2.2|ANTP
Zebrafish|nkx2.9|Nk2.2|ANTP
Human|NKX2-8|Nk2.2|ANTP
Mouse|Nkx2-9|Nk2.2|ANTP
Chicken|Nkx2-9|Nk2.2|ANTP
Sk_Nkx2.2_NP_001158404.1
Human|NKX2-2|Nk2.2|ANTP
Mouse|Nkx2-2|Nk2.2|ANTP
Frog|nkx2-2|Nk2.2|ANTP
Zebrafish|nkx2.2a|Nk2.2|ANTP
Amphioxus|Nkx2-2|Nk2.2|ANTP
Frog|nkx2-11b|Nk2.1|ANTP
Pm_Nkx2.4_PMZ_0055136-P
Zebrafish|nkx2.1c|Nk2.1|ANTP
Pm_nkx2.1_PMZ_0031130-P
Human|NKX2-1|Nk2.1|ANTP
Mouse|Nkx2-1|Nk2.1|ANTP
Chicken|NKX2-1|Nk2.1|ANTP
Frog|nkx2-1|Nk2.1|ANTP
Frog|nkx2-1a|Nk2.1|ANTP
Human|NKX2-4|Nk2.1|ANTP
Mouse|Nkx2-4|Nk2.1|ANTP
Zebrafish|nkx2.1a|Nk2.1|ANTP
Zebrafish|nkx2.1b|Nk2.1|ANTP
Pm_Nkx2.8_PMZ_0054955-P
Sk_Nkx2.1_NP_001158382.1
Amphioxus|Nkx2-1|Nk2.1|ANTP
Frog|nkx2-4|Nk2.1|ANTP
PNANSTVSGESSPDSNISDSPNSNSGDQP
V----------L--------RLTPTQAGV
MSTLGGLTDATK--------PSLGALHSA
-----A-----L--------GDHKMLPTV
:***:***: *. ***:*:. *:**:**


Zebrafish|nkx2.2b|Nk2.2|ANTP
Frog|nkx2-8|Nk2.2|ANTP
Zebrafish|nkx2.9|Nk2.2|ANTP
Human|NKX2-8|Nk2.2|ANTP
Mouse|Nkx2-9|Nk2.2|ANTP
Chicken|Nkx2-9|Nk2.2|ANTP
Sk_Nkx2.2_NP_001158404.1
Human|NKX2-2|Nk2.2|ANTP
Mouse|Nkx2-2|Nk2.2|ANTP
Frog|nkx2-2|Nk2.2|ANTP
Zebrafish|nkx2.2a|Nk2.2|ANTP
Amphioxus|Nkx2-2|Nk2.2|ANTP
Frog|nkx2-11b|Nk2.1|ANTP
Pm_Nkx2.4_PMZ_0055136-P
Zebrafish|nkx2.1c|Nk2.1|ANTP
Pm_nkx2.1_PMZ_0031130-P
Human|NKX2-1|Nk2.1|ANTP
Mouse|Nkx2-1|Nk2.1|ANTP
Chicken|NKX2-1|Nk2.1|ANTP
Frog|nkx2-1|Nk2.1|ANTP
Frog|nkx2-1a|Nk2.1|ANTP
Human|NKX2-4|Nk2.1|ANTP
Mouse|Nkx2-4|Nk2.1|ANTP
Zebrafish|nkx2.1a|Nk2.1|ANTP
Zebrafish|nkx2.1b|Nk2.1|ANTP
Pm_Nkx2.8_PMZ_0054955-P
Sk_Nkx2.1_NP_001158382.1
Amphioxus|Nkx2-1|Nk2.1|ANTP
Frog|nkx2-4|Nk2.1|ANTP
::.

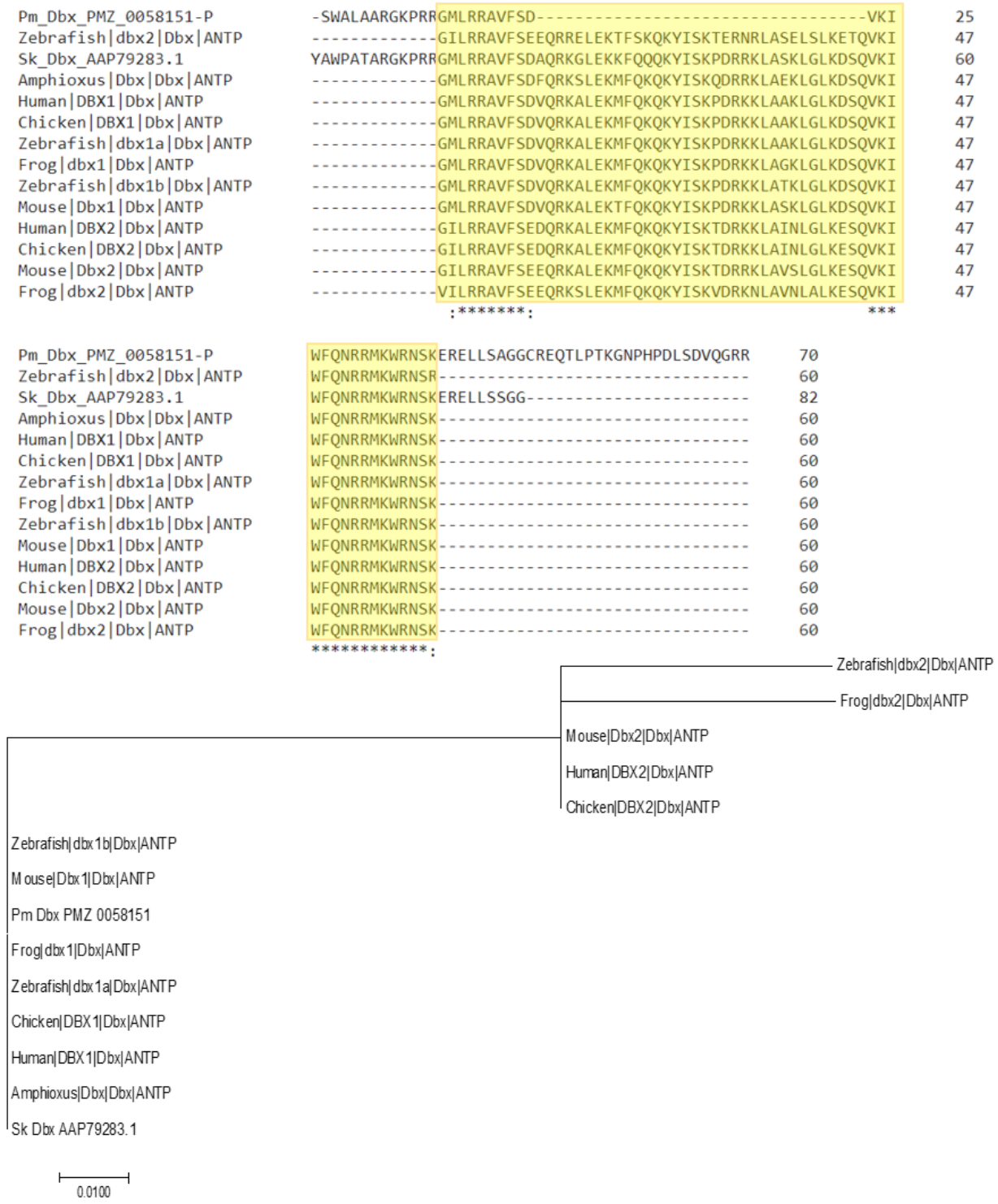


Supplementary Figure 2.7. Molecular phylogenetic analysis of *Nkx2.1* and *Nkx2.2* by maximum likelihood. The upper panel showing alignment with the highlighted region showing the homeodomain which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Pm: *Petromyzon marinus,* lamprey. Accession number follows homologue name.

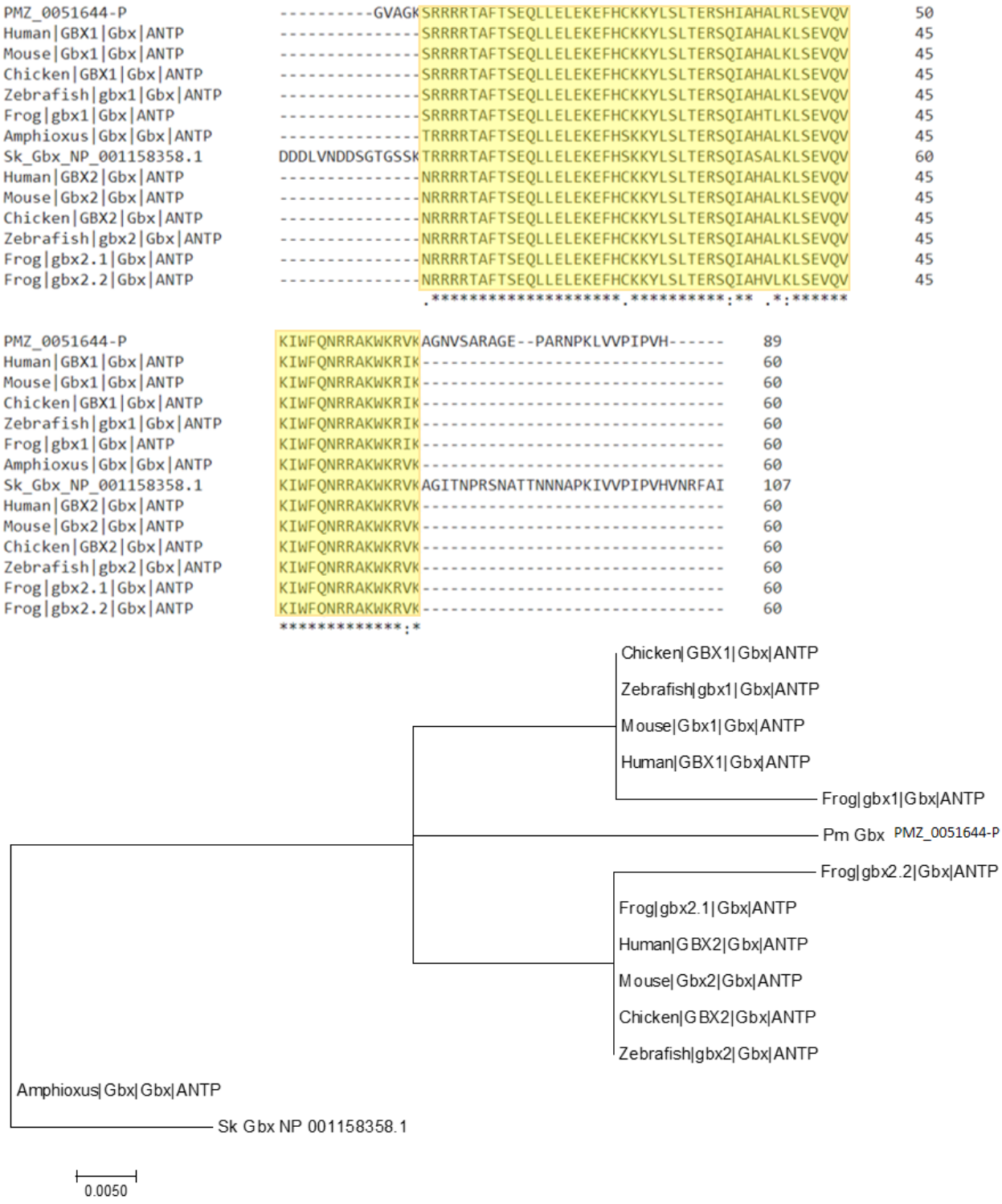

Frog|nkx6-3|Nk6|ANTP
Human|NKX6-3|Nk6|ANTP
Mouse|Nkx6-3|Nk6|ANTP
Chicken|NKX6-3|Nk6|ANTP
Zebrafish|nkx6.3|Nk6|ANTP
Pm_Nkx6B_PMZ_0056113-P
Sk_Nkx6_NP_001161610.1
Human|NKX6-2|Nk6|ANTP
Mouse|Nkx6-2|Nk6|ANTP
Chicken|NKX6-2|Nk6|ANTP
Frog|nkx6-2|Nk6|ANTP
Zebrafish|nkx6.2|Nk6|ANTP
Amphioxus|Nkx6|Nk6|ANTP
Human|NKX6-1|Nk6|ANTP
Mouse|Nkx6-1|Nk6|ANTP
Frog|nkx6-1|Nk6|ANTP
Pm_Nkx6A_PMZ_0052544-P
Zebrafish|nkx6.1|Nk6|ANTP
----------KKHTRPTFTGHQIFALEKTFEQTKYLAGPERARLAFSLGMSESQVKVWF
----------KKHTRPTFTGHQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHTRPTFTGHQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHTRPTFTGHQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHTRPTFSGHQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
VTVTANPVTVTANPTRPTFSGQQIFALEKTFEQTKYLAGPERARLAFSLGMSESQVKVWF
QGVMSLDKDGKRKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYALGMSESQVKVWF
----------KKHSRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHSRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHSRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHSRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------KKHSRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------RKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------RKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------RKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------RKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
SGAVVLDKDGKRKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
----------RKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLAYSLGMTESQVKVWF
: :****:*:************************::***:********
QNRRTKWRKKS----------
QNRRTKWRKKS----------
QNRRTKWRKKS----------
QNRRTKWRKKS----------
QNRRTKWRKKS----------
QNRRTKWRKKHAAELASAKRRRDSEAERLQSWGPSAEGALEGAEEAPRDPGSGSDSD---
QNRRTKWRKRHAAEMATAKKKQEETENMAGDASEAEDG---SENDRHFEEDSEKESEPVF
QNRRTKWRKRH----------
QNRRTKWRKRH----------
QNRRTKWRKRH----------
QNRRTKWRKRH----------
QNRRTKWRKRH----------
QNRRTKWRKKH----------
QNRRTKWRKKH----------
QNRRTKWRKKH----------
QNRRTKWRKKH----------
QNRRTKWRKKHAAEMATAKKKQDSETERLKETSENDEEEDDDEYNKPLDPNSDDEKISQL
QNRRTKWRKRH----------
*********:



Supplementary Figure 2.8. Molecular phylogenetic analysis of *Nkx6* by maximum likelihood. The upper panel showing alignment with the highlighted region showing the homeodomain which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Pm: *Petromyzon marinus,* lamprey. Accession number follows homologue name.



Supplementary Figure 2.9. Molecular phylogenetic analysis of *Dbx* by maximum likelihood. The upper panel showing alignment with the highlighted region showing the homeodomain which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Pm: *Petromyzon marinus,* lamprey. Accession number follows homologue name.



Supplementary Figure 2.10. Molecular phylogenetic analysis of *Gbx* by maximum likelihood. The upper panel showing alignment with the highlighted region showing the homeodomain which was used for tree construction. The tree in the lower panel is drawn to scale, with branch lengths measured in the number of substitutions per site. Sk: *Saccoglossus kowalevskii*, acorn worm (hemichordate); Pm: *Petromyzon marinus,* lamprey. Accession number follows homologue name.

S2.3 Lamprey and Amphioxus probe sequences

Table S2.1. Lamprey gene sequences

Gene Accession no.	Primer sequences	Amplicon sequences
Wnt1 PMZ_0030639_RA	-	ATGTGCCTACGTTCCTCGGGCGCGTATCTGGGCATCGCCCAGTCG GGGACCCCATCCAACCTGCTGCGGGCCCACGGAGGCCCTGCGGC CGCTGCGGAACCCACGGACTTGGCGCTCGCCCCGCTGGGCAAGC GGCAGCGGCGTCTGGTGCGGCACAACCCGGGCGTGCTGCACAG CGCGCTCGCCTCCGTGCACACGGCCGTCCGCGAGTGCAAGTGGC AGTTCCGCCACCGCGCTGGGACCGCCACGGCGCTGGGACTGCCC CACCACGGGCAGCGGTCCCGACATCTTCGGCAAAATCTTGCAGC GAGACACTCGAGAGGCCGCGCGGTGTTTGTCGAGGCGGCAGCG AGGGCAGCATCGCCTCCTGCACCTGCGACTACGGGCGACGTGGC CGTGGAGGTCCCGACTGGGAGTGGGGTGGCTGTAGCGACAACC TGCCTTTCGGAACTCGCTTCGCCCGGGAGCTGGCGGCCGGGGG AGGAGGTGGCCGTGGCCGTGGTGGTGGTGGTGGAGGAGGTCGT GGTGGAGGTCGAGGTCGTGGTGGTGTGGCCGGAGAGCGAGGA GGACGCGACGCACGACTCCGAT
Wnt2 PMZ_0023319-RA	-	ATGGCCATCTGCGTCGACTTCGACAAGTCTCCCAAACCCTTTGAA CTGGAGGAGGTGGAGGAGGAGCCGAGAGGCGGCGTTCGTGTA CGCCATCTCGTCGGCCGGGTGGTGCACGCCATCACGCGCGCCTG CAGCCAGGGCGAGCTCACCTCGTGCTCCTGCGACCCCCGCAAGA AGGGCAAGTGGAGGGACGCGCGCGGCGAGTTCGACTGGGGTG GCTGCAGCGACAACGTGGCGCGCACGGGAGACTTCCTGCGGCG GAGGTACAACGCCGCCATCCAGGTGGCGCTCAACCAGGACGGCA CGGGCTTCCTCGTGGCCAACCGCAACTTCAAGAAGCCCACCAAG AGCGACCTCGTCTACTCCGACACGAGCCCGGACGGTGCCTGCGA GGAGGGCCGGGCGCCTAAGCCAGCGGGATCCCCGGCGTCGCGC GGCACTGACGGCTGCGAGGTGATGTGCTGCGGCCGCGGCTACG ACACGACGCGGGTCACGCGCACCGCCAAGTGCGAGTGCAAGTTC CACTGGTGCTGCTCGGTGCGCTGCAAGGAGTGCCAGGAGAGTCT GGACGTGCACACCTGCAAGGCGCCCAAGCAGGCCGACTGGCTG GACCAGACTTAG
Wnt5_4 PMZ_003197	-	CCCTCACCACCACTCACCAGCTCACTCACCAACTCACCAGCTCACT CACTTACCAGCTCACTCACCAGCTCACCCACCCACACTGCGAGCA ACCCCAGGTTGAAGGCGACGCTGCTCTGGGCGACGACTCGGAG

6_RA		CGTGGAGGGAGTTGTCTCCTCCAGACGGGTGTCCCCAAACCCCG GGTCCCCTGAACCTCGCGGCACCCCGGGACCGCTAGGGACCAAA CAGGACCCCTCAACCACCACTAAACTCGCCTACCCCGGAGCTGAG CCACGACCGCGCCGGCGCAGCGGGCATGGGCTCCCAGGGCGCT GTCGCGTCGATGTGGGTGCGCGCGCTCGCGATCCTCGCGTGCAC GCGCAGTCTCACGGGGGCGAGCACGTGGTGGTCGCTGGCGCTG AGCCCCGTGCAGAGGCCCGAAGTGTACATCCTGGGCGCGCGGCC GCTATGCGCGCAGCTGTCGGGGCTATCCCCGGGCCAGCGCAAGC TGTGCCAGCTGTACCAGGACCACATGGTACACAT
Wnt5_5 PMZ_001477 6-RA	-	TTTGTCTCCCTAGTGGCTACGTTTACACACCGCACCAGGTACCAAT TCGAGGCTGAGTCGACCTAGGGGAGGCCCGGGAACGGTTCATAT AATCGTCCCTGGCTGTTGGCCGCGAGCGGGAATCGAACTCGGGC CGCTGGGGTTGGTAGCAGAGCGCAGCGCTTACCGCGACACTTAC CCAACCGCTCCGCGCTCGCGGGCACAAGCGCAGCTCCACAAACC CCTCGTCCCCCACGATCGTTTTTTGCAACGTGGGCCGACGGTTTT GACGAACGAGGGGAGAGAACGCAGCCCCCCGAAGGTCTCGCCA GCGTGCGTGCCGTCGAACCTCCAAGTCTCCCCCGTCACCTCTTGT TTCGCTTGTATTTCCTTTTTTTATTGTTCGCTCAGACCATCCATCCC GTGGAAAAAGCGCAAGGCGGCTTTGCAGCTCTTGCTTTTTGATC GACGTGTCGTAAATATGACCCCCATTCGGCGAGGCAGTCCCGTCG CCCTCGGCCAAAAGCCGTTGGCTTCCGCTGCATTTCAGTCCTTGG CCCGCAAGCAGACAGCCGGTAATTAATACATCTGCAACTTTAAAC AGTGCACGTGGCGCATTGGCAGAAAACGCCACAATTATGTTGTAA TTTAAGTTGTCCAGGGAGCGGGATTTTATTATTATTATTATTGTGTA CTCCCAATTACAACTGTCATCGTTCGTGTCACGTCCCCGAAAATAA AATTCTGCAGTTCACATGAATGGAAGTCGGGCAGATTAATGCCAT ATAATGAGCTTGGCAATCACACCGTTTGCGTGTTTCACATTAAATA AAATGCAGGTTATGACTCCAAATAATTGGC
Wnt6 PMZ_003064 0-RA	-	ATGCGTGTGGAGTGCAAGTGCCACGGCCTTTCGGGCTCGTGCGC CCTCCGCACGTGCTGGAAGAAGATGCCCCCGTTCCGGGAGGTGG GCGACCGCCTCAAGGATCGCTTCGACGGCGCCTTCAAGGTGATG GGCAGCAACGACGGCAAGACCCTGCTGCCGGTGGGCCACAACA TCAAGCCGCCGGGCAGCCACGACCTCATCTACTCGGCCGAGTCG CCCGACTTCTGCTCGGCGAACCGCAAGACCGGCTCGCTGGGCAC TCGCGGACGGACGTGCAACAGCACGGCGCTCGACATGAGCGGC TGCGACCTCCTGTGCTGCGGCAGAGGAGCTCGTCAGGAGACCCT CACCTTCCAGGAGAACTGCCAGTGCCGCTTCCAGTGGTGCTGCG

		TGGTGCACTGCAAGAAGTGCACGGTGCAGAGGGAGCTCAGCAT CTGCCAGTGA
Wnt7_6 PMZ_0043612-RA	-	TGGTGGAGGGGTGCGGCGGGTGGTGATGGTGGGTTGTGGTAGT GGTGGTGGTGCCAGTGGTGGTACTACAAAACAAGCGACTTAAAA CTATTAACAAAGTTACATCAGCGCTTTTGCGCGCGCTCGTCTCTAT ATATACGAGGCCGCGGCTGCCGCTGCGGCGTCTGCAGACCGAAT GCGGATCACGTCGGACGAAAGAACCGGGGCGACAATCAGCGAG TCATGCGGAGACAGCGCGTGGAGGAGACGACAGAAGAGCGATC TCTCCACCGAGGGACGAAAACTGTTTGAGCAGCGCCGAGCTGCG TGACTGAACCGGGACACAGTCGCGCCCTTAGCAGCACAACGAGA GACAACCCACCCGGCTGCATCATTATCAGCAGCAAGCAAGCATCA AGACACCTCGGACGCTCAACTCGCATCAGCAGTCGACCGGTGCG CGCACACACCGCGAACTGCCACGCGCGAGCTGGTCGGCAACAC GGACAGCATCATGTATGCGCGCATCCGCCGCTGGATACTGCATTG CCTCCTGACTGTTGCGGTCTTCTACTTTCGCATAGGGGCACTTTCC ACGGTCATGGCACTGGGCGCCAACATCATCTGCAACAAGATCCCG GGACTGGCACCACGGCAGCGGGCACTGTGCCAGAGCCGTCCGG ACGCGATCATCGTGATCGGCGAGGGCGCGCAGCTGGGAATCAGC GAGTGCCAGCACCAGTTCCGATTCGGCCGCTGGAATTGCTCGGC GCTGGGTGAGAAGACGGTGTTCGGCCAGGAGCTCAAAGTTGGC AGCCGAGAAGCGGCCTTCACGCACGCGGTGCTCTCGGCCGGAG TCGCGTACGCCATCACGGCGGCGTGCAGCCAGGGCAACCTGAGC GAGTGCGGCTGCGACCGCGCCAAGCAGGGCTTCTACAACGCGC AGGA
Wnt9 PMZ_0030638-RA	-	TGTGCAGGAGGGACCCCGGCGTCGCCGATGTCCTCCTGGGCGCG GCCGAGTTGAGTGCAGCCGAGTGCCGCGCGCAGTTCAGGGGCG AGCGTTTTCGCGAGACCGCCTTCCTCTACGCGGTGTCGGCGGCC GGCCTGGGCCACGCCATGGCGCGGGCGTGCAGCCGCGGGCGTC TGGAGCGCTGCACGTGCGACGAGACGCCGGGCCTGGAGAACCG CCAGGCGTGGCAGTGGGGCGGCTGCGGGGACAACGTTCGCTAC GCCGCCAAGTTCGCGCGCGCCTTCCTCCTGAGACGCGAGGCCAG CAAGGACCTGCGCGCCTGGGGTGGACCGGCACAACACTCACCTG GGCATCAAGTTTTGCTCAGCCTCAATCTGACGTCGCGGCGAACGC GGCGTGGCGAGGTCGCTGGCGCGGGCGCGCCGCCCGTCAGCGA GCTGGTCTACGTCGAGGACTCGCCCAGCTACTGCCGCGCCAGCC GCTTCTCCCCGGGCACCGCCAACCGCAGCTGCCTCAAGGGGCGC AACTGCGACAGCGTGTGCT

Wnt11 PMZ_0033434-RA	-	ATGAGAGCCAACGGGGCGATATCACGGGTGCCCGTGGAGCAGG TGTGGAACCACACGGGGTTGTGCCGGGGGCTGGGCGGCTTCGT GGGCGAGCAGACGCAGCTGTGCCGGCGCCAGCTAGACGTGGCG CCCGTGCTAGTGCAGGCGGCACGCGACACGCTGCGCGCCTGCAC CAAGGCCTTCGCCGACATGCGCTGGAACTGCTCCACCGCCGAGC GGGCGCCCAATTTCTCTCCCGACCTCGAGAGGGGTGAGTGGGCA CGCGGGCGGACAGGCAGGCCTGCCCCCTCGGCCGACTTCCGCTG GGGGGGCTGCGGCGACAACCTGCGGCACGGCCTGGCGGTGAC GCACCGGCACGTGGGCACGCGCCACCAGCTGGTGCCGCGCGAC ATGGACATCCGGCCGGTGCGCGACGACGAGCTGGTGTACACCGT GGGCT
Wnt16 PMZ_0042782-RA	-	ATGCTACCGGCCAGACTCCCAGAGAAGCTGGGCTGCTCGCAAAT TCCCGGACTGCCGAGCCGGCAAATCGAAGTTTGCAAGCAGAAGC CTGACATCCTGGGCAGCATTCAGCGGGGAGCGCGTACGGCCATC CAGGACGCGGGCGAAGTCTCGGGCTGCTCGTGCGACGCCTCGCT GAAGGGCGGCTCGCAGAGCGAGGGCTGGCAGTGGGGCGGCTG CTCGGACGACATCCGCTACGGCCTCACCTTCGGCCGCAAATTCGT GGACGCTGCCGAGAGCGCCGCTGGCGGCCACAACGCGCTGATG AACCTGCACAACGTGGAGGCGGGCCGGCAGAGCTCCGTCCGTG TTCACCGCACGCTCTCGATTATCTCAAGGCGCGCTGACTAA
Pax7 PMZ_0048321-RA	-	GGGAAGTTACTCGCTCGGCTCGCAGCTGAGCGCCGCGCTCGGCT CGCTCCGTGCTCCGTGCGAGGAGCGATTTCCACGATCATAGAGAA CAATAAAGCGACGACAACAACCATAAACAACAACAACCACAAAC AACAACGACGACAAGGCGAGAGAAAGTATTACAACTTTCAGCGC GGTTTATTGAACGCGCAGCAAAAGCAGCAGCATCAGCCAGAGCG ATAGACATCATCAGAAGCAGCATCAGCAGCAGCATCATCAGCAGC ACTTTGCCGGCGCGGTGTTTATGCAGCAGAGTTCTCAGGATGATG CGCCCGGCACCAGGACAGAACTACCCCCGAAGCGGCTTCCCGCT GGAAGTGTCGACCCCTCTGGGCCAGGGTCGCGTGAACCAGCTG GGGGGAGTGTTCATCAACGGGCGCCCCTTGCCCAACCACATCCG GCACAAGATCGTGGAGTTGGCGCACCACGGCGTGAGGCCCTGC GTCATCTCCCGCCAGCTCCGCGTCTCGCACGGCTGCGTCTCCAAG ATCCTGTGCCGATACCAGGAGACCGGATCCATCCGGCCCGGGGC CATCGGAGGCAGCAAGCCCCGGCAGGTGGCGAGCACCGACGTG GAGCGAAAGATCGAGGAGTACAAGCGCGAGAACCCGGGTATGT TCAGCTGGGAGATCCGGGACCGGCTGCTGAAGGACGGCGTGTG TGACCGCGCCTCGGTGCCCTCGGTGAGCTCCATCAGCCGCATCTT

		GCGCGCCAAGTTCGGCAAACGC
Dbx PMZ_0058151-RA	-	TCGACCCACGCGTCCGGAGCTCTCTGCTGCGGTCACCAATTGGCT CACTCCTGTCCTCCTCCGTGTCTGTGTTGCGCAAAAAGCCAGACA CAGCCTGGGGACAGCACCTTCACACCCATGATGTTCCCGAGTGCC CTGGCCAACGTAGCGCTCTACCACCACCACCAGCAGCAGCAGCA GCATCACCAGCAGCAGCAACAGCAGCAGCAGCACCTGCTCCACG CAGTGTACCGGCCCACGCCGACACTGGGATTCCCCGTGTCGGCT GGCAACTCCAGCTCGTCTCAACAGCAGCAGCAGCATCAGCAACC GCAACAACAACATCATCATCAGCATCAGGCAGCGCCGGGCATTCA TCAACCATCGCGCACGCTTGGGGCAAGCTTTCTCATAGACGAGCT GCTCAGGATAAACCGGGGTCCCCCAGGGGGCTACGGGCACGGG CGACAAGGTCCCGCGCCTCCTCCAAACGCGTCCCCTCCGGCGGG GCTGCTACCCCCAGAGCCGCTGTGCTCGGCGGCTGGGGCGTGCA TGCCCGTGACCACGGCGAGCGACGCTGCTACCGCGCTCAAGTTC GGTATGAGGGCCATCCTGTCGTCTCCGCCAAAGTCAGATCCCTCG ACACACTTCATACAAGAGTACTCCACCAAGGCACTTCCCTACTCAT ACTTCGACGGCCACTTTCAACAGTTCATTCGGTCGGCGTATTACCC AGCCTCGGCCGTCATCCCCATGCCCGGCTCTTTCTCGTGGGCCCT GGCAGCCCGCGGCAAGCCGCGCCGGGGCATGCTGCGTC
MsxA PMZ_0039166-RA	-	TGGCGCTGGAGCGCAAGTTCCGGCAGAAGCAGTACCTGTCCATC GCGGAGCGTGCCGAGTTCTCCAACTCGCTCAACCTCACCGAGAC CCAGGTCAAGATCTGGTTCCAGAACCGGCGCGCCAAGGCCAAGC GCCTTCAGGAGGCCGAGCTAGAGAAGCTGAAGATGGCTGCGGC GGCCAAGCCCCTGCTGACGCCCGCTGGGTTCGGGCTACCGTTCC AGTTCGGGGCCCCGCTGTACGCCACGGCCAGCGTGTCCCCCGCC TTCCCGCGGGCAGCGCTGCCCATGGGCCACCTGGGACTCTACTC GGCGCCCATGGGCTACGGCATGTACCCGCTCTCGTGAGGACCCC GGTGGCCCGTGTGGCCGCTTGTTGGACCGTGGAGCCGTCGTCAT CGCCATCGTCGCCGCCGCCGCCGCAGCTGCCACCATCGTCGCGAC CTCGGGAGCGGCGATGGCGTCACGAGTGGCCACGGCTGCGACC GGCACCTGGGCTACGTGTGGCGCAGGGTCCGGCACCAGCCTGG GCGTGTGAGTGCGCGTGGGAGTGTCCGAGGAATCTGTGTACTTG TGTGCGAGTCTGCGCGCGACCTCTGTGTGCCCGTGTATCCGTCTC GATGTCTCATGTGTGTGGCTGTCTGTGCGTTTGTATGCATATGTTTC CGCATTCATGTCTGTCTGCGTATCAGTATGCATGTTTCCGTGTGTCT CTGCGTATACGCATGTCTCTGCGTATATATACGAATGTGTGTGTCTA TTTGTCTTTATCTGTGTCTCCTTGTTTTATTCTTTATGTCTGTGCACG

		TGTACGTGAGTGAGTGTGTGTCCGATTCGCTCACACCGGGACG AGCTACGTACTCGTGGGGCCCCCACGCAGACAGAAAAACTGAAG TCCACGCGGTGCTACGCGCAATGAG
MsxB PMZ_0042426-RA	-	GCATGACTGGTCAATGTCAAAGAGTAGACCAGGTATCTCTCTCGA GAGTTCGCTCGAGTGGAGGTCCTTCGTTCAGCAGTGGTGTGAGC AGCAGCTCCAGGGTCGCATCTCTACCCGCAGCTGTCTCCCCAAAA TTGGAAGGAACACAACTCCGCTGTACACGACACGTGGGCCCACT TTCTGAACAAATGAGCGCAGCTCGTCACGAGACACAAAGTCGCC TCTCGCGTGCGGACGACTCGGGAAGAGACTCCAGTCAGTGAAAC CTCCTCCTGTGGTGAATGCATCTGCCGCTACTGGGTACACGATTTC AACTTCAAGTTACGACGCTACGTTACAATCAGCGGGACGATATTAT TTGTAAAATAAGACCGAGGTTCTTAATTAATTAAAGAATCACCAGG AATGTTAAGCCGAAATTCAGCGACACAAAACTGGAGTTTGGCGT TGTAACTCAAGTGCAGCGCATTTTTCTGCTACCAGGTCGGCTACTT AACCTTACTTCCGCGCATATATAATTATATACACTGACGTGTTTTGA ATGCGGAGAAGCATTTACAGATTTACAGAATAAAAGTGTACATCG AAAGAGACTCTGGGTACTGGGGTGTGAAGGTGTTTGATGATGAT AATTGACGCCATAATGGAATATATTTACCAAGGAAAGAATGAGTCT TAATACTATTTTTTTAGGATGACCCTGTACGCGTTGGGATGTTTGG TCAATTCAATTAATGAGTGTTTGTCACTCCCGTGAGTATTTGGAAC TGCCAAACACTCACGAGAAGTGACCATGGAACATATAGCGACTCA ATGCCCGATTCAAAAATACGAAGGAGGCATCACAACGAGTTACTC CGCTACTCCCGATGCAATCATCTGGCTCAACCACCTCAACGTAAAA TGTTGGTTGTTGCTCCATATCAACATCAGCCGCACGATGA
Nkx2.2 AB583552.1 PMZ_0043038-RA	Fw: 5'-AAACGTCGGGTGCTCTTCTC -3' rv: 5'-AGGATGGCACCGTTCTTCAA -3'	AAACGTCGGGTGCTCTTCTCGAAGGCGCAGACGTTCGAGCTGG AGCGCCGCTTCCGCCACCAGCGCTACCTGTCGGCTCCCGAGCGC GAGCACCTGGCCAGCCTGCTCAAGCTCACGCCGACCCAGGTGA AGATCTGGTTCCAGAACCACCGCTACAAGATGAAGCGAGCGCG CGTCGAGCAGGGGATACACGCGCTGCACGGCCCCGCGTGCTCC GCGCTGTCACCACGCAGGGTCGTGGTGCCCGTGCTGGTCAGAG ACGGAAAGCCGTGCGGGGCAGGCTCCACGGGGATTTTAGGAG GCGGAGGACTTGCAGGTGGTGGAGGAGGAGGAGGTGACGGAC GACTCGATGGCTCACCAACGTCCTCCCTGCATCACCACCACCACC ACAACCACCACAGCCACATGCCATCGGTGGCGGCCGCGGCTGC GGCGGCAGCGGCAGCTGCAGCGGCCGCCGAAAGTCGCCTCCGG TTGCTGGAGGTTCAGCCCGAAGCGTTCCAGACTTGCATCCCATA CGCGCTGTACGGCGCTCACCACGCCACCAAGTACAACCTCCACT

		TTAGCTACTGCGGCGCGGGGGTGGGCAGCCCCGCTTTGAAGAA CGGTGCCATCCT
Nkx2.1 AB052339.1 PMZ_0031130-RA	Fw: 5'-TTCCACCATCTC GCGTTTCA-3' rv: 5'-ACGTCCTGCCGT ATAGGAGA -3'	TTCCACCATCTCGCGTTTCATGAGCACGTCGTCCGGCATGAACAT GGGAGGGATGAACGCGTACGGAGGCATGGGCGGCGACGCCGG CAAGGTCCTGGCTCCGCTCGTCTCGACGCCGCGGCGCAAGCGTC GGGTGCTCTTCTCGCAGGCGCAGGTCTACGAGCTGGAGCGCCG CTTCAAGCAGCAGAAGTACCTGTCTGCGCCCGAGCGCGAGCACC TGGCCAGCATGATCCACCTCACCCCGACGCAGGTCAAGATCTGG TTCCAGAACCACCGCTACAAGATGAAGCGCCAGGCCAAGGACA AGGTGTCGCAACAGCTGCAGCAGGACGGCGCGGGCGGTGGAA GCGGGCAGCAGCAGTCGCCGCGCCGCGTGGCCGTGCCGGTGCT CGTCAAGGACGGCAAGCCGTGCCCCGGCGGAACCGGCGTCTCC GCGCAGGGCACGTCCAACCACCACCAGCAGCAGCAGCAGGGCC ATCACCACAACCACAGCAACCACCACCAGCAGCACCAGCAGCAG CAGCAAGGGGCGGGAGGCGTCATATCGGGCGGCGGCACCGCC TCGTCGCTCATCCAGCAGAC GGAGCAGGGCGTGGAGCTGGTGCAGGTGGACGCGCCGCTCAG CCCCTCCGACCTCCACACGGCCGTGGCGTCGATGGCCTCCATGA CGTCGCCGGTGGCCGAGTACAACATGGTGATGTCGGCGGGCTC CACTCTCCTATACGGCAGGACGT
Pax6β MH778923.1 PMZ_0053727-RA	Fw: 5'-GTTTGCACCAGC GACAACAT -3' rv: 5'-GCTGAAGAGGA AGCTGCAGA -3'	GTTTGCACCAGCGACAACATCCCCAGCGTGAGTGGTGGGGAGG GTGGGGGGAGGTGTGGGGGGTAGCGTGAGTGGGAGTTAGGG GGTTAGGGGATGGAGTGAGTGCGTGCGTGGATCGGTGAGCGT CTGGGCAACTGAGTGGGCGAGTGAGCGTGAGTGAGTGATTGG GGTGAGTGAGCAGGTGAGTAGGTGCGGATGAGCGTGAGGGCG AGCGGACGAGCGAGTGACTGTCCGAGTTGGCGAGCCGATGGCT CCATGAGCAACCGAGCGAAGCATTCCCAGTGGCAGAGGAGTCG CCGCCACAGCGCGCGCGTCTCTCGCCCGCACAGGTCTCCTCCAT CAACCGGGTGCTGCGCAACCTGGCGAGCGAGAAGCAGGGAAT GGGGGGCACCGACGGCGTCTACGACAAACTGCGGCTGCTCAAC GGGGGGCAAGGGGGGGCCGAGTGGGCGGCGCGGTCCGCCTG GTACCCCGGCAGCCCCCTCTCAGCAGGTACGGTGCCCCCCCTCC TTCTACACCCCTCCTACTCCCCCCTCCTACTCCCCCCCTCCTACTCC CCCCCCTCCTACTCACCCCCTCCTACTCACCCCCTCCTACTCACCC CCTCCTACTCACCCCCCTCCTACTCCCCCCCTCCTACTCCGCCCCC CCTCCTACTCACCCCCTCCTACTCACCCCCTCCTACTCCCCCCCTCC TACTCACCCCCTCCTACTCCCCCCCCTCCTACTCCCCCCCTCCTACT

		CCCCCCCCTCCTACCCCCCCGTAATTGACTCCCCTTGGACTCAACC TCGACTCTTAACTCCACTGCACCTCGACTCTGTCTCGACTCTGTCT CGACTCCACCTTGCCCCCAACCTCGACTCCATCTGGACGCCGCCT TGACACTGCACCCCCCCGCACCCCCCGCACCCCCCGCACCCCCCC GCACCCCCCCCCCTCGTCACCGTTTGTTGGCCGCGAGTGGAAAT CGAATTCGGGTCGCGGGGTTCGAGCGCGGTGCGCTAACCGCCG CGCCACCGCTCACCGCGCGTGCAGACGGATGCGTTCAGTCGCAC AAGCTGGGGGGCGGGGCCGGTGGGGGAGGCTCCGCGGGGGG CGGGGCTGGCGGGGTGGGCTCGGCGGGGGGCGGGGCTGGCG GGGGGGGCGAGAACGGAGCTGCGAGTTCCGCGAGCGACGACT CGGAGGAGGCCCAGATGCGGCTGCAGCTGAAGAGGAAGCTGC AGA
Nkx6A XM_032950778.1 PMZ_0041492-RA	Fw: 5'-ATCTTCTGGCCGGGACTCAT -3' rv: 5'-GAGCAGCTGGCTGATCTTCT -3'	ATCTTCTGGCCGGGACTCATCAGACGTCGCCGTGGAGAGACGC GCGCCTCGGCTGCTCCCCGACGTCCGGCGCAGTGGTTCTTGACA AGGATGGGAAGCGAAAGCACACGAGGCCGACCTTCTCGGGCCA GCAGATCTTCGCGCTGGAGAAGACCTTCGAGCAGACCAAGTAC CTGGCGGGTCCCGAGCGCGCCAGACTCGCCTACTCGCTGGGAA TGACGGAGAGTCAGGTTAAGGTCTGGTTCCAGAACCGTCGCAC CAAGTGGCGGAAGAAGCACGCGGCCGAGATGGCCACCGCCAA GAAGAAGCAGGACTCTGAGACGGAGCGTCTCAAGGAGACGTC GTCGGAGAACGACGAGGAGGAGGACGATGACGAGTACAACAA GCCGCTCGACCCCAACTCGGACGACGAGAAGATCAGCCAGCTG CTC
OligA MH020219.1 PMZ_0031563-RA	Fw: 5'-GCGGGGGTGGATCTCGCCCC -3' rv: 5'-TGTTGAGCATGAGGATGTAGT -3'	GCGGGGGTGGATCTCGCCCCGGCTGTCCAAGGAGGACCTGGA CAATCTGCAGCAGCAGCAGGAGGAGCAGCTCGGCGAGCCGTGC GGCCAGGAGCTCAGGCTGAAGATCAACAGCCGCGAGCGCATGA GGATGCAGGATCTGAACGTGGCCATGGACGGCCTCAGGGAGGT GATGCCGTACGCGCACGGGCCCTCGGTCAGGAAGCTCTCCAAG ATCGCCACGCTGCTCCTCGCCAGGAACTACATCCTCATGCTCAAC A
OligB XM_032976813.1 PMZ_0032541-RA	Fw: 5'-ATGCACGACCTCAACCTGGC -3' rv: 5'-GGGTGGAACGCCGAGTGGTG -3'	ATGCACGACCTCAACCTGGCCATGGACGGCCTGCGAGAGGTGA TGCCCTACGCGCACGGACCCTCGGTGAGGAAGCTCTCCAAGATC GCCACGCTGCTACTCGCCAGAAACTACATCCTCATGCTGAGCAG CTCCCTCGAGGAGATGAAGAGGCTCGTGGGCGAGATATACGGC AGCCACCACTCGGCGTTCCACCC

AxinA XM_032960985.1 PMZ_0043301-RA	Fw: 5’-AGTTCATTAAGGATAACGCG -3’ rv: 5’-CCTATATGGTGGGATCCCAT -3’	AGTTCATTAAGGATAACGCGGGCATTGTCCTGAGGCAAATTGGA TCTATCACGAGGACTGGCGTGAAGGAGTGCATTGCCAGCAAGC GCATCGATTCCACCATGTTCGACAGGGCCCAGGAAGAAGTTCAG ACTGCCATGGAGGAAAATGCCTACCCGGCGTTTCTCAGGTCTGA CGTGTACCTGGCATACGCTCGCAGCGGTGGAGAGAGCCCCAGG TGTTTTGCCGACCCTCTCCCCGCTGGCTTCATGGCCATGAAGGG ACTGGACTCGCCTTTGGCCTGCAGACGGAACCCTCTGTGGCCAG GGAATCCAGTGAACCCGTACCACATGAACTCGGGCTACGGACG AGCTCCTGCATCAAGCGCCAACGACAGCGAGCAGAGCTTGTCCA GCGATGCGATGACGGACGACACGATGTCAATGACGGACAGCAG TGTCGATGGGATCCCACCATATAGG

Table S2.2. Amphioxus gene sequences

Gene Accession no.	Primers sequences	Amplicon
Dbx XM_035817632.1	-	ATGTATCAAGGCGTGTTCTCGCCCGTAGCGTTCCCGCAGCCGACAGTTT TCCGGCCGACTATCACCTTCCCTCCGCCGCTGCAGCCTGCCACCACCTGT TCCACCCCGTTCTACGTGGATGACATCCTCCGGGGCTATCACACGCACG CCGGCGGCTGCTACCTCCCCTCTCTCCCGGCCATGTCCACAGCAAGCGG CCCCACGAGCCTACGGAGTTCGCCCTACGAACTGAGTCATAGCAATGGG TGTCTCTCTACCAGGAAATTTGCCGGTGTTAACGGAACATCTCCTCATCA CAGCGGAAATAACGGAACGTACCTTAAGTTTGGTGTAAACGCGATCTTG TCAGGAAAGCAAGGCAGCCCGCCCAGCCCGAGTGTAGAAAGTTGTCG GTGTGCGTTCTACCCCCAGCCCAGGAGTTACCCCGCGGCGGAGACGTA CCTCAGCCCTGTCAACAAGGTGCCTTACACAGAGTCACCGACAACCTTC CCCATCTTCAACGCGTGCCACTGGGGCGCTCCGCCGGGCCTGCACCGC GGCAAGCCCCGCCGGGGGATGCTCCGCCGCGCGGTCTTCTCCGACTTC CAGCGGAAAAGCCTGGAGAAAATGTTCCAGAAACAAAAGTACATCAGC AAACAGGACAGGAGAAAACTGGCGGAGAAGCTGGGGCTGAAGGATT CACAGGTGAAGATCTGGTTCCAGAACCGCCGGATGAAATGGCGGAACT CCAAAGAGCGGGAGCTGCTGTCCACGGGAGGATCACGTGAGTCCACC ATCCCTAACAAGGCCAATCCCAACCCTGATCTCAGTGACGTAGGACAGA CAACAGGTGCATCCTCCCCTCTGATTGGTTCCCCCTCTCCGGGCGTTACG GACCAGAAAACTCCCCCTTCCAGCCCAGAAAGCCCTAGGAACGAGGAG CCAGCTGAAAGTGAGAATAAGAAGTAG
Dlx	Fw: 5’-CGGAATTCA	TTCATGGAGATCCCGCCGGTGCAACAGACACCTGTGTCAATGTCCATG AACCATTCCAGCTACCCCGTCCGATGTCTGCACCCGACCTCGCACACCC

U47058.1	TGGAGATCC CGCCGGTGC -3’ rv: 5’- GGACTAGTT CAAAGTCTC TGTAAGAAG TTGACAG -3’	AGCACGAAAGCGCCTTCAGCACAGCGGCGCCTCACTCCCGCCCCCTCG GGTACCCCTTCTCCATGAACCCGATGGCTAACCATCCAGGCCACCCTTA CGGGAACCCTTACTCCACGGCGGCGGTCACGCCTGTCTCCGAGAGTA GGTTAGGGGGGTTCTCCGACGACGACAACGAGCGGCTATATCACTTTT CCCGGACAGACAAGTCGCCCGACCCGACAGACCAACGCATCAATGGC AAGGGGAAGAAAATGCGGAAACCTCGCACCATCTACACAAGCTTCCA GCTTCAGCAGCTCAACCGAAGGTTCCAGAGAACCCAGTACCTCGCCCT GCCGGAGAGAGCCGAGCTCGCCGCCCAGTTGGGTCTCACTCAAACCC AGGTGAAGATCTGGTTCCAGAACCGTCGTTCTAAGTACAAGAAGCTGA TGAAGCAGGGGGGTGCCCCACCACCCGCCGTCGGTACACCCAACTCCT CCGCCCCGAACCCCGTTCAACAACAGTCCGGTCAGCCCGGCCACGCCA GCCCGCCCGGCGCACCTACAGACCTCAGCGACCA ACAAATGTCTCACGGCGGCCAGCCCAACCTGCCGCCTCGGGCGCCGA GTCAGGGGCAGCCGCAGAACGGCGATGATGACGCCGCCTGGGCAGC AGCAACCTGCCATGACCAGCGCCGCACCAATGCCGACCTCCTCCCCAC CCGTATGGGACGTCAACAGTACCAAAGCCATGGCGAACTACATGAACT CTCACCACTACTCCCCCTCTCCATGGCACTACCAGGAACCGCTGCAGCA ATCCCAACAACTCCTAACATAGAGAGGCAACAACAGTATGTATTGCCG CCTGTCAACTTCTTACAGAGACTTTG
Netrin AJ252166.1	Fw: 5’- TGTAAACAG TGACCCATTC CG -3’ rv: 5’- CACATGGCA TGAAGGTTG A -3’	TGTAAACAGTGACCCATTCCGAGTGAAAATGGCCAGTTGTAGGCTGA GCGGAGGTGCTGGTGAGCGTGCTGGTGTTTGGAGTCCAGCTGTGTTC CCTGGTGCAGCAGACGTCAGGAGGGTTCGACCTCAACATGTTCGCGG CGCACAGAGCCCGCCCGACCCGTGCTACGACGAGAACGGCAAGCCCC GTAGATGCATTCCAGACTTCGTCAACGCGGCGTTCGGGAAGGAGGTG AAGGCGTCGAGCACCTGCGGCCAGCCGGCGAGCCGCTACTGCACCCT GACGGGAAACTCTGCCGACGCCGGGGAGGAACGGGCACGAGTCTGT CACATCTGCGACTCCGCGGCTCCCAAACGGCAGCACCCCCCTTCGTAC CTCACGGATCTCCACAACCCCAACAACGTGACGTGCTGGCAGTCCGAG CCTTACATACAGTACCCCACCAATGTAACGCTCATGCTATCCTTGGGAA AGAAGTTCGAACTGACGTATGTTACTCTCCAGTTCTGTTCCCTGCGACC CGAGTCCATGATCATCTACAAGTCCATGGACTACGGCAAGAGTTGGGT ACCGTACCAGTTCTACAGCACGCAGTGCAAGAAGACGTACCAGAAAC CCAACAGAGCTCTCATCACTAGGGCGAACGAGCAGGAAGCCATCTGC ACCGACGCTCACTCCGACATAGAACCTCTGACGGGCGGGCGCGTCGC CTTCTCGACGCTGGAGGGCCGTCCGTCCGCCTTCAACTTCGACAACAG CCCGGTGCTCCAGGACTGGGTGACGGCCACCGACATCAAGGTGGTCT TTAACCGCCTCAACACGCTCGGCGACGAGGCGGAGGAAGCCTCCAGA

		GACTCCTATTACTACGCGGTGTCTGACTTCGCGGTGGGCGGCCGCTGC AAGTGTAACGGCCACGCGTCGAAGTGTGTCCATGACCGGGAGGGGA GACTGGCCTGCGACTGTAAGCACAACACGGCCGGGGACGAGTGCGAC CACTGCAAGCCCTTCCATTACGACAGGCCCTGGCAACGGGCAACCGCA CGGGACGCCAACGAATGTATAGCATGCAACTGTAACCTACATGCCCGT CGGTGCCGGTTCAACATGGAACTATACAAGTTGTCGGGGCGGAAGGG TGGCGGCGTGTGCCTGAATTGTCGGCACAACACGGCCGGACGCCACT GTCACTACTGCAAGGAGGGGTTCTACCGGGACGAATCCAAACCCATCA CACACCGCAAGGCATGTAAAGAGATTCCGAGACGGCCAACTAGTGTG CCCAGCCCATCCTATCCCAAACCAGATGAACGGTGTCCAGAATGCATG GGAGCCAAGAAAGTTAACAGGAAAAACTACTGCAAGTCAGGCTATGT TATCCAGTGTAGCATACTGTCGAGGGAGATCCTAGGGGAATGGACTA AGTTCACCATAAACGTTGGGACTATATTCAAACGCGAGCAGAGCTTGA AACTTCGGCGCGGAGACCAGTACATGTACGTGCGGCAGTCGGATCTG CAGTGCAAGTGTCCCAAGCTGCGACTCGGGCGCAAATATTTGCTGATG GGTTCGGTGGACAGTAACAACAACAGAGAAGGCCTGACGCTGGACTA CAGAAGTACGTCCATACCCTGGCGGAACAGATGGGGCAAGAGACTCC GTCGGCTCCAGCGGCGCCAGAAAGATGGCAAATGTTGATTCAACCTTC ATGCCATGT
IrxA EU754746.1	Fw: 5’- GGGGTACCA TGTCGTACG CACAGTTCG GT-3’ rv: 5’- GGACTAGTT TACCTCTTTT GTACCGGTT TG -3’	ATGTCGTACGCACAGTTCGGTTACCCCTACCCGTCCTCCTCACAGTTGC TCATGTCCAGCCACGCCGCGTCGTGCTGCGAGGCCGGGCGGCCCATC ATGTCGGACCCGCACACCGGCCAGGCCGTGTGCTCCTGTCAGTACGA GGGCAGGGTGGCCGGGCTGGTGCCGCCACGGGTCGGTGAGATGTCC CTCAGCAGCGTGTACGGCTCGCCGTACGCCGGCCAGGGCTACATCACC TCCTTCGGTGCCGACCCGTCGGCCTTCTATTCTCCTTTGGCGAACCCGT ACGACCCGTCGGCAGGTCTACAGCAGCCCACCGCCTACTACCCGTACG ACCCCGCCACGCTAGCCGCCGGCGCACAGTACCCATACGGCAGCTATG GCGCGATGGACGGAACAAGACGGAAAAACGCCACCCGAGACGCCAC CAGCACTCTCAAAGCCTGGCTGAATGAACACAGGAAGAACCCGTACCC CACTAAGGGAGAGAAGATCATGCTCGCCATCATCACCAAAATGACTCT AACTCAGGTGTCCACATGGTTCGCGAACGCCAGGAGGAGACTCAAGA AGGAAAACAAAATGACCTGGTCTCCAAGAAACAGGTGCGGAGACGA GAGGAGGGAAGACGGCAGCGACAACGAAGATGATGACGATGATGAA GACAGAAAAGATCGAGAAGGAAACAGTGACGACGACAATCACGGAG GACTGGATGGAGACAGAGACAATCAGGAGTTGGATGTGGGAAATCCT TGTGAGGAAAACCGACAGGAAGACAACGCCAGCGGTTCGGTATCGGA CTATCGGCCGCCGCCTCAGTCCTCCAGCCCCGCTGATTCCCTGGGATCT

		ATTAGTGCTGCACCGAGACCGGCCAGTCCCGAGCCCAGTCCCGCCGC GAACTCTGTACCGGTCCCGCCGATTCCTTCCGAAAACAAGCCAAAGAT CTGGTCCTTAGCCGACACAGCGACCTGTAGTAGCCCTGCCCCACGGCC GGACATAATCAAGACGACACTACAAGGCTCCGTACACATGCAATTCTC ACACCATAGACCCTTCGATTCCGCGGTAACCAACCTGCGGCAGTGGGT GGACGGAGTTTTCCACCGGCCGATCGGATCATTGCCTGCGGGCCCGAT GGGACCCCTCCCCGTGGCAAGTCAAGCGTCTGCCTCGGCAACGGTATC GGGAATGCACCCTAGCTCCAATCCGCACACCACTCCCCCAGGGTTCGT GTCGTCCCAGGAGAGGACCCCTGCCCTCCCCATCGCTCTCAGGACTCA TGCTTCGCCAGCCAAAGACGAACCTTGTAGAACGGCTATAGACTTGGG CTGTGAGGCCGACCAACCCATTCGTACAGCTTTCAAACCGGTACAAAA GAGGTAA
Nkx6 XM_002609024.1	Fw: 5’-CGGAATTCATGCTTGTATCGGAGACTGC-3’ rv: 5’-GGACTAGTCTACGTACCGATGTGTGAAG -3’	ATGCTTGTATCGGAGACTGCCGGAATGGATGCATCGCGACAGAGCGC GTTCGTACTCAGTAACCCTCCACTAGCGGCCCTGCACAACATGACGGA GATGCGCGCTCCCCCGGTGTGTCCCAGCCAGTACCCCTCCTTCACGTCA ACGTCCCCCCTCAAGCCCATGTCTCTGTCCACCTTCACCAGTCAAGCCA CCACGCCGTACGGGATATCGGACATTCTCTCCCGACCGGTCACCGTGG CCAACTCGCTCAACTCCGTGAACTCTCTCTTCACTATCCCCAGGCTGAA CATGAACATGAATAACCCTGCCGCGGGCATGTACTTCAACCCGGCGGC GAGGTTTGACAAGGGTCTGGGAGACCTGTCCGGCAGGGCGTCCATGT ACTGGCCGGGCATGATGCAGGGTGCTGGATGGAGAGACAGATTCGC ATGTCCACCTCCTGGCAATGTGACAGTAGACAAGGATGGCAAGAGAA AGCACACACGTCCCACCTTTTCTGGGCAACAAATATTCGCCCTGGAGA AAACTTTTGAGCAGACGAAGTACCTAGCGGGGCCTGAGCGGGCAAGA CTGGCTTATTCCCTGGGCATGACGGAATCCCAAGTCAAGGTATGGTTC CAGAACAGGCGCACAAAATGGCGAAGAAGCACGCAGCCGAGATGGC AACTGCTAAGAAGAAGCAAGAGGACAGCAGAAGCAAATGGCCGCGC AAGACGGCGAGGAAAACTCCGAGCCAGAAGATTTGGAAGACCAAGA CAGGAACAACAATTCCAACGGCAGCAGTAGCGATACAGAGAGGCAAA AAGAGAACACGCCATCGTCTGAAGTGGTGGACACCACACTAGCGCAA ACTTCACACATCGGTACGTAG
Vacht EF118170.1	Fw: 5’-CGGAATTCATGACCGTGCTGGGATTCGA-3’ rv: 5’-	ATGACCGTGCTGGGATTCGAAACACCGTCGCTACAGGGGCTTCGGGA GCGATTCATGGAAAAAATCCAAGACCCCAAGTCTCAACGGAAGCTTGT ACTGGTGATCGTGTGTATAGCTCTCCTCTTAGACAACATGTTGTACATG GTCATCGTTCCCATCATCCCGGACTACCTACACTCCATAGGTGCGTTCG AACCGATTCAAGAGAAGACAGCCTGGAACGGCACGGACAATAAGACG TATTGGAACACCACGACAGTGGGGTACGAAAACGAGGACATTAACAT

	GGACTAGTT CAGTATTCGT CGTCCCACG -3’	TGGGTTTCTGTTCGCATCCAAGGCTATAGTCCAACTAATTGTCAACCCC CTCTCCGGAACACTGATTGATAGGACAGGATACGAGAAGCCAATGGT GATAGGCATCTCAGTCATGTTCATCTCAACAGCTGTTTTTGCTTTCGGT AGTAGTTACACCGTTCTTTTCATCGCTCGTAGTCTCCAAGGTGTCGGAT CCGCCTTCGCCGATTCGGCTGGCTTGGCCATGATCGCAGATCGCTTCA CAGAAGAAGGAGAACGAAGCAAAGCCCTGGGAATAGCACTAGCGTTC ATATCCTTCGGCTGTCTAGTCGCGCCCCCTTTCGGCGGGATTTTGTACC AGTATGCGGGGAAGCGAGTTCCATTCCTGACTCTCTCCTTCATCTGTCT CGTAGACGGCATTCTCCTCCTGTTCGTTACGCGCCCCGATAGGGAATC TGGCAAAGATAACACTCTCGTCGGCACCCCGATGTGGAAACTTCTCAT CGACCCTTACATCGCCATCGCAGGTGGAGCCCTTGTCATGTGCAACGT GTCTCTGGCCTTCCTGGAGCCAACAATCGCCATATGGATGAAGGAAAC CATGCACTCAACCGAGTGGGAAATGGGCATCGTCTGGTTACCCTGCTT CATCCCGTACATCGTGGGAGTGTGTCTCACAGTCTGGTTAGCAGGGAA GTATTGGCACTACCAGTGGCTGCTAGCTCTCGTCGGATTGGTCGTGCA AGGTCTCAGCACCTTCATCGTACCAGAAGCGACCAGCTTCGCTGTCCT CATCCTTCCAGTGGCCGGGATCTGCTTCGGGGAGGCCCTGGTCAGCAC CGCCATGCTCCCAACCCTGGCATACATCGTCGACGTCCGCCACACGTC CGTGTATGGTTCGATCTACGCCATCGCGGACATCTCCTACTCCTTGGCT TACGCGATGGGGCCAATGCTCGCCGGCAAGATCATGCACGACCTCGG CTTCCTCCAGCTCAACATCGGCATCGGGTTAGTCAACATCCTGTACGCA CCTGTTCTTCTCTTCCTGAAGAACATCTACACCATCAAACCGGACTTCC AGGAACACAGCGTGCTCTTGGATGACCCACCAGTGACGGGATCCTAC AAGACGTACCTTCAACAGGCCAATGAGGGGAAAGAGCCGATGGAGAT TCCACCTGAAGAGCAAAACCACGTGGCGCCGCCTCCACAAGTTGAAG ACGAGAGGATCAAAAAGCTTCAAGGCAGCTTCGCGCAAGACTCGTGG GACGACGAATACTGA
Lhx3 XM_002591 792.1	Fw: 5’- ATGATGACC ACAGTTCTA CC-3’ rv: 5’- TCAGTAGTG GCTGTCTAC GTC -3’	ATGATGACCACAGTTCTACCGCTCTCGCCGCCCCCGGAGAACACCAAG CCGGGTACCGCCGTCCCGTCCAGCGCACCGGTCCTCATGGACATTCGC GTGAACTCCCCGAACATCCCGAAGTGTGCGGGGTGTGAACAAGTGAT CCTGGACCGTTTCATCCTGAAAGTCCTGGACAGGTCGTGGCACGCCAA GTGTCTGCAGTGCTCGGACTGCCAGGCCCAGCTCACCGACAAGTGTTT CTCTCGGGACGGTCACGTGTACTGCAAGGACGACTTCTTCAGGCGATT TGGGACGAAATGTGCGGGCTGTGGGCAGGGGATCCCCCCCACCCAGG TGGTCAGGCGCGCACAAGACAAGATCTATCACCTGCAGTGTTTCGCCT GCATCATGTGCAAGAGGCAGCTGGCCACCGGGGACGAGTTCTATCTG ATGGAAGACGCCAAACTGGTCTGCAAGAGTGACTACGAAGCAGCGAA

		ACAGAGAGAAATGGAACTGGAGGGAACACAGAAGAGACCCCGCACA ACTATCACCGCCAAGCAGCTCGAGACCCTGAAACAAGCCTACCAAAAC AGCCCCAAACCCGCCCGCCACGTCAGGGAGCAACTGTCCCAGGAGAC GGGGCTGGACATGAGGGTAGTGCAGGTGTGGTTCCAGAACAGAAGA GCGAAAGAGAAGAGACTGAAGAAGGACGCCGGTCGGGCCCGGTGGG GGCAGTACTTCCGCGGAAACAACAACAACAACAATATCAAACGGAAA GGCTCTCCGGCTAGAAGCGAACAAAATGGCGCTGTGGACGATCTAGA CTCCGATCTCTCTAAGGATATAGTTTATTCAGATGGACAAATGGTTGC GATGAGCCCAGGAAGTGACATCTACCCTGGCGGGACTTTGGAAGGCG ACCCCGCGCTCGGCCCCATGCCTACCAACCCGGGCTACCCGGTAGACA GTAGTCCGTACCCCACACTACAAGGCGGGAACACCTACAACATGTCGC ACTCTCCACCCGTCCCTCTCGGCGGCCCACCAGGCATCCACCCCGGCAT GCCGTACGGGGATGGCGGGCTGGGCAGTCAGAGTTTCGGCCAAGCC ATGAGGGTGATGGTGGGGGGACCCGGCTCGGACCTGTCCAGCGGCA GCAGTCAAGGCGGTTACCCAGACTTCCCTGCCAGTCCCTCCTCGTGGC TTGATGACGTAGACAGCCACTACTGA
Mnxa/Hb9 AF308821.1	Fw: 5'- GTCTCAACA TGATGACGG AAAG-3' rv: 5'- TAATGGAATC ATCTTGTTGT -3'	GTCTCAACATGATGACGGAAAGTATCAAGTCGTTTCGTATCGATGACC TCCTTGCGTCGGACCGGCGCAAGCTGCAGCGCTCCTCGGGGCTGGGT GTGTGGAGCCCCCGCCAGCGAAGCCCTACCGCCTCGGAGCCAAGTCC ACGGGGCCAGCCCGACACGCCGCCCGTCTCACCCCTAGCGGAAGAGA GTCCCAAGATGTCCCCGCCGTCTAGTTTCGTCCCGAGACCGGGACTAC TACACCTGCACCCGTCCGCCGGACTTCCCCTGCACCCGCACCCGGCGT ACGGTCAGCACCACTTCCCGTTGACTGCTAACGGGACGCACCACTTCG GGTTCTCGGGCTTCCATCCACATCACCCCGACCATCTCAAGACGGCCAT GCCCTTCGAGTGGGTCCGTGCAGCGGGAATGATCATGCCAAGAATGG ATTACGGAGTGCACCCGTCCCAGGCCAGCCCCCTGTGTAAGACCCGGC GGCCCCGCACCGCCTTCACCAGCCAGCAGCTCCTGGAGCTGGAGAAG TACTTCAAGGAGAACAAGTACCTGTCCCGCCCCAAACGGTTCGAGGTG GCCACAGCTCTCATGCTCACAGAAACACAGGTCAAAATCTGGTTCCAG AATCGCAGAATGAAGTGGAAGAGAAGCAAGAAGGCCAAAGATGAAG TGACACAAGACAAATCTACAAACGGGAGCTCCAAGGTCACCGAATCCT CCAAAATGGACGACCCCAAAGACACTTACAACATGGCGGACCTTAGTG ACCCGGATGAAGACTGCAGTGAGCATGCGCTGGAGGAGATTGACGTC ACGGACAGTATTGATGGTTCATCAGACAACGAAGAGACCAGGCCAGT GCAGATATCCTTACAGACAAAATCTGACGTCACAACAAGATGATTCCA TTA

Axin XM_035839781.1	Fw: 5’-ATGAGTATTGGTGTACAGGAATAC-3’ rv: 5’-TCACTCTATGCGCTCCACCTT-3’	ATGAGTATTGGTGTACAGGAATACCTCCAGGATACTGGGGGGAGC AAGTTCACCGAGGACTTACCCACGAGTGCACGGCCGCCAGTACCA GGCGAGGAAACCGAATCCAACCATGGAAGCCAGCACAGTGGCTCC AGCAACACAAAGAGTGAATACTCCACAAAATCAGACTACTCAAAG TCGGACTACTCGCATTCGTCAGCGGGTACGCCCAGACGGACAAAC TGTGACTTGGGATTTGAACCAGAAGGGAGTGCATCACCTACGCCA CCGTATCTACGCTGGGCAGAAAACCTGCAAGCATTGCTGTCAGAC AAGGATGGTAGCCAGTTGTTTAAGTTATACCTGGAGAATGAAGAC TCCAAACATCTCTTGGAATTCTACTTTGCTGTCAACGGCTTTAAA AAGATGGATCCCAAGGACGACAAAACACAGAGACTAGCAAATATT GTGTACACCAAGTATATCAAAGAAGGAAGAGGGACCCTCACTCAT AGAGTAAAGCCTGGGACAAAACAGTACTTAGCCAAAAATATTGAT TGTAAGCCATTGGACATATCCATATTTGACCAGGCCCAGTTGGAA ATAGAAAAGTTCATGGAGGAAAATGCATACAGAATGTTCTTGGAG AGTGACATCTACCTTCAGTACTGCAGAAATGGTGGGGAAGTCTCG CCCAACCAAGTAGTGACGGAACTGGCGAACACGTCCGGTAATGTG GGAAACTTGACAACGGGAGTCATATTGGCTCCTTTGCCAACGCTA CATGAAAATGTCGAACTTAGTACCGAAGACACGATGGAGACGCGG ACGACGCTTCCGCTGACGGCAGAACACCTTTTTGCTACCTCACAG ATAAGGCTCAGGAGTGAGAGCAGAGACAGAGAAAGATTGACCAGA GGAGGGAACCCGTATCACCAGACGTACGCACCCGCACCTGCAACA TCGGCGAATGATTCTGAGATCCAGAGCTTGTCCAGTGAACAGACT GATGACACCATGTCGCTGACGGACAGCAGCATTGACTTAGATGGG GTCCCCAACCTGAACCATCGGCACAGAAGGCGGAGACAGCTGAAG GGGATGAGACAAGACGCCCAGCGCAACGGCAACGTCATGTGCTAC CCTCCATTTCCACAACCGAGAACTCAGAGGCTTCCCAAAGAGATG AAAATGATGGAGCCAAAGAAGTTTGCAGAACTGTTAATAGCGAAA CTTGAGAGAGTGAAACAAGACCGGAGAATGAGGAACGACTGGGAG GGAGACAGTGGTGCTAGTCAGGTATCAGACTCTCAGGGAGCGTCA GGATCGGACCCAGCGCAGCCCAGGCTGTTCCCGCTGATGGACGAC ACCACCAGCATCGATGACGACCCCGAGTCTATCCTGGACGAACAC GTGTCCAGAGTACTGCAGACGCCGGGTTGTCGCTCGCCGGGCAGA CACTCTCCCGTCAGACCGAAGGGTCACATCTCAGGGGTTCCCCAT CACATGTCGTACCCGGGCGCGCTCGCCGGCGCCGCAGGCTCGTCC TGGCACAACAGGTCACGACCGAAGGACGGTTCAAAACGGGACAGC GGCATCGGAGACTCTTCTGTCACGGAGGCAGAGACTGTAACTCAT TACCATAAATATTATTACCATCACCACCACCACCATCCCAAGTCA

		AGGGAGGGGTTGGAGAAACAGACAGAGATGGAAACACAGAGGAAA GTACAGAGAGACTCAGAAATGTACACCACGCAGTCCCGGGTACAC CAGGACAGCATGGGCGGCACGGCGGAGTCCCTGCGACGGGAGGGA CGGAGAGATGTTCTGGGGAAGCGTGGCGGCGGGAGGAAAGAGGAG GGCGAACCGTTCCTGGCTCTGCCCGTGGAGACGGATAGATTTGGC AAGGTGTGGGACTGGATGGAGGAAGGTGAGAGGGAGAGGTACAGA GGGGGGATGAACATCAGAGAGCCAAAGTCCAAACGGTCATCCAAG ACAGCTTCCCAAGCAGCCCACAGAAGCTCTGGAAAAGCAAGCTCC AGCTCCGCGAAGAAACCCTTGGCGTACAACACGTCCCGTCCAACA TCGATGGAGCGGCCAGCGGTGACGGCAGAGGTGCGGCCCATGCAG CCCTTCATGCAGGACACACAGATGCCCGTTCTGCCCCAGCCCAAC CCCCTCATCTGCTTGGAGGAAGCGCGCAGACGCTTGGTGGAGAAA GCCAACCCCAAGGCCAGCAGCAGGCATGTTGTGGACAGTGGTGCC AGGAAGGACAAGGGCAAGCTGCCACAGTCCAAGGTCGTCCCTGCC GTGTCCGACATGGAACCAGAAGAAGAGCCCAAGGAGAGGAAACAG CCCATCAAGAAGCCAGCGTCTACAATGTCCGTCTCGGGCGACAAC ACCGTGGTCGGGTATTACTTCTGCGGCGAGCCCATTCCGTACAGA ACGACGCTCCCCGGGAAGGACATCACTCTAGCTATCTTCAAAAAC AGCATCACCAAGAGAGGAAACTACAGATACTTCTTCAAGTGCACA TCAACAGAGTTCGGCTCCAATGCAGTTTACCAGGAGGTGACTGAA GATGAGGACTTGCTGCCACTATGGGAGGGGAAGATCGTCGGAAAG GTGGAGCGCATAGAGTGA
Gbx DQ416766.1	Fw: 5’-AAATACCTCTCGCTAACCGAGCGATC-3’ rv: 5’-TCAAATATAATAATTTAATGGA-3’	AGAAATACCTCTCGCTAACCGAGCGATCCCAGATCGCGCACGCGC TCAAACTCAGCGAGGTGCAGGTGAAGATATGGTTCCAGAACAGGA GAGCGAAGTGGAAACGGGTGAAGGCCGGGAACACCAACACCCGCG GACCCAACCCCAACCCCAAAATCGTGGTGCCGATCCCCGTCCATG TGAACCGCTTCGCCATGAGGAGTCAGCATCAACAGATCGGCAAGA GTCACCAGGGTTTGTGACGGGACGGACACAAGGTAGAACTCCACT GTAAATGTAATATGGACACATGTGTAGTTCACGAGCAATATTAGT TTATAATGATAAATACAGTGCAAGGGCACGGAAAACAGGGGTAAC GTGGCTGGGAACTCTGCTTTCTAAAGTTTATTTACGTCCGGAAAC GGTCTTTTTCTGGTACGCAAAAATATTTCAGGTCGGTTTGCTGAA TTCTAAAATTGTGATGGAGTAGGGTGCGAGCGTTGTGTGAAGAGA ATGAATATGGAAGACGATGTAAACAGTTCGGCCTCTGAAAGCTGT GTTTAGAAATCAAGAAGAACTCTGTAATGAATTTTAACTTTATCC CCGAATGCCATAACATTTCTTCACGTTAAAACGTCGATGTGCATT GTAAACTTACTAAGAAATTATAAAAGGAAACTAAAAACTTTTGAC

		GTTTTGTGAGTGAATGAAAAGTGTCCCGACACAGTCCAGACTACC GGTAGTCCGAGTGTCGGCTGTTGTCGGCTCAGTATCACAAAGGGA AGCCATAAGAATGGCGGCCTGTGAGACACAGGAGTCATGAATACC TTCTCATTTCGCGAAGTGAAAATTGGCATGTTTTTGTGTCGGTTT CGTGACAATGGGCAGAAATTTAGCGCCTTTGTTCTTGTGTTAAAA CGAGTCGCCAAACAACAAGGTCGGCTCGGTTTGATGCGAGGGAAA CCCGCGCGGCCTGACATATGGCTTCACAATGACTAGCCTTCCGTC AAACCGTCTGTGGTATCAAAATACGTTCACATTTCACCACCTCTA AATCAGAATAACTTCTTGATATCTCTTTTATTCTTTGCTGAAATG TTAACAGTAAAATCGCACGGTTGACAGTGTTCCCTGGCCCCAGAG TGGCAGTGAAATCTCCAAGTCACCCCGCTCATTTTACCGACCGCG AATCGCTGACATACTTAGCAGAATTTCACAGCAAACTCCTGCCAA ATTGCCCGGTCCAACCCTGCCACATTCCCCCTGGTTTCCCGAGGG GTAGCAGAGATATTAAAGCGGGTTTGTGAGAGCAGAGGTACTCAT TTCTCCGTTAAACAAGGAGCGAGTTTTTAGCCACGACGTTTCTTT CACAGCAGTTAGAACGGACGGACAATGGTACAGAATAATGGGCGA TGTAAATTTCTATGGTATATTTCCGCAGACAATGTTGTATAGTTG TCACCAGAGTTATTGTTATAGATGACTAATTCCCTAGATTTATTT CGACAGTATTGTATTCGCCTGGTTCGCTTCCATTAAATTATTATA TTTGA
Pax2/5/8 XM_035812698.1	Fw: 5'-GGTGACCCGGAGCAGAAGCG-3' rv: 5'-CATGTCACGGCCTGGCGTCAC-3'	GGTGACCCGGAGCAGAAGCGTTCAACGTTCACACCAGACCAACTAGA AGCCTTGGAACAAGCGTTCAATCGTGGCCACTACCCAACCGACCCCTT TAACAGGGACAACATGTCAAACAAAGTGGACTTGTCACAGACCCGAG TCCAGGATGTAAAACCTAGCATCAGCTGTTCCACCACAAGTGTAGCAA TGACAGACACAGCACCTCACGTCCCCACCGGCCACTATCCAGTTGCTG TTCCAGCCACCACCCCACCCTCAGCCTCAGCTCCATCCAACAGTGCCTC CCCCCCTGGGAGCGCCAAGTCCTCCGAACCCCCTCCCTCCCCCACCGCC ACTACCACCACCACCACCACTGCTGCAACCTCAAACACCGGATCGCTCA CTGTACTGCAGCCAGTGGGGTCGGCATCCTTAGGGGGGTACCCCCTCT CGACCCACGCTGGGTCAGGTACCCCCTCGTTTGGACAGTACACCAGCC AGGGGGTTCTGGCAGGTCTTCCACAGTCGGCAGTGACGCCAGGCCGT GACATG
Zic AJ252245.1	Fw: 5'-CAAGGTGGTGTTTGCATGCA-3'	CAAGGTGGTGTTTGCATGCAAACGGAGAGCACTTTTCTCTGTCTTCCT GCCTCTGCGTTCATTTGCAAATAGAAGCGGCGGGAGCGTCATGAGTG CGGAGGCATTGCCGGCCCAAATTATGGATCACAGCTTTGTCAAGCACG GCCCGGCGCTGCGGTTTGTCGACGTTGGACATCATCAAAATATCAATG

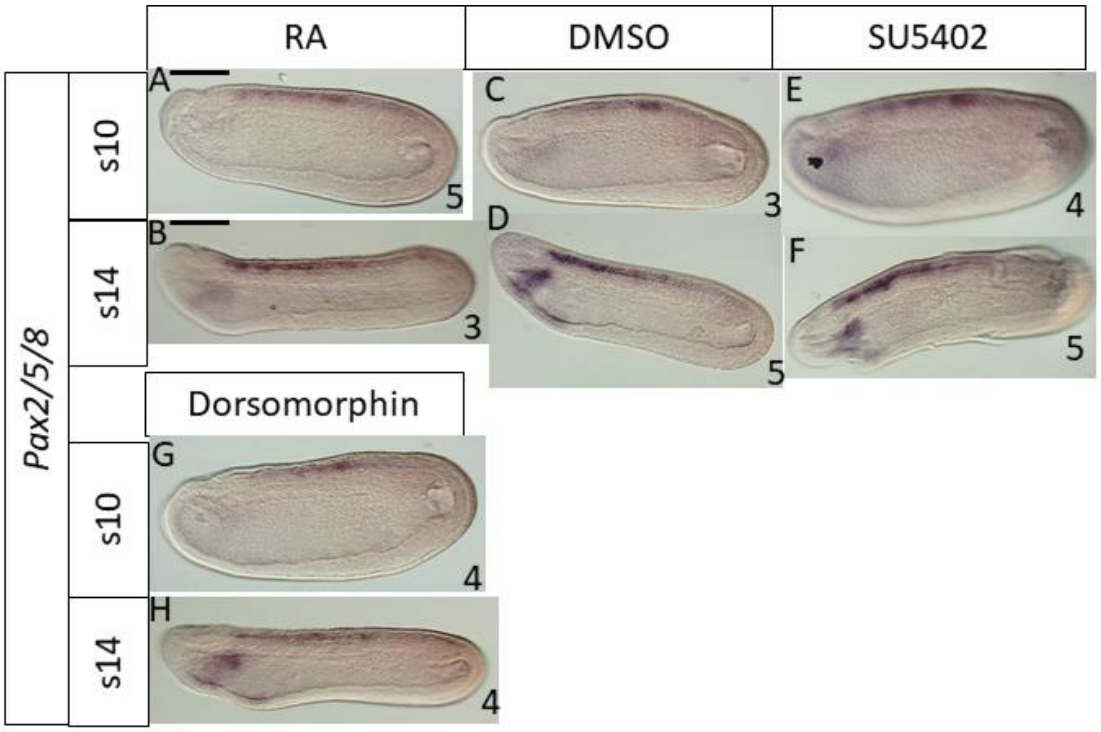
	rv: 5’-CCAAATTTTTACACAACTTTA-3’	GAGTCACAAGTTTCGGATTAGCGCCACATACGCATCATTCCAGTGAGGGAAGTGTGGAGGCCGGCAGCATGAGTCTCGGTCCATTCAGCAGGGACAATTCCCACCACATGGGTGCCCTGAAACTCAGTCCGCCGCATCCCATGAGCGAGATCCCCAGCAGCCAAGCCAACTTCAGCTCGCCGACCAACGGGTACAGTCACGCTCACCACCACGGGCCGGTGTCGGGCTACCCCACGGCAGCAGCGCAGGCCTTCACCGCCAGCAGGGACTTCCTCCTCCGACGAGAGCATCTGGCCGGGACGGTGCCGCTGCCGGGGTCTCTGACGCACGACCACAACACCGCCGGCTCCGGGATGTTCGTGCCAACTTCTGGCAACTTCCACGGCCACGGGCACGCCGACAGCTCCCACCTCATGTTTCCGGGCTTCCACGAGCAGGCCGCCGCCCACCATCCGCAGAACGCGCACCACGTCAACGGACTCAGCCTGCCCGGCGACGGCTACGGCTTCAACCAAGTGTCGAGCCCGCGTCGCGACCATCTCATGTCCAACCAGCTCAACCATCACATGGCGCCCATGAACATGCCACCGCACGGCCCTGCAGCCTTCTTCCGCTACATGAGGCAGCCCATCAAGCAGGAGGCTGTCTGTCTCTGGATAGACCCCGACCAGCCCGAGCCCAAGAAGCCCTGCAACAAGTCCTTCTCCACCATGCACGAGATCGTCACCCACGTCACGGTGGAGCATGTCGGCGGGCCCGAGTGCACCAACCACGCCTGCTTCTGGAAGGACTGTCCGAGAGACGGCAGGGCCTTCAAGGCCAAGTACAAACTGGTCAACCACATACGGGTCCACACGGGCGAGAAGCCCTTCCCCTGTCCCTTTCCAGGCTGCGGCAAGCTCTTCGCAAGATCCGAAAACCTCAAGATACACAAGAGAACACACACAGGCGAAAAACCGTTTAAGTGCGAGTTTGAGGGGTGCGACCGGCGGTTCGCGAACAGCAGTGATCGGAAGAAGCACTCGCACGTCCACACTTCGGACAAGCCCTACAACTGCAAGGTCCGGGGGTGTGACAAGAGCTACACGCACCCCAGTTCGCTACGGAAACACATGAAGGTCCACGGTAAAACTTCTCCTACTCCAGACGACGAGAGTTTCGGGAGCCCGTCCTCCACGAGCTCGGAGACTCACCGGGACTTGCCGCGGATCACGCAGCACCCTCTCCCGACATCTGCCACCCTCGCGCCGACCCCGACACACGCCACCAACCTCAGCGAGTGGTACGTGTGCCAGTCCACCGGAGGGATGCCCACTCCGCCTTCCAACGAACATTCGCCGATATCTTCGTGCGTAAGTCCAGTCCACTTAACAACTACCCCCGTAGCCAGCGGTATTGGTCATTTCTGACTAAGCCAGAGACTCTGACATCACACTTACTCAAACCCAGACAAGACTGGTTACGAAAGAAAAAGAAACTAGACAAGTTCGTCTTGCATAAAATCATACTTTGCCATGTCTGTTATGATTTTACATAGTCAGAACAATCCATGCGAAAGGCCTGGCTGCAGCCAGTCATCGCGTCCAATGTGCCACATTTGTAGAGTTAAAGAATTGAAAGCAATATTTTGTACAGAGGATGTATACTAGAAAGATATCGTTTTTCGGAGAAGGTGCTCTCTGTTGTTATTTTTTGTCGACATACCTCGCTTTAGTTTGTTTCTCAACCGTCGTTCGCTTGTGTAGCTTCGATTTTGTCTTATTAGTGAG

		AGTCATTCGCTGAAAAGTGTGAGGTTTTAGTCTTCTTAATGCGAACAG TGAAACGTGGTCTTTCCCACTATTGTATTGAAAGAACTCTGTGAGAGA AATAATCTTCATATTTCTGAACAAAAAGCATTCCAATTTTAAGCGACAT TGTAGAATAATAGTCTAGTACGCCGCCATAGCCCCTGACTAGGGCTTA TGTAGTCGCTGAGTTATGAACTAAGTAGAGAAAACCTCACGAGAGAG ATCATGATATTTATTGTAAAATATTGACTTTTAACTCTATTTTTTACGTC GCGACATGTGTAAAAAATTCAATATATATTGTTCATAGTTCAGCCTAAG TTTTTGTACGTCCGTACAATAAGCAATGTAGACTCGTGAGTGGAAACA GTGCACTGTCTTTTAGACCGTGGATGGTAGATATAACATATTCTCAGA CCACATGTTAGATACAGAACTGAAATCACAAAAGAAAGCTCAAAGTTT TCGAAGCCCAAGTTTTTAGATTGTCTCAAGCTTGTAAATAACGTAAAAA AAACTTGTTGTCGGAGTTGTTCTCAGTTGCCTAACATTTTCATAATTTTA TGTTTTGTTTATTTATTGATTATGTGCCATTCCCATTCTAATGAACGTAT CGAGAGAAGACAGTTGAAATTCTGTACTAGTTGTTGAGTTCTCCGGTA TTTGCCTGTGTTTGTCGTTCCTACACAAATTGTTTTTTCTTTGTTCATCC GCACACGTGCAGGCCAGATTAAAGTTGTGTAAAAATTTGG
Soxb1a AF271787.1	Fw: 5’- GGCTCGAGT ACGTCCACG CC-3’ rv: 5’- CATAGCTTG ATACCTATTT A-3’	GGCTCGAGTACGTCCACGCCAAGCCGACTGTGTCCTCCAGACTTTCGC GTGAGTAACCCCCCAAAAATGTCTACCAACATGATGATGCACCAAGAG AGGAAGGAAGACGAAGCAGCGAAGAAGGATCATGTGAAGCGGCCCA TGAACGCCTTCATGGTGTGGTCCCGCATCCAGCGGCGGAAGATCGCC GAGGACAACCCGAAGATGCACAACTCGGAGATCAGCAGGCGGCTGG GTGAGATGTGGAAGGAACTGTCGGCGGAAGAGAAGAAGCCGTTCGT CGAGGAGTCGAAGAGACTCCGAGCCGAGCACATGGAGAAGTATCCCG ACTACAAGTACCGCCCGCGCCGCAAGCTCAAAACCATGATGAAGAACT CCAGCTTCCCGACCATGCCGGGCGGCATGCCGCCTAACTACTTCCCGG CGTCCAACTGCAGCCCCCCGCCCGCCTTGCAGATGAGCATGCCGGAGC AGTACGGACACATGAACGGCTACATGAGCGGCTACCACTCCTCGATGA TGCGGGACCCAGGCAACCCGTACAGCCACCAGGTCATCCCGTCTTCTG CAGCGGCGACTACAGCAGGCAGCATGCCGACCAGGTATGACATGGCG GCCCTCACCTACAGCACCTACGCGTCCATGCCGCCTACCACGTCCTACA CCACCAGTAGCTGCGCGCTCCAGCCGTACCCCGTCTCATCCTACCCCAG CACCGACTTCTGCGCCCCGGTCAAGCCCGAGCGGACCCTCCCCGCCAG CTACGACGACCACATCCGACCGTACTACAGCGAACATTGCTGAAATTC TGGATGCGGACATTGTGTTAAGTATCAACTTTACCCGCACTGCAAATC ATAATTTGGGTCCAATTCGAACTCTTTCAACACAAAGTGTGCTTGGGTT AAGGTGACTAGGATATATCGGTATCCGATTTTCAGTGTAGTGTATTTA GCATTTCTTGAAAGTAGATATCAGTTTGTCATCATGTAGGTCTTATATG

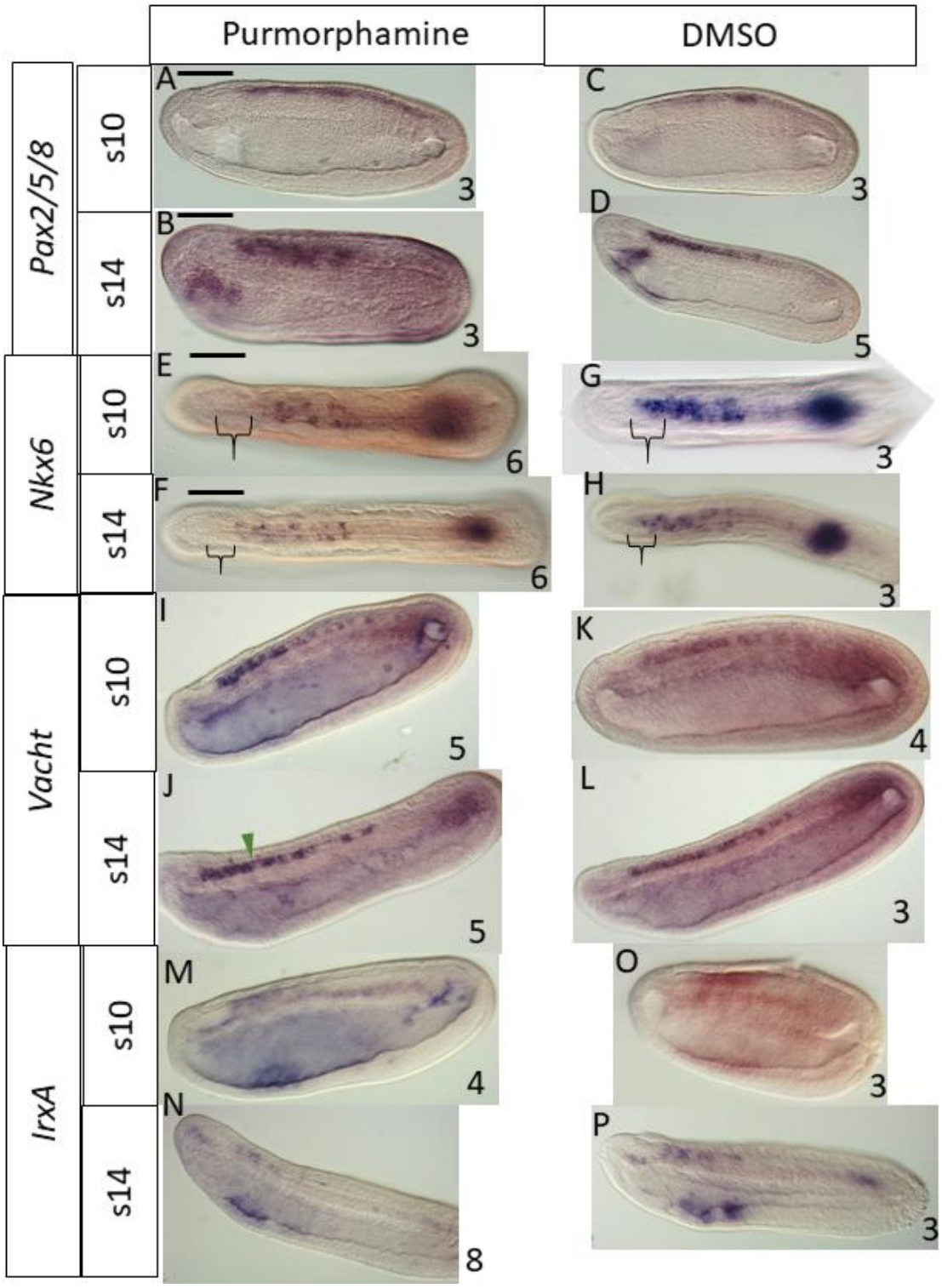
		GCCTCCTTCGGTCAATATTGACCATGGTCTACGTGCTTATATTTTAGAT ATTATGTGCATTTTTTGTGCCACAGGTGGTTTAAAAAAGAAAACAAAC TTTGTTTTCCTATGCCACGAGTTCCAATACTTTTGCTGCTCGGTGCATTT TCTGGCCTTTTTTGACCAGCTAAACTCACATATTTGTTGTGTGGCTTCG AAAAGTTTTGGATGATTCTTTTGTACAGAATTTAGCGGGCAATTATGG ACATTGTTACTAGGCGGACGAGTCCAAATTATGCAGCCATGGGGGGA GATGGGTCAAATCGCCTTTGTCCAAACTTGGTGATGAGAAATTAGCGT AAAAAGTATCAATATCACTGTAGAATCGTTCTGGACCCAATTAATGCA GTAGTTAGTGTCCTATTTTGTAAATGTATTTACACATTGTTTATGTGTA GAAAACTTCGTAAAGATGACGAAGAAAACTGACCTCCTTGTCCTAGAA GGAAATTTTATGTGAAGTGGAAACAGGTTTTCCATTGTAGAAATGATA ATAGCGCATGCGCAATTACGGTGGTCAGAATTTTTCCGACTATGCAAA TCCAAAGAATGCATCCAAACGATCATTGTAACCACCGCTAAGACAGCT TGCGAATTTTGATGTATATGGTGCCAAACTTTGTGAAACAATATATTTG GGCTGAACTTTTTCTACTATGATCAATGTAAATAGTAATGTTAGGGTTT ATAATGAAGTGTCCAGAAGGCAGCTATCACTCTATCTATGTGTAATATT GCAACTGTTGTATAACAGACGAAGTAGTGATTTGACAGAAATTGACGA GCTACTTACCTTTGTAATGTAGCAAAGAGAATAAATAGGTATCAAGCT ATG

Chapter 3 Supplementary data

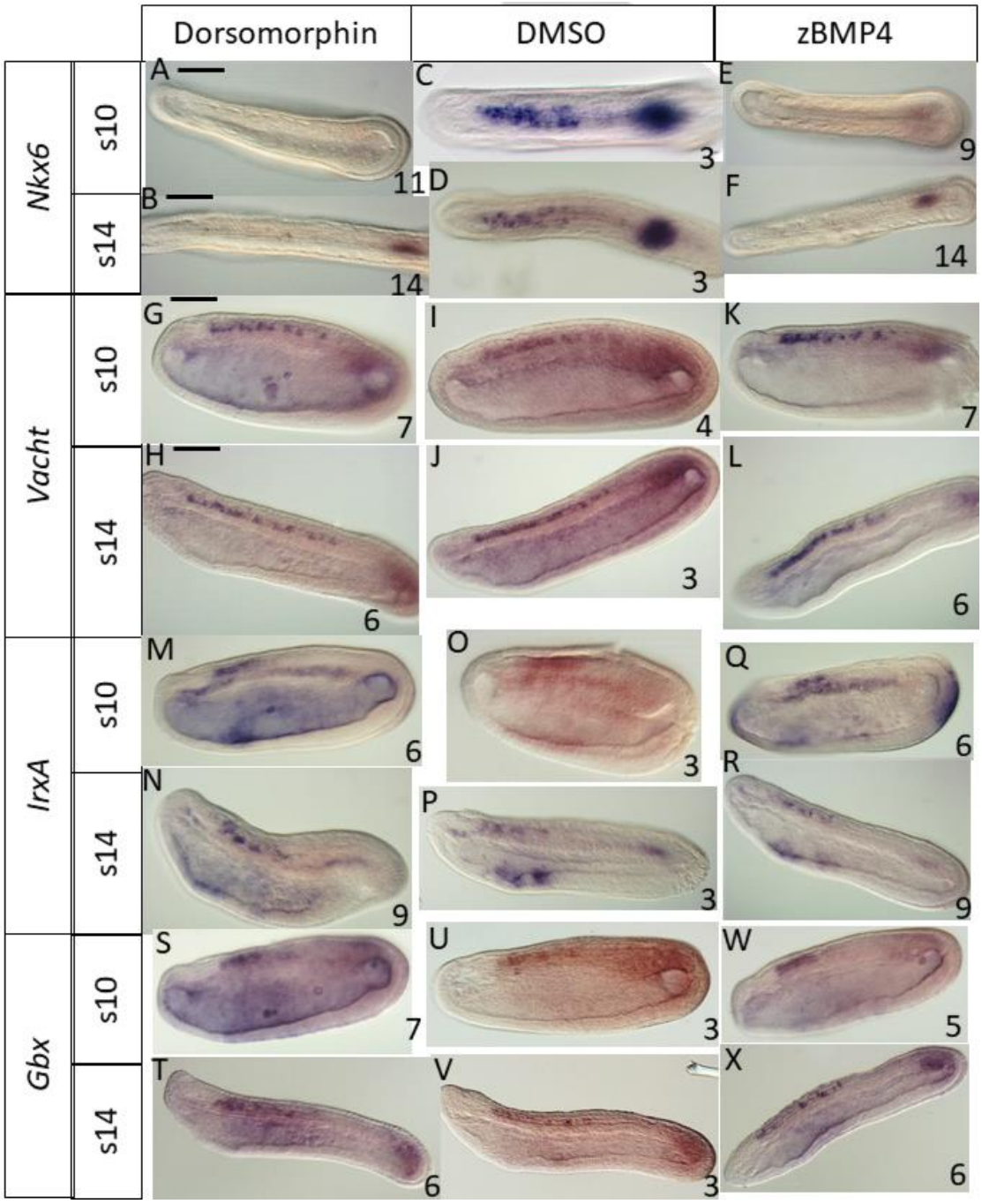
S3.1 Amphioxus drug treated data



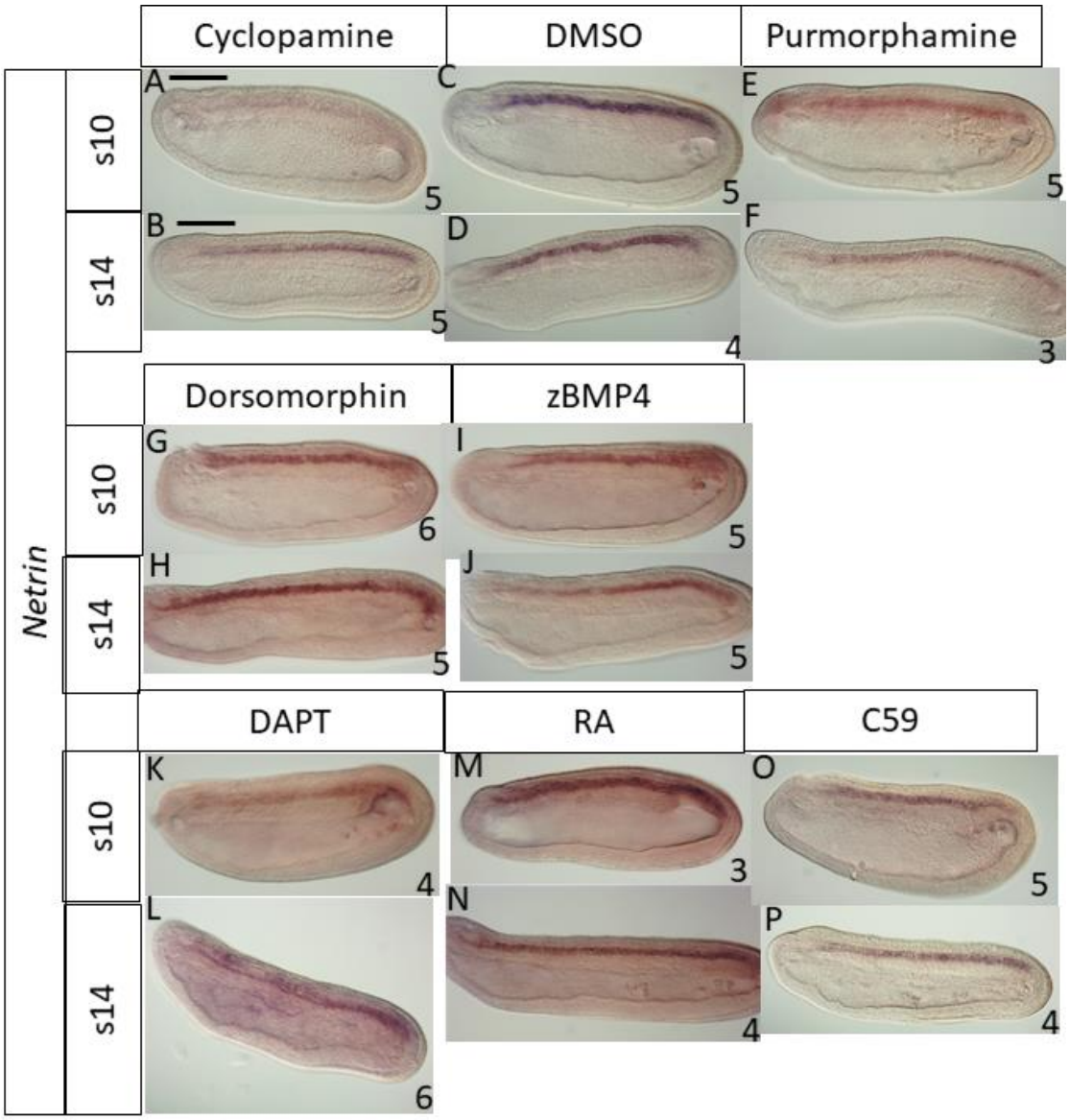
Supplementary Figure 3.1. Effects of RA overactivation, Fgf inhibition, and BMP inhibition on amphioxus *Pax2/5/8* expression. No observable change for *Pax2/5/8* expression. Scale bars: 50um.



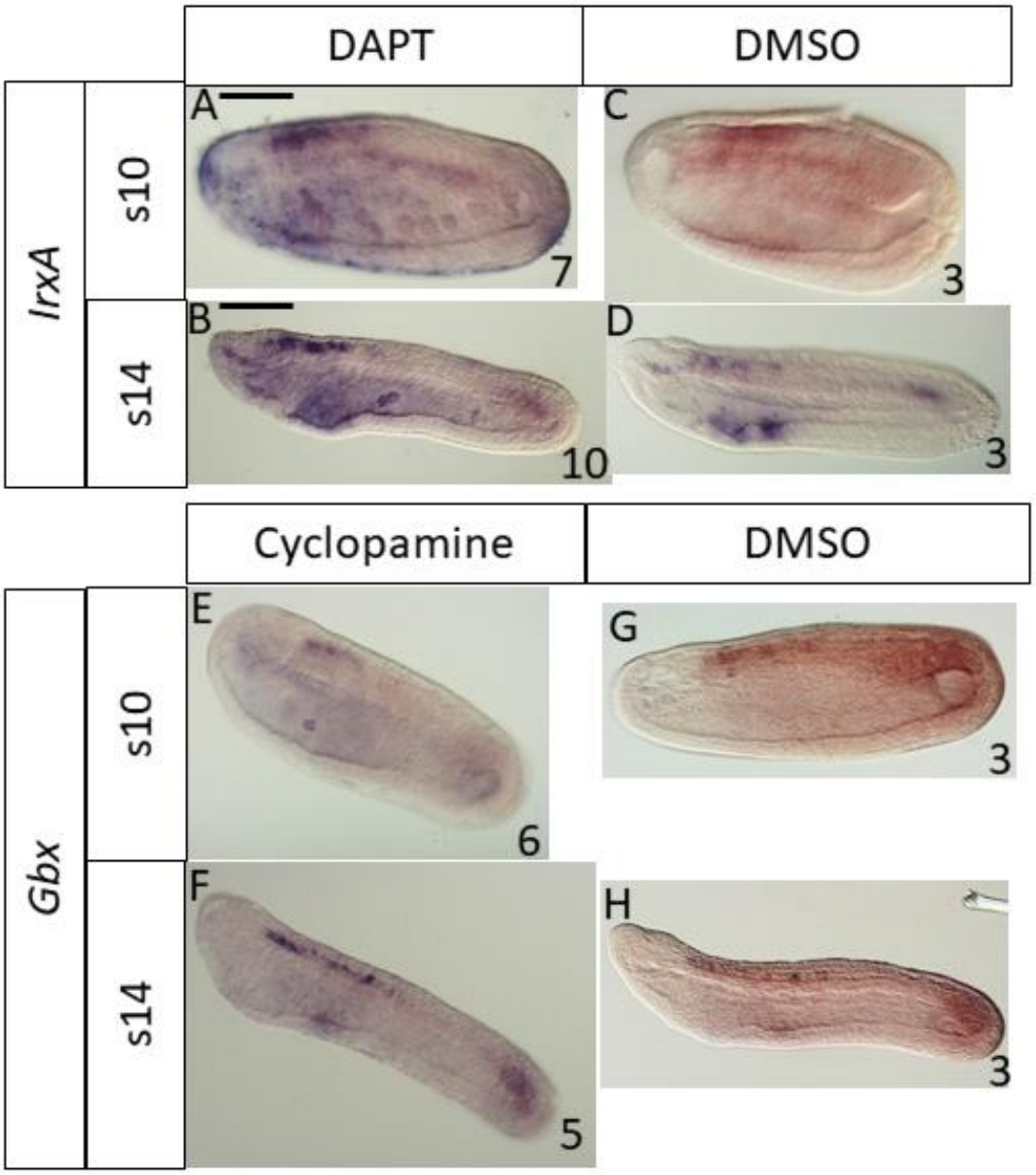
Supplementary Figure 3.2. Effects of Hh overactivation on amphioxus *Pax2/5/8, Nkx6, Vacht,* and *IrxA* expression. No observable change for *Pax2/5/8* and *IrxA.* There is a loss in *Nkx6* neurons around the level of the first somite as marked by black brackets. Green arrowhead indicating a loss in *Vacht* neurons. Scale bars: 50um.



Supplementary Figure 3.3. Effects of BMP signalling manipulation on amphioxus *Nkx6, Vacht, IrxA,* and *Gbx.* For *Nkx6,* there is a varying degree of loss of expression in both up and down regulated embryos in the spinal cord and posterior endoderm. There are no observable change for *Vacht, IrxA,* and *Gbx.* Scale bars: 50um.



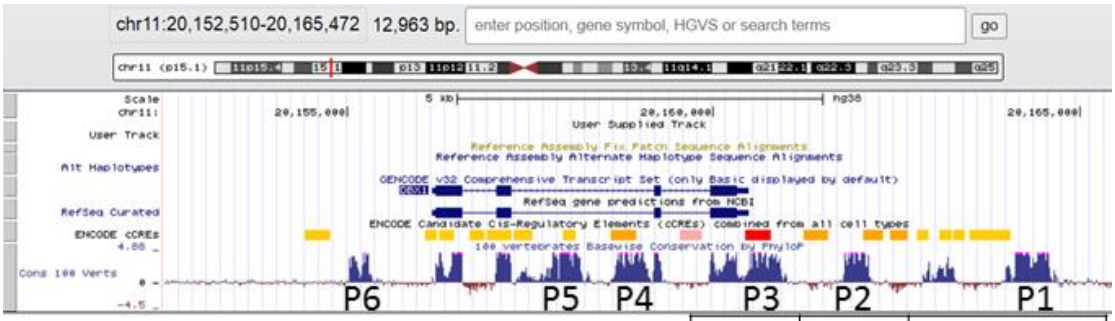
Supplementary Figure 3.4. Effects of Hh, BMP manipulation and Notch, Wnt inhibition on amphioxus *Netrin* expression. There is a partial loss of *Netrin* expression in the Hh inhibited (cyclopamine-treated), BMP overactivated and Notch inhibited (DAPT-treated) embryos. No observable change in RA or Hh overactivated, Wnt or BMP inhibited embryos. Scale bars: 50um.



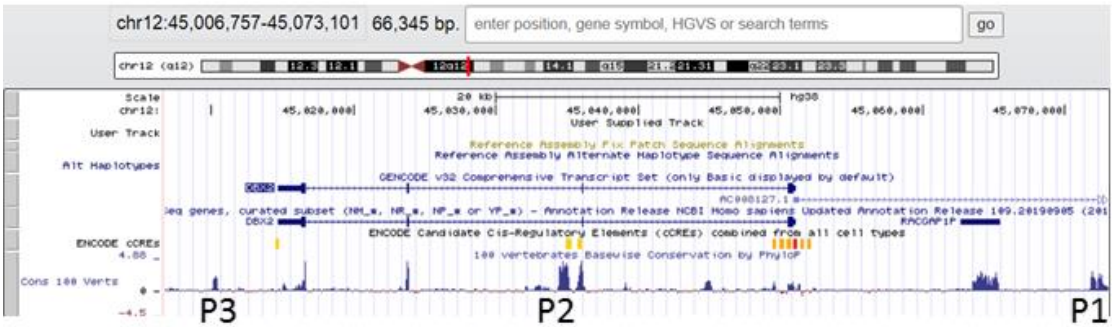
Supplementary Figure 3.5. Effects of Notch inhibition on amphioxus *IrxA* and Hh inhibition on *Gbx* expression. No observable change for *IrxA* and *Gbx*. Scale bars: 50um.

Chapter 4 Supplementary Data

S4.1 Conserved non-coding sequences in 100 vertebrates but not lamprey

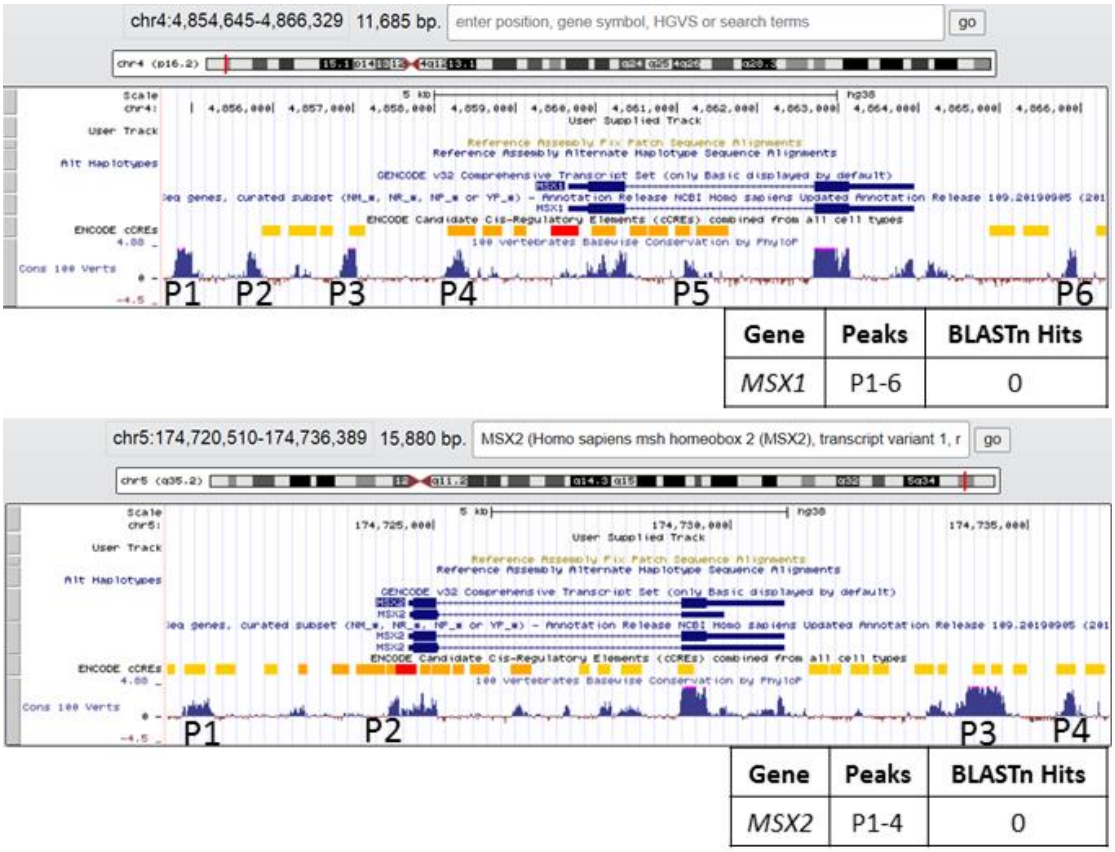


Gene	Peaks	BLASTn Hits
DBX1	P1-6	0



Gene	Peaks	BLASTn Hits
DBX2	P1-3	0

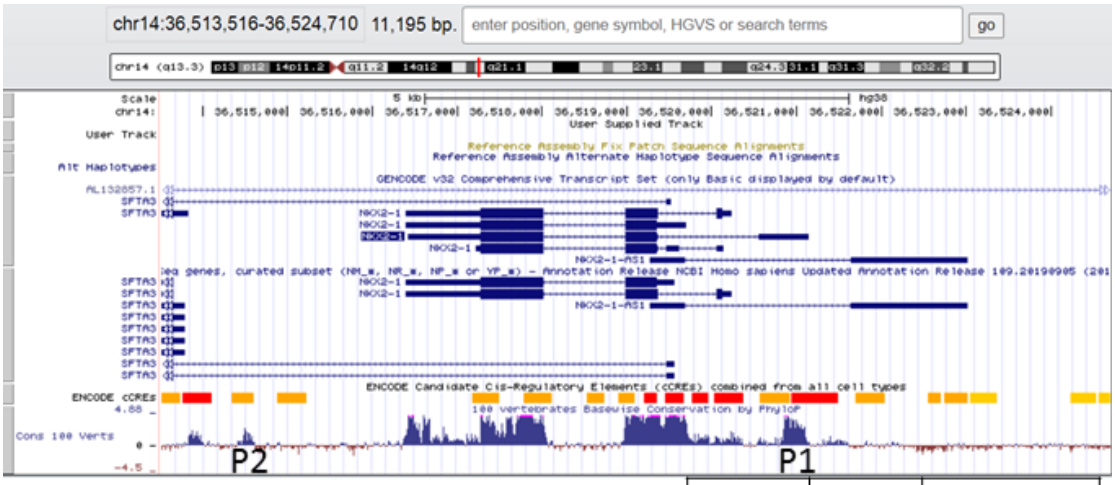
Supplementary figure 4.1. *Dbx* conserved non-coding sequence location in human. The row of peaks in human genome browser shows conservation of sequences in 100 vertebrate genomes



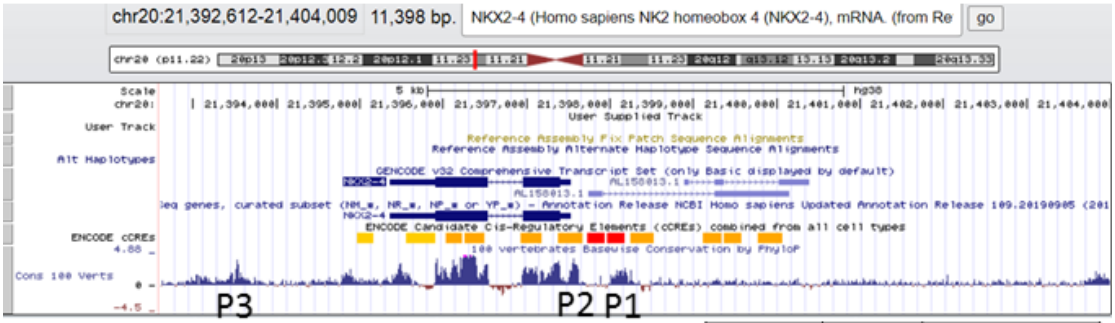
Gene	Peaks	BLASTn Hits
MSX1	P1-6	0

Gene	Peaks	BLASTn Hits
MSX2	P1-4	0

Supplementary figure 4.2. *Msx* conserved non-coding sequence location in human. The row of peaks in human genome browser shows conservation of sequences in 100 vertebrate genomes



Gene	Peaks	BLASTn Hits
NKX2.1	P1-2	0



Gene	Peaks	BLASTn Hits
NKX2.4	P1-3	0

Supplementary figure 4.3. *Nkx2.1* and *Nkx2.4* conserved non-coding sequence location in human.

The row of peaks in human genome browser shows conservation of sequences in 100 vertebrate genomes

S4.2 *Nkx2.2, Nkx6, Pax6,* and *Pax7* conserved non-coding sequences in vertebrates including lamprey

***Nkx2* Conserved non coding sequence used in alignment** Figure 4.20

>Nkx2.2_Pm_10kbupstream_scaf9

GCTGGCGAGACCGGGGGCCCACGTCCTGGGTGGTCGAAGGCACCCGGGCC

AAACCCAAATACAAACCCGATTGCTAAGCTGCACACAAAGAAGGCGATGT

AGACAAATGTCCAGCTCCTGTTGGACGTCTTTGTGGCGGTCCAGTTTTTT

TTGCATTTGTTTTAGTCATGAAATAATGGGTGATTGACGCCCACGGACAA

TGCGCTTTAACTGTTTGGAATAAATCTGAGCTTGTTCCGAGTTGTCATGG

TATTCAGCCGCGTTCAATGG

>petMar3_ct_UserTrack_3545_0 Nkx2.8 range=PIZI01000025v1:7671940-7672317 5'pad=0 3'pad=0 strand=+ repeatMasking=none

TGGTCGCCGCTCCCCAAACTCAAACACAGCTCCGGTCCTAGCTAAGGGCT

AAGCTCGAGTACAATGGGACGAATGTCATGCATTGTGCCGTTCCTGTTGG

AAACCTTTGTTGTCCGGCGGCTTTCGCGCTTGCTTCATTCGCGGCAGAAT

AGGCGATTGACGCCTCGGCACAATGCGAGCGGACTGCTTGGAATAAATAG

AAGGTCGTGCCGAGTTGTCATGACATTCAGTTGCATACGAATGAGCAGTT

TTTCATTCGCTCCCTCGTACCCCACCCTCGCCCCCTCCCCTTCCACCCTC

CTCCGACTGCCTCCGCCCCCTCCCCTCGCCTCTCCCCACCTCGCACCCCC

CCCCCCCCCTCTGAATAAGGAGCGAGTG

>hagfish_nkx2.2

CACAAGAGTGGCGCTTCTGGGCGAGCCCTGAGCTCACGTCCTGGGTGGTCGGAGCGGGAA

CAAACCCAAATACAAGGCCGATTGCTAAGCTGGACACAAAAGCAACGGTGTAGACAAATGTCCGCCTCCGGT

GGGCGCCTTTGTCGCCAGGCAGTTTTTGCATTTGTTTTAGTCAAGAAATAATGGGTGATTGACGCCACAGCAC

AATGTGCTTTAACAGTTTGGAATAAATGTTCGCTTGTTCTGAGTTGTCATGGCATTCGACCGCTTTCACTGTGC

CTTTTTCATTGATTTCTGAGT

>hagfish_nkx2.8

ATCTTGGGTGGTCGTCGCTGCCAAACACAAATACAGGCCGCTGCTGGCTAAGCCGCGCAGAATGGCGCG

AATGTACTCCATTGTGCTTCCCCTGCTGCGGACCTTTGTCACTATCCAGCC

TTCGCGCTTGATTCAGTCGTAGCAGAATAGCAGATTGACGCCCCCTGAACAATGCAAGCC

CACTGTTTGGGATAAATAGAAGGTCGCCTCGAGTTGTCATGACATTCAGCTGCATGCGAA

TGGTCGTTTTTTCCGCCTCTCAATGCAACCCGAATGGACGAGGTGCGGGCGCATCCCGCG

>galGal6_ct_UserTrack_3545_0 range=chr5:36887917-36888389 5'pad=0
3'pad=0 strand=+ repeatMasking=none
TGGCCACCTTCCCGCATATTTCCCCGGGACTGGCCAGGCTCCATATAAAC
CATTCCCAGGCCGTGATTGCCGCAGCCACCTATCAAGCGTGCGGGTGACA
GTAGTTATCTTCCTCGCTCCGAGCGCTGCCTCCCAGCCGTCCTGGGTGGT
CGGAGCGAGGGGAGCACAAATACATGCCAACTGCTAGGCCGTGAACAAAG
GGGAAATGTAGCCCTATGTCCCGGACCTGTTGGAAGCATTTGTTCCAGTC
AAGATTTTGCATTTGTTCAGATCTGGAATAATAGGTGATTGACGCCTTCC
CACAATGTGCTTTAACTCTCCGGGATAAATCGGAGCTTGTTCCTTGTTGT
CATGGTTTTCAATAGGGTTCAATGGGATTGAACCACTATCGACTCTGCTT
TACGTCTCCGGCAGCCCCGTCCATACAAATTTATTAGTTATGTCTTTTTA
AATAAGATTAAATAGAAGCGACG

>mm10a_ct_UserTrack_3545_0 range=chr2:147188191-147188547 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CCAGTATGTGACGTGGGTGACAATGGCCCAGGTTGGAGCGAGCCCCACGT
CGGCGCGTCCTGGGTGGTCGGACCCGGGCAAACACAAATACAAACCGATT
GCTAAGCTGCGGACAATGAGGGAAATGTAGACAAATGTCCCGCTCCTGTT
GGAAGCCTTTGTCCAGGCCCGGTTTTTGCATTTATTTCAGTGGCGAAATA
ATACATGATTGACGCTCTCTTTCAATGTGTCCTAACTGTTTGGAATAAAT
CTAAGGTTGTCCCTAGTTGTCATGGCATTCAACCCTTTTCAATGGACTAT
CTCTTCATTCATTTCTGAATCCGTCCACAATATTAAGGAAACCCCTTCAT
GTCGACG
>mm10_ct_UserTrack_3545_0 range=chr12:56614636-56614969 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CGCGCTGGGTGGTCGGGCGGAGGCTAGCCCACTACAGGCCAACTGCTAGG
CCACAAACAAAGGGGAAATGGAGCTGGGACGCCCGGACCTGTTGGAAGCA
TTTGTTCCAGCAAAGATTTTGCATTTGTTCAGTCCGGAATAATGCGTGAT
TGATGTCTGCTCACAATGCGCTTTAGCTCCCTGGGATAAATCGGAGCTTG
TCCGGTGTTGTCATGGTTTTCAATGGGATTCAATGGGATTGGGAGATCGA
TCGACTCTGCTGGCTGTCTTTCTGTGGTCCTGGCAATAGATTTATTAGTT
ATGTCTTTTTAAATGAGATTAAATTGAAATGACT

>hg38a_ct_UserTrack_3545_0 range=chr20:21515906-21516261 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CCAGTATGTGACGTGGGTGACAATGGCCCAGGTTAGAGCAAGCCCCTCGT
CGGCGCGTCCTGGGTGGTCGGACCCGGGCAAACACAAATACAAACCGATT

GCTAAGCTGCGGACAATGAGCGAAATGTAGACAAATGTCCCGCTCCCGTT
GGAAGCCTTTGTCCCGGCTCGGTTTTTGCATTTATTTCAGTGGCGAAATA
ATACATGATTGACGCTCTCTTTCAATGTGTCCTAACTGTTTGGAATAAAT
CTAAGGTTGTTCCTAGTTGTCATGGCATTCAACCCTTTTCAATGGACTAT
CTCTTCATTCATTTCTGAATCCGTCCACAATATTAAGGAAACCCCTTCAT
GTCGAC
>hg38_ct_UserTrack_3545_0 range=chr14:36584042-36584377 5'pad=0
3'pad=0 strand=+ repeatMasking=none
GCGCGCTGGGTGGTCGGGCGGAGGCAACCCCAATACAGGCCAACTGCTAG
GCCGCGGACAAAGGGGAAATGGAGCTCGACGCCCCGGACCTGTTGGAAGC
ATTTGTTCCAGTCAAGATTTTGCATTTGTTCAGTCCGGAATAATGCATGA
TTGACGTCTGCTCACAATGCGCTTTAACGCGCCGGGATAAATCGGAGCTT
GTCCGGTGTTGTCATGGTTTTCAATGGGGTTCAATGGGATTGGGAGCACG
ATCGACTCTGCTGGCCGTCTTTCCGCGGCCCTGGCAATAGATTTATTAGT
TATGTCTTTTTAAATAAGATTAAACTGAAATGACTT
>calMil1a_ct_UserTrack_3545_0 range=KI635882:5653275-5653593 5'pad=0
3'pad=0 strand=+ repeatMasking=none
TGACAGTGCTGCATGCTCCACAAACTTCCTCACAACATCCTGGGTGGTCG
ATGCTGGGAGAACACAAATGCATGCCAACTGCTAGGCTGCAAACAAAAGA
GGAAATGTAGACCAACGTCCTGGACCTGTTGGAAGCATTTGTTCCAGTCA
AGATTTTGCATTTGTTCAGTCATGAAATAATAGATGATTGACGTCTTCGC
ACAATGTGCTTTAACTTTTTGGGATAAATCCGAGTTTGTTCCTGGTTGTC
ATGGTTTTCAATGGCATTAAAAGGGATTTGGCCTGCTATCGACTCCGCTT
CACGTCCAATGCAGCCAAG
>calMil1_ct_UserTrack_3545_0 range=KI635915:3896573-3896872 5'pad=0
3'pad=0 strand=+ repeatMasking=none
TGGGTGACAATACTGCAGGTTTTACGAACTCCTCAGCGCAACCTGGGTGG
TCGGACCCGGGCAAACACAAATACAAACCGATTGCTAAGCTGCCAACAAT
GGGGGAAATGTAGACAAATGTCCGGCTCCTGTTGGAAGCCTTTGTCCGAG
CCCAGTTTTTGCATTTATTTCAGTCGCGAAATAATAGATGATTGACGCCC
TGGTACAATGTGTTCTAACTGTTTGGAATAAATCTGAGGTTGTTCCTAGT
TGTCATGGCATTCAACTGCTTTCAATGGACTATCTCTTCATTCATTTCTG

Nkx6 conserved non-coding sequences used in alignment Figure 4.14

>hg38_ct_UserTrack_3545_0 range=chr4:84433615-84433929 5'pad=0
3'pad=0 strand=+ repeatMasking=none
TTTTCAGATGGCACACTGACAGAAGATTAATTCAAGTCAGCTGCTGAAAA
TAATTATTGCATTATGGGAGCAGTAATCATTATTTAGAACATGATTAAAT
ATGTCTGCACCTATTGACTCTGCCTATGATTTATGTATTTGTTTAGTCAA
TAGTCTAATGTAAATGATGTAATTAATTATAGATGGTGGTGTCAGGTCAT
TGGTGAAAACAATCTGAGGAGAACAAGGGCTCTGTTTTTTTCATGACAGA
TGCAGGGGGGTGGGGGGCTGAGTTGAGGGAATTCCAGGGGAACTTTTTCA
CGTGTGAATGGCGGC

>Nkx6.2 25kb downstream petMar3_ct_UserTrack_3545_0
range=PIZI01000438v1:56773-57085 5'pad=0 3'pad=0 strand=+
repeatMasking=nonescaf_00015 scaf_00015:3929347..3929659 length=313
AGATGTAGTTAATCTTCCCAGTCCTCATCAAATACTTTTTTTTTTTTGCGGTGACCCGCTGCAAAGCCG
CTCATAATTAATTGGGCAATCTACATTAAACTATTGACTAAACAGCCACATAAATCACAGCCAAAGTCA
ATAGCGATGGACATATTTAATCATGTTCCAAATAAAGATTACCGCGCCTTATAATGCCATAATTATTTT
CGGCCGTTGGGTGGAATTAATCTTGTGGCGCTCGCGAGCGGCGCGTGAAAAATACAGCCGAGAGTCGCT
GGCGCGGCCTCCATTACCCAGAGGTCAAGGTGGCTAT

>hagfish nkx6.1
ATGTTAGGATAGGATGTGGGAAAAGGGGGCAGTGTAAGAATTTGACAAACGTGTTGAGTT
TGTCTGATGAAGATACCGTAAATCTCAGCACCCTCATCAGATACGTCTCTCCCTACGACC
CCGCTCATAATTATTTGCACGATCCGCATTAAACTATTGACTAAACAGCAGCATTAATCA
TTGCCGGAGTCCAATAGCCGAGGACATATTTAATCATGTTTGGAATCGAGATGCGTGCGT
TCTTATATTGCCATAATTATTTTCTGCTCTTGGACGGGAATTAATCTTCTGGCACCGACT
GTGGCATGAAAAATCTCCGAGCAGCCCACATCTACAGTTTTCATTACTTAGAGGTCAAGC
TCGGCATCTTTCTTCTGCCTTTCCTTGTCTTTTTGTCCCCTCCTCCCTGTCTGAATTCTC
CTGCATACACCAATGCTACTCTCATTTTCTATCTTTCTTTATCTCCCACATTATTTCATC
ATTTCTCAATGTGTATCCTTTTCACCAAGCTATGTATCAA

>mm39_ct_UserTrack_3545_0 range=chr5:101738478-101738788 5'pad=0
3'pad=0 strand=+ repeatMasking=none
GACAGAAGATTAATTCAAGTCAGCTGCTGAAAATAATTATTGCATTATGG
GAGCAGTAATCATTATTTAGAACATGATTAAATATGTCTGCACCTATTGA
CTCTGCCTATGATTTATGTATTTGTTTAGTCAATAGTCTAATGTAAATGA

TGTAATTAATTATAGATGGTGGTGTCAGGTCATTGGTGAGAACAATCCGG
GGAGAACAAGGCCTCTGTTTTTTTCATGACAGATGCAGAGAAGGGGGTGG
GCTGAGTTGAGGGAATTCCAGGGGAACTTTTTCACGTGTGAATGGCGGCC
GGGATCTCAAT
>galGal6_ct_UserTrack_3545_0 range=chr4:46780296-46780602 5'pad=0
3'pad=0 strand=+ repeatMasking=none
GATGGCACTTTGACAGAAGATTAATTCAAGTCAGCTGCTGAAAATAATTA
TTGCATTATGGGAGCAGTAATCATTATTTAGAACATGATTAAATATGTCT
GTGTCTATTGACTCTGCCTATGATTTATGTATTTGTTTAGTCAATAGTCT
AATGTAAATGATGTAATTAATTATAGATGGTGGTGTCAGGTCATTGGTAA
AAACAATCTGATGAGAACAAGGCCCCTGTTTTTTCCTGACAGATGCGGAG
GGAGGGCGGGAGGGACGGGCAGAGTTGAGGGAATTCCAGGGGAACTTTTT
CATGTGT
>xenLae2_ct_UserTrack_3545_0 range=chr1L:81956106-81956427 5'pad=0
3'pad=0 strand=+ repeatMasking=none
ACGAGTCCTTGTTCTCTTCAGATTGTTTTCACCAATGACCTGACAACCAC
CATCTATAATTAATTACATCATTTACATTAGACTATTGACTAAACAAATA
CATAAATCATAGACAGAGTCAATAGATGCAGACATATTTAATCATGTTCT
AAATAATGATCACTTCTCCCATAATACAATAATTATTTTCAGCAGGTGAC
TTGAATTAATCTTCTGTCAAAGTGCCATTAAAAAAGAGGAATATTATTTG
ACAGCAACTCTTTTTCAGCCATGATCTTTCTTTTTTTCTTTTGGATTTAC
TTTACCTTTTTGTGTTTTTATG

>danRer11_ct_UserTrack_3545_0 range=chr21:19110101-19110700 5'pad=0
3'pad=0 strand=+ repeatMasking=none
AATAGACATGGGGGTCGAGGAAATCCCAAACCTCCTCTCACATCTGTCAC
ACCCCTCTCACCGACGACCCGATAATTAATTACATCATTTACATTAGACT
ATTGACTAAATAAATACATAAATCACAGGGCCTCGAGTCAATAGACGCAG
ACGTATTTAATCATGTTCTAAATAATGATCACTGTTCCCATAATGCAATA
ATTATTTTCAGCAGCTGACTGGAATTAATCTTGTGCCATCAGAGACACCA
CACACACACACACACACACACACACACACACACACACACACACACACACA
CACACACACACACACACACACACAGTGAGAGAGAGAGAGAGAGAGAGA
GAGAGAGAAGATAAATACAAACTCTTTTGTGAATGACAGCAATAGAGAAA
TTGAAAATGGCGGCTTGTCTCTTTATGCGGAGTGACAGGTGAAAAACGCG

>calMil1_ct_UserTrack_3545_0 range=KI636047:539501-540200 5'pad=0
3'pad=0 strand=+ repeatMasking=none

GGTGCATGCAGATTGAGAGCTGAGCTCTATTCACACATGAGAAATTCTCC
AGGAATTCCCTAAACTCGATCTCACATCTGTCATGAAACAGAGTCCTTGT
TCTCATCAAATTGTTTTCACCAATGACCTGACACCAACATCTATAATTAA
TTACATAATTTACATTAGACTATTGACTAAACAAATACATAAATCATGGG
TAGAGTCAATAGATGCAGACATATTTAATCATGTTCTAAATAATGATTAC
TGCTCCCATAATACAATAATTATTTTCAGCAGCTGACTTGAATTAATCTT
CTGTCAAGGTGCCACAGGAAGCAGAGACGGATGGAAATGCCTCTGAGTTC
AACGCTACCAAATTTCCTAATGCTGAATTGATAATGAACACACTGCATGG
ATCACTGCAAAGGCAATAAATGATTGGAGAGCCATGTAGACGCATGAAAG
CTAAGAAACAAGGACGCTAGATTCAACTGGAAATGGTAACGGAGGGTCTT
AATTTGTGCTATTTCTTGATGGGATCCCTGCATAAGGAATTCTCGTAAAT

Pax7 conserve non-coding sequences used in alignment Figure 4.16

>hg38_ct_UserTrack_3545_0 range=chr1: 18,710,651-18,710,878 5'pad=0
3'pad=0 strand=+ repeatMasking=none
GCCCAGGCAACAAGATCTCATTAACAAGCAGAAAATAAGATTGAAATGGT
GTGTAAAAGAGCTGAAAAATGAATTTGAATGCTGCGTTTGTCCACAATTA
CCCCCTCCTTCACACATATTTTATTGCAATTTTTTTATTATTCTTAATCT
CTCCATTCCTCAAAGCAGATTGGGGACATCCTGGCATTACCAGGGCGTCA
GGAGGAGGTGTGGAGGATTGCCTGCCCCGCTGAAAGGAGACCCTTCTCTCGC ACATTTTCCG
AGGTTGGGTT

>mm39_ct_UserTrack_3545_0 range=chr4: 139,499,932-139,500,191 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CAAGCAGAAA ATAAGATTGA AATGGTGTGT AAAAGAGCTG AAAAATGAATTTGAATGCTG
CGTTTGTCCA CAATTACCCC CTCCTTCACA CATATTTTAT
TGCAATTTTT TTATTATTCT TAATCCCTCC ATTCCTCAAA GCAGATTGGG
GACATCCTGG CATTACCGGG GCACTGGGAG GAGGTGTGAA GGATCTTCTGCGCCCCTGAA
ACGAGACCCT TTCTTGCCAA CATCTTCCAA GGTTGGGATT

>galGal6_ct_UserTrack_3545_0 range=chr21: 4,509,489-4,509,654 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CAAGCAGAAA ATAAGATTGA AATGGTGTGT AAAAGAGCTG AAAAATGAAT
TTGAATGCTG CATTTGTCCA CAATTACCCC CTCGTTCACA CATATTTTAT
TGCAATTTTT TTATTATTCT TAATCTTTCC ATCCCTCAAA GCAGATTGGG
AACATCCTGG CATTACAGTT GCTGAGGGAG GAAGCATTAA GGATTCCCCG
TTTCCCTGAA AGGAGACCCT TCTCTCGCTC GAATTTTAAA GGGTTGGTTT

>xenLae2_ct_UserTrack_3545_0 range=chr7L: 120,842,051-120,842,219 5'pad=0 3'pad=0 strand=+ repeatMasking=none

CAAGCAGAAA ATAAGATTGA AATGGTGTGT AAAAGAGCTG AGAAATGAAT
TTCAATGCTG CATTTGTCCA CAATTACCCC CTCGTTCACA CATATTTTAT
TGCCATTTTT TTTTTATTCT TCTTAATCTC TCCATCCCTC AAAGCAGATT
GGCGACATCC TGGCATTACA GAGGCTGAGG GACGAGAATT AAGGATTTGT TCCTCTGTGA
ATGTCTCAGC TCTCTGTCTG TCTCTTGCTG

>xenLae2chr7s 86,504,228-86,504,393

CAAGCAGAAA ATAAGATTGA AATGGTGAGT AAAAGAGCTG AGAAATGAAT
TTCAATGCTG CATTTGTCCA CAATTACCCC CTCGTTCACA CATATTTTAT
TGCCATTTTT TTATTCTTCT TAATCTCTCC ATCCCTCAAA GCAGATTGGC
GCCATCCTGG CATTACAGAG GCTGAGGGAC GAGAATTAAG GATTTTTTCC
TCTGTGAATG TCGGAGCTCT CGGCCTCTTG CTGCTGCGCG TGGATAACAG

>danRer11a_ct_UserTrack_3545_0 range=chr11: 41,310,571-41,310,670 5'pad=0 3'pad=0 strand=+ repeatMasking=none

CAAGCATAAA ATAAGATTGA AATGGTGTGT AAAGAGTTGA AAATGAATTT
CAATGCTGCA TTTGTCCACA ATTACCCTTT CGTTCACACA TGTTTTATTG
GGGGTTTCCT TTTCTCTCCC CTGCTCTCTC AAAGCAGATT GAGAAACATT
CTTCTGTATT ACCCTCTTTT TTTTCCCAGA GAAGAAAGAA TTAAGGATTT

>danRer11b_ct_UserTrack_3545_0 range=chr23:20905219-20905488 5'pad=0 3'pad=0 strand=+ repeatMasking=none

CAGGCAGAAA ATAAGATTGA AATGGTGTGT AAGGAGTTGA AAATGAAATT
CAATACCGCG TTTGTCCACA ATTACCCTCT TGTTCACACG TTTCATGGCC
ACTTTCCCCC CGCACCAAAT ATTTCTTAAA GCAGATTGGA AACTGGTATT
ACTCTCCATT TTTTTTTCAA AGAGAGGAAA GAATTAGGGA TTTGATGACA
AAAAAGAGAT TTTCTCTTGT ACTGTGGAGT

>calMil1_ct_UserTrack_3545_0 range=KI635912: 4,184,998-4,185,222 5'pad=0 3'pad=0 strand=+ repeatMasking=none

CAAGCAGAAA ATAAGATTGA CAGGGTGTGT GAGGAGCTGT ACATGAATTT
CAATCGGATA TTTGGGCACA ATTACCTCGT TTACACAGAT TTCACTGTAA
ATTGTACATT GCTTGAAGCA CATCTGAAAC AATCGCTATT AGTTTCAGGA

>Pax7 cis-reg ele petMar3_ct_UserTrack_3545_0 range=PIZI01000121v1:624107-624320 5'pad=0 3'pad=0 strand=+ repeatMasking=none Blast Hitscaf_00121 scaf_00121:624108..624320 length=213

GCAATAAAACACCACGGGCGTGCAGAAGCGGCGAAAAATCGCAGACAAACGCGGCGCTGG
AAATCATCTCCAGCGCTGTTTTATGCGCCATTGCAATCTTATTTTCTGCCTGTTAGCGGAGATCG
TGTCGGGCCTTCCGAGTGCCAGGGCTGCGGCCTAATCGCTGTGTAATTGCTTGTAATTTACTTC

TTACTCCGGATCGGCATCGGTTCT

Can’t find sequence in hagfish

Pax6 conserved non-coding sequences used in alignment Figure 4.18

```
>hg38_ct_UserTrack_3545_0 range=chr11:31803943-31804501 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CAATAAATAAGCT
CCAAATATAGAAGAGGGGGGAAAATATGATGAGCCTCTCTGACATTTGTC
TTTAAAATAAAACTAGCTGCACGTCAAGTTTGAGCTTAATTTCTGGAAAT
AGGCGGAAGTCGCTCCGGATCATGCATGGAAGGCTAATTGAAAAGATCAG
TCGGGGCGCTTGTGGCAGCCTTGTCACTTTGTGACAGTGCTCCGCTCCGG
GAAATTGCATCGTCACGACAAA

>mm39_ct_UserTrack_3545_0 range=chr2: 105512633-105512809 5'pad=0
3'pad=0 strand=+ repeatMasking=none
GGGGGAATATGATGAGCCTC TCTGACATTT GTCTTTAAAA TAAAACTAGC TGCACGTCAA
GTTTGAGCTT AATTTCTGGA AATAGGCGGA AGTCGCTCCG GATCATGCAT
GGAAGGCTAA TTGAAAAGAT CAGTCGGGGC GCTTGTGGCA GCCTTGTCAC
TTTGTGACAG TGCTCCGCTC CGGGAAATTG CATCGTCACG ACAAACGGGA

>galGal6_ct_UserTrack_3545_0 range=chr5: 5,453,681-5,453,914 5'pad=0
3'pad=0 strand=+ repeatMasking=none
CAATAAATAA GCTCCAAATA TAGAAGAGGG GGAAAATATG ATGAGCCTCT
GTGACATTTG TCTTTAAAAT AAAACTAGCT GCACGTCAAG TTTGAGCTTA
ATTTCTGGAA ATAGGCGGAA GTCGCCCCGG ATCATGCATG GAAGGCTAAT
TGAAAAGATC AGTTGGAGCA CTTGTGGCAG GCGTGTCACT TTATGACAGT
CCTCTGCTTT TGAAAATTGC ATCGTCACGA CAAATAGTAG
>xenLae2_ct_UserTrack_3545_0 range=chr4L: 13,327,175-13,327,342
5'pad=0 3'pad=0 strand=+ repeatMasking=none
AGGAAA AATATGATGA GTCTTCCCTG ACATTTGTCT TTTAAATAAA ACTAACCGCA
CGTCAAATTT GAGCTTAATT TCTGCAAATA GGAGGAAGTC GCCCGGATCA
TGCTTGGGAG GCTAATTGAA AAGATCAGAT TGAAAGCACT TGTGGCAGAC
TTGTCATTCT GTGACAGTCC TGGGCTTTAG AAAATT
>xenLae2S_ct_UserTrack_3545_0 range=chr4S: 9,858,306-9,858,4965'pad=0
3'pad=0 strand=+ repeatMasking=none
TGA CGATCCTTCCCTGACATTTG TCTTTTAAAT AAAACCAACC AACGCACGTC AAAAAATTTT
```

GAGCTTAATT TCTGCAAATA GGAGGAAGTC GCCCGGATCA TGGTCAGGAG

GCTAATTGAA AAGATCAGAT TGAAGCACTT GTGAAAGGCT TGTCGTTTTG

TGACAGTCCT GTGCTTTTGA AAATTGCATC GTCAGGACAA ATGGATGTGG

>danRer11 chr25: 15040406-15040612

GGGGGAAAAT ATGATGAGCC TGCCTGACAT TTGTCTTTAA AATAAAGCTA

GCTGCACGTC AAGTTTGAGC TTAATTTCTG GAAATAGGCG GAAGTCGCCC

TGGATCATGC ATGGAAGGCT AATTGAAAAG ATCAGTTGGA GCACTTGTGG

CAGGCATGTC ACTTTATGAC ATCGCTCTGC TTTTGAAAAT CGCATCGTCA

CGACAAATAG

>danRer11chr7_ct_UserTrack_3545_0 range=chr7: 15880966-15881174

5'pad=0 3'pad=0 strand=+ repeatMasking=none

GGGGGAAAAT ATGATGAGAC TCTCTGACAT TTGTCTTTAG AATAAAGCTC

GCTGCACGTC AAGTTTGAGC TTAATTTCTG GAAATAGGCG GAGGTCGCCT

CGGATCATGC ATGGAAGGCT CATTGAAAAG ATCAGCTGAA GTACTTGTGG

CAGGACTGTC ACTTTATATG ACACTGCTCT GCTTTTGAAA AACGCATCGT

CACGACAAAT

>calMil1_ct_UserTrack_3545_0 range=KI635862: 11,141,417-11,141,614

5'pad=0 3'pad=0 strand=+ repeatMasking=none

ATGATGAGCT TTCCTGACAT TTGTCTTGAA AATAAAGCTG GCTGCATGTC

AAATTTTGAG CTTAATTTCT GGAAATAAGC GGAAGTCGCA GTGGATCGTG

CACGGAAGGG TAATTGAAAA GATCAGTGAG AGCACTTGTG GCAGGCATGT

CACTTTATGA CAGTGTTCTA CTTTGAAAAA TCGCATCGTC ACGACAAATA

>hagfish

CCGCCAAGCAGCAAATATGCACGAATTTTAAGAAGCGGGTGGGGGGACGCTGAGCTTATT

TGACATTCGTCTCCCGAATAAAGTCACAGAGACGTCAAGTTTGAGTCTAATCTTCGCAAA

TAAGGTTCTAAAGTCGTTCGGCATTACACACGCTACACTTGCCGGAGACGGCCTTGCCAG

CGGTGTCACATCGAGTCACGAATTTCATGAAGTTCTGCCCCGTACGGCACCAGGGTGCGG

GAGGCACACTCTCAGCTTGCTCGACGCCTTGCCTGCCTTTTTCCCCAGGGCTGAGAGCGC

>Pax6 cis-reg ele petMar3_ct_UserTrack_3545_0 range=PIZI01000014v1:14132200-14132399 5'pad=0 3'pad=0 strand=+ repeatMasking=none

AGTGGGGGGGGGGGGGGGAAGCGGAGCCGTGGAACGATGAGGCCGTGTGA

CATTTGTCTTCCGAATAAAGCCATCCAGACGTCAAGTTTGAGCCTAATTT

TCCCAAATAAGCCGAGGTCGTCACGGCGTGTTAGCCACGCTCCAGCGCGC

AGAGGCGGCCTTGCCAGCCTTGTCACATCGCGTCACAAGTTTGCTTCAAG

S4.3 pTK-eGFP plasmid sequence

pTK-eGFP (Chen and Streit, 2015) vector sequence

ggtaccatggatatcatggccaactgactagtggctcgagatctgcgatctaagtaagcttgctgcttcatccccgtgacccgttgccgcgtttg
ctggcggtgtccccggaagaaatatatttgcatgtctttagttctatgatgacacaaaccccgcccagcgtcttgtcattggcgaattcgaacacg
cagatgcagtcggggcggcgcggtcccaggtccacttcgcatattaaggtgacgcgtgtggcctcgaacaccgagcgaccctgcagcgac
ccgcttaacagcgtcaacagcgtgccgcagatcccaagcttgcggccgcttaactgcagaagttggtcgtgaggcactgggcaggtaagtal
caaggttacaagacaggtttaaggagaccaatagaaactgggcttgtcgagacagagaagactcttgcgtttctgataggcacctattggtctt
actgacatccactttgcctttctctccacaggtgtccactcccaggttcaatacagctcttaagcggccgcaagctagcttgccgccaccatggt
gagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccg
gcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcg
tgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggc
tacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaac
cgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctat
atcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgacca
ctaccagcagaacaccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccctgagcaaagaccc
caacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtaagcggcc
gcgactctagagtcggggcggccggccgcttcgagcagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaa
aaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttaacaacaacaattgcattcattttatgt
ttcaggttcaggggaggtgtgggaggttttttaaagcaagtaaaacctctacaaatgtggtaaaatcgataaggatccgtcgaccgatgccctt
gagagccttcaacccagtcagctccttccggtgggcgcggggcatgactatcgtcgccgcacttatgactgtcttctttatcatgcaactcgtag
gacaggtgccggcagcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag
gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaa
ggccgcgttgctggcgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca
ggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttct
cccttcgggaagGgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaac
cccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagcca
ctggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtat
ttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagpggtggtttt
tttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa
ctcacgttaagggattttggtcatgagattatcaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatat
gagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgt
cgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatc
agcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaag
ctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattc
agctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaa
gtaagttggeegcagtgttateaeteatggttatggeageaetgeataattetettaetgtcatgccatccgtaagatgcttttctgtgactggtgag
tactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcaga
actttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgt
gcacccaactgatcttcagcatcttttacttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagg
gcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtat
ttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacctgacgcgccctgtagcggcgcattaagcgcggcgggt
gtggtggttacgcgcagcgtgaccgctacacttgccagcgccciagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccgg
ctttccccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatgg
ttcacgtagtgggccatcgccctgatagacggtttttcgcccttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaaca
acactcaaccctatctcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacg
cgaattttaacaaaatattaacgcttacaatttgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctat
tacgccagcccaagctaccatgataagtaagtaatattaaggtacgggaggtacttggagcggccgcaataaaatatctttattttcattacatct
gtgtgttggttttttgtgtgaatcgatagtactaacatacgctctccatcaaaacaaaacgaaacaaaacaaactagcaaaataggctgtcccca
gtgcaagtgcaggtgccagaacatttctctatcgata