



Lamprey Neural Helix-Loop-Helix (HLH) Genes and the Evolution of the Vertebrate Nervous System

Thesis presented for the degree of Doctor of Philosophy in Zoology

Ricardo Lara-Ramirez

St Anne's College, University of Oxford

Hilary Term 2013

Lamprey Neural Helix-Loop-Helix Genes and the Evolution of the Vertebrate Nervous System

Ricardo Lara-Ramirez

St Anne's College, University of Oxford

Thesis presented for the degree of Doctor of Philosophy in Zoology

Hilary Term 2013

Abstract

Transcription factors of the helix-loop-helix (HLH) gene family are widespread in the animal kingdom. Among them, members of HLH subfamilies such as *ASCL*, *Neurogenin*, *NeuroD*, *COE*, *Atonal*, *Oligo*, *NSCL*, *Hairy/E(spl)* and *Hey* (here referred to as neural HLH genes) have been shown to be fundamental for the development of the nervous system. They are expressed at different time periods of neuronal differentiation, from the specification of ectoderm towards a neural lineage, to the ultimate differentiation of neurons. Few HLH genes have been identified in the lamprey; however, considering the wide diversity of HLH gene subfamilies in metazoans, including vertebrates, it is very likely that lampreys possess a large repertoire of HLH genes in their genome. In the present study, the identification of several HLH genes in the lamprey genome, as well as the isolation and expression of different lamprey neural HLH genes is reported. As expected, a wide repertoire of HLH genes was identified in the sea lamprey (*Petromyzon marinus*) genome. On the other hand, the identification and expression analysis of different neural HLH genes of the *ASCL*, *Neurogenin*, *COE* and *Hairy/E(spl)* in the brook lamprey *Lampetra planeri* showed an overall conservation with other vertebrates, both at the sequence and expression pattern levels. In addition, novel features of the lamprey nervous system are revealed, such as the identification of possible new sensory cranial placodes in pharyngeal arches. Furthermore, these genes can serve as molecular markers for different cranial placodes and dorsal root ganglia (DRG), and their expression also highlights the presence of a ventricular zone in the brain and spinal cord, along with a complementary marginal zone. Finally, with the use of a Notch pathway inhibitor in developing *L. planeri* embryos, the regulation of expression of the isolated genes by the Notch signaling pathway was shown to be generally conserved between lampreys and gnathostomes in the spinal cord. This functional study also revealed that the lamprey spinal cord likely presents an independent developmental programme from the brain. All together, the present study shows that the analysis of neural HLH genes represents an excellent tool to understand the lamprey nervous system.

Acknowledgements

I am largely grateful to my supervisor Dr. Sebastian Shimeld for the opportunity to do my PhD in his lab, for believing in me, and most importantly for his patience and understanding throughout the PhD and the long thesis-viva process. I hope both my research and thesis can honor his great support.

I am also grateful to my parents and family, my brother, my sister, my niece and nephew, who were always with me throughout my PhD and helped me in difficult moments even from the distance.

Also many thanks to my very first labmates Dr. Wei-Chung Chen for teaching me many techniques in the lab, Dr. Nicola Pickup for her help on bioinformatics, to Kumars Riyahi for his help and friendship, Rachel Grant for technical help as well, and Dr. Karl Wotton for his *in situ* protocol for lampreys and technical help on sectioning, even though we didn't get the chance to work together.

Special thanks to Professor Peter Holland who provided excellent guidance on my experiments and direction of my project.

Special thanks as well to Rachel Grant, Israel Salguero, Amy Smith and Robert Belshaw for taking me to New Forest to collect lamprey embryos.

I am also most grateful to Nathan Kenny and Cedric Patthey for their technical help on sequencing and sectioning at the final stage of my experiments that allowed me to complete my analyses.

Also thanks to Dr. Jo Bebgie and Caroline Osborne for access to facilities and equipment and their help for sectioning lamprey embryos.

Many thanks as well to my other lab members which were many and I hope I don't miss anyone: Rebecca Furlong, John Mulley, Thomas Butts, Guang Li, Ying-Fu, Jordi Paps, Jerome Hui, Atsuko Sato, Carmel McDougall, Ana Sara Monteiro, Eva Jiménez Guri, Peter Osborne, Helen Thompson, Kyle Martin and Michael Plugge. They always provided all sorts of help and advice on a daily basis.

Many thanks to the Oxford University Mexican Society and their members who became my friends, who always supported me, and with whom I had amazing Mexican events that made my life in Oxford quite enjoyable.

Last but not least, my cats! Gracias a mis dos Bichas que me acompañaron siempre para entrar a Oxford (mi primera Bicha) y para escribir mi tesis (mi segunda Bicha). También Cairel, sus hermanitos y mi tortuga estuvieron conmigo y mis plantas. Todos me ayudaron a tranquilizarme y poder avanzar.

Contents

1. INTRODUCTION.....	10
1.1 The Lampreys.....	10
1.2 The Agnathans.....	12
1.3 Morphology of the Vertebrate Nervous System.....	20
1.3.1. Vertebrate Central Nervous System.....	20
1.3.2. Vertebrate Peripheral Nervous System.....	26
1.3.3. Vertebrate Autonomic Nervous System.....	28
1.3.4. Vertebrate Cranial Placodes.....	30
1.4. The Lamprey Embryonic Nervous System and its Relation to Gnathostomes...	32
1.4.1. Early Development.....	32
1.4.2. Composition of the Lamprey Central Nervous System.....	34
1.4.3. Composition of the Lamprey Peripheral Nervous System.....	37
1.4.4. The Lamprey Autonomic Nervous System.....	40
1.4.5. Lamprey Cranial Placodes.....	40
1.5. The Helix-Loop-Helix Gene Family.....	44
1.5.1. The bHLH Domain.....	44
1.5.2. The Neural HLH Genes.....	45
1.5.3. HLH Gene Networks in the Nervous System.....	47
1.5.4. Notch Action Over Neural HLH Genes.....	48
1.6. Vertebrate Genome Duplications.....	50
1.7. Aims of the Project.....	52
2. METHODS.....	53
2.1. Computer-based Research.....	53
2.1.1. Genome Databases.....	53
2.1.2. Sequence Analysis and Manipulation.....	53
2.1.3. Phylogenetic Analyses.....	54

2.2. Experimental Research.....	54
2.2.1. Animal Collection.....	54
2.2.2. Fixing Embryos.....	55
2.2.3. RNA Extraction.....	56
2.2.4. cDNA Synthesis.....	56
2.2.5. Polymerase Chain Reaction.....	57
2.2.6. PCR Programme Strategies Used to Clone Each Gene.....	58
2.2.7. Cloning PCR Products.....	62
2.2.8. Sequencing.....	64
2.2.9. <i>In situ</i> Hybridisation.....	64
2.2.10. Riboprobe Synthesis.....	64
2.2.11. Permeabilisation and Hybridisation.....	66
2.2.12. Washing.....	67
2.2.13. Coloring.....	68
2.2.14. Visualisation of Signal and Photography.....	68
2.2.15. Sectioning.....	69
2.2.16. Drug Treatments.....	69
 3. THE LAMPREY REPERTOIRE OF HLH GENES.....	70
3.1. Introduction.....	70
3.2. Aims.....	72
3.3. Hypothesis.....	72
3.4. Objectives.....	73
3.5. Results.....	73
3.6. Discussion.....	91
3.7. Summary of Lamprey HLH Genes.....	93
 4. <i>LAMPETRA PLANERI HAIRY/E(SPL)</i> GENES.....	94
4.1. Introduction.....	94
4.2. Aims.....	104
4.3. Hypothesis.....	104
4.4. Objectives.....	105

4.5. Results.....	106
4.5.1. Isolated Sequences.....	106
4.5.2. Phylogenetic Analysis of Lamprey <i>Hairy/E(spl)</i> Genes.....	113
4.5.3. <i>LpHes-A</i> Expression in Developing <i>L. planeri</i> Embryos.....	118
4.5.4. <i>LpHes-B</i> Expression in Developing <i>L. planeri</i> Embryos.....	120
4.6. Discussion.....	129
4.6.1. Lamprey <i>Hairy/E(spl)</i> (<i>Hes</i>) Gene Sequence.....	129
4.6.2. Expression of Lamprey <i>Hairy/E(spl)</i> (<i>Hes</i>) Genes.....	131
4.6.3. <i>Dec</i> Genes as a New HLH Gene Subfamily.....	134
4.7. Summary of Lamprey <i>Hairy/E(spl)</i> Genes.....	137
 5. LAMPETRA PLANERI NEUROGENIN GENES.....	139
5.1. Introduction.....	139
5.2. Aims.....	146
5.3. Hypothesis.....	146
5.4. Objectives.....	147
5.5. Results.....	148
5.5.1. Cloning of <i>LpNgn</i>	148
5.5.2. Phylogenetic Analysis of <i>LpNgn</i>	152
5.5.3. <i>LpNgn</i> Expression in Developing <i>L. planeri</i> Embryos.....	155
5.6. Discussion.....	165
5.6.1. <i>LpNgn</i> Expression in the Central Nervous System.....	165
5.6.2. <i>LpNgn</i> Expression in the Peripheral Nervous System.....	168
5.7. Summary of Lamprey <i>Neurogenin</i> Genes.....	176
 6. LAMPETRA PLANERI COE GENES.....	177
6.1. Introduction.....	177
6.2. Aims.....	185
6.3. Hypothesis.....	185

6.4. Objectives.....	186
6.5. Results.....	187
6.5.1. Cloning of <i>LpCoe-A</i> and <i>LpCoe-B</i>	187
6.5.2. Phylogenetic Analysis of <i>LpCoe-A</i> and <i>LpCoe-B</i>	202
6.5.3. <i>LpCoe-A</i> Expression in Developing <i>L. planeri</i> Embryos	207
6.5.4. <i>LpCoe-B</i> Expression in Developing <i>L. planeri</i> Embryos.....	215
6.5.5. <i>COE</i> Genes in Vertebrate Lymphocytes.....	224
6.6. Discussion.....	231
6.6.1. <i>LpCoe-A</i> and <i>LpCoe-B</i> Sequence and Gene Structure.....	231
6.6.2. <i>LpCoe-A</i> and <i>LpCoe-B</i> Expression in the Central Nervous System.....	233
6.6.3. <i>LpCoe-A</i> and <i>LpCoe-B</i> Expression in the Peripheral Nervous System.....	237
6.6.4. <i>LpCoe-A</i> and <i>LpCoe-B</i> Expression in non-Neural Tissues.....	238
6.6.5. <i>COE/EBF</i> Expression in Lymphocytes.....	238
6.7. Summary of Lamprey <i>COE</i> Genes.....	241
 7. <i>LAMPETRA PLANERI</i> ASC GENES.....	242
7.1. Introduction.....	242
7.2. Aims.....	250
7.3. Hypothesis.....	250
7.4. Objectives.....	251
7.5. Results.....	252
7.5.1. Cloning of <i>LpAsh</i>	252
7.5.2. Phylogenetic Analysis of <i>LpAsh</i>	254
7.5.3. <i>LpAsh</i> Expression in Developing <i>L. planeri</i> Embryos	255
7.6. Discussion.....	259
7.6.1. <i>LpAsh</i> Expression in the Central Nervous System.....	259
7.6.2. <i>LpAsh</i> Expression in the Peripheral Nervous System.....	260
7.6.3. <i>LpAsh</i> Expression in non-Neural Tissues	261
7.7. Summary of Lamprey <i>ASC</i> Genes.....	262

8. REGULATION OF NEURAL HLH GENES BY THE NOTCH

SIGNALLING PATHWAY IN THE LAMPREY SPINAL CORD.....264

8.1. Introduction.....	264
8.1.1. The Notch Signalling Pathway	264
8.1.2. Notch Expression in the Vertebrate Spinal Cord.....	267
8.1.3. The γ -Secretase Complex.....	268
8.1.4. The Notch Pathway Inhibitor DAPT.....	270
8.2. Aims.....	272
8.3. Hypothesis.....	272
8.4. Objectives.....	273
8.5. Results.....	274
8.6. Discussion.....	287
8.6.1. Morphology of DAPT-treated Embryos	287
8.6.2. Expression Pattern Changes After DAPT Treatment.....	288
8.6.3. Interneurons in the Lamprey Spinal Cord.....	291
8.6.4. The Lamprey Hindbrain-Spinal Cord Transition.....	293
8.7. Summary of Lamprey Notch Signalling.....	295

9. DISCUSSION.....296

9.1. The Lamprey HLH Gene Family.....	296
9.2. Lamprey Neural HLH Genes in the CNS.....	298
9.2.1. Brain Territories.....	298
9.2.2. Spinal Cord.....	299
9.3. Lamprey Neural HLH Genes in the PNS.....	301
9.4. Regulation of Lamprey Neural HLH Genes by the Notch Signalling Pathway...	303
9.5. New Aspects of the Lamprey Nervous System Revealed by the Present Study...	304
9.6. The Ancestral Vertebrate Spinal Cord: A Hypothesis.....	307
9.7. Future Directions.....	310

10. REFERENCES.....	316
APPENDIX A. Ricardo's <i>in situ</i> Protocol for Lampreys.....	349

1. INTRODUCTION

1.1. The Lampreys

Lampreys are enigmatic animals with a very distinct life-cycle. Morphologically, adult lampreys have an eel-like body, no appendages, no lateral fins, but depending on the species they possess different sets of dorsal, anal and tail fins. Perhaps their main characteristic is their possession of an oral hood, a circular, cartilaginous equivalent to a mouth that presents a series of dentitions inside arranged in a geometrical or dispersive manner. Lampreys do not have jaws, therefore they cannot bite, but some species use their oral hood to stick on prey, mostly fish, and suck on their flesh with the aid of a rasping tongue, regardless if they kill the fish or not. For this reason, it is not clear whether lampreys are predators or parasites, given that their objective is not necessarily killing their “prey”. However, in the literature, when lamprey life-cycles are discussed, they are referred to as parasites instead of predators.

According to Hardisty (Hardisty, 2006), thirty-eight lamprey species have been described so far around the world, divided into ten genera. Most lampreys are in the northern hemisphere, and only two genera, *Mordacia* and *Geotria*, are found in the southern hemisphere. There are different lamprey species that are used as models for research, whether morphological, physiological or molecular. The commonly used lamprey species are *Petromyzon marinus*, *Lampetra fluviatilis*, *Lampetra planeri*, *Lampetra japonica* (*Lethenteron japonicum*), and the southern species *Mordacia mordax* and *Geotria australis*.

During the breeding season lampreys mate, and males and females –respectively– release sperm and eggs (Hardisty, 2006). For spawning, lampreys build rudimentary nests by moving pebbles with their oral disks, making discrete depressions (nests) which are specially located in shallow parts of the river where water current is slow, and just before a flow transition into fast, rippling water (Hardisty, 2006). Fertilised eggs sink into the river where they develop. After spawning during spring or early summer, adult lampreys die, which means that they breed only once in their life-time, a characteristic known as a semelparous strategy (Hardisty, 2006). Many lampreys, both male and females, can spawn in the same nest.

Fertilised eggs develop at an optimal temperature range of 12-15 °C (Hardisty, 2006). Initial divisions are in a radial pattern, and they continue until they reach the morula stage (Tahara, 1988). Following Tahara's (1988) staging system, these initial divisions correspond to stages 1 to 7 from 0-24 hr and the morula stage (stage 8) is reached at 28 hr. At stage 9 (32 hr) the blastula is established with the appearance of the blastocoel, and by stage 12 (3 days) the embryo starts gastrulating; the blastopore appears at stage 13 ($3^{1/6}$ days). Neurulation begins at stage 17 (5 days) with the formation of the neural plate. As the embryo develops it starts to have a shape when the head protrudes (stage 21; 7.5 days) and the neck starts growing by stage 22 (8 days). Elongation of the embryo proceeds and –depending on the species– it hatches at stage 24-25 (11-12 days). The embryo keeps elongating, starts swimming, and gradually is transformed into a larva with the appearance of a velum and feeding (stage 27-28; 18-22 days). At stage 30 (31 days), when all the yolk has been consumed and the digestive tract is complete, the lamprey is considered to be a young ammocoete larva.

Young ammocoetes swim downstream to burrow in the mud, where they feed by filtering organic debris (Malmqvist and Brönmark, 1982). Ammocoetes spend several years in the mud before undergoing metamorphosis to become adults (Hardisty, 2006). The time spent as ammocoetes depends on their ultimate adult life-style. There are two types of lampreys according to the adult feeding habits: parasitic (predatory) lampreys which feed on other fish, and non-parasitic lampreys which as adults do not feed at all (Hardisty, 2006). Parasitic lampreys such as *L. fluviatilis* spend more than four years as ammocoetes, and after metamorphosis they live more than two and a half years as adults, spawning at the end of their 7th year of life (Hardisty, 1961, 2006). Between metamorphosis and spawning, *L. fluviatilis* spends about a year and a half feeding on fish (parasitic phase). On the other hand, non-parasitic lampreys such as *L. planeri* metamorphose after six years of larval life and spend some nine months as adults until they spawn at the end of their 7th year of life (Hardisty, 1961, 2006). In either case, metamorphosis involves internal and external morphological changes that set the animal to an active (either parasitic or non-parasitic) adult life-style: pharyngeal pouches are rebuilt, the heart enlarged, eyes are fully developed and fully functional, an entirely new foregut appears, as well as a suctorial disk, new adult head muscles and cartilages, a rasping tongue, horny teeth and salivary glands (Hardisty, 2006; Youson, 1997).

1.2. The Agnathans

Agnathans are a basal group within the subphylum Vertebrata whose defining feature –as its name implies– is the lack of jaws. The subphylum Vertebrata is included in the phylum Chordata which is characterised by the presence of a notochord, myomeres, a dorsal nerve cord, pharyngeal arches and a post-anal tail (Janvier, 1999;

Sansom et al., 2011). There are only two extant agnathan groups: lampreys and hagfishes. After a long controversy over if lampreys and hagfishes form a natural group or not, their joining has now been set by increasing molecular data over the past two decades which upholds their grouping as Cyclostomata. A variety of phylogenetic algorithms and methods using ribosomal RNA genes (Furlong and Holland, 2002b; Mallatt and Sullivan, 1998; Mallatt and Winchell, 2007; Stock and Whitt, 1992; Turbeville et al., 1994; Winchell et al., 2002), nuclear protein-coding gene sequences (Furlong and Holland, 2002b; Kuraku et al., 1999; Takezaki, 2003) and mitochondrial genes (Furlong and Holland, 2002b) have shown that lampreys and hagfishes form a natural group. Compelling evidence has arisen from studies on microRNAs which show that lampreys and hagfishes share unique microRNA families (Heimberg et al., 2010). Furthermore, both lampreys and hagfishes share a similarly unique adaptative immune recognition system that, instead of employing immunoglobulins as gnathostomes do, they use a whole different gene family called variable lymphocyte receptors (VLR) that –essentially– has the exact same role as immunoglobulins in gnathostomes (Pancer and Cooper, 2006).

All together, these findings would replace the traditional group Vertebrata (constituted by lampreys and gnathostomes, joined by their possession of vertebrae) by the group Craniata that also includes hagfishes, which apparently lack any sort of vertebrae but still possess a skull. However, a recent study suggested that hagfishes possess vertebrae-like elements (Ota et al., 2011), which could restore Vertebrata (=Craniata) as a natural, monophyletic, group. In either case, the term “Vertebrates” will be used in the present study, noting that it refers to gnathostomes, lampreys and hagfishes.

Apart from vertebrae-like elements, vertebrates (craniates) possess several unique characters such as a skull, a distinct brain, gill skeleton, otic capsules, paired eyes, fin radials, a heart enclosed in a pericard and zigzag-shaped myotomes (Janvier, 1999; Sansom et al., 2011). Importantly, vertebrates also possess two landmark tissues: placodes and neural crest, although it appears they are not entirely unique to vertebrates and seem to have originated gradually from more basal chordate ancestors (Bronner and LeDouarin, 2012; Donoghue et al., 2008; Holland and Holland, 2001; Jeffery, 2007; Mazet and Shimeld, 2005; Schlosser, 2005). Together, placodes and neural crest provided the earliest vertebrates with paired sensory organs for vision, olfaction, vibration-sensing, electroreception and gustation (Gans and Northcutt, 1983; Lacalli, 2001). These organs focalised in the head and accompanied the emergence of a vertebrate predatory life-style from filter-feeding invertebrate ancestors (Gans and Northcutt, 1983; Lacalli, 2001).

While extant agnathans –Cyclostomes– are quite conservative in shape and diversity, a plethora of agnathan forms existed during the Paleozoic. These extinct groups, collectively known as “ostracoderms” *sensu lato*, are also regarded as “armoured” agnathans by the presence of a mineralised dermal skeleton in the shape of plates (Forey and Janvier, 1993; Janvier, 1996, 2008) (Figure 1-1). This contrasts with Cyclostomes as they lack any sort of dermal or mineralised skeleton (Janvier and Arsenault, 2002). “Ostracoderms” include Anaspids, Arandaspids, Heterostracans, Thelodonts, Pituriaspids, Galeaspids, Osteostracans, and some other genera such as *Astraspis*, *Jamoytius*, *Cornovichthys*, *Euphanerops* and *Eriptychius* (Donoghue et al., 2000; Janvier, 1996; Newman and Trewin, 2001). Another extinct group, which lacked dermal armour but possessed mineralised teeth-like structures, the Conodonts, had an

eel-like body shape similar to Cyclostomes (Aldridge et al., 1993; Donoghue et al., 2000; Purnell and von Bitter, 1992). Phylogenetic analyses have placed them as the sister group to “Ostracoderms” (Donoghue et al., 2000). Cyclostomes, in turn, are placed as the sister group to the Conodont-“Ostracoderm” clade (Conway Morris and Caron, 2014; Donoghue et al., 2000; Gess et al., 2006). Among “Ostracoderms”, Osteostracans are considered to be the sister group to gnathostomes (Conway Morris and Caron, 2014; Gess et al., 2006; Sansom, 2009). Osteostracans share with gnathostomes certain characters such as paired pectoral fins, epicercal tail, cellular bone, an ossified dermo- and endoskeleton, among others (Janvier, 2008; Sansom, 2008). Placoderms, an entire extinct group, are considered to be the most primitive gnathostome group, which in turn constitutes the sister group to crown-group Gnathostomata (Chondrichthyans, Acanthodians and Osteichthyans including Tetrapods) (Zhu et al., 1999). However, even if “Ostracoderms” share some morphological characters with lampreys, mainly the lack of jaws, the “Ostracoderms” are now regarded as stem-gnathostomes rather than primitive agnathans closely related to Cyclostomes (Donoghue et al., 2000). This classification pushes the Cyclostomes as a more basal vertebrate –agnathan– group, further away from crown-group gnathostomes, which has important implications when interpreting data derived from Cyclostomes with respect to gnathostomes.

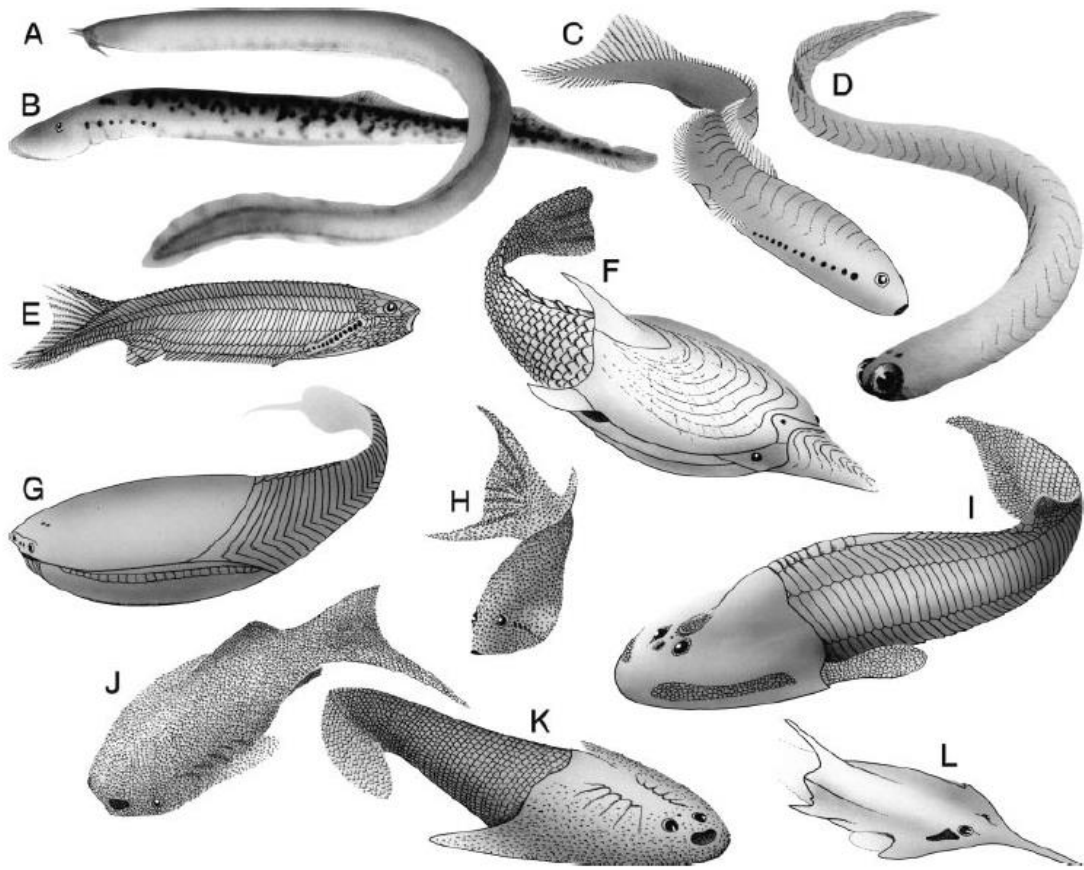


Figure 1-1. Drawings of representative extant and extinct agnathans. (A) *Eptatretus stoutii* (Lockington), a modern hagfish. (B) *Petromyzon marinus* Linnaeus, a modern lamprey. (C) *Jamoytius kerwoodi* White, a naked anaspid-like fish from the Lower Silurian of Scotland. (D) *Clydagnathus windsorensis* (Globensky), a euconodont from the Lower Carboniferous of Scotland. (E) *Pharyngolepis oblongus* Kiaer, a scaled anaspid from the Upper Silurian of Norway. (F) *Errivaspis wayensis* (White), a heterostracan from the Lower Devonian of England. (G) *Sacabambaspis janvieri* Gagnier, an arandaspid from the Upper Ordovician of Bolivia. (H) *Furcacauda heintzei* (Dineley & Loeffler), a deep deep-bodied thelodont from the Lower Devonian of Canada. (I) *Hemicyclaspis murchisoni* Egerton, an osteostracan from the Lower Devonian of England. (J) *Loganellia scotica* (Traquair), a thelodont from the Upper Silurian of Scotland. (K) *Geraspis rara* Pan & Chen, a galeaspid from the Middle Silurian of China. (L) *Pituriaspis doylei* Young, a pituriaspid from the Middle Devonian of Australia. Taken from Donoghue et al. (2000).

Cyclostomes seem to be quite divergent among agnathans in the overall shape. In particular, the lamprey body plan is very ancient and dates as far back as the Devonian period (Gess et al., 2006). The most ancient fossil lamprey discovered so far, dating back to 360 million years ago in the Devonian, *Priscomyzon riniensis* (Figure 1-2A), presents characteristic lamprey features such as a comparatively enormous oral disc

with circumoral teeth, supported by an annular cartilage, and a well-developed branchial basket with possibly seven gill pouches (Gess et al., 2006; Janvier, 2006). Another more recent lamprey fossil, *Hardistiella montanensis* (Figure 1-2B), from the Lower Carboniferous period, could have possessed a modest oral disc similar to that of metamorphosing living lampreys (Janvier and Lund, 1983). *Mayomyzon pieckoensis* (Figure 1-2C) from the Late Carboniferous (300 Ma) period, also resembles metamorphosing living lampreys but does not show an obvious oral sucking disc (Bardack and Zangerl, 1968). On the other hand, the most recent lamprey fossil discovered so far, *Mesomyzon mengae* (Figure 1-2D) from the Early Cretaceous (125 Ma), presents a more noticeable sucking disc and seven gill pouches (Chang et al., 2006). This species was most likely from freshwater, and its body proportions and internal anatomy are more similar to extant lampreys than to the Carboniferous forms (Chang et al., 2006). All together, these fossils show that lampreys have undergone few morphological changes throughout 360 million years of evolution, and suggest a well-established parasitic life-style since the Devonian period as inferred by their conserved branchial apparatus with gill pouches and the presence of an oral disc (Gess et al., 2006).

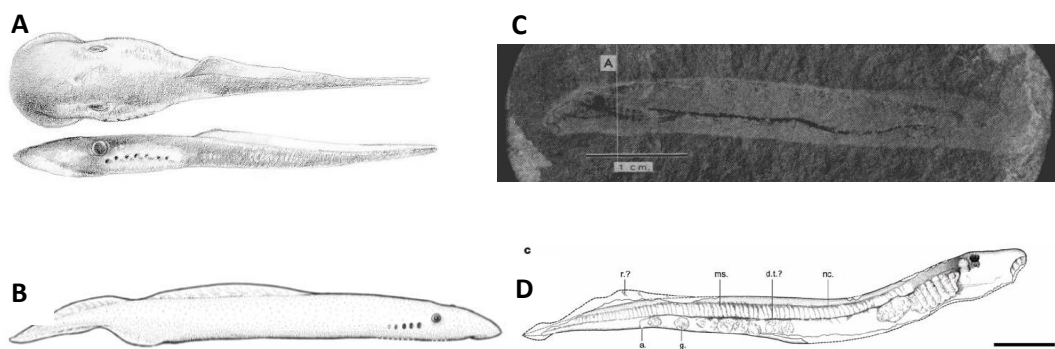


Figure 1-2. Lamprey fossils. (A) Reconstruction of *Priscoomyzon riniensis* from the Late Devonian of South Africa in dorsal (top) and left lateral (bottom) views. (B) *Hardistiella montanensis* from the Lower

Carboniferous of Montana. (C) *Mayomyzon pieckoensis* from the Late Carboniferous of Illinois. (D) *Mesomyzon mengae* from the Early Cretaceous of China. Taken from: (A) Gess et al. (2006); (B) Janvier and Lund (1983); (C) Bardack and Zangerl (1968); (D) Chang et al. (2006).

A particular group of fossils are considered to be very close to the hypothetical ancestor of all vertebrates. *Haikouichthys* and *Mylokunmingia* possess a series of characters that are diagnostic of vertebrates (Shu et al., 1999). They present a skull, gill arches, zigzag-shaped myosepta and fin rays, *Haikouichthys* further presenting vertebrae-like elements (Janvier, 1999; Shu et al., 1999, 2003). *Haikouichthys* and *Mylokunmingia* seem to represent advanced forms of pre-craniate crest animals (i.e. animals with neural crest derivatives such as cartilage of pharyngeal arches, also inferred from fossils), that is, ancestors of true craniates (Chen, 2009). The presence of sensory structures in *Haikouichthys* such as large lateral eyes, a single median nostril with a pair of nasal capsules, and possible otic vesicles, suggests that *Haikouichthys* possessed a well-developed brain like vertebrates (Shu et al., 2003). However, even if *Mylokunmingia* and *Haikouichthys* present many vertebrate characters, their exact phylogenetic position with respect to Cyclostomes is not completely elucidated (Shu et al., 1999, 2003). Phylogenetic analyses using morphological characters have placed *Haikouichthys* as the sister group of lampreys or as the sister group of the lamprey + “Ostracoderms” clade at the exclusion of hagfishes, whereas *Mylokunmingia* has been placed as the sister group of the *Haikouichthys* + lampreys + “Ostracoderms” clade at the exclusion of hagfishes (Shu et al., 1999, 2003) (Figure 1-3).

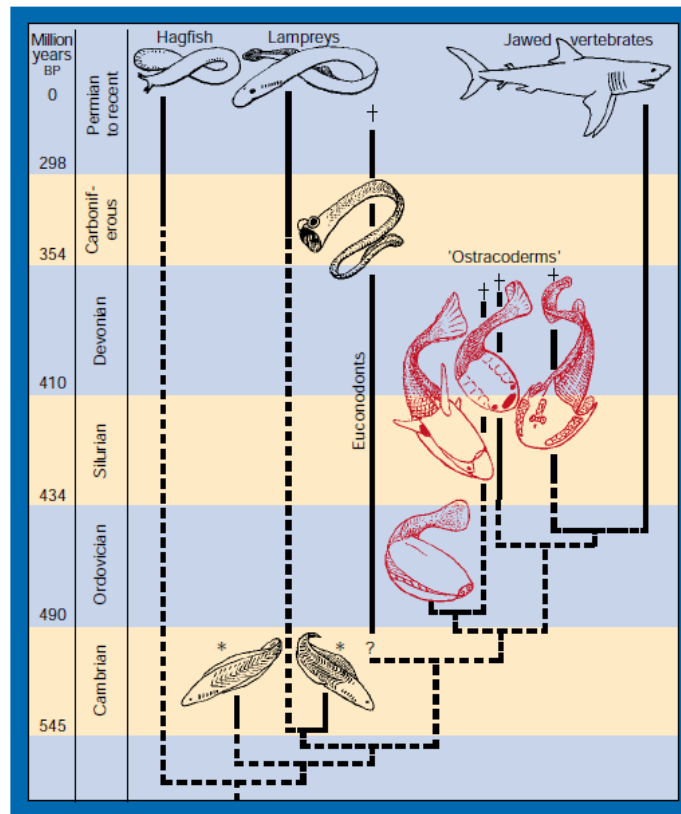


Figure 1-3. Generalised vertebrate phylogeny including representative extant and fossil vertebrates. Phylogenetic tree based on Shu et al. (1999) using morphological characters. According to this phylogeny, “Ostracoderms” (in red) are –in general terms– the sister group of gnathostomes, and are therefore classified as stem gnathostomes. This clade is in turn the sister group of Conodonts (Euconodonts). Lampreys are then shown as the sister group of the Conodont-“Ostracoderm”-Gnathostome clade (this grouping is based on morphological characters only, but current molecular phylogenies have consistently grouped lampreys with hagfishes as the natural –monophyletic– group Cyclostomata). Hagfishes are then shown as the sister group of the Lamprey-Conodont-“Ostracoderms”-Gnathostome clade. Two putative vertebrate fossils, *Haikouichthys* and *Myllokunmingia*, are placed at the base of the vertebrate clade, although their exact phylogenetic position with respect to lampreys, hagfishes and “Ostracoderms” is uncertain. Taken from Janvier (1999).

Lamprey fossils show that lampreys had diverged very early in vertebrate evolution. One study using a combination of mitochondrial and nuclear genes, together with fossil data, suggested that lampreys and hagfishes separated from each other around 390-470 million years ago during the Ordovician-Silurian-Devonian periods (Kuraku and Kuratani, 2006). Additionally, cyclostomes were predicted to have diverged from the gnathostome lineage around 500 million years ago (Kuraku and Kuratani, 2006). In spite of their early divergence, lampreys still have a well-defined and complex internal

anatomy and developmental processes quite similar to extant gnathostomes. These aspects will be described in more detail in the following section, with particular reference to the nervous system.

1.3. Morphology of the Vertebrate Nervous System

The entire vertebrate nervous system can be divided into Central Nervous System (CNS) and Peripheral Nervous System (PNS). The CNS includes the brain and spinal cord (here referred to as the neural tube), whereas the PNS includes cranial and spinal nerves and their associated ganglia (Butler and Hodos, 1996). Here follows a description of the generalised composition of the vertebrate nervous system, mostly based on mouse and chicken as there is a large amount of information from these species and virtually all anatomical structures and compartments can be better distinguished in these species with respect to other vertebrates.

1.3.1. Vertebrate Central Nervous System

The vertebrate brain presents subdivisions known as –from anterior to posterior– forebrain, midbrain and hindbrain (Kiecker and Lumsden, 2005; Murakami et al., 2005; Swanson, 2000). The spinal cord follows the hindbrain posteriorly (Butler and Hodos, 1996). The hindbrain (except roughly the anterior half of the first rhombomere) plus the spinal cord is known as the rhombospinal region of the neural tube (Holland and Holland, 1999; Kozmik et al., 1999). Along the length of the embryonic neural tube, the region facing the lumen is composed of active, proliferating neuroblasts and is called the ventricular zone (Gilbert, 2003; Sanes et al., 2006; Squire et al., 2008). Outside this

zone towards the periphery of the neural tube, the mantle layer is found, which is composed of differentiated neuronal and glial cells (Gilbert, 2003; Sanes et al., 2006). At the internal margin of the mantle layer, there is a thin cell layer called the subventricular zone that corresponds to recently born postmitotic neurons which migrate radially from the ventricular zone (Garel et al., 1997a). Neurons from the mantle layer eventually send axons to the periphery of the neural tube forming a cell-poor marginal zone (Gilbert, 2003; Sanes et al., 2006). However, for convenience with the literature, in the present study the mantle layer plus the marginal zone will be referred to as the marginal zone. Finally, the entire neural tube is composed of, roughly, two halves: the dorsal –alar– plate and the ventral –basal– plate which are separated by the *sulcus limitans* (Squire et al., 2008).

Forebrain: The forebrain is composed of two parts: telencephalon and diencephalon (Gilbert, 2003; Sanes et al., 2006; Squire et al., 2008). The telencephalon is the anterior-most subdivision of the brain; it encompasses the PALLIUM (dorsal part) and SUBPALLIUM (ventral part) (Murakami et al., 2005; Puelles, 2001). In tetrapods, the pallium is further subdivided into dorsal pallium (DP), medial pallium (MP), lateral pallium (LP) and ventral pallium (VP) (Brox et al., 2004; Medina and Abellán, 2009; Puelles et al., 2000). The DP develops into the neocortex which is related to the achievement of complex cognitive capabilities (Medina and Abellán, 2009; Pierani and Wassef, 2009). The MP forms the hippocampus which is involved in the rapid formation of new memories, spatial learning, memory retrieval and emotion processing (Medina and Abellán, 2009; Preston and Eichenbaum, 2013; Sendrowski and Sobaniec, 2013). Finally, the LP and VP give rise to different parts of the olfactory cortex and the claustramygdaloid complex (Medina and Abellán, 2009). The claustrum is involved in

the coordination of widespread –interhemispheric– cortical regions within and across sensory and motor modalities (Crick and Koch, 2005; Smith and Alloway, 2010; Smith et al., 2012), whereas the amygdala regulates cognitive processes such as attention, learning, memory, as well as cardiac and respiratory rhythms and emotional behaviour (Medina and Abellán, 2009; Medina et al., 2004). The ventral pallium also gives rise to the olfactory bulb (Medina and Abellán, 2009; Puelles et al., 2000). The subpallium is composed of the lateral (LGE) and medial (MGE) ganglionic eminences. The LGE gives rise to the striatum, while the MGE originates the pallidum and septum (Puelles et al., 2000). The MGE is the most important source of interneurons in the telencephalon (Medina and Abellán, 2009).

The diencephalon is composed of five major subdivisions in its alar region: PRETECTUM, EPITHALAMUS, DORSAL THALAMUS (or simply THALAMUS), VENTRAL THALAMUS (or PRETHALAMUS), and HYPOTHALAMUS (Medina et al., 2011). The pretectum receives signals from the retina and optic tectum (midbrain) and is involved in visuomotor behaviour (e.g. prey hunting). The epithalamus is composed of the epiphysis which contains the pineal gland, parietal eye and the habenula, and is mainly involved in the circadian rhythm and reproductive cycles in response to light. The dorsal thalamus relays sensory information from the body, retina, ear and optic tectum to the telencephalon. The ventral thalamus processes sensory information from the retina and motor functions of the body, head, eyes and limbs. Finally, the hypothalamus, in its versatile functions, is involved in the regulation of feeding, drinking, sleeping, emotions, temperature regulation, hormone production, sexual behaviour, hibernation and diurnal rhythms, among others.

The Neuromeric Concept: Before proceeding with a general description of the midbrain, it is convenient to make notice of the neuromeric concept. Neuromeres are segments of the brain that possess neuronal populations somewhat separated from one another. Neuromeres in the forebrain are called prosomeres (Figdor and Stern, 1993; Milán and Puelles, 2000; Puelles and Rubenstein, 1993), those in the midbrain are called mesomeres, whereas in the hindbrain they are called rhombomeres (Murakami et al., 2005). This neuromeric concept is more complex and also encompasses nuclei, many small regions within each neuromere (especially in prosomeres), and a separation between alar and basal plates of each neuromere with their corresponding and specific regions (Puelles and Rubenstein, 2003). Given that the present study is dedicated to the embryonic lamprey nervous system (in this case, the embryonic brain), a comparison between the complex gnathostome adult brain regions and the lamprey embryonic brain is not very informative. Instead, only the main brain compartments will be considered.

The neuromeric concept has been used to understand gene expression patterns observed in the vertebrate brain as it visualises all vertebrate brains as segmented, homologous structures, sharing a common Bauplan of equivalent developmental units. From a neuromeric point of view, the forebrain is transversally divided into diencephalon caudally and secondary prosencephalon rostrally (Puelles and Rubenstein, 1993). The ventral part of the secondary prosencephalon consists of the hypothalamus, the dorsal part contains the entire telencephalon, and the optic vesicles are located laterally (Puelles and Rubenstein, 1993). On the other hand, from a neuromeric point of view, the diencephalon is composed of all compartments discussed above except the hypothalamus, which is considered part of the secondary prosencephalon (Medina et al., 2011).

The first prosomeric models in mouse recognised 6 prosomeres, p1-p6 (Puelles and Rubenstein, 1993), three dividing the diencephalon (p1-p3) and three dividing the secondary prosencephalon (p4-p6). However, later examinations considered only p1-p3 in the diencephalon, leaving the secondary prosencephalon as –provisionally– unsegmented (Puelles and Rubenstein, 2003). This latter model (with only three diencephalic prosomeres and the secondary prosencephalon; Figure 1-4) is used as a template in the present study as previous analyses of the lamprey embryonic brain found it difficult to recognise more than three prosomeres (Murakami et al., 2001, 2002; Osorio et al., 2005).

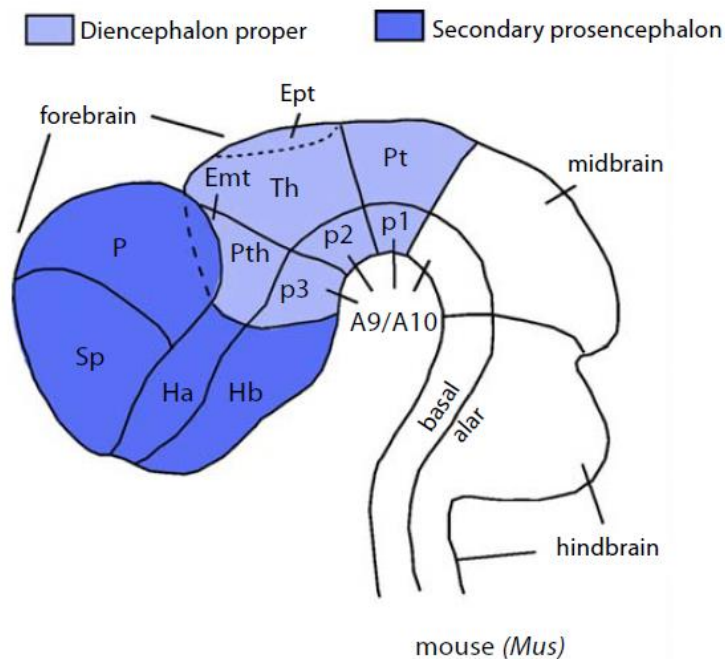


Figure 1-4. Schematic representation of the mouse embryonic brain and its major divisions and subdivisions according to current prosomeric models. The forebrain is divided into the diencephalon proper and the secondary prosencephalon (shown in different colors). The secondary prosencephalon (dark blue) produces the telencephalon dorsally –composed of pallium (P) and subpallium (Sp)–, the optic vesicle laterally (not represented), and the hypothalamus ventrally, including its alar (Ha) and basal (Hb) parts. The diencephalon proper (light blue) is composed in its alar region by the pretectum (Pt), conforming prosomere 1 (p1) together with the corresponding basal part; by the epithalamus (Ept) and thalamus/dorsal thalamus (Th) conforming prosomere 2 (p2) together with the corresponding basal part; and by the eminentia thalami (Emt) (marked behind the pallium by a dashed line) and prethalamus/ventral

thalamus (Pth) conforming prosomere 3 (p3) together with the corresponding basal part. A9/A10: A9 and A10 dopaminergic cell groups. Taken from Medina et al. (2011).

Midbrain: The midbrain is composed of three main parts: the TECTUM, the TEGMENTUM and the ISTHMUS (Butler and Hodos, 1996). The tectum is found at the dorsal side of the midbrain. It is involved in integrating the spatial location of visual, auditory, somatosensory and lateral line inputs to effect spatially oriented motor responses. The tegmentum lies ventral to the tectum and, in general terms, relays sensory information from the hindbrain towards the forebrain to effect motor responses. It also contains the nuclei of cranial nerves III and IV that regulate eye movement. Finally, the isthmus separates the tegmentum from the hindbrain and is a continuation of rostrally projecting tracts coming from the hindbrain and cerebellum.

Hindbrain: The hindbrain is composed of seven or eight segments called rhombomeres. Rhombomeres are thought to be somewhat independent segments and each presents its own proliferation and neurogenesis centres. Each rhombomere is also considered a compartment separated from other rhombomeres by boundaries of reduced proliferation that restrict cell movements (Fraser et al., 1990; Jimenez-Guri et al., 2010). However, it has been suggested that inter-rhombomeric interactions also regulate their own development (Eickholt et al., 2001). Particularly interesting are rhombomeres 3 and 5 which have delayed neuronal differentiation of interneurons and motor neurons (Eickholt et al., 2001).

Spinal cord: The spinal cord is the part of the central nervous system posterior to the hindbrain that processes information coming from, and sent to, different parts of the body, information that is in turn largely controlled by the brain (Butler and Hodos, 1996). Longitudinally, the spinal cord is somewhat segmented, each segment being the region of the spinal cord from which the spinal nerves emerge (see below). Neuroblasts in the ventricular zone exit the cell cycle and then migrate towards the periphery to differentiate into neurons in the marginal zone (Garcia-Dominguez et al., 2003). Eventually, the marginal zone becomes thicker as more neurons differentiate and, consequently, the ventricular zone becomes thinner.

1.3.2. Vertebrate Peripheral Nervous System

Cranial nerves and cranial ganglia: Vertebrates have traditionally been considered to possess twelve –paired– cranial nerves (Table 1-1), which exit at specific positions of the brain. While different cranial nerves contain motor, sensory and/or autonomic axons, only some nerves containing sensory axons (nerves V, VII, VIII, IX and X) present –paired– sensory ganglia external to the brain. In the present study, only such nerves will be considered given that different neural HLH genes are specifically expressed in these ganglia. As a reference, the neuronal bodies of motor and autonomic axons of cranial nerves are located inside the brain. Of note, the trigeminal ganglion possesses two lobes, the ophthalmic (profundal; V₁) and maxillomandibular (trigeminal; V_{2,3}) lobes.

Table 1-1. List of the twelve traditional cranial nerves in vertebrates.

Number	Name	Axon type	Innervation	Ganglion (sensory)
I	Olfactory	Sensory	Olfactory epithelium	Absent
II	Optic	Sensory	Retina	Absent
III	Oculomotor	Motor, Autonomic	Internal and external eye muscles	-
IV	Trochlear	Motor	External eye muscles	-
V	Trigeminal	Motor, Sensory	Jaw muscles, face, snout, oral cavity	External
VI	Abducens	Motor	External eye muscles	-
VII	Facial	Motor, Sensory, Autonomic	Taste buds, facial muscles, salivary and tear glands	External
VIII	Acoustic	Sensory	Cochlea, vestibular organs	External
IX	Glossopharyngeal	Motor, Sensory, Autonomic	Taste buds, pharynx, salivary glands	External
X	Vagus	Motor, Sensory, Autonomic	Taste buds, viscera of thorax and abdomen, larynx, pharynx	External
XI	Spinal accessory	Motor, Autonomic?	Neck and shoulder muscles	-
XII	Hypoglossal	Motor	Tongue	-

An additional series of cranial nerves of the lateral line system are present in anamniote vertebrates. There are primitively six pairs of lateral line nerves in gnathostomes, three pre-otic and three post-otic (Northcutt, 1997). The three pre-otic lateral line nerves are the anterodorsal, anteroventral and otic, whereas the three post-otic lateral line nerves are the middle, supratemporal and posterior lateral line nerves (Northcutt, 1997). All lateral line nerves possess their own sensory ganglion which can be secondarily fused with other cranial sensory ganglia (Northcutt, 1997).

Spinal nerves and ganglia: Similar to the brain which possesses pairs of cranial nerves, the spinal cord also possesses pairs of spinal nerves along its length that exit at regular intervals. Each of these nerves is a grouping of axons coming from a dorsal and a ventral exit point in the spinal cord, named dorsal root and ventral root, respectively. The dorsal root corresponds to incoming sensory axons that enter the spinal cord on the dorsal side, whereas the ventral root corresponds to a combination of motor and

autonomic axons that exit the spinal cord from the ventral side. Both roots join at a certain distance from the spinal cord to form a common spinal nerve. Both the dorsal and ventral roots contain ganglia halfway from the spinal cord to the point they join. Ganglia of the dorsal root are called dorsal root ganglia (DRG) which contain the cell bodies of incoming sensory axons. On the other hand, ganglia in the ventral root are called sympathetic ganglia which contain the cell bodies of outgoing sympathetic axons only (Figure 1-5).

1.3.3. Vertebrate Autonomic Nervous System

The autonomic nervous system (ANS) is a subdivision of the PNS which regulates the activity of visceral organs. It can be classified into the sympathetic nervous system, parasympathetic nervous system (sometimes split into cranial and pelvic parasympathetic nervous system) and enteric nervous system (Funakoshi and Nakano, 2007).

The sympathetic nervous system consists of preganglionic fibres arising at regular intervals from lateral columns in the spinal cord at the cervical or thoracico-lumbar level (Craven, 2007). They connect with postganglionic neurons lateral to the spinal cord forming a chain of paired paravertebral (sympathetic) ganglia or single prevertebral (cervical) ganglia (Asamoto, 2005; Burnstock, 1969). The sympathetic chain thus consists of a bilateral system of paravertebral ganglia joined by longitudinal connectives which, together with cervical fibres and ganglia, constitute the sympathetic trunk (Craven, 2007). In general terms, the sympathetic nervous system regulates the contraction of smooth muscles and the secretion of glands (Craven, 2007; Butler and Hodos, 1996).

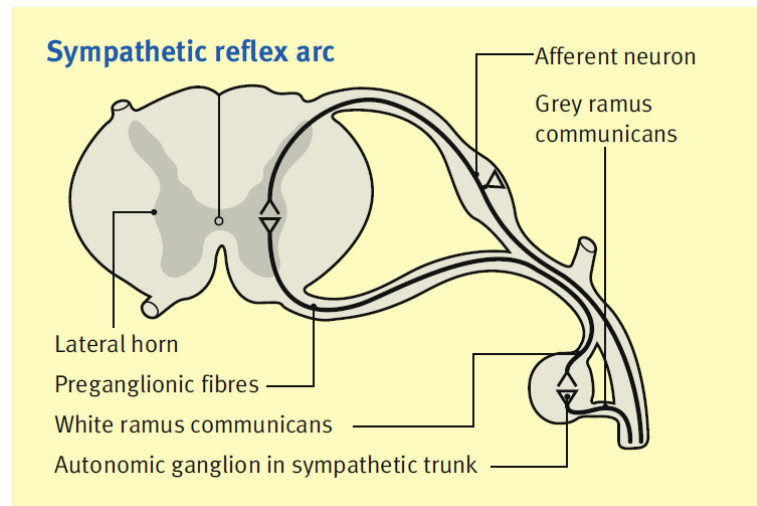


Figure 1-5. Schematic representation of the sympathetic reflex arc to illustrate the anatomical relationship between sympathetic (Autonomic ganglion in sympathetic trunk) and dorsal root ganglia (Afferent neuron) in spinal nerves. Given the repetitive, paired nature of spinal nerves, associated chains of paired sympathetic (ventral) and dorsal root ganglia are observed on each side of the spinal cord. Taken from Craven (2007).

The parasympathetic nervous system is smaller than the sympathetic system and consists of fibres of cranial nerves III, VII, IX and X, and of some sacral spinal nerves (Burnstock, 1969; Craven, 2007). As in the sympathetic nervous system, preganglionic neurons of cranial nerves connect with postganglionic neurons in cranial ganglia which innervate glands such as the lacrimal, nasal and parotid glands. In the sacral region, fibres innervate the bladder, genitalia, the colon and rectum (Burnstock, 1969; Craven, 2007). In general terms, the parasympathetic nervous system maintains the vegetative state of the innervated visceral organs (Craven, 2007).

The enteric nervous system is the intrinsic innervation of the gastrointestinal tract and can function independently of the CNS (Epstein et al., 1996). It controls the motility, exocrine and endocrine secretions, and microcirculation of the gastrointestinal tract, as well as immune and inflammatory processes (Burns and Thapar, 2006; Epstein et al., 1996). Neurons and glial cells of the enteric nervous system originate from vagal

and sacral regions of the neural crest. Neuronal bodies of the enteric nervous system are grouped in ganglia interconnected by bundles of nerve processes forming plexuses, which are located between the external muscle layers of the gut wall (Burns and Thapar, 2006; Epstein et al., 1996).

1.3.4. Vertebrate Cranial Placodes

Cranial placodes are traditionally regarded as ectodermal thickenings that arise in stereotypical positions of the embryonic head (Baker and Bronner-Fraser, 2001; Brugmann and Moody, 2005; Whitlock, 2004). However, not all neurogenic placodes are thickened and some ectodermal thickenings are not considered placodes (Schlosser, 2005, 2006; Schlosser and Northcutt, 2000). Instead, placodes can be defined as specialised areas of the cranial non-neural ectoderm where cells undergo pronounced cell shape changes, resulting in thickening, invagination, and/or cell delamination, and which give rise to various non-epidermal cell types (Schlosser, 2005, 2006). Also, placodes are regions of increased cell proliferation compared to the surrounding ectoderm (Saka and Smith, 2001) and, except for the adenohipophyseal and lens placodes, are neurogenic (Schlosser, 2005).

Different types of placodes have been identified: The adenohipophyseal placode is localized at the anterior part of the head and gives rise to the endocrine adenohipophysis through invagination from the stomodeum. The olfactory placodes also invaginate to give rise to supporting cells, mucus-producing cells and primary sensory neurons of the epithelia of the olfactory and vomeronasal organs. The lens placodes invaginate to form the lens vesicles, which give rise to the crystallin-producing

cells that will form the eye lenses. The profundal and trigeminal placodes (in anamniotes), or the trigeminal (ophthalmic and maxillomandibular) placodes (in amniotes), give rise to neuronal precursors that delaminate to form sensory neurons of the distal parts of ganglia of profundal and trigeminal nerves, respectively. They transmit somatosensory stimuli (touch, pain, temperature) from the face, jaws and teeth. The otic placodes invaginate to form the otic vesicles that produce the secondary mechanosensory hair cells of the inner ear and from which neuronal precursors migrate to form the sensory neurons of the vestibulochoclear ganglion. The lateral line placodes give rise to the peripheral components of the lateral line system in anamniotes, which contains mechanoreceptive (neuromasts) and electroreceptive (ampullary and tuberous organs in teleosts) receptor organs and their associated ganglia that sense disturbances in the surrounding water and electric fields (Baker et al., 2013). The epibranchial placodes (geniculate, petrosal and nodose) give rise to neuronal precursors that migrate from the epithelium to form sensory neurons in the distal ganglia of cranial nerves VII, IX and X, respectively. They transmit stimuli from the taste buds, and concerning heart rate, blood pressure, bronchial irritation, and gut distension. Finally, the neurogenic hypobranchial placodes in amphibians form the hypobranchial ganglia of probable viscerosensory function (Baker and Bronner-Fraser, 2001; Schlosser, 2005, 2006). Recently, a different placode, the paratympanic organ (PTO) placode located immediately dorsal to the geniculate placode, was identified in chicken and shown to be independent of the epibranchial and lateral line placodes (O'Neill et al., 2012). This placode forms the –amniote– PTO and its afferent neurons, an organ which has in turn been suggested to be homologous to the anamniote spiracular organ (O'Neill et al., 2012).

Fate maps of zebrafish, chicken and *Ambystoma* embryos suggest that all cranial placodes originate from a horseshoe-shaped pre-placodal region (Schlosser, 2006; Streit, 2007). This region is located in the outer neural folds and immediately adjacent ectoderm, bordering the anterior neural plate rostrally and the cranial neural crest laterally. However, based on the inductive mechanisms, development, function and evolution of the ectodermal placodes, it has been proposed that placodes do not constitute a single group, and therefore should be considered independently (Begbie and Graham, 2001). This idea might be strengthened by recent findings suggesting the presence of certain types of placodes in tunicates, suggesting that placodes may have different evolutionary origins (Bassham and Postlethwait, 2005; Mazet and Shimeld, 2005; Mazet et al., 2005).

1.4. The Lamprey Embryonic Nervous System and its Relation to Gnathostomes

1.4.1. Early Development

In lampreys, neurulation starts at Tahara (1988) stage 17 with the appearance of the neural plate. A set of genes –*Bmp*, *Wnt*, *Delta*, *Notch*, *Msx*, *Zic*, *Pax*, *Dlx*, *Sox* and *prdm* homologues– and their combined expression patterns allow the identification of a neural plate border at these stages (Nikitina et al., 2011; Sauka-Spengler et al., 2007) (Figure 1-6). Later, a neural groove shows in the middle of the neural plate (Tahara, 1988). Lampreys do not seem to have the three paired stripes of primary neurons as occurs in

vertebrates such as zebrafish and *Xenopus*. In these species, the first wave of neuronal differentiation begins at the open neural-plate stage and the neural plate contains bilateral domains of precursors of motoneurons, interneurons and sensory neurons, which are referred to as primary neurons to indicate their early birth date.

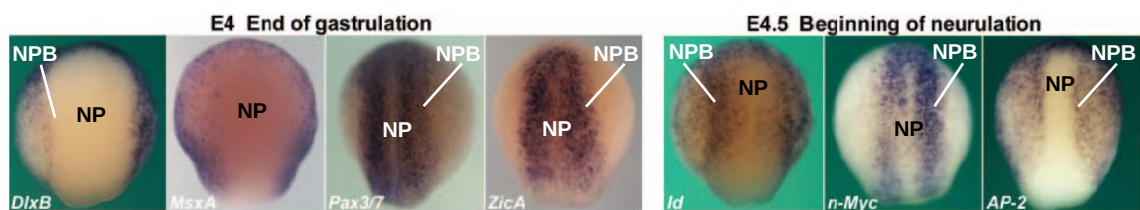


Figure 1-6. Expression of different transcription factor genes that highlight the presence of a neural plate and neural plate border in *Petromyzon marinus* embryos. E: embryonic day. Abbreviations: NP, neural plate; NPB, neural plate border. Taken from Sauka-Spengler and Bronner-Fraser (2008).

Later at Tahara stage 18, the neural folds on both sides of the groove start to elevate as it elongates underneath the ectoderm. By stage 19 the folds elevate even more forming a neural ridge, and later they fuse at the top around stage 20 forming a solid rod, which elevates more, gradually detaching from the epidermis. This neural rod later expands at the anterior as the head protrudes from the yolk around stage 21. Finally, by stage 22 the neural rod cavitates with a longitudinal, modest slit, thus forming the neural tube (Tahara 1988).

The anterior part of the neural tube becomes segmented into forebrain, midbrain and hindbrain at stage 23-24 (Tahara 1988) (Kuratani et al., 1998). The epiphysis appears at stage 24. The lamprey neural tube starts as a two-halved epithelium only one cell-wide each, whose cells are large and placed in a horizontal position forming a very faint cavity in the centre along the dorso-ventral axis (Damas, 1944). Later during

development, cells right next to the cavity start proliferating acquiring a rounded morphology, and the neural tube gradually becomes pseudostratified. With time, these proliferating cells will form a ventricular zone along the neural tube as occurs in gnathostomes, leaving a marginal zone at the periphery (Villar-Cheda et al., 2006).

1.4.2. Composition of the Lamprey Central Nervous System

Brain: The lamprey brain is very much similar to gnathostomes overall. They possess well defined and complex subdivisions, i.e., forebrain, midbrain and hindbrain. Segmentation of the lamprey brain during development is best seen at Tahara (1988) stage 26-27. The forebrain is composed of a clear telencephalon and diencephalon. The telencephalon is further subdivided into pallium and subpallium (Murakami et al., 2005; Sugahara et al., 2011). Expression of lamprey *Dlx1/6* in all of the subpallium correlates this region with gnathostome LGE, but not with MGE, which additionally expresses *Nkx2.1* and *sonic hedgehog (shh)* (Murakami et al., 2002, 2005). Therefore, lampreys seem to lack an equivalent of the MGE (Medina and Abellán, 2009; Murakami et al., 2005). The lamprey pallium is marked by expression of *Pax6*, and a dorsal pallium region is additionally differentiated by co-expression of *Pax6* and *Emx*, possibly corresponding to the combined gnathostome MP, DP and LP (Murakami et al., 2001, 2002, 2005). The region in the pallium expressing *Pax6* but devoid of *Emx* expression thus corresponds to the gnathostome VP (Figure 1-7).

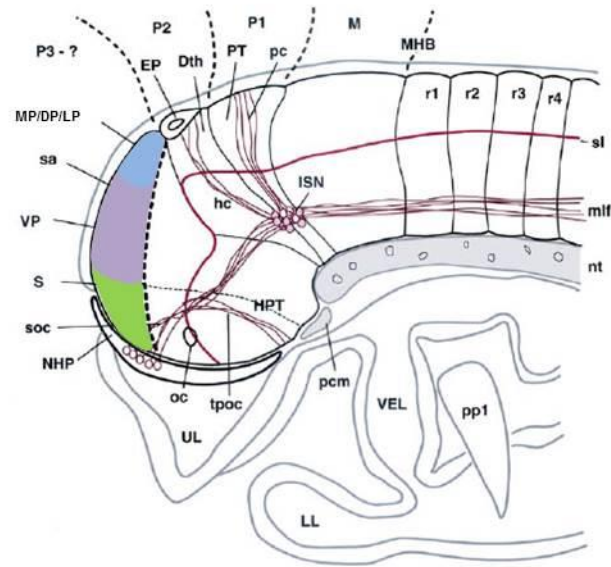


Figure 1-7. Schematic representation of the main lamprey brain subdivisions at stage 26. Gene expression patterns and nerve tracts help to define the neuromeric organisation of the lamprey brain. The telencephalon is composed of the pallium and subpallium (confirmed by other studies, e.g. Murakami et al, 2005; Sugahara et al., 2011). Expression of different transcription factors suggest a subdivision of the pallium into a ventral pallium (VP, purple) and a dorsal region corresponding to combined gnathostome medial + dorsal + lateral pallium (MP/DP/LP, blue), also defined by expression of lamprey *Emx*. The lamprey subpallium likely corresponds to the gnathostome lateral ganglionic eminence (LGE, green) as suggested by expression of lamprey *Dlx1/6*. The diencephalon contains a hypothalamus (HPT), a dorsal thalamus (Dth), a pretectum (PT) and the epiphysis (EP). PT is part of prosomere 1 (P1) and Dth part of prosomere 2 (P2). Further prosomere subdivisions of the rostral diencephalon (which would correspond at least to P3) are uncertain. The thick dotted line demarcates the telencephalon-diencephalon boundary that corresponds to the anterior intraencephalic sulcus (sa), whereas the thin dotted lines at the top represent inferred prosomere boundaries. Additionally, the hindbrain is composed of rhombomeres as in gnathostomes. Abbreviations: hc, habenular commissure; ISN, interstitial nucleus; LL, lower lip; M, midbrain; MHB, midbrain-hindbrain boundary; mlf, medial longitudinal fasciculus; NHP, nasohypophysial plate; nt, notochord; oc, optic chiasm; pc, posterior commissure; pcm, prechordal mesoderm; pp1, pharyngeal pouch 1; r1-4, rhombomeres; S, striatum; sl, sulcus limitans; soc, supraoptic commissure; tpoc, postoptic tract; UL, upper lip; VEL, velum. Modified from Murakami et al. (2002, 2005).

In the dorsal diencephalon, the epiphysis lies above the habenular complex and is in direct contact with the telencephalon (Kawano-Yamashita et al., 2007). Within the diencephalon, a region homologous to the gnathostome hypothalamus is suggested by the ventral expression of transcription factor *TTF1* (*Nkx2.1*) (Murakami et al., 2001; Ogasawara et al., 2001). Also, *Dlx1/6* expression in the diencephalon suggests a region homologous to the gnathostome thalamus that is separated from the hypothalamus by the *sulcus limitans* (Murakami et al., 2001). Additionally, *Pax6* and *OtxA* expression in

the diencephalon suggest the presence of a dorsal thalamus and pretectum (Murakami et al., 2002). Morphologically, the boundary between forebrain and midbrain corresponds to the position of the posterior commissure, whereas the boundary between telencephalon and diencephalon corresponds to the anterior intraencephalic sulcus (Murakami et al., 2001).

Adult lampreys have been described to present a neuromeric (prosomeres P1-P6) forebrain (Pombal and Puellas, 1999). In lamprey embryos, however, only three to four prosomeres (P1-P4) have been suggested so far by expression of a number of transcription factors (Murakami et al., 2001, 2002; Osorio et al., 2005). However, two models differ considerably in the boundaries of each prosomere, particularly with respect to the epiphysis (Murakami et al., 2001, 2002; Osorio et al., 2005). As a consequence, there is not a consensus definition of embryonic lamprey prosomeres so far. In the present study, however, the model proposed by Murakami et al. (2001) (Figure 1-7) will be used as a template as it agrees well with the results presented here as will be discussed in further chapters.

Lampreys also present a morphologically recognisable midbrain-hindbrain boundary (MHB) marked by the expression of *Pax2/5/8*, *Fgf8/17* and *Engrailed* (Holland et al., 1993; Murakami et al., 2001, 2002; Sugahara et al., 2011; Takio et al., 2007).

The lamprey hindbrain is composed of rhombomeres as occurs in gnathostomes. So far, six rhombomeres have been identified in the lamprey hindbrain with certainty, but a 7th and 8th rhombomeres have also been suggested to be present although with no clear definition (Meléndez-Ferro et al., 2003; Osorio et al., 2005). Morphologically, they can

be recognised at stage 23 (Kuratani et al., 1998), but molecular markers such as *Krox20* show that the lamprey hindbrain is already segmented much earlier (Jimenez-Guri and Pujades, 2011) and keeps being segmented at later stages (Murakami et al., 2004).

Spinal cord: Lampreys possess a well defined spinal cord and its general anatomy is nearly identical to gnathostomes. As stated before, they present a ventricular zone of neural precursors and a marginal zone of differentiating cells. Motor neurons are found on the ventral region of the spinal cord, sensory neurons in the dorsal part, and interneurons in the middle along the dorso-ventral axis (Viloria et al., 2008).

The transition of the hindbrain and spinal cord can be recognised by the presence of cerebrospinal fluid-contacting (CSF-c) GABA immunoreactive cells (Kolmer-Agduhr cells) (Meléndez-Ferro et al., 2003). Additionally, expression of *zic-1* at the dorsal-most part of the lamprey spinal cord suggests that lampreys have a roof plate, and further expression of *raldh2* reveals the presence of dorsal interneurons (dIs) (Castillo et al., 2010).

1.4.3. Composition of the Lamprey Peripheral Nervous System

Cranial nerves and ganglia: As in gnathostomes, lampreys have a virtually complete repertoire of cranial nerves. They possess an olfactory nerve (nerve I) formed by axons of the olfactory sensory neurons located in the olfactory epithelium that project to the olfactory bulb in the telencephalon (Frontini et al., 2003; Zaidi et al., 1998). Lampreys also have optic nerves (nerve II) and motor nerves III, IV and VI that innervate the extrinsic eye muscles (Richardson et al., 2010). They also have cranial nerves V (with the ophthalmic, V₁, and maxillomandibular, V_{2,3}, branches), VII, IX and

X with both sensory and motor components, as well as lateral line nerves branching from nerves VII and X (Fritzsche and Northcutt, 1993; Richardson et al., 2010) (Figure 1-8). Cranial nerves VIII and XII are also present, and they only appear to lack an accessory nerve XI (Richardson et al., 2010). As in gnathostomes, the neuronal bodies of sensory axons of cranial nerves V, VII, VIII, IX, X and the lateral line branch from the vagus are also arranged in paired sensory ganglia (Hammond and Whitfield, 2006; Kuratani et al., 1997) (Figure 1-9).

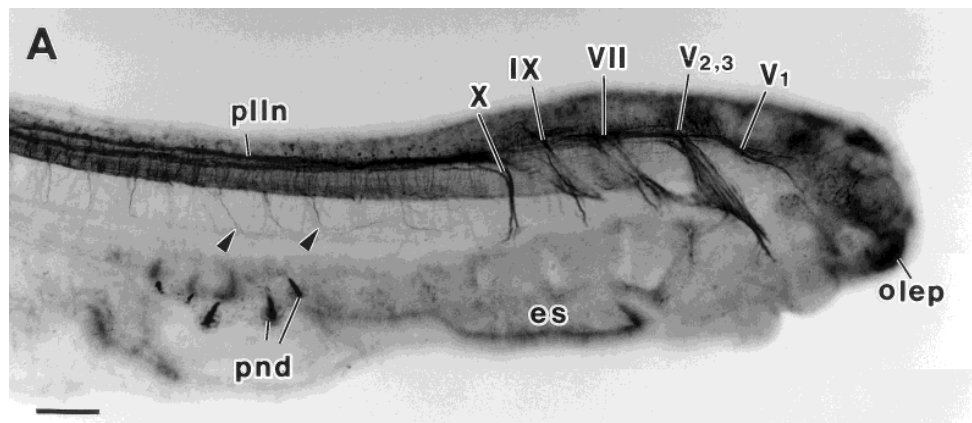


Figure 1-8. Cranial nerves of developing *Lampetra japonica* embryos. Stage 25 *L. japonica* embryos stained with anti-acetylated tubulin antibody showing branchiomic nerve roots of cranial nerves V₁, V_{2,3}, VII, IX and X as indicated. Abbreviations: es, endostyle; olep, olfactory epithelium; pll, posterior lateral line nerve; pnd, pronephric duct; V₁, ophthalmicus profundus nerve root; V_{2,3}, maxillomandibular nerve root; VII, facial nerve root; IX, glossopharyngeal nerve root; X, vagus nerve root. Taken from Kuratani et al. (1997).

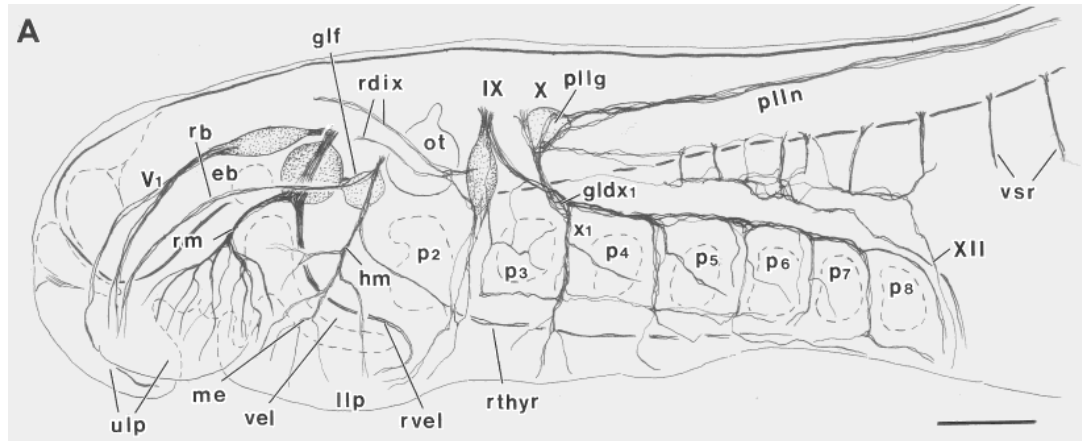


Figure 1-9. Cranial nerves and cranial ganglia of developing *Lampetra japonica* embryos. Graphical illustration of developing neurites based on a stage 27 *L. japonica* embryo stained with anti-acetylated tubulin antibody showing both nerves and cranial ganglia. Abbreviations: eb, eyeball; gldx₁, epibranchial ganglion of the first vagus nerve; glf, facial ganglion; hm, hyomandibular nerve; ot, otocyst; llp, lower lip; me, external mandibular ramus; olep, olfactory epithelium; p2-8, pharyngeal pouch 2-8; pllg, posterior lateral line ganglion; plln, posterior lateral line nerve; pnd, pronephric ductus; rb, buccal ramus of the facial nerve; rdix, dorsal ramus of the glossopharyngeal nerve; rm, maxillary ramus of the maxillomandibular nerve; rthyr, thyroid ramus of the facial nerve; rvel, velar ramus of the maxillomandibular nerve; ulp, upper lip; vel, velum; vsr, ventral spinal nerves; x1, first vagal branchial branch; V₁, ophthalmicus profundus nerve; IX, glossopharyngeal nerve; X, vagus nerve; XII, hypoglossal nerve. Taken from Kuratani et al. (1997).

Spinal nerves and dorsal root ganglia: As in gnathostomes, lampreys also possess a series of neural crest-derived DRG located on both sides of the spinal cord, which first appear anteriorly at stage 24 in the intersomitic spaces and continue growing in number as the embryo elongates (Damas, 1944; Häming et al., 2011; Hirata et al., 1997; Nakao and Ishizawa, 1987). The spinal roots in lampreys have the peculiarity of being asymmetrical, where the right side occurs slightly caudal to the left side (Fritsch and Northcutt, 1993). Also, while dorsal roots are intersomitic, ventral roots are mid-somitic in position (Richardson et al., 2010).

1.4.4. The Lamprey Autonomic Nervous System

The lamprey autonomic nervous system is poorly understood. Autonomic fibres in the lamprey head exit the brain through cranial nerves VII, IX and X (Nilsson, 2011). In addition, a consensus is emerging that lampreys do not have a chain of sympathetic ganglia as in gnathostomes, although scattered ganglia along the dorsal aorta have been claimed (Häming et al., 2011; Nilsson, 2011). While they lack a complex sympathetic nervous system, lampreys have chromaffin-like cells along blood vessels, heart and cloaca (Funakoshi and Nakano, 2007; Häming et al., 2011). These cells have been considered analogous to postganglionic neurons (Häming et al., 2011). However, preganglionic fibres that exit the CNS directly innervate their target tissues forming a plexus (Häming et al., 2011; Johnels, 1956). In contrast to gnathostomes, where autonomic pathways exit the spinal cord via ventral roots only, in lampreys, autonomic axons leave the spinal cord via both dorsal and ventral roots (Johnels, 1956; Nilsson, 2011). Finally, lampreys have neural crest-derived enteric ganglia (Häming et al., 2011).

1.4.5. Lamprey Cranial Placodes

As in gnathostomes, lampreys have a full placodal repertoire (Table 1-2). They possess a well defined olfactory and adenohipophyseal placode, which develop from a common anlage known as the nasohypophyseal plate (NHP), morphologically visible at stage 23 (Kuratani et al., 2001; Uchida et al., 2003). This plate later splits into the olfactory placode, which eventually gives rise to the olfactory epithelium, and the adenohipophyseal placode, which develops into the adenohipophysis (Uchida et al., 2003). Lampreys also present a lens placode which develops into a lens as in

gnathostomes (Damas, 1944; Tahara, 1988; (Meléndez-Ferro et al., 2002). They also have a trigeminal placode marked by *pax-7* and an otic placode marked by *pax-2* that forms sensory neurons of the otic vesicle (McCauley and Bronner-Fraser, 2002). Lampreys also present a series of epibranchial placodes as in gnathostomes: facial (geniculate) placode, glossopharyngeal (petrosal) placode, and vagal (nodose) placode (Damas, 1944). There can be more than one vagal placode, one at the dorsoventral position in each posterior pharyngeal arch.

Lateral line placodes have not been described in lampreys as such, even though lampreys have an anterior and a posterior lateral line ganglion as well as an intracapsular ganglion associated with the anterior lateral line nerve (Koyama et al., 1990; Kuratani et al., 1997; Modrell et al., 2014; Ronan and Northcutt, 1987; Takio et al., 2007). This suggests the presence of at least an anterior and a posterior lateral line placode in lamprey embryos. Given that the facial and anterior lateral line ganglion form a single complex as they develop (Modrell et al., 2014), it is possible that the facial and a putative anterior lateral line placode also form a single complex. It is therefore unknown whether lampreys possess a set of lateral line placodes as occurs in zebrafish, *Xenopus* or *Ambystoma*. *Xenopus* and *Ambystoma* possess five lateral line placodes, namely, anterodorsal, anteroventral, middle, supratemporal and posterior lateral line placodes, having lost the otic lateral line placode (Schlosser, 2002; Schlosser and Ahrens, 2004; Schlosser and Northcutt, 2000). On the other hand, four lateral line placodes have been described in zebrafish, namely, anterodorsal, anteroventral, middle and posterior (Andermann et al., 2002). Larval and adult lampreys possess many well-developed lines of neuromasts in the head and body, forming a fully-functional lateral

line system (Gelman et al., 2007, 2008). This makes it reasonable to think that lampreys may have more lateral line placodes still to be identified.

The developmental origin of all placodes during lamprey development has not been elucidated as has been in –for example- zebrafish and *Xenopus*, where a general pan-placodal field has been described from which all placodes eventually segregate (Schlosser, 2006; Streit, 2007). Lineage-tracing, combined with expression patterns of a number of genes (for example, *six*, *eyes absent* and *dachshund* homologues), will allow establishing whether lamprey placodes also come from a common pan-placodal field similar to other vertebrates.

Placodes/Ganglia	Stage 21	Stage 22	Stage 23	Stage 24	Stage 25	Stage 26
Olfactory			Nasohypophyseal plate	Olfactory placode	Olfactory epithelium	Olfactory epithelium
Adenohypophyseal			Nasohypophyseal plate	Adenohypophyseal placode		
Lens	Future lens placode		Lens placode	Lens placode		
Trigeminal	Small thickening fused with otic placode	Future ophthalmic/profundal placode Trigeminal placode	Ophthalmic/profundal placode/ganglion Trigeminal ganglion	Both trigeminal ganglia nearly detached from ectoblast		
Otic	Small thickening fused with trigeminal placode/otic field marked by <i>Pax-2</i>	Otic placode	Otic vesicle	Otic capsule completely detached. Cells begin to delaminate to form the statoacoustic (VIIIth) ganglion	Condensing statoacoustic ganglion	Statoacoustic ganglion formed
Facial			Facial placode	Facial ganglion		
Glossopharyngeal	Glossopharyngeal placode					
Vagal			Vagal placode	Vagal placode/ganglion attached to ectoblast		
Lateral line	Lateral line placode					

Table 1-2. List of lamprey cranial placodes described to date. The main developmental steps per embryonic stage (Tahara, 1988) are indicated for each placode.

1.5. The Helix-Loop-Helix Gene Family

1.5.1. The bHLH Domain

The basic helix-loop-helix (bHLH) domain, as its name implies, is composed of a basic region and two helices separated by a loop. There are, however, certain proteins that lack the basic portion of this domain, such as Id and COE proteins. The basic region is the part of the domain that binds to DNA in the major groove, whereas the HLH region interacts with HLH domains of other HLH proteins forming a left-handed four-helix bundle, thus constituting HLH homo- or heterodimers (Bertrand et al., 2002; Chien et al., 1996; Longo et al., 2008). The bHLH domain is approximately 60 amino acids long with conserved residues that are specific for several or particular HLH subfamilies (Bertrand et al., 2002).

HLH homo- or heterodimers bind to DNA in specific orientations depending on the binding partner as well as on the distribution of phosphate backbone contacts with the DNA (Longo et al., 2008) (Figure 1-10). The core hexanucleotide DNA sequences bound by HLH dimers are different for each HLH gene subfamily (Bertrand et al., 2002). Certain heterodimers, specifically between tissue-specific HLH proteins in the nervous system with ubiquitous HLH proteins, are favoured with respect to homodimers of tissue-specific proteins as these are unstable and/or with low affinity to DNA (Aguado-Llera et al., 2010; Longo et al., 2008). bHLH domains have also been reported to change from a mainly unfolded to an α -helical form upon binding to DNA (Aguado-Llera et al., 2010; Meierhans et al., 1995). Additionally, specific residues of particular tissue-restricted HLH subfamilies make contact with specific parts of the DNA (bases

or the phosphate backbone) that are different from contacts by other HLH subfamilies, allowing only certain 3D conformations (Longo et al., 2008).

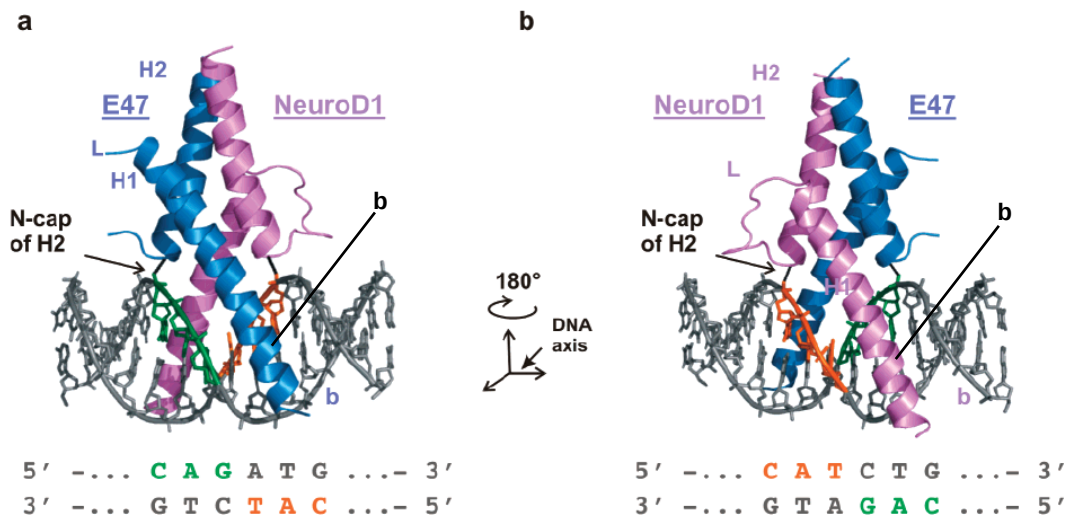


Figure 1-10. Structure of dimers between basic helix-loop-helix proteins. (a and b) Ribbon drawing of the ubiquitous E47 (blue) and the neurogenic NeuroD1 (violet) bHLH domains bound to DNA as heterodimers. In (b) the structure is rotated horizontally 180° with respect to (a). The hexanucleotide E box sequence is shown at the bottom of each structure rotated 180° accordingly. E47 interacts with the CAT half-site (orange) while NeuroD1 interacts with the CAG half-site (green). In both half-sites the amino terminal dipole of helix 2 is capped by the phosphate oxygen of amino acid Cys1. The basic (b), helix 1 (H1), loop (L), and helix 2 (H2) regions are labelled for E47 in panel (a) and for NeuroD1 in panel (b). Taken from Longo et al. (2008).

1.5.2. The Neural HLH Genes

The nervous system encompasses a wide variety of neuronal cell types. However, it is important to make a distinction between two concepts that sometimes are thought to be the same: neural and neuronal. In general terms, the nervous system is composed of two main cells types, neurons and glial cells (e.g. astrocytes, oligodendrocytes, Schwann cells). These cells can be thought of as “accessory” cells to neurons, although it is now well known that they display a wide variety of functions fundamental for the functioning of neurons. Therefore, the term “neural” refers to both neurons and glial

cells, whereas the term “neuronal”, as is obvious, refers to neurons only. Accordingly, the term “glial” refers only to glial cells.

In the present study, the term “neural HLH genes” will refer to a set of HLH gene subfamilies expressed mainly in the nervous system and that play an important role in neuronal and glial development. The genes that are here referred to as “neural HLH genes” are members of the *ASC*, *Neurogenin*, *NeuroD*, *Atonal*, *COE*, *Hairy/E(spl)*, *Hey*, *Oligo* and *NSCL*. However, this distinction is not intended to be strict as other HLH genes of other subfamilies might also be expressed in the nervous system.

One of the clearest examples of correlation between neural HLH gene function, their timing of activation along the neuronal differentiation programme, and their expression patterns is seen in cross-sections of the neural tube, particularly in the spinal cord. The so called pro-neural genes of the *ASC* and *Neurogenin* subfamilies are expressed in the ventricular zone where proliferating neuroblasts are located (Jasoni et al., 1994; Ma et al., 1996). Rat *Hes1* and chicken *Hes5-1*, *Hes5-2*, *Hes5-3* and *Hes6-2* are also expressed in the ventricular zone (Fior and Henrique, 2005; Sasai et al., 1992; Vilas-Boas and Henrique, 2010). Further HLH genes are restricted to the subventricular zone of postmitotic, differentiating neurons, and excluded from the ventricular and marginal zones such as *NSCL1* and *NSCL2* (Begley et al., 1992; Göbel et al., 1992; Li et al., 1999a; Perez et al., 1999; Theodorakis et al., 2002). Chicken *NSCL1* has also been suggested to be involved in preventing postmitotic neurons from reentering the cell cycle (Li et al., 1999a). Other HLH subfamilies are expressed in the subventricular and marginal zone of differentiating post-mitotic neurons such as chicken and mouse *Ebf1*

and *Ebf3* genes (Garcia-Dominguez et al., 2003; Garel et al., 1997a) and, in particular, in motor neurons (Garel et al., 1997a). These expression patterns are in agreement with the requirement of *COE/Ebf* genes in neuronal differentiation and migration towards the marginal zone, in close relationship with cell cycle exit (Garcia-Dominguez et al., 2003).

1.5.3. HLH Gene Networks in the Nervous System

The different neural HLH transcription factors are not only expressed in the nervous system, but they also act in concert for the differentiation of a particular neuronal lineage. In this way, they are expressed in a sequential pattern in the same tissue/organs and in most cases they establish transcription factor networks among themselves. These networks are fundamental for the development and differentiation of neurons and glial cells and their alteration can change the fate of a particular cell lineage. Networks of neural HLH transcription factors are found throughout the nervous system. For example, in the primary stripes of neurogenesis (dorsal, medial and lateral) in the neural plate of *Xenopus* embryos, a network led by *Xenopus neurogenin2* activates *coe2* and *neuroD*, which in turn activates *coe3* to induce neuronal differentiation (Pozzoli et al., 2001). *Xenopus coe2* also activates *neuroD* and appears to be involved in a positive feedback loop with *neurogenin2* (Dubois et al., 1998).

In the chicken spinal cord, *ngn2* (expressed in neuroblasts) and *NeuroM* (expressed in post-mitotic neurons) each induce both *Ebf1* and *Ebf3* (Garcia-Dominguez et al., 2003). Conversely, *Ebf1* induces expression of *ngn1* and *ngn2*, but not *Cash1*, a member of the ASC HLH gene subfamily (Garcia-Dominguez et al., 2003). Other

notable sites of neural HLH networks are, for example, in the olfactory epithelium and retina (Cau et al., 1997, 2002).

1.5.4. Notch Action Over Neural HLH Genes

In addition to the expression patterns and interaction networks of neural HLH genes in the nervous system, they are well-known to be regulated by the Notch signalling pathway (explained in detail in Chapter 8), and, together, they constitute one of the most prominent mechanisms that regulate the development of the entire vertebrate nervous system. This section thus presents a brief description of experimental evidence that reveals the role of the Notch signalling pathway in the regulation of expression of different vertebrate HLH gene subfamilies expressed in the neural tube, given that the neural tube is one of the best examples of neural HLH gene and Notch signalling activity. Subfamilies considered in this section include only those to which the genes cloned in the present study belong, namely, *Hairy/E(spl)*, *ASCL*, *Neurogenin* and *COE* subfamilies, as explained in subsequent chapters.

- ***Hes genes:*** In mouse, both *Hes1* and *Hes5* expression is downregulated in the spinal cord in *Notch1* mutants (de la Pompa et al., 1997; Yang et al., 2006) and upregulated when *Notch* is misexpressed (Ohtsuka et al., 1999). In chicken, Notch signalling upregulates expression of *Hes5-1*, *Hes5-2* and *Hes5-3*, but downregulates expression of *Hes6-2* (Fior and Henrique, 2005). This evidence suggests that *Hes* genes are generally upregulated by the Notch signalling pathway.

- **ASCL genes:** In mouse *Notch1*-null and *Notch1* conditional knock out (cKO) mutants, *Mash1* expression is upregulated in the midbrain, hindbrain and spinal cord (de la Pompa et al., 1997; Yang et al., 2006), suggesting that *ASCL* genes are generally downregulated by the Notch signalling pathway in the spinal cord.

- **Neurogenin genes:** Expression of *neurogenin1* and *neurogenin2* is upregulated in the spinal cord in mouse *Notch1* cKO mutants (Yang et al., 2006), suggesting that *neurogenin* genes are generally downregulated by the Notch signalling pathway in the spinal cord.

- **COE genes:** Of the two *Xenopus* *COE* genes identified so far, expression of *Xcoe2* is downregulated by the Delta/Notch signalling pathway at neural plate stages, whereas expression of *Xebf3* is insensitive to this pathway in the same time period (Dubois et al., 1998; Pozzoli et al., 2001).

1.6. Vertebrate Genome Duplications

When analysing large families such as the HLH gene family, especially in a vertebrate species, it is important to consider the evolution of the genome at the origin of vertebrates. In 1970, Ohno suggested for the first time that two genome-wide duplications might have occurred sometime in early vertebrate history, one duplication in the lineage leading to amphioxus and vertebrates, and a second genome expansion in fish or amphibians (Ohno, 1970). Subsequent molecular data (most prominently homeobox gene families including the Hox gene clusters) provided more evidence for two rounds of whole-genome duplications at the origin of vertebrates (Holland et al., 1994). The timing of these duplications, however, were suggested to have occurred one after the divergence of cephalochordates, but before the origin of lampreys, and the second after the divergence of lampreys close to the origin of jawed vertebrates (Holland et al., 1994). Regardless of the exact timing of these duplications, it is now generally accepted that two genome-wide duplications occurred before the gnathostome radiation and after the divergence of cephalochordates. This hypothesis is now known as the “2R hypothesis”.

A number of subsequent studies of gene phylogeny and genomic organisation were consistent with the 2R hypothesis (Dehal and Boore, 2005; Panopoulou and Poustka, 2005; Zhang and Cohn, 2008). With the sequencing of the amphioxus genome and analyses of synteny in chordate linkage groups, it was found that there is a quadruple conserved synteny between amphioxus and gnathostomes, providing more substantial evidence for the 2R hypothesis (Putnam et al., 2008). This study was followed by an extensive phylogenetic analysis with several cyclostome gene sequences and synteny,

which concluded that the two whole-genome duplications predated the cyclostome-gnathostome split (Kuraku et al., 2008). Sequencing of the lamprey genome has provided more evidence for two rounds of whole-genome duplications in the common ancestor of lampreys and gnathostomes (Smith et al., 2013). While it is still a matter of debate, accumulating evidence suggests that the 2R events occurred before the divergence of lampreys.

The cytogenetic basis for the two whole-genome duplications in the vertebrate stem has been proposed to have been by two rounds of tetraploidy (Furlong and Holland, 2002a). It has also been suggested that both genome-wide duplication events occurred in rapid succession after the divergence of cephalochordates, and that both events were preceded by an intense genome rearrangement, their occurrence being symptoms of existing genome instability (Hufton et al., 2008). Also, duplications were followed by a rapid diploidization event and by lower rates of genome rearrangement. Of note, another genome-wide duplication event occurred in the teleost lineage (Hufton et al., 2008; Semon and Wolfe, 2006).

1.7. Aims of the Project

This project is designed to study the development of the lamprey nervous system. An effective way of characterising the internal structure and developmental sequence of differentiation of the nervous system is by studying the HLH gene family of transcription factors (in particular the “neural HLH genes”), members of which are well-known to regulate the development of the nervous system in gnathostomes. The “neural HLH genes” not only mark specific regions or structures of the vertebrate nervous system, but they also act at different stages of neuronal differentiation, thereby representing a very useful tool for characterising the sequence of formation and differentiation of the lamprey nervous system.

The genes analysed in the present study were selected to study all major subdivisions of the vertebrate nervous system: central nervous system (brain and spinal cord), peripheral –sensory– nervous system (cranial sensory placodes/ganglia), and the autonomic nervous system (mainly sympathetic and enteric ganglia). In addition to this, given the importance and size of the HLH gene family in metazoans, studying the gene content and expression patterns of members of this family in lampreys adds substantial information to the evolution of the HLH gene family itself. To complete the picture, this project also analyses the role of one of the main pathways known to regulate the expression of the “neural HLH genes”, the Notch pathway, in the lamprey spinal cord.

Therefore, this project is composed of three main fronts: 1) the developmental/neurological front, 2) the genetic front, and 3) the signalling pathways front.

2. METHODS

2.1. Computer-based Research

2.1.1. Genome Databases

Predictions of genes of the HLH class of transcription factors from the lamprey *Petromyzon marinus* were extracted from the available genome sequences in the Ensembl (version 7) website for lamprey. Contigs were retrieved by blasting the HLH domain amino acid sequences of one member of each of the human HLH families and the full set of *Branchiostoma floridae* HLH proteins against the lamprey genome. The human and amphioxus HLH domain amino acid sequences were obtained from (Simionato et al., 2007). Gene predictions containing a HLH domain were isolated and independently blasted specifically against the human genome to find the best hit with human HLH proteins. Once predicted sequences containing a HLH domain were isolated, predictions with high sequence similarity to gnathostome HLH families known to be expressed in the nervous system such as *Neurogenin*, *NeuroD*, *Atonal*, *ASC*, *Hes* and *COE* gene families were identified for subsequent gene cloning.

2.1.2. Sequence Analysis and Manipulation

Sequence analysis and manipulation were performed using Bioedit software (Hall, 1999). Multiple sequence alignments were carried out using MAFFT software version 6.864b for windows (Katoh et al., 2002; Katoh and Toh, 2008). The parameter for strategy for MAFFT alignments was set as “auto”. All other parameters were unchanged

and used as default. Depending of each alignment, sequences were trimmed in Bioedit and used for phylogenetic analyses.

2.1.3. Phylogenetic Analyses

For phylogenetic tree construction, the Neighbor Joining (NJ) algorithm was used and tree construction was conducted using MEGA version 5 (Tamura et al., 2011). For all NJ trees, the Jones-Taylor-Thornton (JTT) amino acid substitution matrix was employed, gamma distribution with 4 categories, and 1000 bootstrap replicates to obtain support values at each node. Maximum likelihood and Bayesian inference algorithms were also used, but given that trees by these methods had very low support values and were highly polytomic (therefore not very informative), only NJ trees were considered in the present study.

2.2. Experimental Research

2.2.1. Animal Collection

Embryos from the brook lamprey *Lampetra planeri* were collected from a shallow river in the New Forest National Park with permission from the Forestry Commission. The brook lamprey spawns every year in April, and the animals can be easily seen spawning in transitions of slow and rippling water in the shallow-most areas of the river. Lampreys build nests in these transitions by removing small stones to make a slight depression where they lay and fertilise the eggs. Eggs were collected by digging at the bottom and surrounding areas of the nest; this allows the eggs to float which can

then be easily collected with a trapping net. Quite often, nests could be identified just by looking at depressions of small stones even if there were not any adults swimming around.

After collecting, embryos were brought to the laboratory and placed in Petri dishes with filtered river water from the same river where they were caught. They were kept at 13-15 °C and later fixed (see below) at different stages of development following the staging system of Tahara (1988). When necessary, embryos were dechorionated with fine forceps before fixation.

2.2.2. Fixing Embryos

Embryos were fixed in 4% PBS-buffered paraformaldehyde (PFA). To prepare 50 ml, 2 g of PFA were weighed in a 50 ml Falcon tube, to which 35 ml of water were added and 2 ml of 1M NaOH. The mixture was then heated at 65 °C in a water bath until all PFA dissolved with occasional shaking by hand. Once dissolved, 2 ml of 1 M HCl and 5 ml of 10X PBS (1.37M NaCl, 27mM KCl, 100mM Na₂HPO₄, 17.6mM KH₂PO₄) were added. The pH was then measured with a pH strip and adjusted if necessary with 1M NaOH or 1M HCl until set to ~pH7.5. PFA was cooled on ice before use and stored at 4 °C for short-term storage or at -20 °C for long-term storage in 14 ml aliquots. When frozen at -20 °C, PFA was thawed only once and then kept at 4 °C but never re-frozen. Embryos were fixed in a ~10X excess volume of 4% PFA-PBS with respect to river water at 4 °C overnight or longer.

After fixation, embryos were washed twice in DEPC-treated 1X PBS for 10 minutes each, and then dehydrated through a graded series of 1X PBS:Methanol (i.e. 25%, 50%, and 75% of Methanol in 1X PBS) once for 10 minutes each. Finally, they were washed twice in 100% methanol for 10 minutes each and stored in fresh methanol at -20 °C.

2.2.3. RNA Extraction

Total RNA was extracted from stages 17-20, 21-23, 24-26 and 28-30 embryos. Approximately 20 embryos per stage range were placed in 1.7 ml eppendorf tubes, then all river water removed, and quickly added 400 µl of TRI-reagent. Samples were stored at -70 °C immediately. To extract RNA, chloroform was added in a 1:5 proportion, i.e. 80 µl per 400 µl of TRI-reagent. Tubes were then shaken vigorously for 15 seconds, then incubated at room temperature (RT) for 15 minutes, and centrifuged at 12,000 Gs for 15 minutes at 4 °C. The aqueous phase that forms at the top was then transferred to a clean tube, then 200 µl of isopropanol were added and vortexed for 10 seconds. The mixture was then incubated at room temperature (RT) for 10 minutes and centrifuged at 12,000 Gs for 8 minutes at 4 °C. All supernatant was discarded and 800 µl of 75% ethanol were added to the pellet. This was then centrifuged at 7,500 Gs for 5 minutes and then the ethanol removed. This step was repeated to remove any residual ethanol. The pellet was air-dried for 5 minutes and resuspended in 50 µl of deionised water.

2.2.4. cDNA Synthesis

All extracted RNA was used to obtain different cDNA pools by reverse transcription. cDNA was synthesised using SuperScriptTM III reverse transcriptase from

Invitrogen following the manufacturer's instructions. These pools were employed to perform subsequent PCR reactions.

2.2.5. Polymerase Chain Reaction

Genes were isolated by the polymerase chain reaction (PCR) using either genomic DNA (gDNA) or complementary DNA (cDNA) as templates. Standard PCR reactions were carried out starting with the following reaction mixtures:

a) For PCR using gDNA:

Reagent	Final concentration	Quantity, for 100µl of reaction mixture
Primer 1 (10 µM)	0.1 µM	1 µl
Primer 2 (10 µM)	0.1 µM	1 µl
25 mM dNTP mix	0.2mM each	0.8 µl
10X PCR buffer	1X	10 µl
5u µl ⁻¹ <i>Taq</i> DNA Polymerase	5u/100µl	1 µl
50mM MgCl ₂	1.5 mM	3 µl
Sterile deionized water	-	83.2 µl
Total	-	100 µl

This mix was then aliquoted in 20 µl samples and 0.5 µl of gDNA were added per reaction to obtain a final working volume of 20.5 µl.

b) For PCR using cDNA:

Reagent	Final concentration	Quantity, for 100µl of reaction mixture
Primer 1 (10 µM)	0.1µM	1 µl
Primer 2 (10 µM)	0.1µM	1 µl
25 mM dNTP mix	0.2mM of each	0.8 µl
10X PCR buffer	1X	10 µl
5u µl ⁻¹ <i>Taq</i> DNA Polymerase	5u/100µl	1 µl
50mM MgCl ₂	1.5 mM	3 µl
Sterile deionized water	-	68.2 µl
Total	-	85 µl

This mix was then aliquoted in 17 µl samples and 3 µl of cDNA were added per reaction to obtain a final working volume of 20 µl.

2.2.6. PCR Programme Strategies Used to Clone Each Gene:

- For *Lampetra planeri* *Hairy/E(spl)* genes:

Fragments of two *Hes* homologues, named *LpHes-A* and *LpHes-B*, were isolated from *L. planeri* genomic DNA. Isolation of these sequences was based on two full-length *P. marinus* *Hes*-homolog cDNA sequences kindly provided by Dr. Tatjana Sauka-Spengler. Primers were designed based on these cDNA sequences delimiting a region of a single exon contained in a single contig from the *P. marinus* Ensemble genome database. The primers used are:

Hes-A sense primer: 5'-GTCAACGTCCAAACTGGAGAAGGC-3'

Hes-A antisense primer: 5'-CCACTGTACAACATTCATCGCTTGCAC-3'

Hes-B sense primer: 5'-CCCCGCTGCCCACGGCAA-3'

Hes-B antisense primer: 5'-GCTTTTGTGAGACATTGGCTTTTATTGACATTC-3'

The PCR programme used was:

<u><i>LpHes-A</i></u>		<u><i>LpHes-B</i></u>
94 °C 4 minutes		94 °C 4 minutes
94 °C 1 minute	} 35 cycles	94 °C 1 minute
59.7 °C 1 minute		62.5 °C 1 minute
72 °C 1:30 minutes		72 °C 1:30 minutes
72 °C 5 minutes		72 °C 5 minutes

- For *Lampetra planeri* *LpNgn*:

A fragment of a gene homologous to vertebrate *neurogenin* genes, *LpNgn*, was isolated by RACE-PCR. Primers were designed using an initial sequence of 99 bp kindly provided by Dr. Rebecca Furlong as a template. Different primer combinations were tested, both for 5' and 3' RACE-PCR. The primer combination designed for 5' RACE-PCR that amplified the isolated *LpNgn* fragment was:

LpNgn-GSP1-rev: 5'-GTTGTAGGCGAAGCGCAGCGTCTCGATCT-3'

LpNgn-NGSP1-rev: 5'-TGGTGAGCCGAGCGTCGTCCGGCAGGGT-3'

An initial PCR reaction using the *LpNgn*-GSP1-rev primer was carried out with the following programme:

5 cycles of:	5 cycles of:	25 cycles of:
94 °C 30 seconds	94 °C 30 seconds	94 °C 30 seconds
57 °C 30 seconds <i>then</i>	55 °C 30 seconds <i>then</i>	53 °C 30 seconds
72 °C 3 minutes	72 °C 3 minutes	72 °C 3 minutes

After this first round of PCR, RACE products were diluted 1:49 in Tricine-EDTA buffer and used for a second round of PCR (nested RACE-PCR) using the *LpNgn-NGSP1*-rev primer following the programme:

20 cycles of:

94 °C	30 seconds
53 °C	30 seconds
72 °C	3 minutes

In this way, a fragment of 654 bp containing the 5' UTR and part of the coding sequence could be isolated for further analysis (see Chapter 5).

- For *Lampetra planeri* *LpCoe-A* and *LpCoe-B* genes:

In order to isolate *L. planeri* COE homologs, the amphioxus COE amino acid sequence was blasted against the *Petromyzon marinus* genome by using the Pre!Ensemble website (now replaced by Ensembl version 7 for *Petromyzon marinus*). Contigs of different sizes with high sequence similarity to different parts of COE proteins were retrieved. Small different fragments of two COE-like genes could be identified in different contigs. However, these fragments could not be assembled into an

entire sequence. Therefore, gene-specific primers at the start of the DNA-binding domain (DBD, see Chapter 6) and downstream of the HLH domain of each gene fragment were designed to perform PCR with a cDNA pool from stage 24-26 *L. planeri* embryos. Different combinations of these primers allowed the isolation of two cDNA fragments of 1155 bp and 1090 bp, named *LpCoe-A* and *LpCoe-B*, respectively, and their corresponding predicted amino acid sequences are of 385 and 363 amino acids, respectively.

The PCR programme used for both fragments was:

94 °C	4 minutes	
94 °C	30 seconds	} 35 cycles
53 °C	30 seconds	
72 °C	1:30 minutes	
72 °C	5 minutes	

- For *Lampetra planeri* *LpAsh*:

From gene predictions identified in Chapter 3, a predicted sequence containing a HLH domain with high similarity to vertebrate *ASCL* homologues was identified. Primers were designed within the predicted open reading frame and a fragment of 472 bp was isolated using gDNA as a template. Its predicted amino acid sequence is 157 amino acids long. The primers used to clone this fragment are:

LpAsh sense primer: 5'-CACGAGTACGCGGGCTTCGA-3'

LpAsh antisense primer: 5'-TGCACACACAAGGAGCCATCCC-3'

The PCR programme to isolate *LpAsh* was:

94 °C	4 minutes	
94 °C	1 minute	} 35 cycles
59.7 °C	1 minute	
72 °C	1:30 minutes	
72 °C	5 minutes	

2.2.7. Cloning PCR Products

PCR products were run on an agarose gel and visualised with ethidium bromide under UV light. Bands of the correct size were excised using a clean razor blade and purified with the QIAEX II gel extraction kit from QIAGEN via their protocol. Extracted bands were subsequently ligated into pCRII vector from the Dual Promoter TA Cloning Kit from Invitrogen with the Rapid DNA Ligation Kit from Roche. Ligations were carried out with a modified version of the Rapid DNA Ligation Kit protocol. First, a mix of the PCR product, pCRII vector (Invitrogen), T4 DNA Ligation Buffer (Roche), DNA Dilution Buffer (Roche) and T4 DNA Ligase (Roche) was made as follows:

pCRII vector	0.5 µl
DNA Dilution Buffer	2 µl
T4 DNA Ligation Buffer	5 µl
DNA band	3 µl
Ligase	1 µl
Total	11.5 µl

This ligation mix was incubated at room temperature (RT) for 15 minutes to allow the DNA band integrate into the pCRII vector. Resultant ligated constructs with the PCR insert were then transformed into NovaBlue competent cells from Novagen and selected by blue-white screening with ampicillin as antibiotic. First, a mixture of 50 μ l of X-gal (5-Bromo-4-chloro-3-indolyl- β -D-galactoside) from a 20 mg/ml stock solution, and 10 μ l of IPTG (Isopropylthio- β -D-galactoside) from a 200 mg/ml stock solution, was spread onto Luria Agar plates (LB-Amp plates) already made with 0.1% ampicillin of an initial 50 mg/ml stock solution. Plates were dried at 37 °C for ~30 minutes. Later, 6 μ l of the ligation mix were mixed with 25 μ l of competent cells very gently. This cell-ligation mix was placed on ice for 5 minutes, then heat-shocked at 42 °C for 30 seconds, and placed back on ice for 5 minutes. 80 μ l of 37 °C pre-warmed SOC medium from Novagen were then added to the transformed cells, which later were spread onto the above-mentioned agar plates. These plates were in turn incubated at 37 °C overnight and examined for positive white colonies following the sequence structure of the pCRII plasmid, information available from Invitrogen.

Identified positive white colonies were selected and transferred to an LB-Amp masterplate with the aid of a sterile toothpick. This toothpick was then dropped into a 15 ml tube containing 3 ml of LB Broth from Sigma with 0.1% ampicillin. The tube was incubated at 37 °C overnight with shaking.

To obtain a pellet of cultured transformed bacteria, 1.5 ml of the overnight culture was centrifuged at 6,000 rpm for 3 minutes and the supernatant removed. Plasmids from transformed bacteria were purified using the QIAprep spin miniprep kit from QIAGEN

following the supplied protocol. Once purified, plasmids were digested with EcoRI restriction enzyme and run on an agarose gel to test whether they contained the PCR insert. Plasmids showing a band of the expected size according to the length of the insert were subsequently sequenced.

2.2.8. Sequencing

Sequencing was performed on an ABI PRISM 3100, a fluorescence-based DNA analysis system.

2.2.9. *In situ* Hybridisation

Most of the solutions employed for *in situ* hybridisation were treated with diethyl dicarbonate (DEPC) to destroy RNases by adding it at 0.1% concentration and left overnight, followed by autoclaving to destroy the DEPC. Solutions treated with DEPC are indicated below.

2.2.10. Riboprobe Synthesis

After plasmid sequencing, inserts were analysed to identify their 5' and 3' ends – therefore their orientation– with respect to the plasmid and its promoter sequences. Before probe synthesis, the insert was amplified by PCR using plasmids derived from minipreps as templates. The same standard initial mix specified in section 2.2.5 for genomic DNA was prepared using M13 reverse and M13 forward primers (located upstream and downstream of the polylinker region, respectively; more information

about the polylinker region sequence of the pCRII vector can be found in the Invitrogen website), and aliquoted in 50 μ l per probe. Minipreps were diluted 1:2000 in deionised water, and 2 μ l of this dilution were added to each 50 μ l aliquot. All samples were submitted to the following PCR programme:

Step	Temperature in °C	Time
1. Initial denaturation	94 °C	4 minutes
2. Denaturation	94 °C	30 seconds
3. Annealing	53 °C	30 seconds
4. Elongation	72 °C	90 seconds
Repeat steps 2-4 35 times		
5. Final elongation	72 °C	5 minutes

1 μ l of PCR products was later run on an agarose gel to visualise the quality of the bands. Successful reactions were then used for riboprobe synthesis.

To synthesise probes, T7 and Sp6 RNA polymerases from Roche were chosen to synthesise either sense or antisense riboprobes according to insert orientation. These polymerases bind to their respective promoter sequences located inside the region flanked by the M13 promoters in the plasmid sequence. A starting reaction was made as follows:

DNA fragment direct from PCR	10 µl
10X RNA polymerase buffer	2 µl
0.1 M DTT	2 µl
Dig UTP labelling mix	2 µl
RNAse inhibitors	1 µl
RNA polymerase (T7 or Sp6)	1 µl
PCR water	2 µl
Total	20 µl

This mix was later incubated at 37 °C for 2-4 hours, and then placed on ice to stop the reaction. 1 µl of the final synthesis was run on an agarose gel to visualise the integrity of the riboprobes. Successfully synthesised riboprobes were then precipitated with a mix of 100 µl of DEPC-water, 10 µl of 4M LiCl, and 300 µl of 100% Ethanol overnight at -20 °C. After precipitation, samples were centrifuged at 13,000 rpm for 10 minutes, then 200 µl of 70% Ethanol were added, and centrifuged again at 13,000 rpm for 5 minutes. The resulting supernatant was removed, and pellets were air-dried for 10 minutes and resuspended in 50 µl of DEPC-water. Precipitated riboprobes were also checked by running 1 µl on an agarose gel and were stored at -70 °C in 10 µl aliquots.

2.2.11. Permeabilisation and Hybridisation

Embryos were first treated with a mixture of hydrogen peroxide and methanol in a 1:1.5 ratio overnight, and later washed three times with PBS-DEPC for 5 minutes each rocking. Embryos were then treated with 0.2 N hydrochloric acid in PBS-DEPC for 10 minutes static, followed by a treatment with 5µg/ml of Proteinase K (Roche) in PBS-

DEPC, static, at 37 °C for 18–19 minutes depending on the stage. Samples were post-fixed with 4% PFA, 0.2% glutaraldehyde in PBS-DEPC for 20 minutes on ice, then washed three times with PBS-DEPC 10 minutes each rocking at room temperature, and twice in hybridisation buffer (50% formamide, 5X SSC pH 7.0, 0.1% tween 20, 0.1% CHAPS, 1 mg/ml total yeast RNA, 100 µg/ml heparin in DEPC-H₂O) 10 minutes each rocking at 65 °C. After that, embryos were prehybridised overnight at 65 °C in hybridisation buffer rocking, and later incubated with 600-700 ng/ml of DIG-labeled sense or antisense riboprobe in pre-warmed hybridisation buffer at 65 °C for 48-60 hours.

2.2.12. Washing

After hybridisation, embryos were washed twice with wash A (50% Formamide, 5X SSC pH 4.5, 1% SDS in DEPC-H₂O) for 30 minutes each at 65 °C, then substituted gradually with 50% and 75% TST (10 mM Tris-HCl pH 7.5, 0.5 M NaCl, 0.1% Tween 20) in wash A 10 minutes each, and twice with 100% TST for 10 minutes each at 65 °C. RNase A was then added at a 0.05 mg/ml concentration in TST for 30 minutes rocking at room temperature, followed by washes with wash B (50% Formamide, 2X SSC pH 4.5), wash C (2X SSC pH 4.5, 0.3% CHAPS), and wash D (0.2X SSC pH 4.5, 0.3% CHAPS) rocking at 65 °C for 30 minutes each twice. For immunological detection, samples were blocked with heat-treated sheep serum (incubated at 65 °C for 30 min) combined with TST in a 4:1 ratio overnight at 4 °C, followed by incubation with anti-digoxigenin-AP Fab fragments antibody from Roche overnight at 4 °C. Embryos were later washed with TST five times for 60 minutes each at RT.

2.2.13. Coloring

After being washed, embryos were transferred to sterile 4-well Nunc dishes with 5 embryos per well and added 6 µl/ml of NBT (Nitro blue tetrazolium chloride)/ BCIP (5-Bromo-4-chloro-3-indolyl phosphate) stock solution from Roche in NTMT buffer (100 mM NaCl, 100 mM Tris-HCl pH 9.5, 50 mM MgCl₂, 1% Tween 20) to detect alkaline phosphatase activity. Embryos were left in NBT/BCIP for 2-3 days depending on the gene and probe length. Once specific staining had developed, embryos were washed with PBT (1X PBS-DEPC, 0.1% tween 20) three times for 10 minutes each at room temperature, then fixed overnight with 4% PFA in PBS-DEPC at 4 °C, and washed again with PBT three times 10 minutes each at room temperature. Finally, embryos were cleared with a graded series of glycerol 25%, 50%, 75% in PBS-DEPC, and 99% at room temperature, static. They were stored in 99% glycerol at 4 °C indefinitely.

2.2.14. Visualisation of Signal and Photography

To observe the embryos under the microscope, they were first mounted on a glass microscope slide immersed in 99% glycerol and observed with a Carl Zeiss Axioskop 2 plus optic microscope. Pictures were taken with a Carl Zeiss AxioCam HRc camera with the microscope reflector turret set to 2, and Differential Interference Contrast (DIC) set to II. All images were processed using the AxioVision version 4 software.

2.2.15. Sectioning

Embryos were embedded in gelatine/albumin (0.45% gelatine, 25% albumin, 20% sucrose in PBS) and fixed with 2.5% glutaraldehyde. Embedded embryos were then sectioned at 50 μm thickness using a vibratome.

2.2.16. Drug Treatments

Live embryos were treated with DAPT (N-[N-(3,5-Difluorophenacetyl-L-alanyl)]-S-phenylglycine t-Butyl Ester), a drug that inhibits the Notch signalling pathway (see Chapter 7). After collecting, embryos were placed in Petri dishes with filtered water from the same river where they were collected from, and kept in an incubator at 13-15 °C. Embryos that reached developmental stage 24 according to Tahara (1988) staging classification were placed in 4-well Nunc dishes and added 100 μM DAPT (from a 10 mM stock solution dissolved in DMSO) in filtered river water. All embryos were allowed to develop until control embryos treated with DMSO only reached stage 26 at 13-15 °C. DMSO in control embryos was added at the same concentration as with DAPT-treated embryos. Once control embryos reached stage 26, all embryos were fixed immediately in a ~10X excess volume of 4% PFA-PBS at 4 °C overnight or longer. They were later washed and dehydrated in the same way as with normal embryos as specified in section 2.2.2. for further *in situ* hybridisation.

3. THE LAMPREY REPERTOIRE OF HLH GENES

3.1. Introduction

The HLH gene family encodes transcription factors that have important roles in metazoan development (Ledent and Vervoort, 2001; Simionato et al., 2007). In most members of this family, the HLH domain is preceded by a basic region and together they constitute the basic HLH (bHLH) domain which is conserved across metazoans. This family is also present and widely diversified in plants and fungi (Carretero-Paulet et al., 2010; Sailsbery et al., 2012). Based on their sequence similarity, metazoan HLH genes have been classified into higher-order groups, namely, groups A, B, C, D, E and F (Atchley and Fitch, 1997; Ledent and Vervoort, 2001; Ledent et al., 2002; Simionato et al., 2007).

Groups A and B include bHLH proteins that bind to hexameric core DNA sequences referred to as “E boxes” (CANNTG). Group A includes tissue-specific bHLH proteins such as MyoD, Twist, Achaete-Scute, and the ubiquitously expressed E12/Daughterless-type transcription factors, all of which bind to CACCTG or CAGCTG core sequences. The tissue-specific homodimers are frequently inactive, and require the formation of heterodimers with the E12/Daughterless proteins to initiate the transcriptional activation of the target gene. Group B includes bHLH proteins with different functions, for example Myc, Max, USF, SREBP, MITF, that bind to CACGTG or CATGTTG core DNA sequences. Some group-B proteins contain an additional motif known as the Leucine Zipper also involved in protein dimerisation. Group C includes proteins that, in addition to the bHLH domain, contain an additional domain known as

the “PAS domain” and bind to ACGTG or GCGTG core sequences. The PAS domain is ~260-310 amino acids long and allows the dimerisation between PAS proteins, the binding of small molecules (e.g., dioxin), and interactions with non-PAS proteins. The bHLH-PAS proteins are involved in neurogenesis, tracheal and salivary duct formation, toxin metabolism, circadian rhythms, and response to hypoxia (Ledent and Vervoort, 2001).

Group D corresponds to HLH proteins that lack the basic domain and are hence unable to bind DNA (Ledent and Vervoort, 2001; Simionato et al., 2007). This group includes the Id and Extramacrochaete (Emc) proteins, which act as antagonists of group A bHLH proteins possibly by binding to, and sequestering the ubiquitously expressed “E-box proteins” like E12/47 (Liu and Harland, 2003). This in turn prevents the heterodimerisation of E proteins with tissue-specific bHLH proteins and in this way they inhibit differentiation and promote proliferation throughout development (Meulemans et al., 2003). Group E includes proteins of the Hairy/E(spl) and Hey subfamilies. These proteins bind preferentially to core DNA sequences (CACGCG or CACGAG) referred to as “N boxes” and are generally involved in the inhibition of neurogenesis by suppressing expression of group A genes (Ledent et al., 2002; Simionato et al., 2007). Group-E proteins also contain two characteristic domains in addition to the bHLH, the “Orange” domain and a “WRPW” (Hairy/E(spl)) or “YRPW” (Hey) motif in their C-terminal part. Group F corresponds to the COE family proteins, which lack the basic domain but instead possess the COE DNA-binding domain (DBD) involved both in dimerisation and in DNA binding (Ledent et al., 2002; Simionato et al., 2007).

3.2. Aims

In this section I intend to characterise the diversity of HLH genes in the lamprey genome. The analyses in this section are targeted to answer the following questions:

- How many HLH genes do lampreys have?
- Do they have a similar repertoire of HLH gene subfamilies as in gnathostomes?
- Do they have HLH gene subfamilies belonging to all higher-order groups, i.e. groups A-F?
- Do they have members of HLH gene subfamilies considered neural in this study?

3.3. Hypothesis

The HLH gene family is widespread in the animal kingdom with a repertoire of major groups (A, B, C, E and F) already present in basal metazoans such as sponges and cnidarians (Simionato et al., 2007). Besides, in deuterostomes, the sea urchin and the cephalochordate *Branchiostoma floridae* possess a nearly complete set of all HLH gene subfamilies identified so far in metazoans and, in particular, in the vertebrate lineage (the exception being the urochordate *Ciona intestinalis* that has suffered a considerable loss of many HLH subfamilies). Therefore, it is predicted that lampreys, as in other vertebrates such as human, mouse, chicken, *Xenopus*, zebrafish, and other deuterostomes such as the sea urchin and *Branchiostoma floridae*, also possesses a similar wide repertoire of HLH gene subfamilies. The number of members of each HLH subfamily and of the entire HLH family as a whole is also predicted to be larger than non-vertebrate deuterostomes but lower or similar than gnathostomes considering the two genome-wide duplications at the origin of vertebrates.

3.4. Objectives

- To blast members of each HLH subfamily in human against the *Petromyzon marinus* genome in the Pre!Ensembl and Ensembl website.
- Collect all sequences of predicted models containing a putative HLH domain from the BLAST searches.
- Perform phylogenetic analyses to identify which predicted models group with particular HLH gene subfamilies.
- Identify lamprey homologs of gnathostome HLH genes of subfamilies with an important role in the development of the nervous system (i.e. *Hairy/E(spl)*, *ASC*, *Atonal*, *Neurogenin*, *NeuroD*, *COE*, *Oligo*, *NSCL*) for further cloning in *Lampetra planeri*.

3.5. Results

Sequences of the HLH family of transcription factors from the sea lamprey *Petromyzon marinus* were extracted from the publicly available genome database Pre!Ensembl and Ensembl version 7 for this species. Models were retrieved by blasting the HLH domain amino acid sequences of at least one member of each of the human HLH subfamilies. Human HLH domain amino acid sequences were obtained from (Simionato et al., 2007). In this way, 51 models containing a predicted HLH domain were retrieved (Table 3-1). Gene prediction amino acid sequences containing a HLH domain were then independently blasted against the human genome to find the best hit with human HLH proteins (Table 3-1). Only curated sequences were considered for best hits.

Lamprey sequences were later compared with already published HLH domains of human (Simionato et al., 2007), mouse (Ledent and Vervoort, 2001), rat (Zheng et al., 2009), chicken (Wang et al., 2009 – supplementary material; Liu and Zhao, 2010), zebrafish (Wang et al., 2009 – supplementary material) and sea urchin (Simionato et al., 2007). Comparisons were made for each HLH subfamily separately. For lamprey sequences, the best BLAST hit with human (Table 3-1) was used as a guide for comparison by each HLH subfamily. Lamprey sequences were then trimmed according to HLH domains of the above-mentioned species for each subfamily to obtain predicted amino acid sequences of predicted lamprey HLH domains. It is important to note that these lamprey HLH domain predictions are solely based on alignments; these predictions are not based on structural analyses.

Table 3-1 (see next page for table). Predicted *Petromyzon marinus* bHLH domains. First column: predicted amino acid sequences of predicted bHLH domains obtained from Pre!Ensembl and Ensembl genome databases for *Petromyzon marinus*. This table is based mostly on predictions obtained from the current Ensembl assembly, complemented with some predictions from the previous assembly in Pre!Ensembl for members of subfamilies for which no prediction model was found in the current assembly, but were predicted in the previous assembly. Second column: accession numbers from the current assembly. Gene prediction numbers from the previous Pre!Ensembl assembly are in parentheses. Third column: best (curated) hit of each lamprey model against human. Fourth column: based on BLAST searches and phylogenetic analyses (see below), shown are gene subfamilies to which the identified *P. marinus* HLH gene predictions belong to. Note that in the present study *Dec* genes are considered as an independent subfamily from the *Hairy/E(spl)* and *Hey* subfamilies (discussed in detail in section 4.6.3). Fifth column: Bootstrap values obtained from the tree in Figure 3-2 (see below) supporting the inclusion of each lamprey model in respective HLH subfamilies. Sixth column: higher-order group (A-F) to which each subfamily belongs to.

bHLH domain protein sequence*	Ensembl accession code	Best –curated- hit with human	Gene family	Support value	Group
RRVKANARERNRMHGLNAALDNLRKVVPCYSKTQKLSKIETLRLAKNYIWALS	ENSPMAP00000008309	NeuroD1 (NP_002491.2)	NeuroD	85	A
RRLKANARERNRMHGLNAALDGLRRVVPCHSRTQKLSKIETLRLARNYIGALS	(98653)	NeuroD1 (NP_002491.2)	NeuroD	85	A
RRMKANDRERNRMHNLNDALDALRGVLPPLDDARLTKEITLRFAYNYIWVLT	ENSPMAP00000006813	Neurogenin2 (NP_076924.1)	Neurogenin	98	A
RRLESNERERQRMHGLNDAFQGLREAIPHVRRGRKLSKLETTLAKNYIEALT	(75235)	Mist1 (NP_803238.1)	Mist	96	A
RRLAANARERRRMQGLNVAFERLRQVVPQWGQDRKLSKYETLQMAISYIAALS	(39660)	ATOX1 (NP_660161.1)	Atonal	92	A
RERQRTQSLNEAFSSLRQIIPTLPSDKLSKIQTLLKATRYIDFLY	ENSPMAP00000011396	Twist2 (NP_476527.1)	Twist	97	A
QRFVANVRERQRTQSLNEAFASLRQIIPTLPSDKLSKIQTLLKATRYIDFLY	ENSPMAP00000004354/ ENSPMAP00000004357	Twist2 (NP_476527.1)	Twist	97	A
QRVLANVRERQRTQSLNEAFASLRKIIPTLPSDKLSKIQTLLKASRYIDFLY	ENSPMAP00000011107	Twist2 (NP_476527.1)	Twist	97	A
LRQAANQRERRRMQSINAAFEGLRAHIPTLPYEKRLSKVDTLRLAIGYISFLG	ENSPMAP00000004055	PTF1A (NP_835455.1)	PTFb	64	A
QRTAANVRERRRMLSINSAFEELRCHVPTFPYEKRLSKIDTLRLAIAIYIALLR	ENSPMAP00000000075	PTF1A (NP_835455.1)	PTFb	64	A
QRRANARERDRTQSVNSAFTALRTLIPTEPADRKLSKLETLASSYIAHLG	ENSPMAP00000007312	Paraxis (NP_004600.2)	Paraxis	90	A
QRNAANARERARMRVLSKAFSRLKTTLPWVPADTKLSKLDTLRLASSYIGHLR	ENSPMAP00000000421	TCF21 (NP_003197.2)	MyoRa	95	A
KESSGANRERSVRAMRAAFLALQKSLPAVPPDTKLSKLDVLLLASAYIAHLS	ENSPMAP00000010377	TCF23 (NP_786951.1)	MyoRb	98	A
RRIFTNSRERWRQNVNGAFSELRRLLPHTPPDRKLSKNEILRLAMRYIDFLV	ENSPMAP00000002458	TAL1 (NP_003180.1)	SCL	97	A
RRARKNEREKLRMRLKTRALRALQGILPPHLAAVGRPLSKIQTLLRTIRYIAQLS	(93980/31380)	Mesogenin1 (NP_001099039.1)	Mesp	43	A
FIRKRNERRERQVKCVNEGVARLRDHLDPDEFSDKRLSKVETLRAAIRYIKHLQ	ENSPMAP00000008927	ASCL4 (NP_982260.2)	ASCb	98	A
RRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAQAVILSLE	ENSPMAP00000001063	TCF12 (NP_996923.1/NP_996919.1)	E12/E47	98	A
RRRNHNFLERQRREDLKRSFRALRDEVPELAGNEKAAKVILRKAESAAT	ENSPMAP00000004198	N-myc (NP_005369.2)	Myc	86	B
RRRTHNILERQRRDGLRSSFVTLRDSVPELRANERAAKVILRKAELARSLG	ENSPMAP00000003372	N-myc (NP_005369.2)	Myc	86	B
RRRTHNILERQRRDGLRSSFLGLRDVPELCRNDKAAKVILRKAELARSLV	ENSPMAP00000009801	N-myc (NP_005369.2)	Myc	86	B
KRRTHNILERQRRNDLKNSFFWLRDHIPELAHNDKAAKVILKKAEMEYSRTLQ	ENSPMAP00000011193	N-myc (NP_005369.2)	Myc	86	B
KRAHHNALERKRRDHIKDSFHSRLRDSVPSLQGEKVNKKASRAQILNKATEYIQYMR	ENSPMAP00000010427	MAX (NP_660087.1)	Max	99	B
KRAHHNALERKRRDHIKDSFHGLRDSIPAIQGEKVCHRASRALILNKATDYIQHMK	ENSPMAP00000009950	MAX (NP_002373.3)	Max	99	B
RRQAHTQAEQKRDAIKKGYEDLQAIVPTCMQQDLIGSQKMSKATVLRQSIDYIHFLH	ENSPMAP00000005560	TCFL4 (NP_937847.1)	TF4	96	B

NMSTHNELEKNRRAHLRKCLERLKEMVPLEAECsrHTTLGLLNKAkNHIKRL	ENSPMAP00000001203	MAD2 (NP_005953.4)	Mad	96	B
TREVHNKLEKSRRRAHLKECFDALKKIIPSMEEKKTSNLTILRGALRYIQVLK	ENSPMAP00000004888	MNT (NP_064706.1)	Mnt	83	B
RRAQHNEVERRRRDKINNWIVELSKYIPDCANDHTKSVQSKGGVLSKACEYIQELR	ENSPMAP00000010532 / ENSPMAP00000010533	USF2 (NP_003358.1/NP_997174.1)	USF	98	B
KKDNHNLIERRRRFNINDRIKELGTLIPKSNPDMDRWNGTILKASVDYIRRLQ	ENSPMAP00000005401	MITF (NP_937821.2)	MITF	96	B
KKDNHNLIERRRRFNINDRIKELGTLIPKSSDPDTRWNKGTLKASVDYIRRMQ	ENSPMAP00000001118/ ENSPMAP00000001120	MITF (NP_006713.1)/ TFE3 (NP_006512.2)	MITF	96	B
KKESHNLIERRRRFNINDRIKELGALITKSSDPDIRWNKGTLKASVDYIRRLQ	ENSPMAP00000009490	MITF (NP_937821.2)	MITF	96	B
KRTAHNAIEKRYRSSINDKIVELKDMIAGSEGKLNKSSVLRKAIEYIRYLQ	ENSPMAP00000001561	SREBF1 (NP_004167.3)	SREBP	50	B
RRSSHNAIERRYRCsINDRIGEMRSLVGPHNKLNSAVLRKAIDYIKFLQ	ENSPMAP000000010317	SREBF1 (NP_004167.3)	SREBP	50	B
GPSVKCTPERQWREqENKYIEELAEELISANISDIDSLNVKPKCAVLRETVNQIRQIK	ENSPMAP00000008090	NCOA2 (NP_006531.1)	SRC	87	B
EKLRRQENKYIEELAEELISANISDIDSLNVKPKCAILKETVQNIRQIK	ENSPMAP00000002809	NCOA2 (NP_006531.1)	SRC	87	B
SSSPQEHMEKRRRREQTSRYIEELAEELISANMGDIDGLGVKPDSCAVLRETVSQRQIK	ENSPMAP00000007705	NCOA2 (NP_006531.1)	SRC	87	B
KRAQRIRSEKRRRKQFNELISELGALLPDGQSRMDKPTVLQRTLHFFSSH	ENSPMAP00000002971	CLOCK (NP_004889.1)	Clock	88	C
RRENHSEIERRRRNKMTAYITELSDMVPTCSALARKPDKLTILRMAVSHMKSLR	(130764)	ARNT (NP_848514.1)	ARNT	95	C
SKEAHSQIEKRRRDKMNSYIDELASMVPTCSAMSRKLDKLTVLRMAVQHMKTLR	ENSPMAP00000004560	ARNTL (NP_001025444.1)	Bmal	38	C
RKEKSRDAARSRRGKENFEFYELAKMLPLPGAITSQLDKASHRLTISYLMKRD	ENSPMAP00000008398	NPAS3 (NP_071406.1)	Trh	39	C
RKEKSRDAARSRRGKENLEFYELGKMLPLPGAITGQLDKASLVRLTISHLRMRD	ENSPMAP000000010632	NPAS3 (NP_001159365.1)	Trh	39	C
RKEKSRDAARCRRSNETEVFYELANELPLPHSVTSHLDKASIMRLAISYLRMRK	ENSPMAP00000000148	HIF1A (NP_851397.1)	HIF	81	C
RKEKSRNAARCRRSKETEVFCELARELPLPQSVSASLDKASVMRFAISFLMRK	ENSPMAP000000010199	EPAS1 (NP_001421.2)	HIF	81	C
SEGIHNNASKRHRDRLNSEMEDLAAALPLPPAAITKDKLSVLRLSVGYLRARA	ENSPMAP00000006554	AHR (NP_001612.1)	AHR	94	C
SRLSQLVPTLATTGRKASRMEILQHVIDYILDQ	ENSPMAP00000001089	Id2 (NP_002157.2)	Emc	71	D
RKVRKPLVEKQRRDRINNLEHLKSIVLDSSQHEHASTSKLEKADILELTINYLCKVQ	ENSPMAP00000004457	Hes-5 (NP_001010926.1)	Hes	17	E
RKVRKPLVEKQRRDRINNLEHLKSII LDSSQHEHASTSKLEKADILELTINYLCKVQ	ENSPMAP00000004455	Hes-5 (NP_001010926.1)	Hes	17	E
RKSTKPIMEKRRRARINDSLNQLKALILETLRKDSSRHSKLEKADILELTVKHLRSLH	ENSPMAP00000001555 / ENSPMAP00000001558	Hes-1 (NP_005515.1)	Hes	17	E
YKLPHRLIEKKRRDRINECISQLKELLPDHLKQTTLGHLEKAVVLELTLKYMQLH	ENSPMAP00000007486	Dec1 (NP_003661.1)	Dec	99	E
YKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLSTLGHLEKAVVLELTLKHVKTLT	ENSPMAP00000004014	Dec2 (NP_110389.1)	Dec	99	E
ALNEPTIDYGFQRLQKLIPRHPGDPERLPKVRLRIHADWTETLIRLPNHSSEIILKRAADIAEALY	ENSPMAP00000003591	COE3 (NP_001005463.1)	COE	91	F

ALNEPTIDYGFQRLQKVIPRHGDPERLPKEVILKRAADLVEALYGLPHYNQEILKRAADIAEALY	ENSPMAP00000004636	COE3 (NP_001005463.1)	COE	91	F

Gene name	Family	Group	Accession number	Gene name	Family	Group	Accession number
NeuroD1	NeuroD	A	NP_002491.2	MNT	Mnt	B	NP_064706.1
NeuroD2	NeuroD	A	NP_006151.3	USF1	USF	B	NP_996888.1
NeuroD4	NeuroD	A	NP_067014.2	USF2-1	USF	B	NP_003358.1
NeuroD6	NeuroD	A	NP_073565.2	USF2-2	USF	B	NP_997174.1
Ngn1	Neurogenin	A	NP_006152.2	MITF	MITF	B	NP_000239.1
Ngn2	Neurogenin	A	NP_076924.1	TFEB	MITF	B	NP_009093.1
Ngn3	Neurogenin	A	NP_066279.2	TFEC	MITF	B	NP_001018068.1
Beta3	Beta3	A	NP_689627.1	TFE3	MITF	B	NP_006512.2
Beta4	Beta3	A	NP_542173.1	Unknown	MITF	B	AAS02009.1
Olig1	Oligo	A	NP_620450.2	SREBF1	SREBP	B	NP_004167.3
Olig2	Oligo	A	NP_005797.1	SREBF2	SREBP	B	NP_004590.2
Olig3	Oligo	A	NP_786923.1	NCOA1	SRC	B	NP_003734.3
Mist1	Mist	A	NP_803238.1	NCOA2	SRC	B	NP_006531.1
ATOH1	Atonal	A	NP_005163.1	NCOA3	SRC	B	NP_006525.2
ATOH7	Atonal	A	NP_660161.1	NPAS2	Clock	C	NP_002509.2
ATOH8	Net	A	NP_116216.2	Clock	Clock	C	NP_004889.1
Twist1	Twist	A	NP_000465.1	ARNT	ARNT	C	NP_848514.1
Twist2	Twist	A	NP_476527.1	BMAL1	Bmal	C	NP_001025443.1
N-Twist	PTFa	A	NP_690862.1	BMAL2	Bmal	C	AAL50340.1
PTF1a	PTFb	A	NP_835455.1	NPAS3	Trh	C	NP_775182.1
Paraxis	Paraxis	A	NP_004600.2	SIM1	Sim	C	NP_005059.2
Scleraxis	Paraxis	A	NP_001008272.1	SIM2	Sim	C	NP_005060.1
Musculin	MyoRa	A	NP_005089.2	HIF1A	HIF	C	NP_001521.1
TCF21	MyoRa	A	NP_003197.2	HIF3A	HIF	C	NP_690008.2
TCF23	MyoRb	A	NP_786951.1	NPAS1	HIF	C	NP_002508.2
TCF24	MyoRb	A	NP_001180431.1	EPAS1	HIF	C	NP_001421.2
HAND1	Hand	A	NP_004812.1	AHR	AHR	C	NP_001612.1
HAND2	Hand	A	NP_068808.1	AHRR	AHR	C	NP_065782.2
TAL1	SCL	A	NP_003180.1	Id1	Emc	D	NP_002156.2
TAL2	SCL	A	NP_005412.1	Id2	Emc	D	NP_002157.2
LYL1	SCL	A	NP_005574.2	Id3	Emc	D	NP_002158.3
NSCL1	NSCL	A	NP_005589.1	Id4	Emc	D	NP_001537.1
NSCL2	NSCL	A	NP_005590.1	Hey1	Hey	E	NP_036390.3
Msgn1	Mesp	A	NP_001099039.1	Hey2	Hey	E	NP_036391.1
Mesp1	Mesp	A	NP_061140.1	Hey3	Hey	E	NP_055386.1
Mesp2	Mesp	A	NP_001035047.1	Hes1	Hairy/E(spl)	E	NP_005515.1
ASCL1	ASCa	A	NP_004307.2	Hes2	Hairy/E(spl)	E	NP_061962.2
ASCL2	ASCa	A	NP_005161.1	Hes3	Hairy/E(spl)	E	NP_001019769.1
ASCL3	ASCb	A	NP_065697.1	Hes4	Hairy/E(spl)	E	NP_066993.1
ASCL4	ASCb	A	NP_982260.2	Hes5	Hairy/E(spl)	E	NP_001010926.1
ASCL5	ASCb	A	NP_001035047.1	Hes6	Hairy/E(spl)	E	NP_061115.2
Myogenin	MyoD	A	NP_002470.2	Hes7	Hairy/E(spl)	E	NP_115969.2
MyoD1	MyoD	A	NP_002469.2	Dec1	Dec	E	NP_003661.1
Myf5	MyoD	A	NP_005584.2	Dec2	Dec	E	NP_110389.1
Myf6	MyoD	A	NP_002460.1	HELT	Dec	E	NP_001025058.1
TCF4	E12/E47	A	NP_003190.1	COE1	COE	F	NP_076870.1
TCF12	E12/E47	A	NP_996920.1	COE2	COE	F	NP_073150.2
E12	E12/E47	A	NP_003191.1	COE3	COE	F	NP_001005463.1
E47	E12/E47	A	NP_001129611.1	COE4	COE	F	NP_001103984.1
FIGLA	Figa	B	NP_001004311.2				
Myc	Myc	B	NP_002458.2				
MYCL1	Myc	B	NP_001028253.1				
MYCL2	Myc	B	P12525.1				
N-myc	Myc	B	NP_005369.2				
Max	Max	B	NP_660092.1				
TCFL4	TF4	B	NP_937848.1				
MondoA	Mlx	B	NP_055753.3				
MondoB	Mlx	B	NP_116569.1				
AP4	AP4	B	NP_003214.1				
MAD1	Mad	B	NP_002348.1				
MAD3	Mad	B	NP_112590.1				
MAD4	Mad	B	NP_006445.1				
MXI1b	Mad	B	NP_569157.2				
MXI1c	Mad	B	NP_001008541.1				

Table 3-2. Full set of human bHLH genes used in this study classified according to subfamilies and higher order groups (i.e. A-F) as described by Simionato et al. (2007). Note that in the present study *Dec* genes are considered as an independent subfamily from the *Hairy/E(spl)* and *Hey* subfamilies (discussed in detail in section 4.6.3).

Once predicted HLH domains were isolated from each lamprey model, they were aligned with the full set of human (Table 3-2) and sea urchin HLH (Table 3-3) domains obtained from Simionato et al. (2007) and SpBase, respectively, although several sequences were either deleted or updated (see Figure 3-1 for alignment). This alignment was used to perform subsequent phylogenetic analyses using human as a reference and sea urchin as an outgroup. Sea urchin was preferred as an outgroup rather than amphioxus or tunicates given that it contains at least one member of all HLH subfamilies (except of the *Figa* subfamily) and their relationships are well established according to Simionato et al. (2007), whereas grouping of several amphioxus (*Branchiostoma floridae*) sequences is ambiguous, particularly sequences of the *Hes/Hey* and *Sim/Trh*, *HIF* subfamilies, and tunicates have lost several HLH subfamilies (Simionato et al., 2007).

Gene name	Subfamily	Group	Accession number	Gene name	Subfamily	Group	Accession number
Sp-NeuroD1	NeuroD	A	SPU_024918	Sp-Mitf	MITF	B	SPU_008175
Sp-Ngn	Neurogenin	A	SPU_007147	Sp-Srebp	SREBP	B	SPU_007253
Sp-Beta3a	Beta3	A	SPU_026601	Sp-Ncoa	SRC	B	SPU_021430
Sp-Olig3	Oligo	A	SPU_002627	Sp-Clock	Clock	C	SPU_017407
Sp-Mist	Mist	A	SPU_027623	Sp-Arnt	ARNT	C	SPU_000129
Sp-Ato_1	Net	A	SPU_017983	Sp-Bmal	Bmal	C	SPU_027935
Sp-Twi	Twist	A	SPU_030059	Sp-Trh	Trh	C	SPU_014249
Sp-NtL	PTFa	A	SPU_014401	Sp-Sim_1	Sim	C	SPU_013962
Sp-Ptf1a	PTFb	A	SPU_002677	Sp-Birc6	HIF	C	SPU_001262
Sp-Par1	Paraxis	A	SPU_016650	Sp-Ahr_1	AHR	C	SPU_005022
Sp-Myor2	MyoRa	A	SPU_012008	Sp-Id	Emc	D	SPU_015374
Sp-Myor4	MyoRb	A	SPU_016445	Sp-Hey	Hey	E	SPU_009465
Sp-Hand	Hand	A	SPU_017287	Sp-Hairy2/4	Hairy/E(spl)	E	SPU_006814
Sp-Scl	SCL	A	SPU_028093	Sp-Hairy	Hairy/E(spl)	E	SPU_006813
Sp-Ncsl	NSCL	A	SPU_009231	Sp-HesC	Hairy/E(spl)	E	SPU_021608
Sp-Sage	Mesp	A	SPU_002448	Sp-Hey4	Hairy/E(spl)	E	SPU_015712
Sp-Ac/Sc3	ASCa	A	SPU_022554	Sp-Ebf3	COE	F	SPU_004702
Sp-Ac/Sc	ASCa	A	SPU_028148				
Sp-Ash3L	ASCb	A	SPU_030149				
Sp-MyoD	MyoD	A	SPU_021119	Sp-Atol1	Atonal-like?	?	SPU_000990
Sp-MyoD2	MyoD	A	SPU_006232	Sp-Atl7L	Atonal-like?	?	SPU_030150
Sp-E12	E12/E47	A	SPU_016343	Sp-Atol2	Orphan?	?	SPU_003681
Sp-Myc	Myc	B	SPU_003166				
Sp-Max	Max	B	SPU_022163				
Sp-Mlx	TF4	B	SPU_005787				
Sp-Mlx/Ip	Mlx	B	SPU_008845				
Sp-AP4	AP4	B	SPU_003179				
Sp-Mad	Mad	B	SPU_006583				
Sp-Mnt	Mnt	B	SPU_026205				
Sp-Usf	USF	B	SPU_014332				

Table 3-3. Full set of *Strongylocentrotus purpuratus* HLH genes used in this study classified according to subfamilies and higher order groups (i.e. A-F) as described by Simionato et al. (2007).

	10	20	30	40	50	60	70	80
ENSMAP00000008309	R-----	RVKANARERIRMHGLNAAALDNRKVV	PCY-----	S-KTQK-----	LSK	ETLRLAKNIYIALS		
98653	R-----	RLKANARERIRMHGLNAAALDNRKVV	PCH-----	S-RTQK-----	LSK	ETLRLAKNIYIALS		
ENSMAP00000006813	R-----	RKEANDRERIRMHGLNAAALDNRKVV	PTL-----	P-DDAR-----	LT	KIETLRFAYNIYIALS		
75235	R-----	RLESNERERIRMHGLNAAALDNRKVV	PHV-----	R-RGRK-----	LSK	ETLRLAKNIYIALS		
39660	R-----	RLAANARERIRMHGLNAAALDNRKVV	PQW-----	G-QDRK-----	LSK	ETLRLAKNIYIALS		
ENSMAP000000011396	-----	RERORTQSLNEAFSSRQII	PTL-----	P-SD-K-----	LSK	QTLKLATRYIDFLY		
ENSMAP00000004354	Q-----	RFVANVERORTQSLNEAFSSRQII	PTL-----	P-SD-K-----	LSK	QTLKLATRYIDFLY		
ENSMAP000000011107	Q-----	RVLANVERORTQSLNEAFSSRQII	PTL-----	P-SD-K-----	LSK	QTLKLATRYIDFLY		
ENSMAP00000004055	L-----	RQAANVERORTQSLNEAFSSRQII	PTL-----	P-YEKR-----	LSR	DTLRLAIGYISFLG		
ENSMAP00000000075	Q-----	RTAANVERORTQSLNEAFSSRQII	PTF-----	P-YEKR-----	LSK	DTLRLAIGYISFLG		
ENSMAP000000007312	Q-----	RRAANARERIRMHGLNAAALDNRKVV	PTE-----	P-ADRK-----	LSK	ETLRLASSYIAFLG		
ENSMAP00000000421	Q-----	RNAANARERIRMHGLNAAALDNRKVV	PWV-----	P-ADTK-----	LSK	DTLRLASSYIAFLG		
ENSMAP000000010377	K-----	ESSGANRERIRMHGLNAAALDNRKVV	PAV-----	P-PDTK-----	LSK	DTLRLASSYIAFLG		
ENSMAP00000002458	R-----	RIFTNSRERIRMHGLNAAALDNRKVV	PTH-----	P-PDRK-----	LSK	ETLRLAMKYNIALS		
93980	R-----	RARKNERERIRMHGLNAAALDNRKVV	PHL-----	AAVGRP-----	LSK	QTLRLTIRYIACLS		
ENSMAP000000008927	F-----	IRKNERERIRMHGLNAAALDNRKVV	PDE-----	F-SDKR-----	LSR	ETLRLAIRYIKFLQ		
ENSMAP00000001063	R-----	RMANNARERIRMHGLNAAALDNRKVV	QLH-----	LKSDKP-----	QTK	LLIHLQAVAVILSL		
ENSMAP000000004198	R-----	RRNHNFERIRMHGLNAAALDNRKVV	PBL-----	A-GNEK-----	AAR	VTLRKAAESAVAAT		
ENSMAP000000003372	R-----	RRTHNIERIRMHGLNAAALDNRKVV	PBL-----	R-ANER-----	AAR	VTLRKAAESAVAAT		
ENSMAP000000009801	R-----	RRTHNIERIRMHGLNAAALDNRKVV	PBL-----	C-RNDK-----	AAR	VTLRKAAESAVAAT		
ENSMAP000000011193	K-----	RRTHNIERIRMHGLNAAALDNRKVV	PBL-----	A-HNDK-----	AAR	VTLRKAAESAVAAT		
ENSMAP000000010427	K-----	RAHHNAERIRMHGLNAAALDNRKVV	PSL-----	QGEKVNLRKAAESAVAAT	QTK	LLIHLQAVAVILSL		
ENSMAP000000009950	K-----	RAHHNAERIRMHGLNAAALDNRKVV	PAT-----	QGEKVNLRKAAESAVAAT	QTK	LLIHLQAVAVILSL		
ENSMAP000000005560	R-----	RQAHTQERIRMHGLNAAALDNRKVV	PTC-----	MSATVLRKAAESAVAAT	QTK	LLIHLQAVAVILSL		
ENSMAP000000001203	N-----	MSHNEERIRMHGLNAAALDNRKVV	PL-----	EAECSS-----	HTT	GLVLRKAAESAVAAT		
ENSMAP000000004888	T-----	REVHNKLEKRRRAHLKCEFDALCKII	PSM-----	EEKK-----	TSN	LTIRGALRYIQVLR		
ENSMAP000000010532	R-----	RAQHNEERIRMHGLNAAALDNRKVV	PDC-----	ANDHTK-----	SVQ	SKGVLKACEYIQELR		
ENSMAP000000005401	K-----	KDNHNLERIRMHGLNAAALDNRKVV	PKS-----	NDPDMR-----	WKK	ETLRLASSYIAFLG		
ENSMAP000000001118	K-----	KDNHNLERIRMHGLNAAALDNRKVV	PKS-----	SDPDTR-----	WKK	ETLRLASSYIAFLG		
ENSMAP000000009490	K-----	KESHNLERIRMHGLNAAALDNRKVV	TKS-----	SDPDTR-----	WKK	ETLRLASSYIAFLG		
ENSMAP000000001561	K-----	RTAHNAERIRMHGLNAAALDNRKVV	PKS-----	AGSEK-----	WKK	ETLRLASSYIAFLG		
ENSMAP000000010317	R-----	RSSHNAERIRMHGLNAAALDNRKVV	PKS-----	AGSEK-----	WKK	ETLRLASSYIAFLG		
ENSMAP000000008090	G-----	PSVKCTERIRMHGLNAAALDNRKVV	PKS-----	AGSEK-----	WKK	ETLRLASSYIAFLG		
ENSMAP000000002809	-----	EKLREKENKYIEEAEI	-----	SANISDIDSLNVK-----	PDK	CAVLRKAAESAVAAT		
ENSMAP000000007705	S-----	SSPQEHKRRERIRMHGLNAAALDNRKVV	-----	SANISDIDSLNVK-----	PDK	CAVLRKAAESAVAAT		
ENSMAP000000002971	K-----	RAQRIRERIRMHGLNAAALDNRKVV	-----	SANISDIDSLNVK-----	PDK	CAVLRKAAESAVAAT		
130764	R-----	RENHSEIRERIRMHGLNAAALDNRKVV	-----	SANISDIDSLNVK-----	PDK	CAVLRKAAESAVAAT		
ENSMAP000000004560	S-----	KEAHSQERIRMHGLNAAALDNRKVV	-----	SANISDIDSLNVK-----	PDK	CAVLRKAAESAVAAT		
ENSMAP000000008398	R-----	KEKSRAAARSRGKENFEYEDAKML	PLP-----	GAITSQ-----	LDK	ASVLRKAAESAVAAT		
ENSMAP000000010632	R-----	KEKSRAAARSRGKENFEYEDAKML	PLP-----	GAITSQ-----	LDK	ASVLRKAAESAVAAT		
ENSMAP000000000148	R-----	KEKSRAAARSRGKENFEYEDAKML	PLP-----	GAITSQ-----	LDK	ASVLRKAAESAVAAT		
ENSMAP000000010199	R-----	KEKSRAAARSRGKENFEYEDAKML	PLP-----	GAITSQ-----	LDK	ASVLRKAAESAVAAT		
ENSMAP000000006554	S-----	EGIHNNASKRRDRINLEMEDEAAL	PLP-----	PAAITK-----	LDK	ASVLRKAAESAVAAT		
ENSMAP000000001089	-----	SRUSQLV	PTL-----	ATTGRK-----	ASR	MEILQHVIVYDILQ		
ENSMAP000000004457	R-----	KVRKPLVERIRMHGLNAAALDNRKVV	-----	LDSSQHEHASTSK-----	LEK	ADIELTINYLCKVQ		
ENSMAP000000004455	R-----	KVRKPLVERIRMHGLNAAALDNRKVV	-----	LDSSQHEHASTSK-----	LEK	ADIELTINYLCKVQ		
ENSMAP000000001555	R-----	KSTKPLVERIRMHGLNAAALDNRKVV	-----	LDSSQHEHASTSK-----	LEK	ADIELTINYLCKVQ		
ENSMAP000000007486	Y-----	KLPHRLVERIRMHGLNAAALDNRKVV	-----	LDSSQHEHASTSK-----	LEK	ADIELTINYLCKVQ		
ENSMAP000000004014	Y-----	KLPHRLVERIRMHGLNAAALDNRKVV	-----	LDSSQHEHASTSK-----	LEK	ADIELTINYLCKVQ		
ENSMAP000000003591	-----	ALNEPTIDYGFQRLQKVI	PRH-----	PGDPER-----	LFR	RLRIHADWTETLI		
ENSMAP000000004636	-----	ALNEPTIDYGFQRLQKVI	PRH-----	PGDPER-----	LFR	RLRIHADWTETLI		
Hs NeuroD1 (NP_002491.2)	R-----	RVKANARERIRMHGLNAAALDNRKVV	PCY-----	S-KTQK-----	LSK	ETLRLAKNIYIALS		
Hs NeuroD2 (NP_006151.3)	R-----	RQKANARERIRMHGLNAAALDNRKVV	PCY-----	S-KTQK-----	LSK	ETLRLAKNIYIALS		
Hs NeuroD4 (NP_067014.2)	R-----	RVKANARERIRMHGLNAAALDNRKVV	PCY-----	S-KTQK-----	LSK	ETLRLAKNIYIALS		
Hs NeuroD6 (NP_073565.2)	R-----	RQKANARERIRMHGLNAAALDNRKVV	PCY-----	S-KTQK-----	LSK	ETLRLAKNIYIALS		
Hs Ngn1 (NP_006152.2)	R-----	RVKANDRERIRMHGLNAAALDNRKVV	PSF-----	P-DDTK-----	LT	KIETLRFAYNIYIALS		
Hs Ngn2 (NP_076924.1)	R-----	RLKANARERIRMHGLNAAALDNRKVV	PTF-----	P-EDAK-----	LT	KIETLRFAYNIYIALS		
Hs Ngn3 (NP_066279.2)	R-----	RKKANDRERIRMHGLNAAALDNRKVV	PTF-----	P-DDAK-----	LT	KIETLRFAYNIYIALS		
Hs Beta3 (NP_689627.1)	L-----	RLNINARERIRMHGLNAAALDNRKVV	PYA-----	HSP-SVRK-----	LSK	ATLRLAKNIYIALS		
Hs Beta4 (NP_542173.1)	L-----	RLSINARERIRMHGLNAAALDNRKVV	PYA-----	HSP-SVRK-----	LSK	ATLRLAKNIYIALS		
Hs Olig1 (NP_620450.2)	L-----	RLKINSRERIRMHGLNAAALDNRKVV	PYA-----	HGP-SVRK-----	LSK	ATLRLAKNIYIALS		
Hs Olig2 (NP_005797.1)	L-----	RLKINSRERIRMHGLNAAALDNRKVV	PYA-----	HGP-SVRK-----	LSK	ATLRLAKNIYIALS		
Hs Olig3 (NP_786923.1)	L-----	RLKINSRERIRMHGLNAAALDNRKVV	PYA-----	HGP-SVRK-----	LSK	ATLRLAKNIYIALS		
Hs Mist1 (NP_803238.1)	R-----	RLESNERERIRMHGLNAAALDNRKVV	PHV-----	R-ADKK-----	LSK	ETLRLAKNIYIALS		
Hs ATOH1 (NP_005163.1)	R-----	RLAANARERIRMHGLNAAALDNRKVV	PSF-----	N-NDKK-----	LSK	ETLRLAKNIYIALS		
Hs ATOH7 (NP_660161.1)	R-----	RLAANARERIRMHGLNAAALDNRKVV	PQW-----	G-QDKK-----	LSK	ETLRLAKNIYIALS		
Hs ATOH8 (NP_116216.2)	R-----	RLANARERIRMHGLNAAALDNRKVV	PCY-----	S-YGQK-----	LSK	ADIRIACNIYIALS		
Hs Twist1 (NP_000465.1)	Q-----	RVMANVERORTQSLNEAFSSRQII	PTL-----	P-SD-K-----	LSK	QTLKLATRYIDFLY		
Hs Twist2 (NP_476527.1)	Q-----	RILANVERORTQSLNEAFSSRQII	PTL-----	P-SD-K-----	LSK	QTLKLATRYIDFLY		
Hs N-Twist (NP_690862.1)	Q-----	RQAANVERORTQSLNEAFSSRQII	PTF-----	A-YEKR-----	LSR	DTLRLAIGYISFLG		
Hs PTF1a (NP_835455.1)	L-----	RQAANVERORTQSLNEAFSSRQII	PTL-----	P-YEKR-----	LSR	DTLRLAIGYISFLG		
Hs Paraxis (NP_004600.2)	Q-----	RHTANARERIRMHGLNAAALDNRKVV	PTE-----	P-VDRK-----	LSK	ETLRLASSYIAFLG		
Hs Scleraxis (NP_00100827)	Q-----	RHTANARERIRMHGLNAAALDNRKVV	PTE-----	P-ADRK-----	LSK	ETLRLASSYIAFLG		
Hs Musculin (NP_005089.2)	Q-----	RNAANARERIRMHGLNAAALDNRKVV	PWV-----	P-PDTK-----	LSK	DTLRLASSYIAFLG		
Hs TCF21 (NP_003197.2)	Q-----	RNAANARERIRMHGLNAAALDNRKVV	PWV-----	P-PDTK-----	LSK	DTLRLASSYIAFLG		
Hs TCF23 (NP_786951.1)	A-----	SPENARERIRMHGLNAAALDNRKVV	PAV-----	P-PDTK-----	LSK	DTLRLASSYIAFLG		
Hs TCF24 (NP_001180431.1)	P-----	AAANARERIRMHGLNAAALDNRKVV	PSV-----	P-PDTK-----	LSK	DTLRLASSYIAFLG		
Hs HAND1 (NP_004812.1)	R-----	RGSFKERIRMHGLNAAALDNRKVV	PWV-----	P-ADTK-----	LSK	ETLRLATSYIAFLM		
Hs HAND2 (NP_068808.1)	R-----	RGTANARERIRMHGLNAAALDNRKVV	PWV-----	P-ADTK-----	LSK	ETLRLATSYIAFLM		
Hs TAL1 (NP_003180.1)	R-----	RIFTNSRERIRMHGLNAAALDNRKVV	PTH-----	P-PDKK-----	LSK	ETLRLAMKYNIALS		

Hs TAL2 (NP_005412.1)	R-----KIFTNTRERVRQDMVNSAFARLRLI----	PTH-----P-PDKK-----LSRNETRLAMRYINILV
Hs LYL1 (NP_005574.2)	R-----RVFTNSRERVRQDMNGAFARLRLI----	PTH-----P-PDKK-----LSRNETRLAMRYINILV
Hs NSCL1 (NP_005589.1)	Y-----RTAHATREIRVR/EAFLNLAFAERLRLI----	PTL-----P-PDKK-----LSKTEIRLAICYISILN
Hs NSCL2 (NP_005590.1)	Y-----RSAHATREIRVR/EAFLNLAFAERLRLI----	PTL-----P-PDKK-----LSKTEIRLAICYISILN
Hs Msmg1 (NP_001099039.1)	R-----RKASERERKRMRTLADALHTLRN/LP----	PVY-----SQRGQP-----LTKITQIKYTIKIGTGL
Hs Mesp1 (NP_061140.1)	Q-----RQSASERERKRMRTLARALHELRRLP----	PSV-----APAGQS-----LTKITETRLAIRYIGLS
Hs Mesp2 (NP_001035047.1)	Q-----RQSASERERKRMRTLARALHELRRLP----	PSL-----APAGQS-----LTKITETRLAIRYIGLS
Hs ASCL1 (NP_004307.2)	A-----VARRNERERVR/KLVNLGFATLRHV----	PNG-----A-ANKK-----MSR/ETIRSAVEYIRALQ
Hs ASCL2 (NP_005161.1)	A-----VARRNERERVR/KLVNLGFATLRHV----	PHG-----G-ASKK-----LSR/ETIRSAVEYIRALQ
Hs ASCL3 (NP_065697.1)	F-----TRKNERERVR/KCVNEGVARLRHHL----	PBE-----Y-LEKR-----LSR/ETIRRAAIKINILQ
Hs ASCL4 (NP_982260.2)	F-----LRKNERERVR/KCVNEGVARLRHHL----	PRE-----L-ADKR-----LSR/ETIRRAAIDYIKELQ
Hs ASCL5 (Q7RTU5.2)	F-----IQKNERERVR/KCVNEGVARLRHHL----	PGA-----L-AEKR-----LSR/ETIRRAAIRIKILQ
Hs Myogenin (NP_002470.2)	R-----BRAATLREKRLKKVNEAFELKRSTL--LP----	NQR-----NQR-----LFR/ETIRSAIQYIEFLQ
Hs MyoD1 (NP_002469.2)	R-----RKAATMRERRLSKVNEAFETLKRCTS--SLP----	NQR-----NQR-----LFR/ETIRNAIRYIEGLQ
Hs Myf5 (NP_005584.2)	R-----RKAATMRERRLKKVNEAFETLKRCTT--TLP----	NQR-----NQR-----LFR/ETIRNAIRYIESLQ
Hs Myf6 (NP_002460.1)	R-----RKAATLREKRLKKVNEAFELKRRTV--ALP----	NQR-----NQR-----LFR/ETIRSAISYIEFLQ
Hs TCF4 (NP_003190.1)	R-----RMANNARERVR/RDINEAFKELGRMV----	QLH-----LKSDEP-----OTKLLILHQAVAVILSL
Hs TCF12 (NP_996920.1)	R-----RMANNARERVR/RDINEAFKELGRMC----	QLH-----LKSDEP-----OTKLLILHQAVAVILSL
Hs E12 (NP_003191.1)	R-----RVANNARERVR/RDINEAFKELGRMC----	QLH-----LKSDEP-----OTKLLILHQAVSVILSL
Hs E47 (NP_001129611.1)	R-----RMANNARERVR/RDINEAFKELGRMC----	QMH-----LKSDEP-----OTKLLILHQAVQVILSL
Hs FIGLA (NP_001004311.2)	R-----RRVANARERVR/KLVNLGFATLRKALV----	PFL-----P-QSRK-----PSR/ETIRKATEYIQVLS
Hs Myc (NP_002458.2)	K-----RRTHNVLERVR/RDINEAFKELGRMV----	PBL-----E-INNEK-----APR/ETIRKATYILSVQ
Hs MyCL1 (NP_001028253.1)	K-----RKQHNFLERVR/RDINEAFKELGRMV----	PTL-----A-SCSK-----APR/ETIRSKALEYIQVLS
Hs MyCL2 (P12525.1)	K-----KKYHSYLERVR/RDINEAFKELGRMV----	PAL-----A-SCSK-----VSR/ETIRKATEYILHSL
Hs N-myc (NP_005369.2)	R-----RRNHNILERVR/RDINEAFKELGRMV----	PBL-----V-KNEK-----APR/ETIRKATEYIVHSL
Hs Max (NP_660092.1)	R-----RAHNALEKRRRDKIKRSGFDDLDQIV----	PSL-----QGEK-----ASRAQIDKATEYIQMR
Hs TCF14 (NP_937848.1)	R-----RRAHTQAEKRRRDKIKRSGFDDLDQIV----	PTCQQQDF--SIGSQK-----LSKAIMQKTIYIQQLH
Hs MondoA (NP_055753.3)	R-----QMKHISAEKRRRDNKMCDFDNLISLI----	SNNSKL-----TSHAITDKTVEYITFLQ
Hs MondoB (NP_116569.1)	R-----RITHISAEKRRRDNKMGFDTLHGLV----	STLS-----AQPSLK-----VSKATTDKTAEYIILQ
Hs AP4 (NP_003214.1)	R-----REIANSNERVR/RDINEAFKELGRMV----	PHT-----DGEK-----LSKAIMQKTIYIFSL
Hs MAD1 (NP_002348.1)	S-----RSTHNEKERRRAHLRLCLERLKLGLV----	PL-----GPSSSR-----HTTSLSLTKAKHLIKKLE
Hs MAD3 (NP_112590.1)	G-----RSVHNEKERRRAHLRLCLERLKLGLV----	PL-----GADCAR-----YTTLSLARRAHMIQKLE
Hs MAD4 (NP_006445.1)	N-----RSSHNEKERRRAHLRLCLERLKLGLV----	PL-----GPDSTR-----HTTSLSLTKAKHLIKKLE
Hs MXI1b (NP_569157.2)	SSGSSNTSTANRSTHNEKERRRAHLRLCLERLKLGLV----	PL-----GPDCTR-----HTTSLSLTKAKHLIKKLE
Hs MXI1c (NP_001008541.1)	S-----SGSSNTSTANRRAHLRLCLERLKLGLV----	PL-----GPDCTR-----HTTSLSLTKAKHLIKKLE
Hs MNT (NP_064706.1)	T-----REVHNKLEKERRRAHLRLCLERLKLGLV----	PNV-----DDKK-----TSNLSVRLTALRYIQSLK
Hs USF1 (NP_996888.1)	R-----RAQHNEVERRRDKINNWIVQSKII----	PDC-----SMESTK-----SGQSGGILSKACDYIRELR
Hs USF2-1 (NP_003358.1)	R-----RAQHNEVERRRDKINNWIVQSKII----	PDC-----NADNSK-----TGASKGGILSKACDYIRELR
Hs USF2-2 (NP_997174.1)	R-----RAQHNEVERRRDKINNWIVQSKII----	PDC-----NADNSK-----TGASKGGILSKACDYIRELR
Hs MITF (NP_000239.1)	K-----KDNHNLDERRRFNDIRIKELGTLLI----	PKS-----NDPDMR-----WNKGTILKASVDYIRRLQ
Hs TFEB (NP_009093.1)	K-----KDNHNLDERRRFNDIRIKELGTLLI----	PKA-----NDLDVR-----WNKGTILKASVDYIRRLQ
Hs TFEC (NP_001018068.1)	K-----KDNHNLDERRRFNDIRIKELGTLLI----	PKS-----NDPDMR-----WNKGTILKASVDYIRRLQ
Hs TFEB (NP_006512.2)	K-----KDNHNLDERRRFNDIRIKELGTLLI----	PKS-----SDPEMR-----WNKGTILKASVDYIRRLQ
Hs Unknovm (AAS02009.1)	K-----KDNHNLDERRRFNDIRIKELGTLLI----	PKS-----NDPDMR-----WNKGTILKASVDYIRRLQ
Hs SREBF1 (NP_004167.3)	K-----RTAHNALEKRRRSSINDKILKLGLV----	VGTEAK-----LNKSAVIRKADYIRRLQ
Hs SREBF2 (NP_004590.2)	R-----RTTHNILEKRRRSSINDKILKLGLV----	MGDAK-----MHKSGVIRKADYIKILQ
Hs NCOA1 (NP_003734.3)	C-----DTLASSTEKRRRQENKYIEELAEALL----	SANISDIDSLSVK-----PDKKILKKTVDQIQMLK
Hs NCOA2 (NP_006531.1)	G-----PSPKRNTEKRRRQENKYIEELAEALL----	FANFNDIDNFMFK-----PDKKILKKTVDQIQMLK
Hs NCOA3 (NP_006525.2)	G-----QGLTCSGEKRRRQENKYIEELAEALL----	SANISDIDNFMFK-----PDKKILKKTVDQIQMLK
Hs NPAS2 (NP_002509.2)	K-----RASRNKSEKRRRQENKYIEELAEALL----	P-----GNTRK-----MDKSTVLRKSIDFLRRHK
Hs Clock (NP_004889.1)	K-----RVSRNKSEKRRRQENKYIEELAEALL----	P-----GNTRK-----MDKSTVLRKSIDFLRRHK
Hs ARNT (NP_848514.1)	A-----RENHSEDERRRRNKMTAYITELSDMV----	PTC-----SALARK-----PDKKILKKTVDQIQMLK
Hs BMAL1 (NP_001025443.1)	A-----REAHSQEKRRRDRNMFIDEASAMI----	PTC-----NAMSRR-----LDKLTVMMAVQHMRLR
Hs BMAL2 (AAL50340.1)	S-----REAHSQEKRRRDRNMFIDEASAMI----	PQC-----NPMARK-----LDKLTVMMAVQHMRLR
Hs NPAS3 (NP_775182.1)	R-----KEKSRDAARSRRKESEVFEYLAHQL----	PLP-----AAITSQ-----LDKASIIRLTISYLRMR
Hs SIM1 (NP_005059.2)	M-----KEKSQAARTREKENSEFEYLAHQL----	PLP-----SAITSQ-----LDKASIIRLTISYLRMR
Hs SIM2 (NP_005060.1)	M-----KEKSQAARTREKENSEFEYLAHQL----	PLP-----SAITSQ-----LDKASIIRLTISYLRMR
Hs HIF1A (NP_001521.1)	R-----KEKSRDAARSRRKESEVFEYLAHQL----	PLP-----HNVSSH-----LDKASVIRLTISYLRMR
Hs HIF3A (NP_690008.2)	R-----KEKSRDAARSRRKESEVFEYLAHQL----	PFA-----RGVSAH-----LDKASIIRLTISYLRMR
Hs NPAS1 (NP_002508.2)	R-----KEKSRDAARSRRKESEVFEYLAHQL----	PLP-----GAISSQ-----LDKASVIRLTISYLRMR
Hs EPAS1 (NP_001421.2)	R-----KEKSRDAARSRRKESEVFEYLAHQL----	PLP-----HSVSSH-----LDKASVIRLTISYLRMR
Hs AHR (NP_001612.1)	A-----EGIKSNPSKRHRDLNTELDRLASLL----	PPP-----QDVINK-----LDKLSVIRLTISYLRMR
Hs AHR1 (NP_065782.2)	V-----GAEKSNPSKRHRDLNTELDRLASLL----	PPP-----PDIISK-----LDKLSVIRLTISYLRMR
Hs Id1 (NP_002156.2)	-----SRKELV-----P-QNRK-----VSR/ETIRHVIDYIRLQ	PTL-----P-QNRK-----VSR/ETIRHVIDYIRLQ
Hs Id2 (NP_002157.2)	-----SRKELV-----P-QNRK-----VSR/ETIRHVIDYIRLQ	PSI-----P-QNRK-----VSR/ETIRHVIDYIRLQ
Hs Id3 (NP_002158.3)	-----SRKELV-----P-QNRK-----VSR/ETIRHVIDYIRLQ	PGV-----P-RGTO-----LSQVEIRQRVIDYIRLQ
Hs Id4 (NP_001537.1)	-----SRKELV-----P-QNRK-----VSR/ETIRHVIDYIRLQ	PTI-----P-PNKK-----VSR/ETIRHVIDYIRLQ
Hs Hey1 (NP_036390.3)	R-----KRRRGIDKRRDRINNSLSELRRLV----	PSAFERQ-----GSAK-----LEKAEILQMTVDHLKRLH
Hs Hey2 (NP_036391.1)	R-----KRRRGIDKRRDRINNSLSELRRLV----	PTAFERQ-----GSAK-----LEKAEILQMTVDHLKRLH
Hs HeyL (NP_055386.1)	R-----KRRRGIDKRRDRINNSLSELRRLV----	PTAFERQ-----GSSK-----LEKAEILQMTVDHLKRLH
Hs HELT (NP_001025058.1)	T-----FVSHKIDKRRDRINNSLSELRRLV----	PMALALQ-----SSGK-----LEKAEILQMTVDHLKRLH
Hs Hes1 (NP_005515.1)	R-----KSKPILKRRDRINNSLSELRRLV----	LDALKDSSRHSK-----LEKADILEMTVRLQBLR
Hs Hes2 (NP_061962.2)	R-----KSKPILKRRDRINNSLSELRRLV----	LPLLGRENNSCK-----LEKADILEMTVRLQBLR
Hs Hes3 (NP_001019769.1)	-----MEKRRARINVSLEQKSL-----EKHYSHQIRKRK-----LEKADILEMTVRLQBLR	
Hs Hes4 (NP_066993.1)	R-----KSKPILKRRDRINNSLSELRRLV----	LDALKDSSRHSK-----LEKADILEMTVRLQBLR
Hs Hes5 (NP_001010926.1)	N-----RLRKPVEKRRDRINNSLSELRRLV----	EQEFARHPNSK-----LEKADILEMTVRLQBLR
Hs Hes6 (NP_061115.2)	R-----KARKPLVEKRRDRINNSLSELRRLV----	AGAEVQAK-----LEKADILEMTVRLQBLR
Hs Hes7 (NP_115969.2)	P-----KMLKPLVEKRRDRINNSLSELRRLV----	LERTRDQNLNPK-----LEKADILEMTVRLQBLR
Hs DEC1 (NP_003661.1)	Y-----KLPHRLDEKRRDRINNSLSELRRLV----	PEHLKLTTLGH-----LEKADILEMTVRLQBLR
Hs DEC2 (NP_110389.1)	Y-----KLPHRLDEKRRDRINNSLSELRRLV----	PEHLKLTTLGH-----LEKADILEMTVRLQBLR
Hs COE1 (NP_076870.1)	-----ALNEPTIDYGFQRLQKVI-----PRH-----PGDPER-----LFR/ETIRKADILEMTVRLQBLR	
Hs COE2 (NP_073150.2)	-----ALNEPTIDYGFQRLQKVI-----PRH-----PGDPER-----LFR/ETIRKADILEMTVRLQBLR	

Hs COE3 (NP_001005463.1)	-----ALNEPTIDYGFQRLQKVI-----	PRH-----PGDPER-----LPRFVLKRAADLVEAL
Hs COE4 (NP_001103984.1)	-----ALNEPTIDYGFQRLQKVI-----	PRH-----PGDPER-----LPRFVLKRAADLVEAL
Sp-NeuroD1 (SPU_024918)	R-----RVKANTERRRMHGLNDALDLRVV-----	PCY-----S-STQK-----LSKETETLRANKYIHA
Sp-Ngn (SPU_007147)	R-----RLKANDERRRMHNLNIALDGLREVL-----	PNF-----P-DDTK-----LTKTETLRFAHNYIWA
Sp-Beta3a (SPU_026601)	I-----RLNINARERRRMHLDLADLDGRVI-----	PYA-----HSP-SVRK-----LSKATLTLAKNYILMO
Sp-Olig3 (SPU_002627)	L-----RLKINSERRRMHNLKALDGLREVM-----	PYA-----HGP-SVRK-----LSR4STLLAKNYILML
Sp-Mist (SPU_027623)	R-----RLSENERERRRMHTLINDAFQNLRII-----	PHV-----R-AERK-----LSKETETLRANKYIHA
Sp-Ato11 (SPU_000990)	K-----RESVMERRRRMHQLRDAPFLRLRV-----	PRY-----P-VRQK-----CSR4DTLLAQDYILTQ
Sp-Ato12 (SPU_003681)	R-----RLLTNHERRRRMSLINDADRLNIV-----	PHY-----P-SKRR-----LSR4ETLLAQSYIMAL
Sp-Ato17L (SPU_030150)	K-----RTEANARERRRVTLNGGFQRLRVL-----	PWA-----YRGGR-----VSR4DTLRAISYIQFL
Sp-Ato1 (SPU_017983)	R-----RIVANARERRRHTISSAFEGURRAV-----	PSY-----S-HNQK-----LSK4ATLRIACSYILA
Sp-Twi (SPU_030059)	Q-----RVLANVERRRRTQSINDAFTNRKII-----	PTL-----P-SD-K-----LSK4OTKLASRYIDFL
Sp-Ntl (SPU_014401)	R-----RTAANVERRRRMFLNDADFNRKEV-----	PKF-----K-HEKR-----LSRIETLRALYIETLA
Sp-Ptf1a (SPU_002677)	Q-----RQANLERRRRMSSINDAFGRREHI-----	PTL-----P-YEKR-----LSR4DTLRALIGYINLA
Sp-Par1 (SPU_016650)	Q-----RGAANARERRRTHSVNSAFVQRLDI-----	PTE-----P-RDRK-----LSKETETLRATSYINELG
Sp-Myor2 (SPU_012008)	Q-----RNAANARERRRMVLSKAFSKATLSL-----	PWV-----P-PDTK-----LSK4DTLRASSYISEL
Sp-Myor4 (SPU_016445)	N-----DPMNAERERRMTLRDAPFLRLRV-----	PSV-----P-PDTK-----LSK4DVLATTYISELM
Sp-Hand (SPU_017287)	E-----AVMGNKERRRRRTQSINTAFADURECI-----	PNV-----P-ADTK-----LSK4KTTLRLATSYIAILT
Sp-Sc1 (SPU_028093)	R-----RIFTNERRRRMQMNSAFSELRKLL-----	PCH-----P-VDKK-----LSK4ETLRLTIRYINELM
Sp-Ncs1 (SPU_009231)	Y-----RTAHATERRRIR/EA/NVAPTEURKLL-----	PTL-----P-PDKK-----LSKETETLRALICYIAIN
Sp-Sage (SPU_002448)	Q-----RTASDRERRRRMDINAFETLRILAKHQP-----	P-----GKK-----MSK4QALRPAFYIYNLE
Sp-Ac/Sc3 (SPU_022554)	L-----VQRNARERRRQLVNDGFIRLRKI-----	PTE-----P-RNKK-----LSR4KTTLRSAINYILELQ
Sp-Ac/Sc (SPU_028148)	F-----VARNERRRRR/KLVNHGFANRQQL-----	PNG-----A-INKK-----MSR4ETLRSAVSYIRQLQ
Sp-Ash3L (SPU_030149)	A-----IRRNERRRRR/RM/NIEG/VARREHL-----	PCD-----N-PEKR-----MSR4ETLRMAIRYIKLQ
Sp-MyoD (SPU_021119)	K-----RKAATLERRRRLRKVNFAFEALKRHTC-----	ANP-----NQR-----LPR4ETLRNAIEYIERLE
Sp-MyoD2 (SPU_006232)	K-----RRAATQERRRR/LCKVNSAFELKQRTC-----	SIP-----EQR-----MPR4TILRNAIQYIERLQ
Sp-E12 (SPU_016343)	R-----RHANNARERRR/RDINEAFKELGRMC-----	QLH-----LKQDK-----QTK4TILHSVSVITTE
Sp-Myc (SPU_003166)	R-----RANHVLERRRRRLDTSFFLRDQV-----	PBL-----A-SQER-----AAK4VTLKKATDYVHHL
Sp-Max (SPU_022163)	K-----RAHNAUERERRRDIKDSFSLRDSV-----	PNL-----QGEK-----ASRAQTLNKATDYIQMR
Sp-Mlx (SPU_005787)	R-----RNAHTAERRRRDAIKKGIEDQLVL-----	PTCQPPDQ-V-GSQK-----LSK4ATLQRSIDYIQILI
Sp-Mlx/Ip (SPU_008845)	G-----KLSHIAPERRRRNIRSGFDTSNMI-----	PSLA-----QNSAK-----ISK4NMVKGVEYTRQLQ
Sp-Ap4 (SPU_003179)	R-----REIANSERRRRMQSINAGFQSKLL-----	PHN-----EGEK-----LSK4ATLQQTADYICRME
Sp-Mad (SPU_006583)	S-----RSTHNEURRRRAHLRTCLERLKM-----	PL-----DGMPR-----HTLGLTLNKAQFIVDL
Sp-Mnt (SPU_026205)	I-----REVHNKLEKRRRAHLKDCPENRNTV-----	PNM-----EDKKAK-----TSNLSTLGRKALFQIVLT
Sp-Usf (SPU_014332)	R-----RATHNEVERRRRRDKINNVILDSKII-----	POC-----NIDHSK-----QGQSGGLSRKATDYIHLR
Sp-Mitf (SPU_008175)	K-----KDNHMDERRRRFNINDRIKELSTLI-----	PKH-----LDPQR-----QNK4TILKSSDYIRKML
Sp-SrebP (SPU_007253)	K-----RKSHNALEKRRYRKSINGRIEDKLMI-----	-----FKENKK-----VSKSTLQKCIDHVLGR
Sp-Ncoa (SPU_021430)	E-----KGRFPQKRRRCEKRSSTYIEELVEL-----	-----TASLNPESLNMK-----PDK4ALQETLSQIKRIK
Sp-Clock (SPU_017407)	K-----RKSRNAEKRRRDQFNVLQBELCSM-----	S-----TKTRK-----LDK4SAVLRATIHFLKAHN
Sp-Arnt (SPU_000129)	A-----RENHSEURRRRRNKMTAYITELSDM-----	PSC-----SALARK-----PDK4TLIRMAVSHMKLR
Sp-Bmal (SPU_027935)	A-----KQNHSELEKRRREKNTYIQELSAMV-----	PTC-----SSMSK-----LDK4TLIRMAVQHMKTLR
Sp-Trh (SPU_014249)	R-----KERSRDAASRRGKENYEFYDAKLL-----	PLP-----AAITSQ-----LDK4SIIRLTFLIGHMR
Sp-Sim1 (SPU_013962)	M-----KEKSKNAARTREKENAEFFELAKLL-----	PLP-----SAITQ-----LDK4SIIRLTFLYLMRA
Sp-Birc6 (SPU_001262)	R-----KEKSRNAASRRGKETEIFYDAHTL-----	PLA-----HNACQ-----LDK4GIMRLILSYLRSK
Sp-Ahr1 (SPU_005022)	N-----GTVRTNPSKRRDRDLNTELDRLASLL-----	PDF-----SDTVTK-----LDK4SLIRLSVSYLRNKS
Sp-Id (SPU_015374)	-----ERK4QLV-----	PTI-----P-RNRK-----VTR4VELLQHVIDYIQDLQ
Sp-Hey (SPU_009465)	R-----KKRRGILKRRRRDRINNSLTELRLRV-----	PAAFEQK-----GSAK-----LEK4EILQMTVDBHLKLIQ
Sp-Hairy2/4 (SPU_006814)	K-----KSSKPMLEKRRRARINDSLGLKALI-----	-----LEATNKDSSRHSK-----LEK4ADILEMTVKHLRNH
Sp-Hairy (SPU_006813)	F-----TSLKPMLEKRRRARINDSLGLKQNLV-----	-----LDALNKINPRHSK-----LEK4ADILEMTVYLRSLH
Sp-HesC (SPU_021608)	P-----RTAKHLTERRRRARINDSLGLKSM-----	FFV/KKD/ISRHPK-----MEK4ADILEMTVYLRKDVQ
Sp-Hey4 (SPU_015712)	R-----KSSKPMLEKRRRARINDSLGLKQTL-----	-----DALNPENNSTRQNK-----REK4ADILEQTVKRLVLQ
Sp-Ebf3 (SPU_004702)	-----SLTEPTIDYGFQRLQKVL-----	PRH-----PGDPR-----LPRFVLKRAADLVEAL

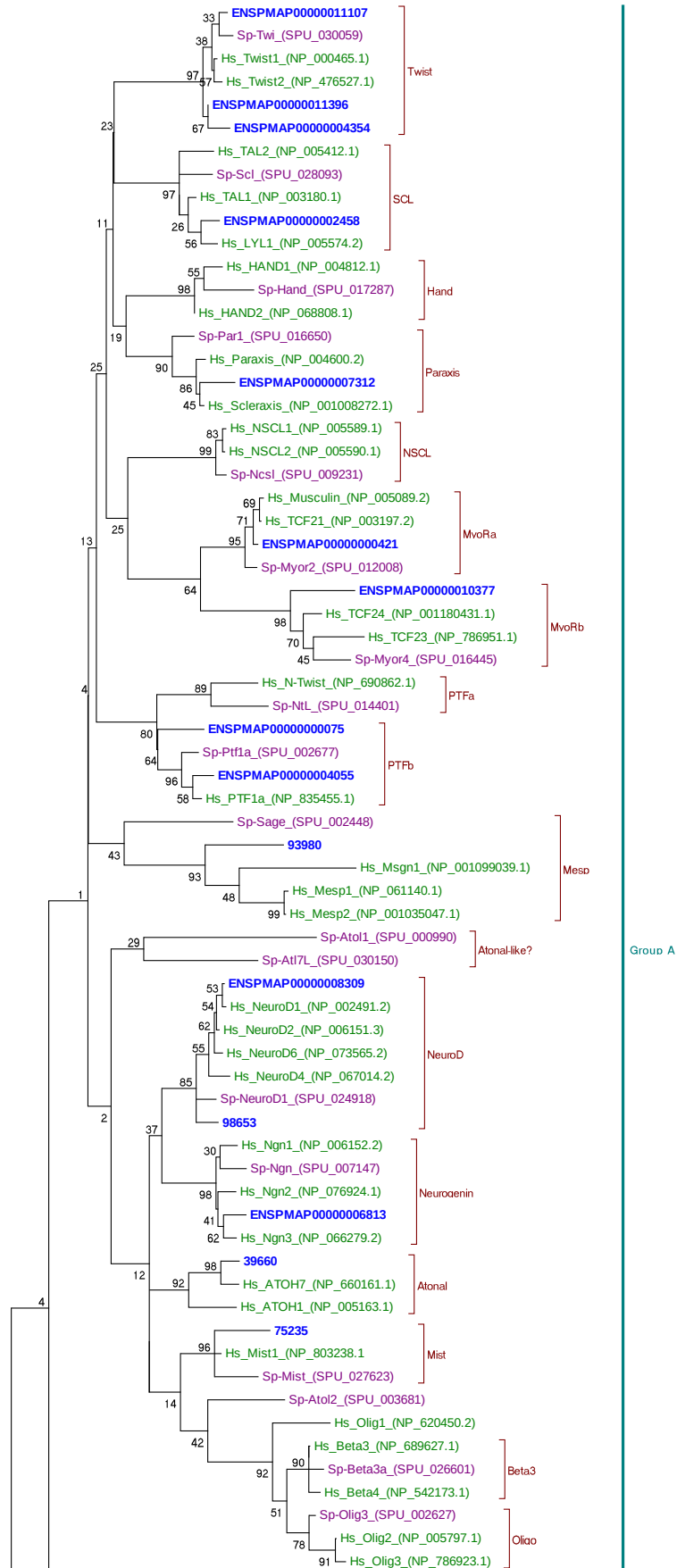
	90	100
ENSPMAF00000008309	-----	-----
98653	-----	-----
ENSPMAF00000006813	-----	-----
75235	-----	-----
39660	-----	-----
ENSPMAF000000011396	-----	-----
ENSPMAF000000004354	-----	-----
ENSPMAF000000011107	-----	-----
ENSPMAF000000004055	-----	-----
ENSPMAF000000000075	-----	-----
ENSPMAF000000007312	-----	-----
ENSPMAF000000000421	-----	-----
ENSPMAF000000010377	-----	-----
ENSPMAF000000002458	-----	-----
93980	-----	-----
ENSPMAF000000008927	-----	-----
ENSPMAF000000001063	-----	-----
ENSPMAF000000004198	-----	-----
ENSPMAF000000003372	-----	-----
ENSPMAF000000009801	-----	-----
ENSPMAF000000011193	-----	-----
ENSPMAF000000010427	-----	-----
ENSPMAF000000009950	-----	-----
ENSPMAF000000005560	-----	-----
ENSPMAF000000001203	-----	-----
ENSPMAF000000004888	-----	-----
ENSPMAF000000010532	-----	-----

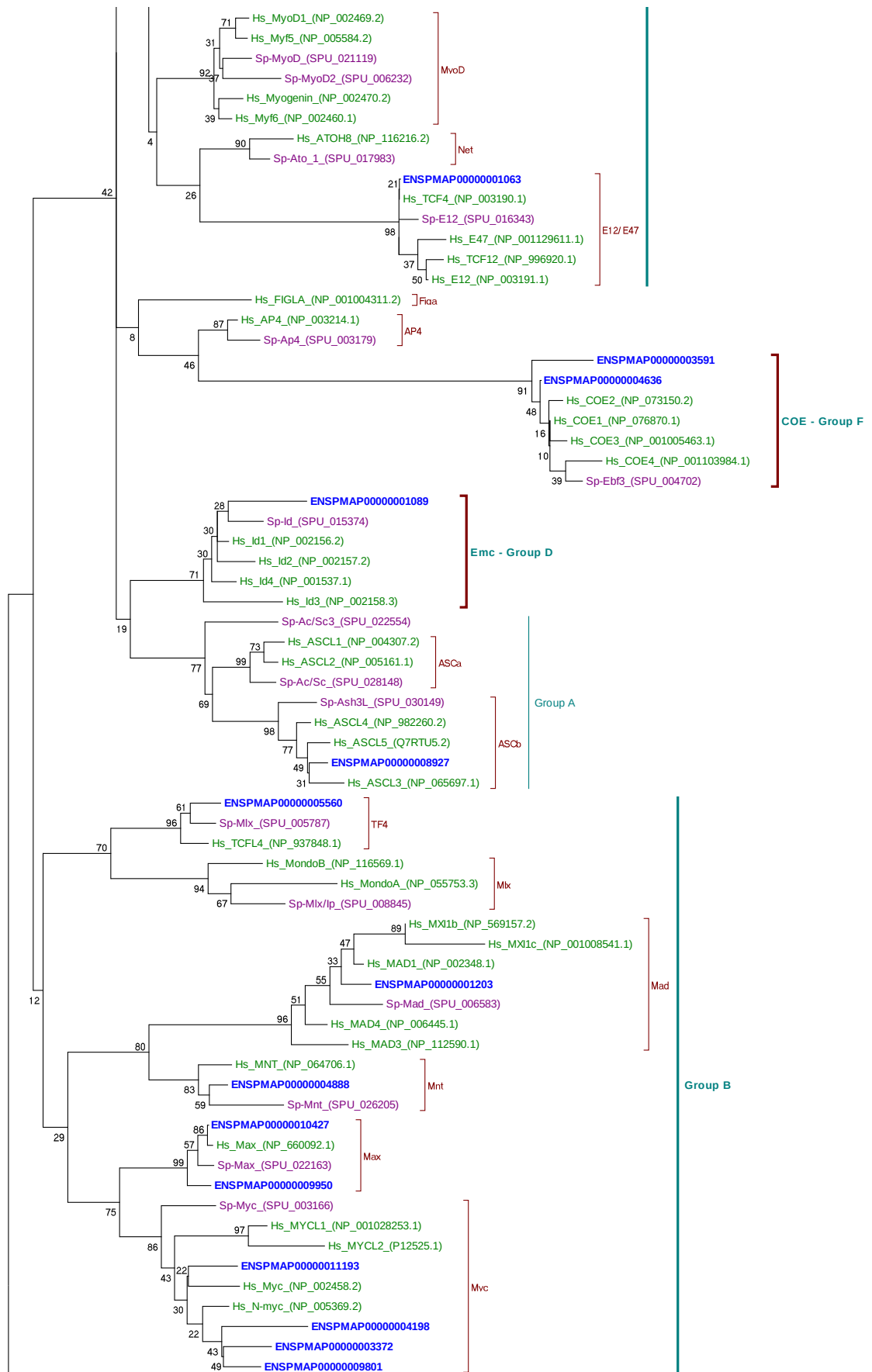
Figure 3-1. MAFFT alignment of *Petromyzon marinus*, human, and sea urchin (*Strongylocentrotus purpuratus*) HLH domains using sequences from Tables 3-1, 3-2, and 3-3 to construct the Neighbour Joining tree in Figure 3-2. Positions where at least 75% of amino acids are identical are shaded in blue.

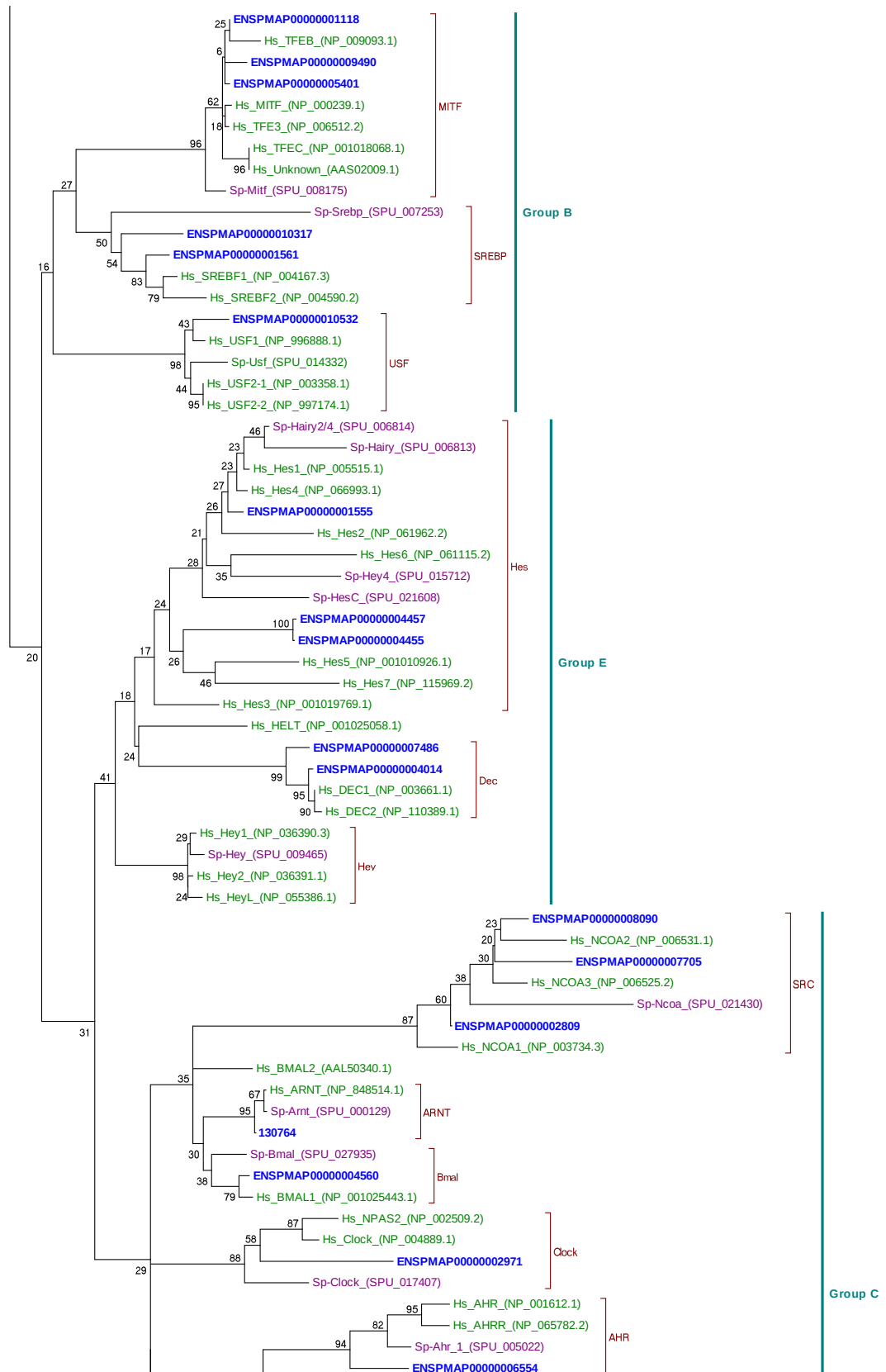
After the putative lamprey bHLH domains were identified, a Neighbor-joining (NJ) tree was constructed as described in Methods with the lamprey sequences and the full set of human and *Strongylocentrotus purpuratus* bHLH domains as described by Simionato et al (2007) (Figure 3-2).

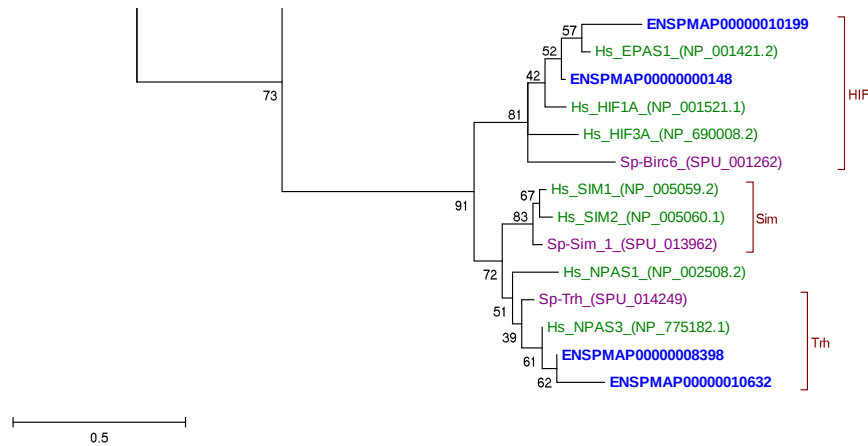
The tree recovered some higher-order groups of bHLH genes (i. e. groups A-F) as described by Ledent and Vervoort (2001), Ledent et al. (2002) and Simionato et al. (2007), although with very low support. For example, group A was recovered as a single clade (but with a bootstrap support value of 4) except for subfamilies *ASCa* and *ASCb*, which nonetheless grouped together with somewhat high support (bootstrap value of 69, Figure 3-2). They were recovered as a sister clade to group D (*Emc*) genes, but again, with very low support. On the other hand, group B was broken into two big clades, one of them containing *Myc*, *Mad*, *Mnt*, *Max*, *Mlx* and *TF4* subfamilies, and the other clade containing *USF*, *MITF* and *SREBP* subfamilies. *SRC*, *Figa* and *AP4* subfamilies, considered members of group B by Ledent et al. (2002), Simionato et al. (2007), Wang et al. (2009) and Zheng et al. (2009) were recovered outside the two main groups each in separate. All group C was recovered in a single clade but with a bootstrap value of 29. Similarly, group D and group E were recovered each as a single clade with bootstrap support values of 71 and 41, respectively. Finally, group F was also recovered as a single clade with a high support value of 91.

Figure 3-2. Phylogenetic analysis of isolated HLH domain predicted amino acid sequences from the *Petromyzon marinus* genome obtained from the Ensembl genome database. Unrooted NJ consensus tree showing the relationships of lamprey sequences (Ensembl or Pre!Ensembl accession numbers in blue) with respect to human (gene names and corresponding accession numbers in green) and sea urchin (gene names and corresponding SpBase accession numbers in purple) HLH amino acid sequences. Tree based on the Jones-Taylor-Thornton (JTT) matrix amino acid substitution model, 1000 bootstrap replicates and gamma distribution. Scale bar, amino acid substitutions per position.









In contrast to the very low support values for the higher-order groups (A-F), support values for each HLH gene subfamily are generally very high, reaching values of even 98-99 in subfamilies such as *Neurogenin*, *NSCL*, *ASCa* or *ASCb* (Table 3-4). However, clades of some subfamilies present very low bootstrap support values such as *Hairy/E(spl)*, *Mesp*, *SREBP*, *Bmal* and *Trh* with values of 17, 43, 50, 38 and 39, respectively (Table 3-4). In any case, most lamprey sequences clustered with specific subfamilies with high bootstrap support values (Table 3-1).

While a wide range of HLH subfamilies was identified, members of other subfamilies could not be retrieved (see Table 3-5). However, the tree in Figure 3-2 shows a fairly good distribution of lamprey HLH sequences among the different subfamilies of HLH transcription factors in all higher-order groups, suggesting that the retrieved sequences are somewhat close to the full set of *Petromyzon marinus* HLH genes.

Group	Subfamily	Bootstrap support value	Petromyzon Gene Prediction	Lamprey	Sea Urchin	Amphioxus	Human
A	NeuroD	85	ENSPMAP00000008309 (98653)	2	1	1	4
A	Neurogenin	98	ENSPMAP00000006813	1	1	1	3
A	Beta3	90	-	0?	1	1	2
A	Olig	78	-	0?	1	2	3
A	Mist	96	(75235)	1	1	1	1
A	Atonal	92	(39660)	1	0?	1	2
A	Net	90	-	0?	1	1	1
A	Delilah	-	-	0?	0	1	0
A	Twist	97	ENSPMAP00000011107 ENSPMAP00000011396 ENSPMAP00000004354	3	1	1	2
A	PTFa	89	-	0?	1	1	1
A	PTFb	64	ENSPMAP00000000075 ENSPMAP00000004055	2	1	3	1
A	Paraxis	90	ENSPMAP00000007312	1	1	2	2
A	MyoRa	95	ENSPMAP00000000421	1	1	4	2
A	MyoRb	98	ENSPMAP00000010377	1	1	1	2
A	Hand	98	-	1 ^a	1	1	2
A	SCL	97	ENSPMAP00000002458	1	1	1	3
A	NSCL	99	-	0?	1	1	2
A	Mesp	43	(93980)	1	1	1	3
A	ASCa	99	-	1 ^b	1(2?)	3	2
A	ASCb	98	ENSPMAP00000008927	1	1(2?)	1	3
A	MyoD	92	-	0?	2	4	4
A	E12/E47	98	ENSPMAP00000001063	1	1	1	4
A	Atonal-like?	29	-	0?	2	0	0
B	Figa	-	-	0?	0	1	1
B	Myc	86	ENSPMAP00000011193 ENSPMAP00000004198 ENSPMAP00000003372 ENSPMAP000000009801	4	1	1	4
B	Max	99	ENSPMAP00000010427 ENSPMAP000000009950	2	1	1	1
B	TF4	96	ENSPMAP00000005560	1	1	0	1
B	MLx	94	-	0?	1	1	2
B	AP4	87	-	0?	1	1	1
B	Mad	96	ENSPMAP00000001203	1	1	1	5
B	Mnt	83	ENSPMAP00000004888	1	1	1	1
B	USF	96	ENSPMAP00000010532	1	1	1	3
B	MITF	96	ENSPMAP00000001118 ENSPMAP000000009490 ENSPMAP000000005401	3	1	1	5
B	SREBP	50	ENSPMAP00000010317 ENSPMAP00000001561	2	1	1	2
B	SRC	87	ENSPMAP00000008090 ENSPMAP000000007705 ENSPMAP00000002809	3	1	1	3
C	Clock	88	ENSPMAP00000002971	1	1	1	2
C	ARNT	95	(130764)	1	1	1	1
C	Bmal	38	ENSPMAP00000004560	1	1	1	2
C	Trh	39	ENSPMAP00000008398 ENSPMAP00000010632	2	1	1	1
C	Sim	83	-	0?	1	1	2
C	HIF	81	ENSPMAP00000010199 ENSPMAP00000000148	2	1	1	4
C	AHR	94	ENSPMAP00000006554	1	1	1	2
D	Emc	71	ENSPMAP00000001089	1	1	1	4
E	Hey	98	-	0?	1	1	3
E	Hairy/E(spl)	17	ENSPMAP00000001555 ENSPMAP00000004457 ENSPMAP00000004455	3	4	11	7
E	Dec	99	ENSPMAP00000007486 ENSPMAP00000004014	2	0	?	2(3?)
F	COE	91	ENSPMAP00000003591 ENSPMAP00000004636	2	1	1	4
-	Orphan	-	-	0?	1?	6	0

Table 3-4. Summary of *Petromyzon marinus* HLH gene predictions. First column: Higher-order groups; Second column: HLH gene subfamilies; Third column: Bootstrap values supporting each HLH gene subfamily obtained from the NJ tree in Figure 3-2; Fourth column: Gene predictions from Ensembl or PreEnsembl (numbers in parentheses); Fifth-Eighth column: Number of HLH genes present in *Petromyzon marinus* (Lamprey), *Strongylocentrotus purpuratus* (Sea urchin), *Branchiostoma floridae* (Amphioxus) and *Homo sapiens* (Human), respectively, according to each HLH subfamily. “a” and “b”, genes not identified in Ensembl or PreEnsembl but described by other authors.

^a (Cerny et al., 2010)

^b (Häming et al., 2011)

A	B	C	D	E	F
Beta3	Figa	Sim		Hey	
Oligo	AP4				
Net	Myx				
Hand					
NSCL					
MyoD					

Table 3-5. List of HLH subfamilies for which no lamprey gene prediction was found either in PreEnsembl or Ensembl for *Petromyzon marinus* separated according to the higher-order group to which they belong to as described by Simionato et al. (2007).

3.6. Discussion

Analysis of gene prediction models derived from both Pre!Ensembl and Ensembl for *Petromyzon marinus* revealed that lampreys have a large repertoire of HLH gene subfamilies. Gene prediction models, however, are not completely accurate, as exemplified by models of *COE* genes (cloned in the present study from *Lampetra planeri*, described in detail in chapter 6). As expected, lampreys have a wide diversity of HLH gene subfamilies belonging to all main groups (i.e. A-F), with the exception of eleven subfamilies for which no models were recognised (Table 3-5). However, the fact that models for particular subfamilies were not retrieved is not enough evidence for the complete absence of those subfamilies in the lamprey (*P. marinus*) genome. Rather, the most likely explanation is the incomplete assembly of the lamprey genome. This can be best seen when one assembly (Pre!Ensembl) predicts models of genes belonging to distinct subfamilies such as *Mist*, *Atonal*, *Mesp* and *ARNT* that were not recognised by the current Ensembl assembly (Table 3-1, numbers in parentheses). Conversely, other subfamilies such as *COE* and *Neurogenin* were only predicted by the current Ensembl assembly but not from Pre!Ensembl. In either case, notable exceptions relevant for the present study in the context of the vertebrate nervous system are the lack of predictions from either assembly for members of the *Oligo* and *NSCL* subfamilies which are found in many metazoans (Simionato et al., 2007, 2008; Table 3-5) and have very specific roles in other vertebrates such as the specification of oligodendrocyte and motor neuron lineages (*Oligo*) or regulation of cerebellum development and patterning of rhombomere boundaries (*NSCL*) (Park et al., 2002).

The generally high bootstrap support values defining the grouping of most subfamilies also highlights the use of sea urchin as an outgroup. While support values grouping different subfamilies are very low, the interrelationships of all the subfamilies and higher-order groups are beyond the scope of this study. Instead, this study focuses on particular subfamilies whose members are expressed in different regions of the gnathostome nervous system. In spite of the low support values between different subfamilies obtained by Neighbour-Joining, this algorithm was preferred for phylogeny reconstruction rather than Maximum Likelihood (ML) or Bayesian Inference (BI) given that, for the large number of sequences used, it offers a fast –yet trustworthy– method for phylogeny reconstruction with high support values relevant to the present study (i.e. those values defining each HLH subfamily and their internal relationships). On the other hand, ML or BI are much slower algorithms and, while they are recognised as somewhat better algorithms than Neighbour-Joining, they do not necessarily improve support values with the type of alignment used in this chapter (i.e. with few positions).

As described in sections 1.5.2. and 1.5.3. of the Introduction, some of the most important subfamilies with a central role in nervous system development and neuronal differentiation are *ASC*, *Neurogenin*, *NeuroD*, *Atonal*, *Oligo*, *NSCL*, *Id*, *Hes*, *Hey* and *COE*. From this, genes belonging to these subfamilies were looked for for subsequent cloning from *Lampetra planeri* as described in the following chapters.

3.7. Summary of Lamprey HLH Genes

The lamprey genome has a wide repertoire of HLH genes belonging to specific animal HLH gene subfamilies that, in turn, belong to all higher-order groups of HLH genes, namely groups A-F. 51 models containing a predicted HLH gene were retrieved from *Petromyzon marinus* genome databases publically available. Among these, HLH gene predictions include genes belonging to *NeuroD*, *Neurogenin*, *Atonal*, *ASCL*, *Hairy/E(spl)*, *Dec* and *COE*, which are well-known in vertebrates to regulate the development of the nervous system. However, other HLH gene subfamilies, including *Oligo*, *NSCL* and *Hey*, important for the development of the nervous system could not be retrieved, most likely due to an incomplete assembly of the lamprey genome. Together, the *in silico* identification of several neural HLH genes in the lamprey facilitates the further characterisation of their expression patterns to better understand the lamprey nervous system.

4. *LAMPETRA PLANERI HAIRY/E(SPL) GENES*

4.1. Introduction

Structure and function of *Hairy/E(spl)* genes: As mentioned earlier, Group E HLH genes include genes of the *Hairy/E(spl)* (also called *Hes* genes) and *Hey* subfamilies. In the present study, *Dec* genes are proposed to constitute an independent subfamily from *Hes* and *Hey* genes (discussed in detail in section 4.6.3). Group E proteins have three conserved regions that confer transcriptional functions, namely the bHLH domain, the Orange domain and the diagnostic WRPW or YRPW motif located at the C-terminal end of *Hes* or *Hey* proteins, respectively (Figure 4-1A) (Kageyama et al., 2007, 2008). Given that no *Hey* or *Dec* genes were experimentally analysed in the present study, *Hes* genes are the main focus of this chapter.

As with other bHLH genes, the basic region in *Hes* proteins serves to bind DNA and it contains a conserved proline residue in the middle. The HLH region functions in dimerisation with other HLH factors forming homo- or heterodimers with HLH proteins of the closely-related *Hey* subfamily (Figure 4-1B) (Kageyama et al., 2007, 2008). As homo- or heterodimers, *Hes* proteins bind to the so-called “N boxes” with a consensus sequence CACNAG or to “C sites” with the consensus sequence CACG(C/A)G in promoter regions of target genes (Akazawa et al., 1992; Kageyama et al., 2007, 2008; Sasai et al., 1992; Takebayashi et al., 1994). However, *Hes1* also recognises “E boxes” though only weakly, and *Hes2* preferentially binds to “E boxes” than “N boxes” (Ishibashi et al., 1993; Sasai et al., 1992). Binding of dimers to DNA triggers an “active repression” mechanism whereby *Hes* homodimers or *Hes/Hey* heterodimers recruit

Groucho (*Drosophila*)-related co-repressors (TLE1-4 in mammals) by –at least for Hes proteins– the conserved WRPW motif (Figure 4-1B) (Fisher et al., 1996; Grbavec and Stifani, 1996). This is in contrast with a “passive repression” mechanism where Hes proteins dimerise with bHLH activators such as mouse Mash1 and E47 preventing them from binding to “E boxes” in promoter regions of target genes to activate their transcription (Figure 4-1C) (Kageyama et al., 2007, 2008). In the absence of Hes factors, Mash1 and E47 can dimerise and promote transcription of target genes (Figure 4-1D).

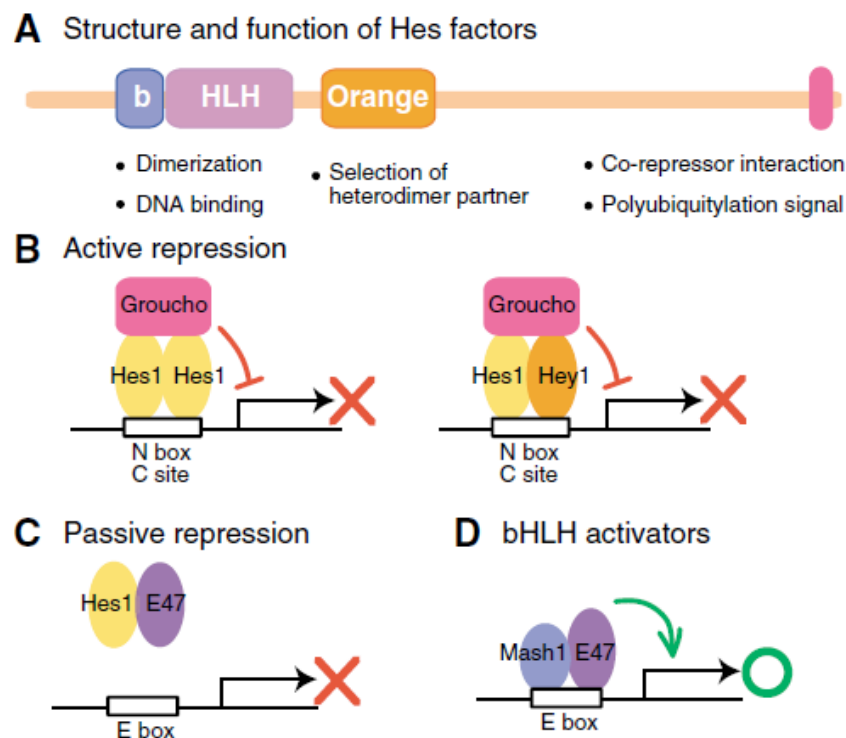


Figure 4-1. Structure and function of Hes transcription factors. (A) Basic structure of Hes transcription factors showing four conserved regions: the basic domain (blue) which functions in DNA binding, the HLH domain (purple) that serves in dimerisation with other bHLH factors, the Orange domain (orange) involved in selection of dimerisation partners, and the WRPW motif (pink) that binds to Groucho-type co-repressors of transcription and possesses degradation signals for the proteasome. (B) Hes factors can form homo- or heterodimers with members of the closely related Hey subfamily. As dimers, they bind to “N boxes” or “C sites” in promoter regions of target genes and recruit Groucho-type co-repressors to actively inhibit their transcription. (C) Additionally, Hes factors can also dimerise with bHLH transcription activators such as E47 and Mash1 (mouse) preventing them from binding to their target DNA sequences known as “E boxes”. (D) In the absence of Hes factors, these activators can dimerise and

activate transcription of target genes to specify particular cellular lineages. Taken from Kageyama et al. (2007).

The Orange domain is adjacent to the bHLH domain downstream and has two amphipathic helices (Eastwood et al., 2011; Kageyama et al., 2007). Studies on human Hey1, a group E-family protein of the Hey subfamily that also possesses an Orange domain (see Chapter 3), have shown that the Orange domain is able to form dimers in a double-hairpin conformation (Eastwood et al., 2011). Its exact role is not well understood. Early studies suggested that it has an important role in dimerisation (Taelman et al., 2004). However, recent studies have suggested that the Orange domain has a structural role rather than in selection of bHLH heterodimer partners (Eastwood et al., 2011; Kageyama et al., 2007), making it an extended dimerisation domain additional to the HLH domain. Additionally, it may also serve to interact with other proteins that regulate its activity (Van Wayenbergh et al., 2003).

Expression of Hes genes in the vertebrate nervous system: Given the large number of members of the *Hes* gene subfamily in vertebrates, only expression of murine *Hes* genes, namely, *Hes1*, *Hes2*, *Hes3*, *Hes5*, *Hes6* and *Hes7* (with the addition of *XHes2* as explained below) will be described in this section to be later used as a reference for information from other vertebrate *Hes* homologues in other sections. Additionally, only expression related to the nervous system is considered, even though several *Hes* genes are also expressed in some mesodermal and endodermal tissues.

Rat *Hes1* is expressed in the ventricular zone of the entire developing CNS, but is excluded from outer layers of differentiating/differentiated neurons and glial cells (Sasai et al., 1992). Also, mouse *Hes1* is expressed in retinal progenitor cells, olfactory placodes/epithelium, cranial ganglia and DRG (Cau et al., 2000; Ishibashi et al., 1995; Tomita et al., 1996).

Rat *Hes2* expression has not been well characterised; negligible –if any– expression has been observed in the brain and spinal cord of rat embryos (Ishibashi et al., 1993). However, expression of a *Xenopus laevis* HLH gene, *XHes2*, which possesses highest sequence similarity to mouse and human *Hes2*, has been described (Sölter et al., 2006). Since the beginning of neural plate formation (stage 13.5) *XHes2* is expressed in the otic placodes, and its expression becomes much stronger as development proceeds, persisting in the dorsal region of the otic vesicles as far as stage 38 (Sölter et al., 2006). From stage 21, expression appears in presumptive territory of the retina, which considerably intensifies at later stages as the retina develops (Sölter et al., 2006). Weak expression is also found in the olfactory placodes. Additionally, *XHes2* is expressed superficially in two lateral and two medial parallel stripes of proliferating neuronal precursors in the posterior neuroectoderm, the latter stripes also continuing –and expanding– into the anterior neuroectoderm between the midline of the embryo and the otic placodes (Sölter et al., 2006). Later, the posterior medial stripes fuse at the midline constituting a posterior (hindbrain) domain, which separates from an anterior (forebrain) domain where these stripes do not fuse (Sölter et al., 2006). At this time (stage 38), *XHes2* expression appears in the lateral line system. Overall, *XHes2* is mainly expressed in presumptive and developing sensory organs.

Mouse *Hes3* is particularly interesting as it is expressed in the hindbrain in rhombomeres 1, 2, 4, 6 and 7, leaving rhombomeres 3 and 5 unmarked at E8.5. *Hes3* is also expressed in the dorsal midbrain and in the neural plate in E8.5 embryos (Lobe, 1997). Later at E9.5, *Hes3* expression extends along the entire length of the neural tube and now appears in the forebrain (Lobe, 1997). At E11.5, *Hes3* is expressed in the dorsal and –weakly– ventral forebrain as well as in the lateral walls as occurs in the midbrain. In the midbrain transcripts are also seen in the dorsal region. In the hindbrain and spinal cord expression is found in the floorplate, and in the spinal cord *Hes3* is also expressed in the ventral marginal zone and later in the dorsal-most region. All expression sites in the CNS are mostly seen in the ventricular zone. At this stage, *Hes3* expression extends to the peripheral nervous system into the olfactory epithelium, retina, trigeminal, facial and acoustic ganglia, otic vesicle and dorsal root ganglia among others (Lobe, 1997).

Mouse *Hes5* is strongly expressed in the forebrain, midbrain and hindbrain but absent from the forebrain-midbrain boundary and MHB (Hatakeyama et al., 2004). It is also expressed in the ventral half and the dorsal-most part of the spinal cord (Ishibashi et al., 1995). It is also expressed in the olfactory placodes/epithelium (Cau et al., 2000). On the other hand, rat *Hes5* is found in the entire CNS in the ventricular zone and in the neural retina, as well as in the DRG (Akazawa et al., 1992).

Finally, mouse *Hes7* seems to be expressed only in presomitic and somitic mesoderm (Bessho et al., 2001).

Developmental roles of *Hes* genes: One of the most important roles of *Hes* genes in nervous system development is the maintenance of neural stem cells by actively inhibiting neurogenesis (Ohtsuka et al., 2001). *Hes1* has been shown to inhibit progression of the neuronal and glial lineage by maintaining proliferating neural precursors in the ventricular zone and inhibiting their differentiation and migration towards the marginal zone (Ishibashi et al., 1994, 1995). Also, *Hes* genes regulate the appropriate timing of cell differentiation by delaying it until sufficient numbers and cell types are generated (Hatakeyama et al., 2004). *Hes* genes are well known to function as repressors of transcription and this inhibition usually entails the inhibition of activity of HLH members of the ASC (Akazawa et al., 1992; Ishibashi et al., 1993, 1995; Sasai et al., 1992; Takebayashi et al., 1994), *Neurogenin* (Cau et al., 2000; Geling et al., 2004; Murai et al., 2011; Takke et al., 1999), *NeuroD* (Murai et al., 2011), *Atonal* and *COE* (Geling et al., 2004) gene families, as well as *E47* (Akazawa et al., 1992; Sasai et al., 1992) which is required by at least ASC, Neurogenin and Atonal proteins to form heterodimers (Akazawa et al., 1995; Gradwohl et al., 1996).

Some *Hes* genes are known to be direct effectors of Notch signalling and their expression is generally upregulated by this pathway. One such gene, *Hes5*, is upregulated by Notch signalling in the mouse spinal cord. However, in contrast to common belief, not all *Hes* genes are upregulated by Notch signalling. *Hes1* and *Hes3* are not direct effectors of Notch signalling (Lütolf et al., 2002; de la Pompa et al., 1997). Some zebrafish *Hes* genes are also upregulated by, and are effectors of, Notch signalling such as *Hes5e* (Takke et al., 1999). However, other zebrafish *Hes* genes are not upregulated by Notch, and instead are downregulated, such as *Her3*, a *Hes3* homolog (Hans et al., 2004).

A different *Hes* gene, *Hes6*: Among the *Hes* genes, *Hes6* shows distinct properties in its sequence, expression patterns and function. There are two subgroups of *Hes6* genes, *Hes6-1* and *Hes6-2*. One gene of each *Hes6* subgroup has been identified in chicken and *Xenopus* genomes, while only one *Hes6* gene (corresponding to *Hes6-1*) seems to be present in mouse and human (Vilas-Boas and Henrique, 2010). In contrast, four zebrafish *Hes6* genes, two of each subgroup, have been identified (Chung et al., 2011), most likely as a consequence of an independent whole-genome duplication in the teleost lineage as explained in section 1.6. All *Hes6* genes have the Orange domain, the WRPW motif at the C-terminal end and the conserved proline residue in the basic region (Bae et al., 2000; Koyano-Nakagawa et al., 2000; Pissarra et al., 2000; Vasiliauskas and Stern, 2000). However, the loop region of *Hes6-1* proteins is shorter than all other *Hes* proteins including *Hes6-2*, which also seems to affect their DNA binding properties as they do not recognise neither N nor E boxes (Bae et al., 2000; Koyano-Nakagawa et al., 2000; Vilas-Boas and Henrique, 2010).

In contrast to other *Hes* genes, *Hes6* expression in the nervous system is correlated with neurogenesis and is generally absent from neuronal precursors. Mouse *Hes6* is expressed in the dorsal telencephalon, diencephalon, midbrain, hindbrain (seen as longitudinal stripes of varying intensities along the dorsal-ventral axis) and spinal cord (most prominently in the ventral half and dorsal-most spinal cord), in the trigeminal, otic (otic vesicle), geniculate, petrosal and nodose placodes/ganglia, in the dorsal root ganglia, sympathetic ganglia, olfactory placode/epithelium and in the optic stalk, lens and retina (Bae et al., 2000; Koyano-Nakagawa et al., 2000; Pissarra et al., 2000; Suzuki et al., 2003; Vasiliauskas and Stern, 2000). Expression in the CNS is first observed in the ventricular zone and later in the subventricular and marginal zones,

unlike *Hes1* which is only expressed in precursor cells in the ventricular zone (Bae et al., 2000; Koyano-Nakagawa et al., 2000; Pissarra et al., 2000). In addition, *Hes6* expression usually follows that of *neurogenins*, overlaps with that of *NeuroD*, and precedes that of neuronal differentiation markers such as *N-tubulin* (Koyano-Nakagawa et al., 2000).

Chicken *cHes6-1* is expressed throughout the nervous system (Vilas-Boas and Henrique, 2010). In the spinal cord, *cHes6-1* is initially expressed both in the ventricular and marginal zones, but later its expression is restricted to the subventricular zone (Vilas-Boas and Henrique, 2010). Expression in the spinal cord generally progresses from the basal to the alar plate (Vilas-Boas and Henrique, 2010). Also, as in mouse, expression of *cHes6-1* generally follows that of *neurogenins*, coincides with *NeuroD*, but precedes that of the neuronal differentiation marker Tuj-1 (Vilas-Boas and Henrique, 2010). *cHes6-1* is also expressed in cranial ganglia and DRG (Vilas-Boas and Henrique, 2010). On the other hand, *cHes6-2* is initially expressed in the ventricular zone of the spinal cord but it is later expressed in neurons that start to differentiate (Fior and Henrique, 2005; Vilas-Boas and Henrique, 2010, 2010).

Xenopus Hes6-1 is expressed along the entire neural tube and in the presumptive eye (Koyano-Nakagawa et al., 2000). In the CNS, *Xenopus Hes6-1* is mildly expressed in the ventricular zone but strongly in the subventricular zone along the entire dorsal-ventral axis, while it is absent from the marginal zone (Koyano-Nakagawa et al., 2000).

In agreement with their expression patterns, mouse and *Xenopus Hes6* promote – rather than inhibit– neurogenesis. *Hes6* acts in an indirect positive-feedback loop with neurogenin and achaete-scute by binding to other *Hes* proteins that inhibit proneural genes and, in this way, inactivating their inhibitory role (Koyano-Nakagawa et al., 2000). *Mash1* also binds to the *Hes6* promoter (Poitras et al., 2007). Mammalian *Hes6* has also been shown to inhibit astrocyte differentiation in the cerebral cortex and independently promotes cortical neurogenesis (Belanger-Jasmin et al., 2007; Gratton et al., 2003; Jhas et al., 2006). In agreement with its inability to bind N- or E- boxes, *Hes6* DNA binding is not necessary for its neurogenic and anti-gliogenic roles in the cerebral cortex (Jhas et al., 2006). Furthermore, *Hes6* is able to suppress the ability of *Hes1* to mediate transcriptional repression by forming heterodimers with *Hes1*, thereby preventing *Hes1* from binding to N boxes and having a poor interaction with Groucho/TLE co-repressors (Bae et al., 2000; Gratton et al., 2003). *Hes6* may also render *Hes1* prone to proteolytic degradation upon dimerization (Gratton et al., 2003). In addition, *Hes6* prevents *Hes1* from binding to E boxes, which in normal conditions would inhibit transcription promoted by *Mash1-E47* heterodimers, hence acting indirectly as a positive (neurogenic) regulator of *Mash1-E47* (Bae et al., 2000). Thus, *Hes6* seems to counteract the inhibitory function of *Hes1*. By inhibiting *Hes1*-mediated transcriptional repression, *Hes6* promotes the differentiation of cortical neurons, a process normally inhibited by *Hes1* (Gratton et al., 2003). Also, *Hes6* is not a target of Notch signalling as other –inhibitory– *Hes* genes (Koyano-Nakagawa et al., 2000), and in chicken *cHes6-1* is rather downregulated by this pathway (Vilas-Boas and Henrique, 2010).

As mouse *Hes6*, chicken *cHes6-1*, which is coexpressed with *neurogenin1* and *neurogenin2* in some cells of the spinal cord, (Vilas-Boas and Henrique, 2010). *cHes6-1* regulates expression of *cHes5-1*, *cHes5-2* and *cHes5-3* by binding to their promoter sequences and preventing their auto-inhibition. In contrast, *cHes6-2* actively represses transcription of *cHes5-1*, *cHes5-2* and *cHes5-3* relieving differentiating neurons from Notch activity. The difference in their activities seems to lie in the shorter loop of *cHes6-1* compared to *Hes6-2* proteins. Also, forced expression of *cHes6-2* rescues loss of neurogenesis in the spinal cord upon inhibition of *cyclin D1* (Lukaszewicz and Anderson, 2011).

Finally, in *Xenopus*, inhibition of *Hes6-1* by morpholino oligonucleotides causes a failure of neuronal differentiation and reduction of transcription of *NeuroD* and *neurogenin2* (Murai et al., 2011). Conversely, injection of *Hes6-1* mRNA increases expression of *N-tubulin* and *NeuroD*, indicating that it is required for primary neurogenesis (Murai et al., 2011). Furthermore, *Xenopus* *Hes6-1* is required for the neurogenic activity of *neurogenin2* and *NeuroD* (Murai et al., 2011). Promotion of neurogenesis does not require DNA binding by *Hes6-1* and *Hes6-1* blocks the antineurogenic function of *Xhair1* by direct protein-protein interaction (Murai et al., 2011). Groucho/TLE proteins are dispensable for the neurogenic function of *Xenopus* *Hes6-1* to some extent only as full neurogenic activity of *Hes6-1* requires Groucho/TLE binding (Murai et al., 2011).

4.2. Aims

The aim of this chapter is to characterise the sequence and expression patterns of *Lampetra planeri* genes belonging to the *Hairy/E(spl)* (*Hes*) gene subfamily based on their identification *in silico* in the *Petromyzon marinus* genome (Chapter 3). This chapter is intended to answer the following questions:

- Does *L. planeri* have *Hes* genes as *P. marinus*? Do they have the diagnostic domains and residues of the *Hes* subfamily in their sequence?
- Provided the recovery of *L. planeri Hes* genes, where are *L. planeri Hes* genes expressed? Are they found in the nervous system as in gnathostomes? Are *L. planeri Hes* genes expressed in the ventricular zone of the CNS as in gnathostomes?
- Are *L. planeri Hes* genes more similar to “traditional” *Hes* genes (i.e. *Hes1*, *Hes3*, *Hes5*) or are they more similar to gnathostome *Hes6* genes both in their sequence and expression patterns?

4.3. Hypothesis

Vertebrate *Hes* genes are known to be well conserved in gnathostomes both at the sequence level and in their expression patterns in the nervous system. Therefore, predictions are that lampreys have more than one *Hes* homolog and that these genes will present –apart from the HLH domain– the conserved Orange domain downstream of the HLH domain, and the diagnostic WRPW motif at the C-terminal end and a conserved proline in the basic region of the HLH domain. At the expression level, lamprey *Hes*

homologues are predicted to be expressed mostly in the ventricular region of the neural tube, although they can be also expected to be expressed in different parts of the peripheral nervous system or in tissues derived from mesoderm or ectoderm as in mouse *Hes1*. The time frame of expression is also predicted to occur throughout neurogenesis.

4.4. Objectives

- Isolate homologous sequences of gnathostome *Hes* genes from the *L. planeri* genome.
- Perform phylogenetic analyses to study the phylogenetic relationships of the isolated sequences with respect to other animal (Deuterostome) *Hes* genes.
- Characterise the expression patterns of the isolated sequences in *L. planeri* embryos at different developmental stages by *in situ* hybridisation.
- Combine both expression and phylogenetic data to identify differences and similarities of expression patterns mapped into a phylogenetic context and in this way identify evolutionary trends of this family in vertebrates.

4.5. Results

4.5.1. Isolated Sequences

Fragments of two *Hairy/E(spl)* (*Hes*) homologues, named *LpHes-A* and *LpHes-B*, were isolated from *L. planeri* genomic DNA. The *LpHes-A* fragment is 937 bp long and has a near-identical sequence to -provisionally named- *PmHes-A* (Figure 4-2). The first 253 bp (~27% of total length) of the *LpHes-A* fragment correspond to coding sequence and the stop codon when compared to *PmHes-A*, whereas the remaining 684 bp (~73% of total length) correspond to a portion of the 3' UTR sequence of *PmHes-A*. On the other hand, the *LpHes-B* fragment is 700 bp long and has a near-identical sequence to -provisionally named- *PmHes-B* (Figure 4-3). The first 163 bp (~23.3% of total length) of the *LpHes-B* fragment correspond to coding sequence and the stop codon when compared to *PmHes-B*, whereas the remaining 537 bp (~76.7% of total length) correspond to a portion of the 3' UTR sequence of *PmHes-B*.

The fact that both fragments possess near-identical sequences to UTR portions of *P. marinus* cDNAs indicates that these fragments correspond to *L. planeri* cognate *Hes* homologs. Furthermore, when both *L. planeri* fragments are aligned with both *P. marinus* *Hes* sequences it can be observed that all four sequences present near-identical sequences in the coding region, whereas sequences of *PmHes-A/LpHes-A* and *PmHes-B/LpHes-B* are substantially different in the UTR region (Figure 4-4). This also shows that probes synthesised from both *L. planeri* fragments for *in situ* hybridisation (see below) are highly unlikely to recognise each other.

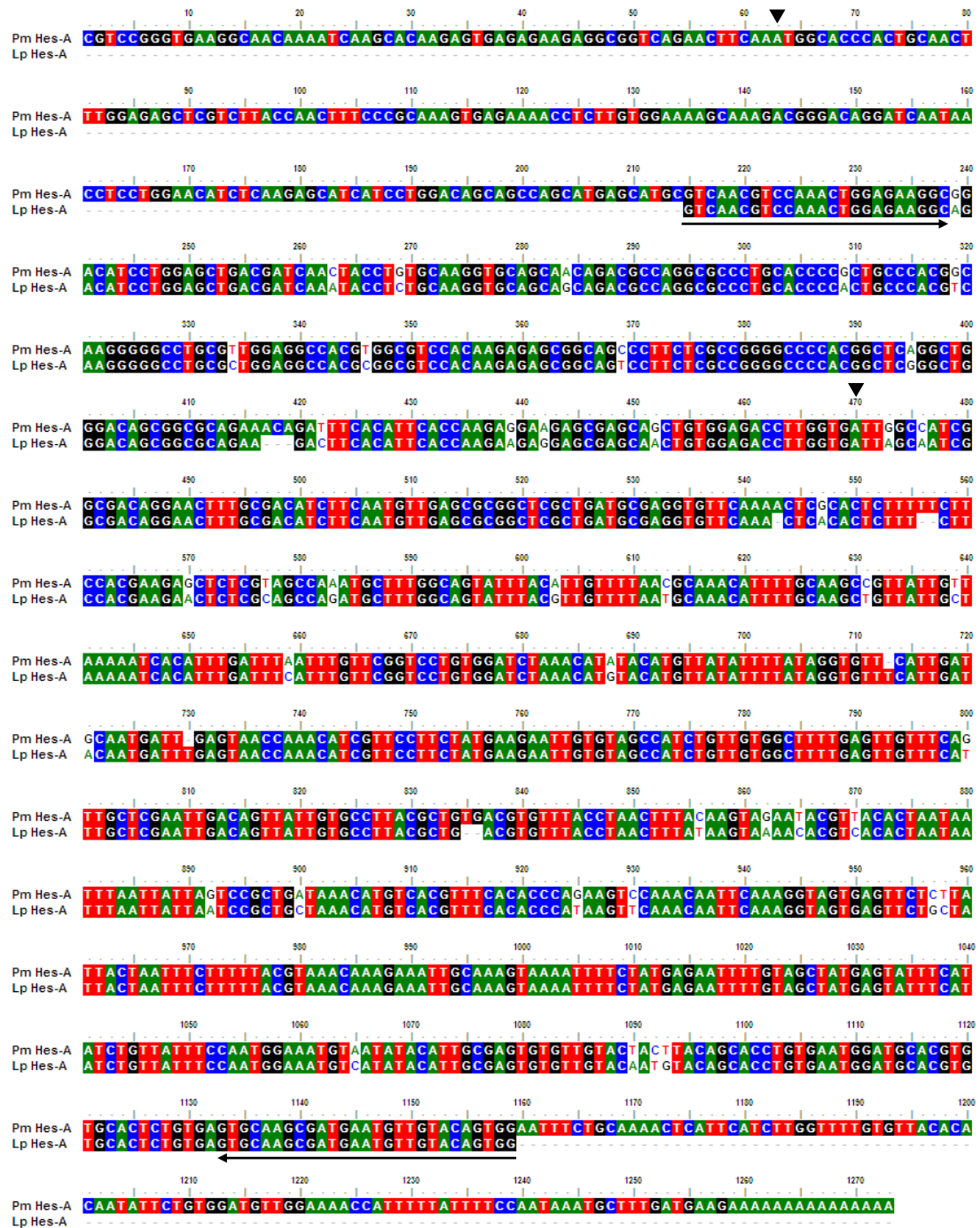


Figure 4-2. Nucleotide alignment of the cloned fragment of *LpHes-A* with *PmHes-A*. Arrowheads indicate the boundaries of the open reading frame (ORF) of *PmHes-A*. Arrows mark the primer sequence according to *PmHes-A*. Note the near-exact sequence of *LpHes-A* with *PmHes-A* both at the coding region and the 3' UTR.

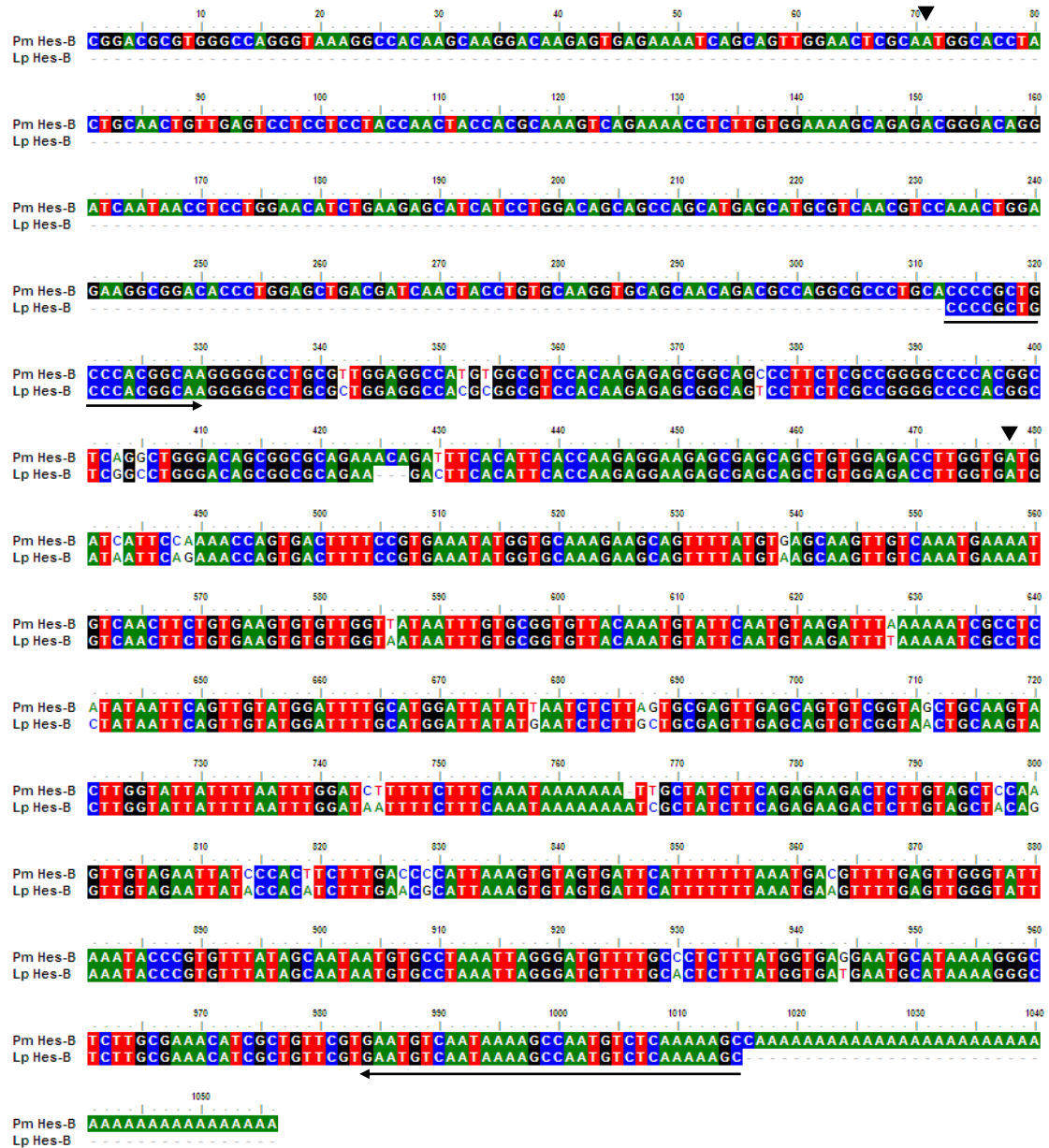


Figure 4-3. Nucleotide alignment of the cloned fragment of *LpHes-B* with *PmHes-B*. Arrowheads indicate the boundaries of the ORF of *PmHes-B*. Arrows mark the primer sequence according to *PmHes-B*. Note the near-exact sequence of *LpHes-B* with *PmHes-B* both at the coding region and the 3' UTR.

		10	20	30	40	50	60	70	80
Pm Hes-A	-----	CGTCGGG	GTGAAGCA	CAATCAAG	CAACAGTCACA	ASAGCGGT	CAAACTTCA	AATGGCACCA	A
Lp Hes-A	-----	CGTCGGG	GTGAAGCA	CAATCAAG	CAACAGTCACA	ASAGCGGT	CAAACTTCA	AATGGCACCA	A
Pm Hes-B	-----	CGTCGGG	GTGAAGCA	CAATCAAG	CAACAGTCACA	ASAGCGGT	CAAACTTCA	AATGGCACCA	A
Lp Hes-B	-----	CGTCGGG	GTGAAGCA	CAATCAAG	CAACAGTCACA	ASAGCGGT	CAAACTTCA	AATGGCACCA	A
		90	100	110	120	130	140	150	160
Pm Hes-A	-----	CTGCAACT	TTGAGAG	CTCTCT	TACCAACTTT	CCCGCAAGT	ACAAAACCT	CTTGTGCAAAAGCA	ACACGGCACAGG
Lp Hes-A	-----	CTGCAACT	TTGAGAG	CTCTCT	TACCAACTTT	CCCGCAAGT	ACAAAACCT	CTTGTGCAAAAGCA	ACACGGCACAGG
Pm Hes-B	-----	CTGCAACT	TTGAGAG	CTCTCT	TACCAACTTT	CCCGCAAGT	ACAAAACCT	CTTGTGCAAAAGCA	ACACGGCACAGG
Lp Hes-B	-----	CTGCAACT	TTGAGAG	CTCTCT	TACCAACTTT	CCCGCAAGT	ACAAAACCT	CTTGTGCAAAAGCA	ACACGGCACAGG
		170	180	190	200	210	220	230	240
Pm Hes-A	-----	ATCAATAACCT	CCTGCAACATCT	CAACAGCATC	ATCCTGACAGCAGCCAGCATCAGCATGCGTCAACGTCCAAACTGGA				
Lp Hes-A	-----	ATCAATAACCT	CCTGCAACATCT	CAACAGCATC	ATCCTGACAGCAGCCAGCATCAGCATGCGTCAACGTCCAAACTGGA				
Pm Hes-B	-----	ATCAATAACCT	CCTGCAACATCT	CAACAGCATC	ATCCTGACAGCAGCCAGCATCAGCATGCGTCAACGTCCAAACTGGA				
Lp Hes-B	-----	ATCAATAACCT	CCTGCAACATCT	CAACAGCATC	ATCCTGACAGCAGCCAGCATCAGCATGCGTCAACGTCCAAACTGGA				
		250	260	270	280	290	300	310	320
Pm Hes-A	-----	CAAGGCGCACAT	CCTGAGCTCACCATCA	TAACCTTGCAAGGTGCAGCA	ACACAGCCAGGCGCCCTGCACCCCGCTG				
Lp Hes-A	-----	CAAGGCGCACAT	CCTGAGCTCACCATCA	TAACCTTGCAAGGTGCAGCA	ACACAGCCAGGCGCCCTGCACCCCGCTG				
Pm Hes-B	-----	CAAGGCGCACAT	CCTGAGCTCACCATCA	TAACCTTGCAAGGTGCAGCA	ACACAGCCAGGCGCCCTGCACCCCGCTG				
Lp Hes-B	-----	CAAGGCGCACAT	CCTGAGCTCACCATCA	TAACCTTGCAAGGTGCAGCA	ACACAGCCAGGCGCCCTGCACCCCGCTG				
		330	340	350	360	370	380	390	400
Pm Hes-A	-----	CCCACGGCAAGGGGGCCTGCGT	TGAGGCCACGGGCGTCCACAACACAGCGGAGC	CCTTCTCGCCGGGGCCCCACGGC					
Lp Hes-A	-----	CCCACGGCAAGGGGGCCTGCGT	TGAGGCCACGGGCGTCCACAACACAGCGGAGC	CCTTCTCGCCGGGGCCCCACGGC					
Pm Hes-B	-----	CCCACGGCAAGGGGGCCTGCGT	TGAGGCCACGGGCGTCCACAACACAGCGGAGC	CCTTCTCGCCGGGGCCCCACGGC					
Lp Hes-B	-----	CCCACGGCAAGGGGGCCTGCGT	TGAGGCCACGGGCGTCCACAACACAGCGGAGC	CCTTCTCGCCGGGGCCCCACGGC					
		410	420	430	440	450	460	470	480
Pm Hes-A	-----	TCAGGCTGGCACAGCGGCGCACAAACACA	TTACATTACCAACACACA	CAGCCAGCAGC	CTGTGCACACCTTGGTCATT				
Lp Hes-A	-----	TCAGGCTGGCACAGCGGCGCACAAACACA	TTACATTACCAACACACA	CAGCCAGCAGC	CTGTGCACACCTTGGTCATT				
Pm Hes-B	-----	TCAGGCTGGCACAGCGGCGCACAAACACA	TTACATTACCAACACACA	CAGCCAGCAGC	CTGTGCACACCTTGGTCATT				
Lp Hes-B	-----	TCAGGCTGGCACAGCGGCGCACAAACACA	TTACATTACCAACACACA	CAGCCAGCAGC	CTGTGCACACCTTGGTCATT				
		490	500	510	520	530	540	550	560
Pm Hes-A	-----	GGCCATCGGCCACAG--CAAGTTTTCGACATCTTCAATGTTTACG	GCAGCTCGCTGATGCAAGT	TTTCAAACTCGGA					
Lp Hes-A	-----	GGCCATCGGCCACAG--CAAGTTTTCGACATCTTCAATGTTTACG	GCAGCTCGCTGATGCAAGT	TTTCAAACTCGGA					
Pm Hes-B	-----	GGCCATCGGCCACAG--CAAGTTTTCGACATCTTCAATGTTTACG	GCAGCTCGCTGATGCAAGT	TTTCAAACTCGGA					
Lp Hes-B	-----	GGCCATCGGCCACAG--CAAGTTTTCGACATCTTCAATGTTTACG	GCAGCTCGCTGATGCAAGT	TTTCAAACTCGGA					
		570	580	590	600	610	620	630	640
Pm Hes-A	-----	CTCTTTTCTTCCACCAAGAGCTCTGCTAGCCAAATG	TTTGGCAGTATTACAT	TGT--TTTAAAGAAACATTTTGCA					
Lp Hes-A	-----	CTCTTTTCTTCCACCAAGAGCTCTGCTAGCCAAATG	TTTGGCAGTATTACAT	TGT--TTTAAAGAAACATTTTGCA					
Pm Hes-B	-----	CTCTTTTCTTCCACCAAGAGCTCTGCTAGCCAAATG	TTTGGCAGTATTACAT	TGT--TTTAAAGAAACATTTTGCA					
Lp Hes-B	-----	CTCTTTTCTTCCACCAAGAGCTCTGCTAGCCAAATG	TTTGGCAGTATTACAT	TGT--TTTAAAGAAACATTTTGCA					
		650	660	670	680	690	700	710	720
Pm Hes-A	-----	GCGGTTATTGTIAAAATCAGATTTTCATTTAATTTGT	TCGCTCCTGTGATCTTAACATATACATGTTATATTTTATAG						
Lp Hes-A	-----	GCGGTTATTGTIAAAATCAGATTTTCATTTAATTTGT	TCGCTCCTGTGATCTTAACATATACATGTTATATTTTATAG						
Pm Hes-B	-----	GCGGTTATTGTIAAAATCAGATTTTCATTTAATTTGT	TCGCTCCTGTGATCTTAACATATACATGTTATATTTTATAG						
Lp Hes-B	-----	GCGGTTATTGTIAAAATCAGATTTTCATTTAATTTGT	TCGCTCCTGTGATCTTAACATATACATGTTATATTTTATAG						
		730	740	750	760	770	780	790	800
Pm Hes-A	-----	GT-GTT-CATTGATGCAATGATT	GAGTAACCAAAACATCGTTTCCTTCTA	TTAAATTTGTGTAGGCATCGTTTGTGCT					
Lp Hes-A	-----	GT-GTT-CATTGATGCAATGATT	GAGTAACCAAAACATCGTTTCCTTCTA	TTAAATTTGTGTAGGCATCGTTTGTGCT					
Pm Hes-B	-----	GT-GTT-CATTGATGCAATGATT	GAGTAACCAAAACATCGTTTCCTTCTA	TTAAATTTGTGTAGGCATCGTTTGTGCT					
Lp Hes-B	-----	GT-GTT-CATTGATGCAATGATT	GAGTAACCAAAACATCGTTTCCTTCTA	TTAAATTTGTGTAGGCATCGTTTGTGCT					

(continued on next page)

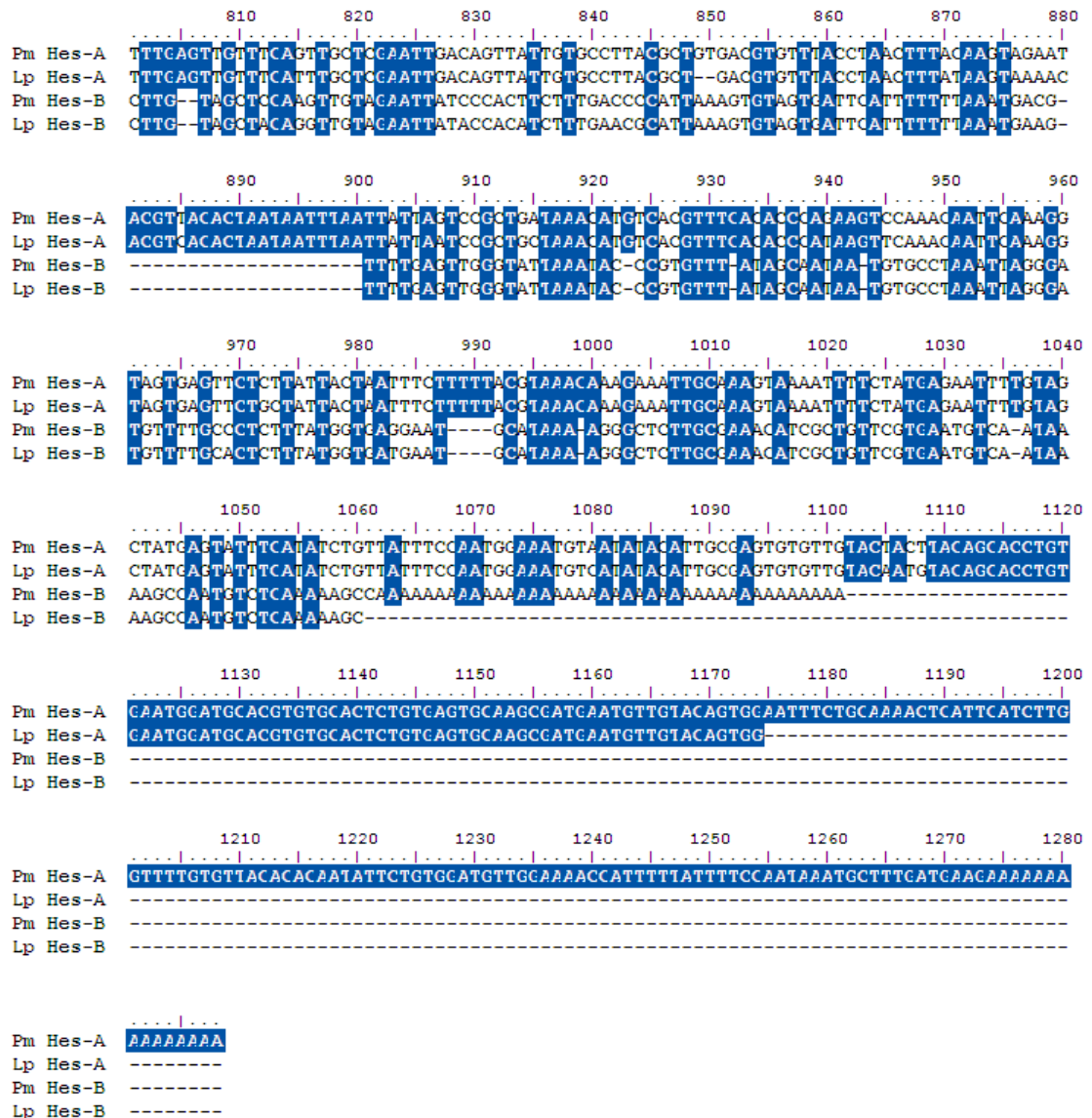


Figure 4-4. Nucleotide alignment of all four lamprey *Hes* sequences. Arrowheads indicate the boundaries of the ORF. Note the near-exact sequence similarity of all four sequences at the coding region, whereas *Hes-A* and *Hes-B* homologues diverge considerably at the 3' UTR region.

Given that both *L. planeri* fragments contain an incomplete coding region (Figure 4-5), only *P. marinus* sequences were used for further comparison at the sequence level. An alignment between predicted amino acid sequences of the entire coding region (i.e. excluding both 5' and 3' UTRs) of *PmHes-A* and *PmHes-B* shows that both genes differ only in three positions (Figure 4-6). Both sequences possess a bHLH domain, the diagnostic WRPW motif at the C-terminal end, and the conserved proline residue in the

basic region (Figure 4-6). However, there is no clear evidence for the presence of an Orange domain when compared to mouse *Hes* genes, as the region of the lamprey *Hes* genes that aligns with the Orange domain of mouse sequences is considerably short, and with a big gap in the middle (Figure 4-7).

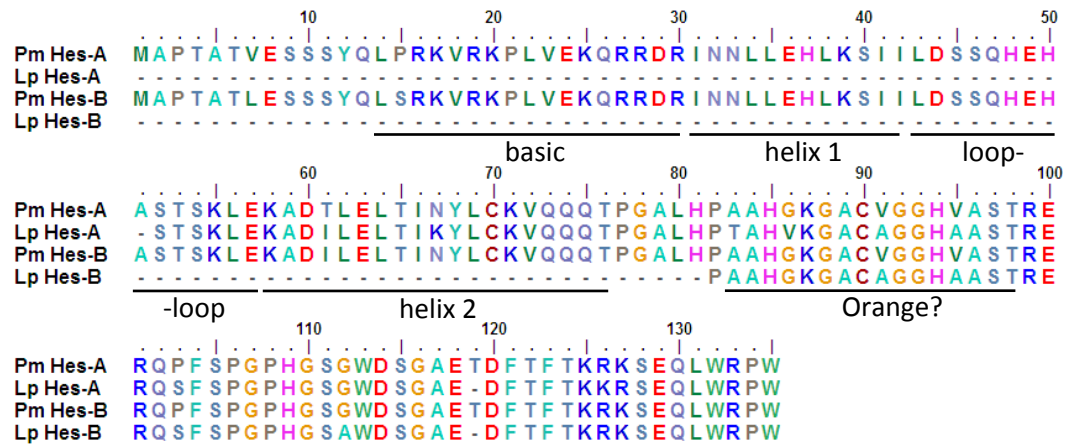


Figure 4-5. Alignment of predicted amino acid sequences of the cloned fragments of *LpHes-A* and *LpHes-B* together with *PmHes-A* and *PmHes-B*. The bHLH and Orange domains are indicated, along with the conserved WRPW motif at the C-terminal end (underlined). The *LpHes-A* fragment encompasses part of the loop region downstream to the WRPW motif in the coding region. On the other hand, the *LpHes-B* fragment covers only the Orange domain downstream to the WRPW domain.

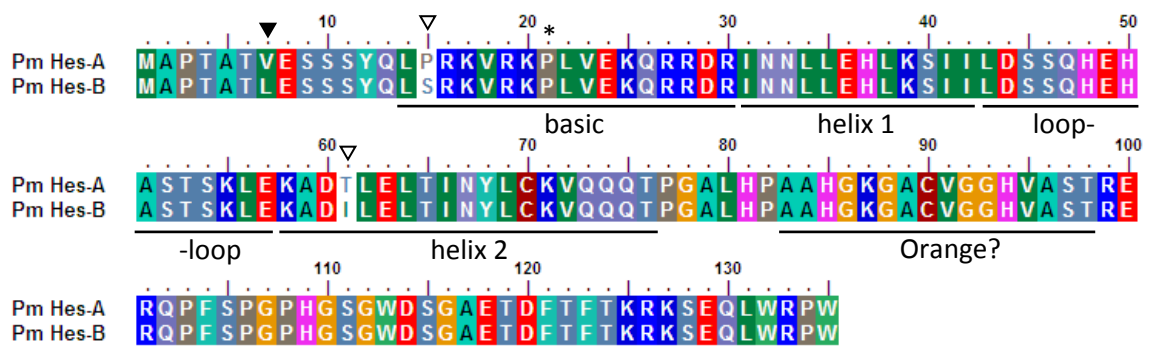


Figure 4-6. Alignment of predicted amino acid sequences of the entire coding region of *PmHes-A* and *PmHes-B* (sequences kindly provided by Dr. Tatjana Sauka-Spengler). The bHLH domain and Orange domain are indicated, as well as the highly conserved WRPW motif at the C-terminal end (underlined). Both sequences differ only in three positions (arrowheads). The conserved proline residue in the basic region is indicated by an asterisk. Boundaries of regions are according to Davis and Turner (2001) when aligned to mouse *Hes* homologues (see Figure 4-7).

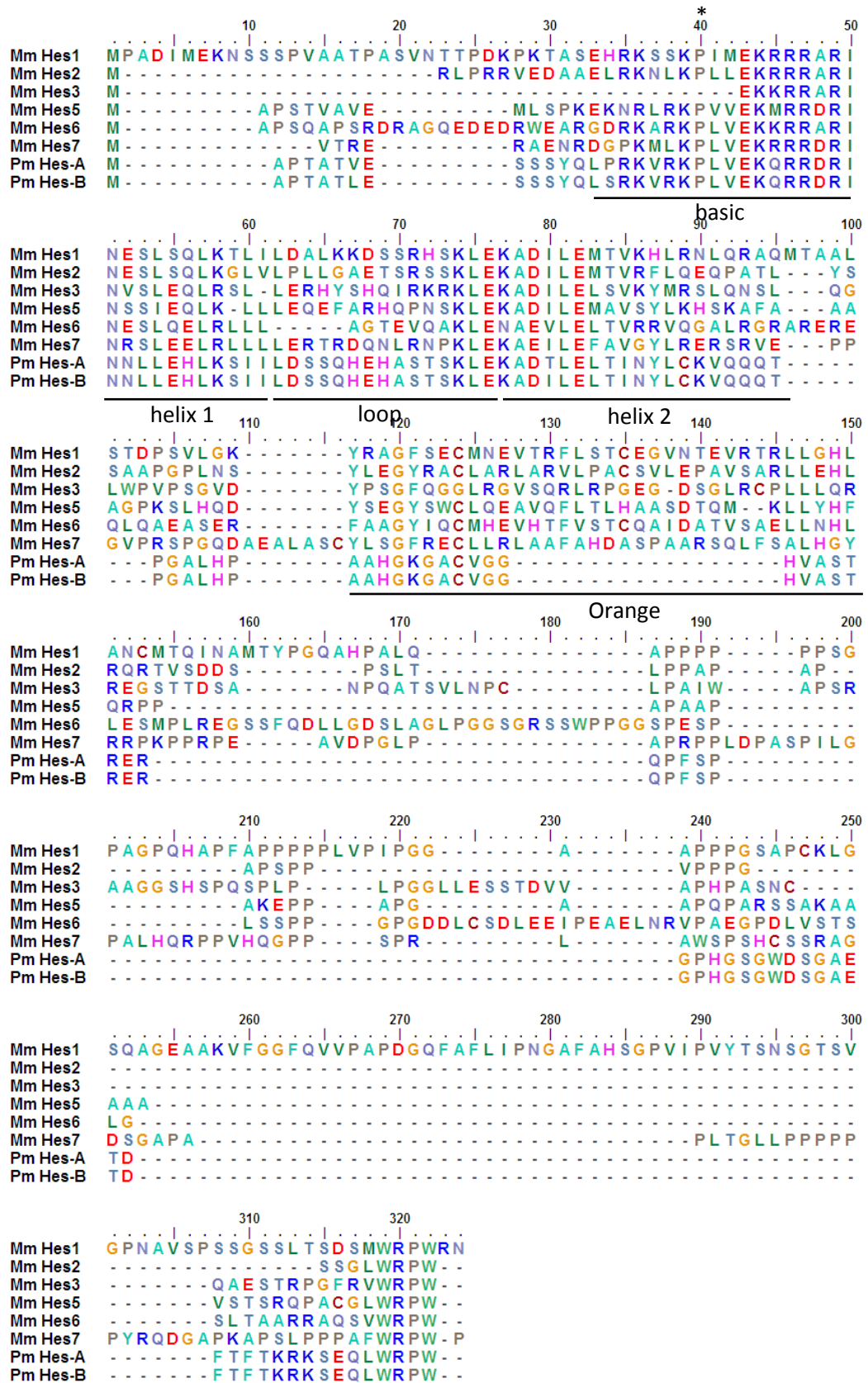


Figure 4-7. Alignment of predicted amino acid sequences of *PmHes-A* and *PmHes-B* with mouse Hes proteins. The bHLH and Orange domains are indicated, along with the conserved WRPW motif at the C-terminal end (underlined). The conserved proline residue in the basic region is marked by an asterisk. Notice the large gap in both lamprey sequences that align with the Orange domain of mouse Hes proteins.

4.5.2. Phylogenetic Analysis of Lamprey *Hairy/E(spl)* Genes

Amino acid sequences of genes belonging to *Hes*, *Hey* and *Dec* gene subfamilies from representative chordate species (Table 4-1) were aligned as described in Methods (Figure 4-8). This alignment was used to construct a NJ tree as explained in Methods. The three subfamilies are recovered and both *Petromyzon marinus* *PmHes-A* and *PmHes-B* cluster as sister groups within the *Hes* clade, which is supported by a bootstrap value of 64 (Figure 4-9). They, however, do not cluster with any particular *Hes* sub-group (i.e. *Hes1-7*) with significant support. Nonetheless, considering this NJ tree, the sequence of their bHLH domains and the presence of a WRPW motif, both sequences seem to be true members of the *Hes* subfamily.

Species	Gene name	Accession number	Species	Gene name	Accession number
<i>Mus musculus</i>	<i>MmHes1</i>	NP_032261.1	<i>Homo sapiens</i>	<i>HsDec1</i>	NP_003661.1
<i>Mus musculus</i>	<i>MmHes2</i>	NP_032262.2	<i>Homo sapiens</i>	<i>HsDec2</i>	NP_110389.1
<i>Mus musculus</i>	<i>MmHes3</i>	NP_032263.2	<i>Mus musculus</i>	<i>MmDec1</i>	NP_035628.1
<i>Mus musculus</i>	<i>MmHes5</i>	NP_034549.1	<i>Mus musculus</i>	<i>MmDec2</i>	NP_077789.1
<i>Mus musculus</i>	<i>MmHes6</i>	NP_062352.1	<i>Gallus gallus</i>	<i>GgDec1</i>	XP_414439.1
<i>Mus musculus</i>	<i>MmHes7</i>	NP_149030.2	<i>Xenopus tropicalis</i>	<i>XtDec1</i>	XP_002938358.1
<i>Gallus gallus</i>	<i>GgHes1</i>	NP_989803.1	<i>Xenopus tropicalis</i>	<i>XtDec2</i>	NP_001027504.1
<i>Gallus gallus</i>	<i>GgHes5a</i>	NP_001012713.1	<i>Danio rerio</i>	<i>DrDec1</i>	NP_997844.2
<i>Gallus gallus</i>	<i>GgHes5b</i>	XP_417552.2	<i>Danio rerio</i>	<i>DrDec2</i>	NP_001034196.1
<i>Gallus gallus</i>	<i>GgHes6-1</i>	BI393243.1			
<i>Gallus gallus</i>	<i>GgHes6-2</i>	XP_422641.3			
<i>Danio rerio</i>	<i>DrHes1a</i>	NP_571948.1	<i>Mus musculus</i>	<i>MmHey1</i>	NP_034553.2
<i>Danio rerio</i>	<i>DrHes1b</i>	NP_571154.1	<i>Mus musculus</i>	<i>MmHey2</i>	NP_038932.1
<i>Danio rerio</i>	<i>DrHes2b</i>	NP_001038818.1	<i>Mus musculus</i>	<i>MmHeyL</i>	NP_038933.2
<i>Danio rerio</i>	<i>DrHes3</i>	NP_571155.1	<i>Mus musculus</i>	<i>MmHelt</i>	NP_776150.2
<i>Danio rerio</i>	<i>DrHes5c</i>	NP_991182.1	<i>Gallus gallus</i>	<i>GgHey1</i>	XP_425926.3
<i>Danio rerio</i>	<i>DrHes5d</i>	NP_571165.3	<i>Gallus gallus</i>	<i>GgHeyL</i>	XP_001234406.2
<i>Danio rerio</i>	<i>DrHes5e</i>	NP_001096598.1	<i>Danio rerio</i>	<i>DrHey1</i>	NP_997726.1
<i>Danio rerio</i>	<i>DrHes6</i>	NP_919381.2	<i>Danio rerio</i>	<i>DrHey2</i>	NP_571697.2
<i>Danio rerio</i>	<i>DrHes7</i>	NP_571153.1	<i>Danio rerio</i>	<i>DrHeyL</i>	NP_859425.1
<i>Danio rerio</i>	<i>DrHes8</i>	NP_001159638.1	<i>Danio rerio</i>	<i>DrHelt</i>	NP_996948.1
<i>Petromyzon marinus</i>	<i>PmHes-A</i>	This study			
<i>Petromyzon marinus</i>	<i>PmHes-B</i>	This study			
<i>Ciona intestinalis</i>	<i>CiHesa</i>	NP_001071989.1			
<i>Ciona intestinalis</i>	<i>CiHesb</i>	NP_001071684.1			
<i>Ciona intestinalis</i>	<i>CiHesc</i>	NP_001071685.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyA</i>	AAQ93667.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyB</i>	AAQ93668.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyC</i>	AAQ93669.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyD</i>	AAQ93670.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyE</i>	AAQ93671.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyF</i>	AAQ93672.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyG</i>	AAQ93673.1			
<i>Branchiostoma floridae</i>	<i>AmphiHairyH</i>	AAQ93674.1			

Table 4-1. Names of Hes, Hey and Dec-related amino acid sequences from representative chordate species employed to construct the phylogenetic tree shown in Figure 4-8. GenBank accession numbers for each sequence are indicated.

	10	20	30	40	50	60	70	80	90
MmHes1	SSKPI	MEKRRRARIN	ESLSQLKTLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
MmHes2	NLKP	LLEKRRRARIN	ESLSQLKGLVLPILGAE	TSRSSKLEKADILEMTVRF	---	QEQP			
MmHes3	----	MEKRRRARIN	VSLEQLRSLLERHYSHQ	IRKRKLEKADILELSVKYMRSLQNSL					
MmHes5	LRKP	VVEKMRDRIN	SSIEQLKLLLEQEFARH	QPN	SKLEKADILEMAVSYLKHSKAF				
MmHes6	ARKP	LVKEKRRRARIN	ESQLRLLLAGTEV	---	QAKLENAEVL	ELTVRRVQ	GALGR		
MmHes7	MLKP	LVKEKRRDRIN	RSLEELRLLLETRDQ	NLRNPKLEKAEILEFAVG	YLER	SRVE	---	PP	G
GgHes1	SSKPI	MEKRRRARIN	ESLQKMLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
GgHes5a	LRKP	IVKEKRRDRIN	SSIEQLKLLLEKEFQRH	QPN	SKLEKADILEMTVS	YLKYSRAFA			
GgHes5b	LRKP	VVEKMRDRIN	SSIEQLKLLLEKEFQRH	QPN	SKLEKADILEMAVSYLKQSSHQ				
GgHes6-1	ARKP	LVKEKRRRARIN	ESLRELRLLAGG	---	EAKLENAEVL	ELTVRRVQ	AVLERR		
GgHes6-2	LRKP	LLEKRRERIN	NCLDQLKETVVGAFHLD	---	QSKLEKADILEMTVKHLQNIQNSK				
DrHes1a	SSKPI	MEKRRRARIN	ESLQKTLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
DrHes1b	SSKPI	MEKRRRARIN	ESLQKTLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
DrHes2b	TLKP	LLEKRRRARIN	DSLRLKALILPLTGKD	NCRYSKLEKADILEMTVRF	FLTDIQTTP				
DrHes3	VSKP	IMEKRRRARIN	CKNLQKLSLESACSNN	---	IRKRKLEKADILEMTVKHLRNLQNTK				
DrHes5c	LRKP	IVKEKRRDRINT	CIDQLKSLLKEFHSH	DP	TKLEKADILEMTV	SKHLQ	TIQQ		
DrHes5d	LRKPM	VEKIRRRERIN	SSIEKLKTLTLLAQEFVKQ	QPD	SRQEKADILEMTD	FLRRSQSS			
DrHes5e	LRKPM	VEKIRRRERIN	SSIEKLKTLTLLAQEFIKQ	QPD	SRQEKADILEMTD	FLRRSQSS			
DrHes6	TRKP	LVKEKRRRARIN	ESQLRLLLAGDPA	---	QVKMNAEVL	ELMTVKRVE	ILQNK		
DrHes7	ILKP	IVKEKRRDRIN	QRLLEELRLLLDNTLDS	RLQNPKLEKAEILELAVE	YIRTKTATARDQ	GDSSKDTHDP	KPPPELLSRPQMP	CAS	
DrHes8	LRKP	LLEKRRERIN	NCLDQLRETVVAVFKPD	---	QSKLEKADILEMTVKHLQNIQSSR				
PmHes-A	VRKP	LVKEKRRDRIN	NLLEHLKSIILDSQHE	HASTSKLEKADILEMTV	RLNQVQQQT				
PmHes-B	VRKP	LVKEKRRDRIN	NLLEHLKSIILDSQHE	HASTSKLEKADILEMTV	RLNQVQQQT				
CiHesA	TKNP	IMEKRRERIN	CKLEDLKSIVLTAVAE	---	SRNPKLEKADILEMTVRYLKS	LQGGK			
CiHesb	ASKP	IMEKRRERIN	CKSLNELKSILLHALRKD	QSTCHSKLEKADILEMTVRYL	RGIQRR				
CiHesc	SSKPI	MEKRRRARIN	SSNELKSILLEALKKD	STRHSKLEKADILEMTVKYLNVERQR					
AmphiHairyA	SSKPI	MEKRRRARIN	DSLNLKTLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
AmphiHairyB	SSKPI	MEKRRRARIN	DSLNLKALILADALKD	SS	HSKLEKADILEMTVKHLRNLQRAQ				
AmphiHairyC	SSKPI	MEKRRRARIN	DSLNLKTLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
AmphiHairyD	SSKPI	MEKRRRARIN	ESLTELKTLILEALNKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
AmphiHairyE	SSKPI	MEKRRRARIN	DSLNLKTLILDALKKD	SSRHSKLEKADILEMTVKHLRNLQRAQ					
AmphiHairyF	SSKPI	MEKRRRARIN	ESLNLKTLILDALKKD	SSRQNKLEKADILEMAVRYL	LDIQRHQ				
AmphiHairyG	SSKPI	MEKRRRARIN	SSNELKNLIDTYKND	STHHSKLEKADILEAVKHVR	SLQRAQ				
AmphiHairyH	SSKPI	MEKRRRARIN	SSNELKNLILGTVKDDINAPH	SHSKLEKADILEI					
HsDec1	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQQKQIIALQS		G
HsDec2	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQHQKIIALQN		G
MmDec1	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQQKQIIALQS		G
MmDec2	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQQKQIIALQS		G
GgDec1	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQQKQIIALQS		G
XtDec1	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQQKQIIALQS		G
XtDec2	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQHQKIIALQN		G
DrDec1	LPHRL	IEKRRDRIN	ECIAQLKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVKALTNLID	QQQKQIIALQN		G
DrDec2	LPHRL	IEKRRDRIN	ECIQKDLLPEHLKLT	T	LGHLEKAVVLELT	LKHVNALTNLID	QQHQKIIALQN		G
MmHey1	RRRG	IEKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	AKLEKAEILQMTVDHLKMLH	ATAG		
MmHey2	KRRGI	EKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	AKLEKAEILQMTVDHLKML	ATAG		
MmHeyL	KRRGI	EKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	SKLEKAEVLQMTVDHLKML	HASG		
MmHelt	VSHK	IVKEKRRDRIN	RCNELGKTVPMALAKQ	S	SGKLEKAEILQMTVDHLKML	HASD			
GgHey1	RRRG	IEKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	AKLEKAEILQMTVDHLKML	HATAG		
GgHeyL	KRRGI	EKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	SKLEKAEILQMTVDHLKML	HATAG		
DrHey1	RRRG	IEKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	AKLEKAEILQMTVDHLKML	HAG		
DrHey2	KRRGI	EKRRDRIN	SSLSLELRLVPSAFEKQ	---	GS	AKLEKAEILQMTVDHLKML	ATAG		
DrHeyL	KRRGI	EKRRDRIN	SHSLELRLVPSAFEKQ	---	GS	SKLEKAEILQMTVDHLKML	HAMG		
DrHelt	VSHK	IVKEKRRDRIN	RCNELGKTVPMALAKQ	N	SGKLEKAEILQMTVDHLKML	HASD			

	100	110	120	130	140	150	160
MmHes1	MTAALST-DPSV--LGKYRAG-----	FSECMNEVTRFL-STCEGVNT---	EVRTLLGHLANCMTQINAMTYP				
MmHes2	ATLYSSA-APGP--LNSYLEG-----	YRACLARLARVL-PACSVLEP---	AVSARLLHHLRQRTVSDSPSLT				
MmHes3	---QGLWFPVS--GVDYPSG-----	FQGGLRGVSQLR-PPGEGDSG---	LRCPLLLQRRREGST---				
MmHes5	---AAAG-PKSL--HQDYSEG-----	YSWCLQEAQFL-TLHA-A-S---	DTQMKLLYHQRPPAPAPAKEP				
MmHes6	AREREQL-QAEA--SERFAAG-----	YIQCMHEVHTFV-STCQAIDA---	TVSAELLNHLLESMPLREGSSFQ				
MmHes7	VPRSPGQDAEAL--ASCYLSG-----	FRECLRLAFAFA---HDASP---	AARSQLFSAHGYRRPKPPPEA				
GgHes1	MAAALSA-DPSV--LGKYRAG-----	FNECMNEVTRFL-STCEGVNA---	DVRARLLGHLSACLGGIVAMNYL				
GgHes5a	---AS---AKSL--QQDYCEG-----	YAWCLKEALQFL-SLHS-ANT---	ETQMKLLCHFQSQAMPKDSGSP				
GgHes5b	MKTSGSF-HKSY--QDFREG-----	YSRCLQEAFFHL-SLHK-VRT---	ETQTKLLSHFQKSQSAAPVFS				
GgHes6-1	ALEGGRL-HREA--SERFAAG-----	YIQCMHEVHTFV-SGCPGIDA---	TTAAELLNHLLESMPLSEGGLR-				
GgHes6-2	LMADSKV-GLEA--QQRYSTG-----	YIQCMHEVHLL-LTCEWMDK---	TLGARLLNHLKSLPRSSDTSK				
DrHes1a	MSAALSA-DTNV--LSKYRAG-----	FNECMNEVTRFL-STCEGVNT---	EVRSLNHLSGCMGQMMAMNYP				
DrHes1b	MTAALNT-DPTV--LGKYRAG-----	FSECMNEVTRFL-STCEGVNT---	EVRTLLGHLSACMTQINAMNYP				
DrHes2b	-----SKDT--AVSFTEG-----	YTTCLQRVSAFL-PQTS-LDA---	ETRHRVNDFFQSRVMPKTPACQN				
DrHes3	RGLSKACD-----SAEYHAG-----	YRSCNTVSHYL-RASD-TDR---	DSRSIMLTNTLSGLNHNRPVDFS				
DrHes5c	-----QQIP--QRDFNEG-----	YSHCWRESVHFL-SLHSNAGE---	LQHLHSGPKTNSMTGST---				
DrHes5d	-----AAGDG-----	RSRCVQEAVSFL-SQCP-VQT---	QSHTRLMKFLHMQTPADQHTRV				
DrHes5e	-----AAGDG-----	RSRCVQEAVSFL-SQCP-VQT---	QSHTRLMKFLHMQTPADQHTRV				
DrHes6	AKEADSV-NREA--NERFAAG-----	YIQCMHEVHTFV-SSCPGIDA---	TIAADLLNHLLECMPLNDEERFQ				
DrHes7	IPESIQTHNSPS--NPIYKAG-----	FKECISRSASFI-DCVEPSQR---	DSFVQGLCHHLDSYSSALPHGRV				
DrHes8	VSDPVL--NTGA--RQRYSTG-----	YQCMQEVHLL-HSCDWMDK---	TLGSRLNHLFKSLPLSAKDCPR				
PmHes-A	---PGAL-HPAAHGKGCACVGG-----	HVASTRERQPFPS-PGPHGSGW---	DSGAETDFTFTK-----				
PmHes-B	---PGAL-HPAAHGKGCACVGG-----	HVASTRERQPFPS-PGPHGSGW---	DSGAETDFTFTK-----				
CiHesa	PIDAGIL-SLSR--TSDYDDSDDDTKSSMTESYQC--	DVITML-NRSD-VPL---	HFRITQFLGHLAGKHNVEPRKKPQ				
CiHesb	MNAATIL-DPSV--VSKYRNG-----	YVECKNEVSHFLENSENHVP---	DVKSRLLNHLGNLPHQFVHQLP				
CiHesc	LSVSLTL-DPAE--INQYKAG-----	FNECRNEVMHFL-STCEGVTV---	DVRTLLNHLCLHSLIQKQSN				
AmphiHairyA	IADAVIR-DPVA--LSKYRAG-----	YSECMTEVSRFL-TGSDGVDD---	QVQQRLLGHLAGSCCQTVDNIAV				
AmphiHairyB	LTAANT-NPAL--PGQYRAG-----	FNECLMEVNRFL-GASDSVDI---	QVRQRLNHLAGACSPARPGTYP				
AmphiHairyC	LTAANT-NPSL--PGQYRAG-----	FNECLMEVNRFL-GASDSVDI---	QVRQRLNHLAGACSPARPGTYH				
AmphiHairyD	MAAAVST-DPSL--LGKYRSG-----	FSECMTEVSRVM-GSMDGVDS---	QVKQRLLNHLATHQPATNVMSG				
AmphiHairyE	VAAAVST-DPNV--LGQYRAG-----	FNECMTEVSSFL-GSASTVET---	EVHQRLNHLAGCCHTIVSPVST				
AmphiHairyF	LTVSAST-EPGT--HARYRAG-----	FNHCTAEVSRFT---EGMDP---	PVRQRLSHLAGLCQTVDEVG				
AmphiHairyG	MAAAVST-DPAA--LGRYRAG-----	YSRCRAEVARFM-GTVDGVDP---	QVRQRLNHLDGCSQRVENAAHS				
AmphiHairyH							
HsDec1	LQAGELSGRNVEVGQEMFCSG-----	FQTCAREVLQYL-AKHENTRD---	LKSSQLVTHLHRVVSELLQGGTS				
HsDec2	ERSLK---SPIQSDLDLAFHSG-----	FQTCAREVLQYL-SRFESWTP---	REPRCQQLNHLHVAATQFLPTPQL				
MmDec1	LQAGDLSGRNLEAGQEMFCSG-----	FQTCAREVLQYL-AKHENTRD---	LKSSQLVTHLHRVVSELLQGGAS				
MmDec2	ERSLK---SPVQADLDLAFHSG-----	FQTCAREVLQYL-SRFESWTP---	REPRCAQLVSHLHVAATQL---				
GgDec1	LQAGDLSARSLDSSQEMFRSG-----	FQLCAKELLQYV-AKHNDN-KE---	LKAAQLVGHHLHRVAELLPGG--				
XtDec1	SPSE-----ENRISAEEMFHS-----	FQLCAEEALRFL-QGGE-----	RKELVAHLHRVASKLGPASFP				
XtDec2	EHALK---SPIQTDLDLAFHSG-----	FQTCAREVLQYL-SRFESWTS---	RDQRCQQLNHLHHTVSSQILSSQL				
DrDec1	LQIGEQGNGPSENSEEMFRSG-----	FHLCAKEVLQFL-ANQETMRD---	LTTAHIEHLQKVASELIQSPPS				
DrDec2	ERSLK---SSLQADLDLAFHSG-----	FQACAREVLQYL-NKVENWTA---	REQRCQRLNHLHKVSAQFPPTGI				
MmHey1	GKGYFDA-HALA--MDYRSLG-----	FRECLAEVARYL-SIEGLDA---SDPLRVRLVSHLNNYASQREASGA					
MmHey2	GKGYFDA-HALA--TDFMSIG-----	FRECLTEVARYL-SSVEGLDP---SDPLRVRLVSHLSTCASQREAAVMT					
MmHeyL	GTGFFDA-RALA--VDFRSIG-----	FRECLTEVIRYL-GVLEGPSHADPVRIRLLSHLKSAAEMEPSPTT					
MmHelt	FPRGREK-ELLAEFANYFHYG-----	YHECMKNLVHYL-TTVERMET---KDTKYARILAFQSKARLGAETFFP					
GgHey1	GKGYFDA-HALA--MDYRSLG-----	FRECLAEVARYL-SIEGLDA---SDPLRVRLVSHLNNYASQREASGA					
GgHeyL	GAGFLDA-RALA--VDYRSIG-----	FRECLTEVVRYL-GVLEGQST---ADPILRLLSHLNNYVAEMEPSAA					
DrHey1	GKGYFDA-HALA--MDYRSLG-----	FRECLAEVARYL-SIEGLDN---TDPLRIRLVSHLNNYASQREA----					
DrHey2	GKGYFDA-HSLA--MDFLSIG-----	FRECLTEVARYL-SSVEGLDS---SDPLRVRLVSHLSSCASQREAAAMT					
DrHeyL	GKGYFDA-RALA--VDYRILG-----	FRECVGEVVRYL-SSLEGVES---SDPIGARLVSHLSHCASELDPQLQS					
DrHelt	FPRGREKGELLTEFANYFHYG-----	YHECMKNLVHYL-TTVERMET---KDTKYARILAFQSK--VTEPEVFG					

Figure 4-8. MAFFT alignment of *Hairy/E(spl)* (*Hes*) sequences (shown in Table 4-1) used to construct the phylogenetic tree in Figure 4-9.

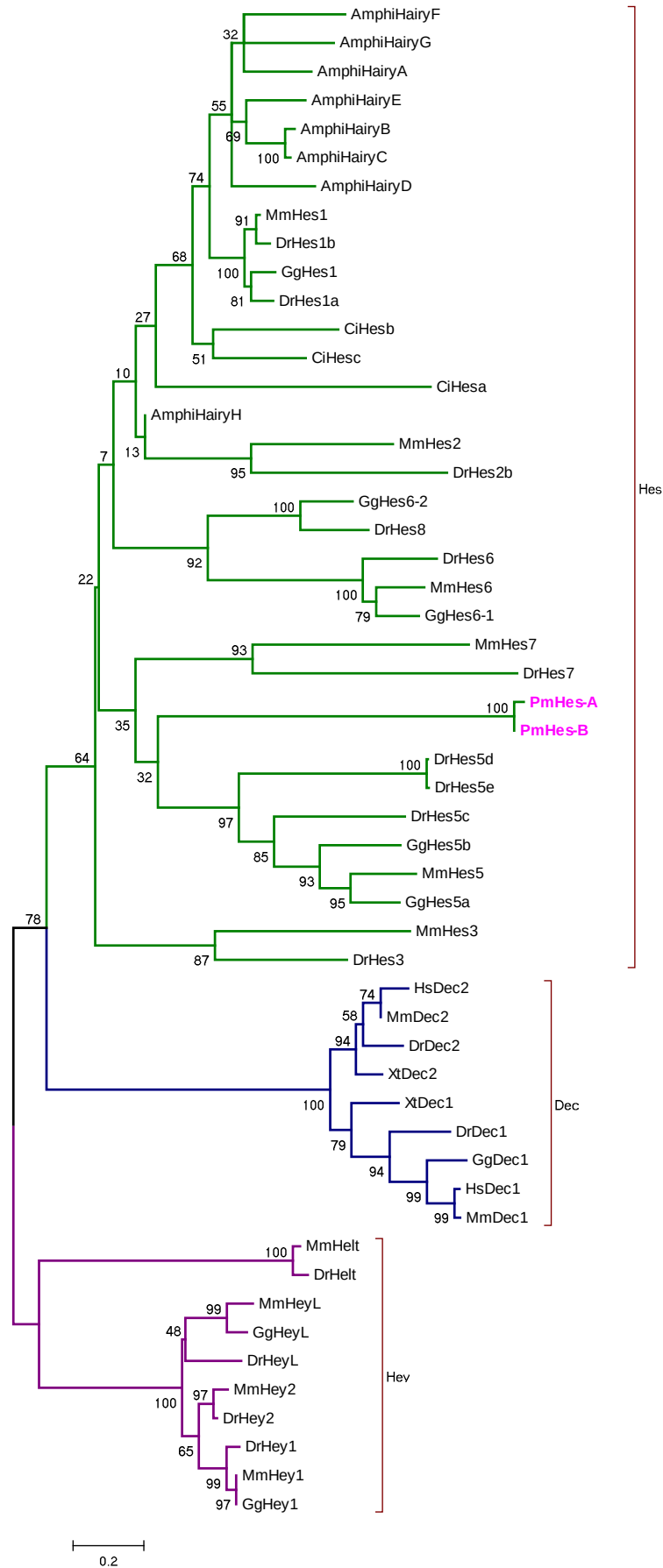


Figure 4-9. NJ phylogenetic tree using the alignment shown in Figure 4-8. Tree based on the Jones-Taylor-Thornton (JTT) matrix amino acid substitution model, 1000 bootstrap replicates and gamma distribution. GenBank accession numbers for each sequence are given in Table 4-1. Scale bar, amino acid substitutions per position.

4.5.3. *LpHes-A* Expression in Developing *L. planeri* Embryos

Owing to technical problems, expression of *LpHes-A* was only analysed at stages 23-24. At stage 23, expression is seen in the ventral midbrain and hindbrain, whereas in the anterior spinal cord transcripts are observed covering more dorsal domains (Figure 4-10A). Dorsal views show that expression is restricted to the ventricular zone of the neural tube (Figure 4-10B). At stage 24, expression increases both in the brain and spinal cord (Figure 4-10C), and in the brain it also expands dorsally (Figure 4-10D). Transcripts are now clearly observed in the diencephalon and telencephalon (Figure 4-10C). While almost all the brain is stained, clear gaps of expression occur at the MHB and at the level of the telencephalon-diencephalon boundary (Figure 4-10D). Dorsal views show that, as in stage 23 embryos, expression is completely restricted to the ventricular zone of the neural tube (Figure 4-10E).

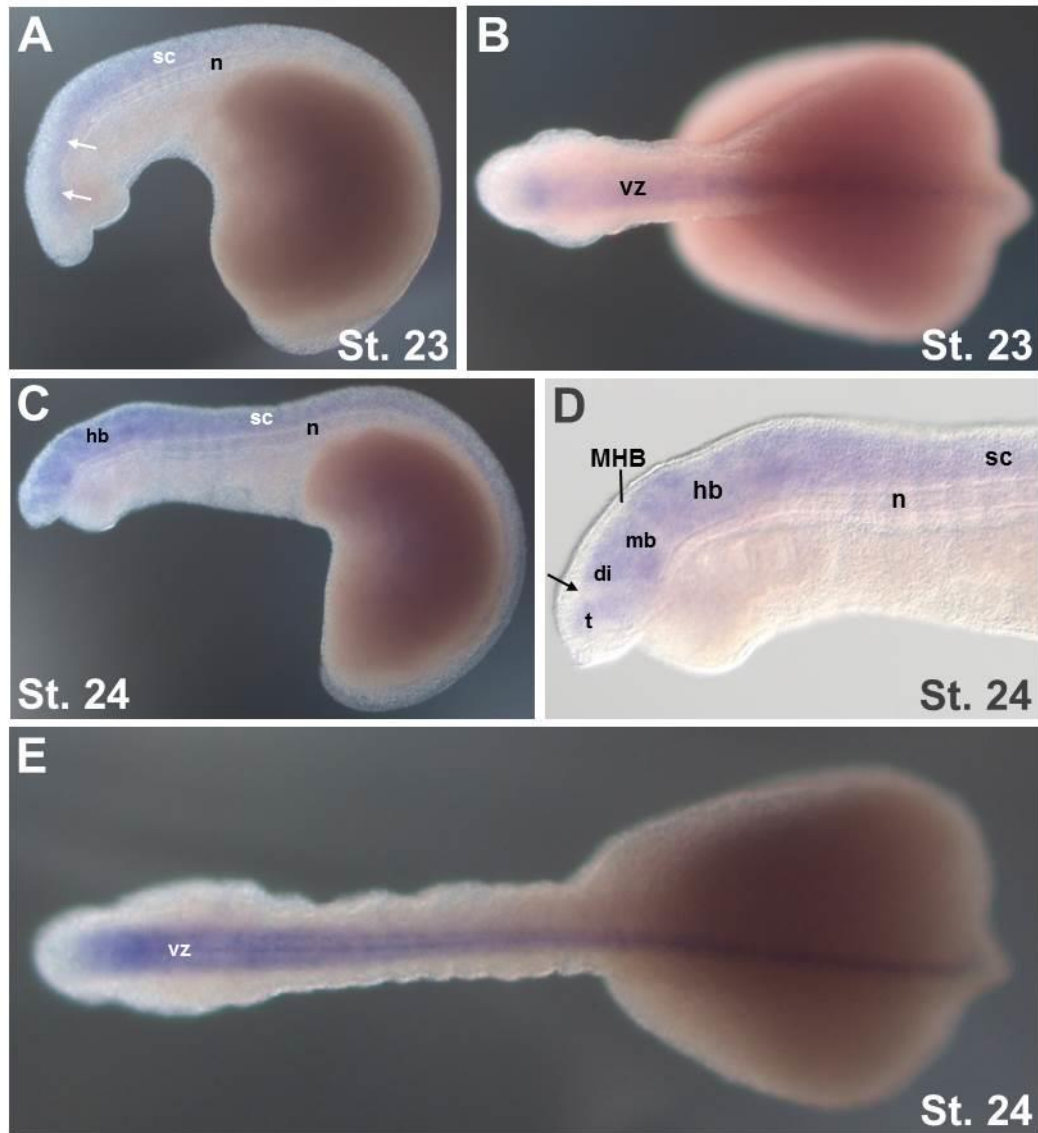


Figure 4-10. Expression of *LpHes-A* in developing *Lampetra planeri* embryos at stages 23 and 24. (A, C, D) are lateral views. (B, E) are dorsal views. Anterior is to the left in all images. (A) At stage 23, weak expression is observed in ventral midbrain and hindbrain (white arrows) and anterior spinal cord. (B) Dorsal view of a stage 23 embryo shows expression in the middle part of the neural tube corresponding to the ventricular zone. (C-E) At stage 24, expression increases both in the brain and spinal cord covering the entire dorsoventral extent of the neural tube (C, D), and expression is now observed in the forebrain (D). Dorsal views show expression in the ventricular zone as in stage 23 embryos (E). Note how expression is completely restricted to the ventricular zone of the CNS at these stages and gaps of expression occur at the MHB and telencephalon-diencephalon boundary (D, black arrow). Abbreviations: di, diencephalon; hb, hindbrain; mb, midbrain; MHB, midbrain-hindbrain boundary; n, notochord; sc, spinal cord; t, telencephalon; vz, ventricular zone.

4.5.4. *LpHes-B* Expression in Developing *L. planeri* Embryos

LpHes-B expression was analysed in *Lampetra planeri* embryos from stage 21 to stage 28. At stage 21, *LpHes-B* is expressed solely in the neural tube (Figure 4-11A, B). Two strong expression sites are seen in the anterior and posterior hindbrain in presumptive rhombomeres 1/2 and 6, respectively (Figure 4-11A, B). On the other hand, faint expression is observed in presumptive midbrain and the middle part of the hindbrain (Figure 4-11A, B). A gap of expression is seen at the MHB (Figure 4-11A). Small clusters of faint expression are also seen in the forming spinal cord (Figure 4-11A).

At stage 22, expression in the midbrain becomes stronger and expression in rhombomere 1/2 and posterior hindbrain is maintained (Figure 4-11C). Faint expression is also seen at the ventral-most part of rhombomeres 3-5. Expression in the spinal cord is also considerably increased with clusters at the anterior and a slightly more homogenous expression towards the posterior (Figure 4-11C). Expression is generally biased towards the ventral part of the neural tube. Slightly later, expression increases in the same domains and the neural tube is united by a ventral expression domain covering the midbrain, hindbrain and spinal cord (Figure 4-11D). At this stage, expression in rhombomere 1/2 and from rhombomere 6 towards the spinal cord notably progresses dorsally. Expression in presumptive rhombomere 4 becomes stronger and also expands dorsally, leaving gaps of expression in the dorsal part of rhombomeres 3 and 5 (Figure 4-11D). From a dorsal view (Figure 4-11E), all expression in the brain is seen as two lateral stripes leaving the centre generally unstained. In the spinal cord however,

expression is seen in the middle. Up to this stage, no expression was observed in the forebrain.

At stage 23, expression in the midbrain is further expanded and the MHB is clearly discerned, although this expression domain is still restricted to the ventral half (Figure 4-11F). Expression starts to appear in the ventral diencephalon and telencephalon, whereas in the hindbrain, expression in rhombomere 1/2 becomes even stronger along the dorso-ventral axis (Figure 4-11F). The rest of the hindbrain is strongly stained except for the dorsal half of rhombomeres 3 and 5, with faint expression in dorsal rhombomere 4 (Figure 4-11F). Strong expression is also observed along the spinal cord (Figure 4-11G). From a dorsal view, expression in the brain expands medially as compared to stage 22 embryos (compare Figure 4-11E and H). At this particular stage, transient expression is observed in cranial ganglia (Figure 4-11H). Expression in cranial ganglia was not observed at any other developmental stages examined.

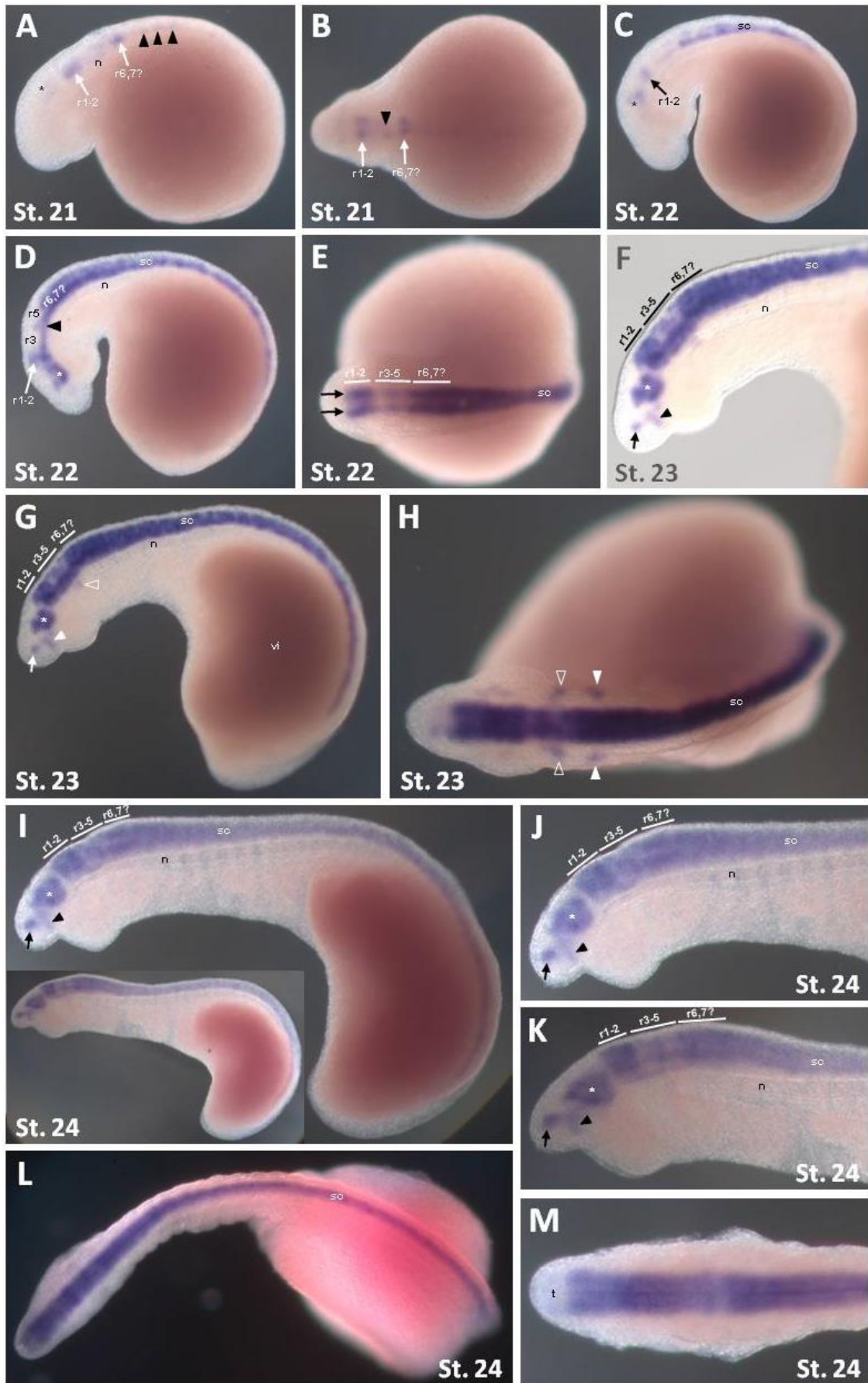


Figure 4-11. Expression of *LpHes-B* in *Lampetra planeri* developing embryos at stages 21-24. (A, C, D, F, G, I-K) are lateral views. (B, E, H, L and M) are dorsal views (M, dorsal view of the head). Anterior is to the left in all images. (A, B) At stage 21, *LpHes-B* is expressed in the forming neural tube. Strongest expression is seen in presumptive rhombomeres 1-2(r1-2) and 6-7(r6-7). Weaker expression is observed in r4 (black arrowhead in B), presumptive midbrain (black asterisk in A) and spinal cord (sc; black arrowheads in A). (C) At stage 22, expression increases in the midbrain (black asterisk) and considerably in the spinal cord (sc). Expression in the hindbrain is maintained. Note how expression is generally restricted to the ventral side of the neural tube. (D) In a slightly older stage 22 embryo, expression further intensifies in the same expression domains, particularly in the midbrain (white asterisk) and spinal cord where expression extends posteriorly. Expression in r4 progressing dorsally starts to demarcate unstained r3 and r5. (E) From a dorsal view at the same stage, expression in the brain is seen as two lateral stripes, while in the spinal cord expression occurs medially. (F, G) At stage 23, expression has covered most of the rhombospinal region except in the dorsal half of rhombomeres 3 and 5. Expression is also distinguished in a large territory in the ventral midbrain (white asterisks). Notably the MHB is not stained. At this stage expression first appears in a discrete domain in the telencephalon (arrows) and ventral diencephalon (solid arrowheads). (H) At this stage, transient expression is observed in two pairs of cranial placodes/ganglia, possibly corresponding to otic (open arrowheads in G and H) and lateral line (solid arrowheads in H) placodes/ganglia. Expression in the brain expands medially. (I, J) At stage 24, r3 and r5 are almost completely stained, while expression in the midbrain (white asterisks) progresses dorsally. Expression in the telencephalon (arrows) slightly intensifies whereas in the ventral diencephalon (arrowheads) expression remains virtually unchanged. (K, inset in I) In other stage 24 embryos however, r3 and r5 still show weak expression. (L and M) At this stage, all expression along the neural tube is located medially in the ventricular zone (the small expression site in the telencephalon is not observed in this dorsal view). n: notochord, r: rhombomere; sc: spinal cord, t: telencephalon.

At stage 24, expression in the telencephalon and diencephalon is maintained, and expression in the midbrain expands dorsally forming a bulge demarcating the entire midbrain (Figure 4-11I, J). A clear gap of expression can still be observed at the MHB. Expression in the hindbrain and spinal cord is maintained strong and nearly homogenous, except at the level of rhombomeres 3 and 5 where staining is weaker (Figure 4-11J). In slightly younger stage 24 embryos, where rhombomeres 3 and 5 specifically show very faint staining, expression in the midbrain has not progressed dorsally as much as older stage 24 embryos (Figure 4-11I, inset; compare J and K). A dorsal view shows that expression is restricted to the middle part of the neural tube which corresponds to the ventricular zone (Figure 4-11L). As mentioned above, expression of *LpHes-B* in cranial ganglia is not observed at this stage (Figure 4-11M).

At stage 25, expression is preserved in the entire neural tube with slight differences (Figure 4-12A). In the brain, expression has considerably expanded in the dorsal telencephalon and in nearly all the diencephalon (Figure 4-12B). The midbrain is completely stained and expression is now seen much closer to the MHB, which is virtually all stained (Figure 4-12B). It is also clear that transcripts are not found in the forming epiphysis (Figure 4-12B). In the hindbrain, all the dorsal half of each rhombomere is stained, and instead, expression starts to disappear on the ventral side (Figure 4-12B). In the spinal cord, expression becomes faint in the anterior half, whereas in the posterior half it is maintained strong, reaching its posterior end (Figure 4-12A and D).

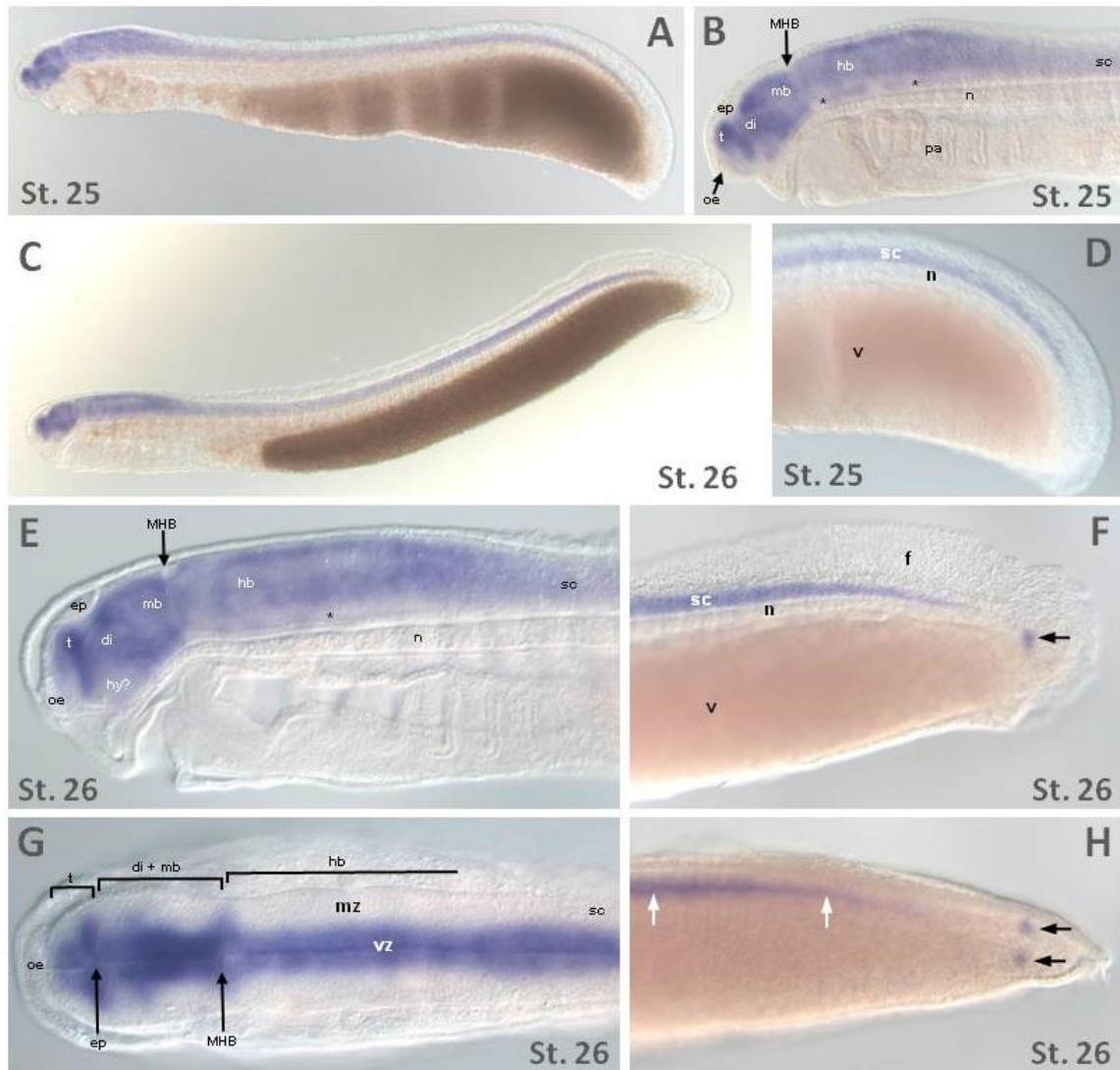


Figure 4-12. Expression of *LpHes-B* in stage 25 and stage 26 *L. planeri* embryos. (A-F) are lateral views. (G) and (H) are dorsal views. Anterior is to the left in all images. (D), (F) and (H) are photographs of the tail. (A) At stage 25, *LpHes-B* is expressed in the entire neural tube reaching the end of the spinal cord (D). (B) In the brain, expression is seen in the dorsal telencephalon (t), diencephalon (di) excluding the epiphysis (ep), dorsal and ventral midbrain (mb) and the entire hindbrain (hb). However, note how expression in the hindbrain starts fading from the ventral side (asterisks). At this stage, expression now covers the midbrain-hindbrain boundary (MHB) with respect to stage 24 embryos where this region is largely unstained (B; cf. Figure 4-11I, J and K). (C) At stage 26, expression is maintained all along the neural tube. (E) In the brain, all expression sites are the same as in stage 25. Expression is clearly absent in the ventral telencephalon, epiphysis, hypothalamus (hy), dorsal-most part of the diencephalon and MHB, and in the ventral hindbrain. (G) From a dorsal view of the head, expression is clearly seen in the medial part of the neural tube delimiting the extent of the ventricular zone (vz) and the peripheral marginal zone (mz) of the neural tube. (F, H) In the tail region, expression reaches the end of the spinal cord and in some embryos an additional expression domain is observed in the tailbud seen as a pair of dots from a dorsal view (black arrows). White arrows in H indicate expression in the spinal cord. Abbreviations: ep: epiphysis, di: diencephalon; f: caudal fin, hb: hindbrain, hy: hypothalamus, mz: marginal zone, mb: midbrain, MHB: midbrain-hindbrain boundary, n: notochord, oe: olfactory epithelium, pa: pharyngeal arches, sc: spinal cord, t: telencephalon, vz: ventricular zone, v: vitellum.

Later in development, at stage 26, expression is maintained in the same regions (Figure 4-12C). Also, the basic brain segments can be clearly observed at this stage (Figure 4-12E). Transcripts are observed in the dorsal telencephalon and in the anterior intracephalic sulcus. In the diencephalon, expression seems to be excluded from the hypothalamus, or at least from its ventral part, and from a thin domain at its dorsal part between the epiphysis and the midbrain corresponding to the pretectum (Figure 4-12E). The midbrain is also largely stained except for a thin layer at the dorsal-most side as in the pretectum, and this gap of expression expands at the MHB (Figure 4-12E). In the hindbrain, a dorsal and a medial band of expression with respect to the dorso-ventral axis show stronger expression than the rest of the hindbrain, leaving a certain gap between both bands (Figure 4-12E). These two stripes are not observed in rhombomeres 1/2 where expression is fainter. Between the hindbrain and spinal cord, these two bands converge forming a big patch of stronger staining (Figure 4-12E). Notably, expression in the entire ventral hindbrain continues disappearing. From a dorsal view, expression is clearly seen in the middle of the neural tube corresponding to the ventricular zone, although it expands more laterally at the level of the epiphysis, the MHB and posterior hindbrain/anterior spinal cord (Figure 4-12G). Expression in the spinal cord is maintained with a stronger signal in the posterior half that reaches the end of the spinal cord (Figure 4-12C and F). At this stage, a very distinct pair of dots of staining is observed in the tailbud (Figure 4-12F and H).

At stage 28, *LpHes-B* expression is maintained along the entire neural tube (Figure 4-13A). One of the most striking features of *LpHes-B* expression is its downregulation in the ventral hindbrain (Figure 4-13B). Expression in other parts of the brain is essentially maintained with respect to stage 26, with the notable absence of expression in the epiphysis (Figure 4-13B). From a dorsal view, a difference with respect to earlier stages is the narrowing of expression in the ventricular zone along the neural tube, in particular in the spinal cord (Figure 4-13C and E; cf. Figure 4-12G). However, it is clear that the ventricular zone is still expanded at the level of the epiphysis, the MHB, and the transition between the hindbrain and spinal cord (Figure 4-13C). In the spinal cord, expression is observed in the middle part with respect to the dorso-ventral axis (Figure 4-13F). *LpHes-B* expression also reaches the end of the spinal cord where it covers a slightly expanded territory with respect to the dorso-ventral axis (Figure 4-13D). At this stage, no expression was observed in the tailbud in any of the embryos examined (n=5; Figure 4-13D). *LpHes-B* is also expressed in a discrete domain in the olfactory epithelium at this stage (Figure 4-13B).

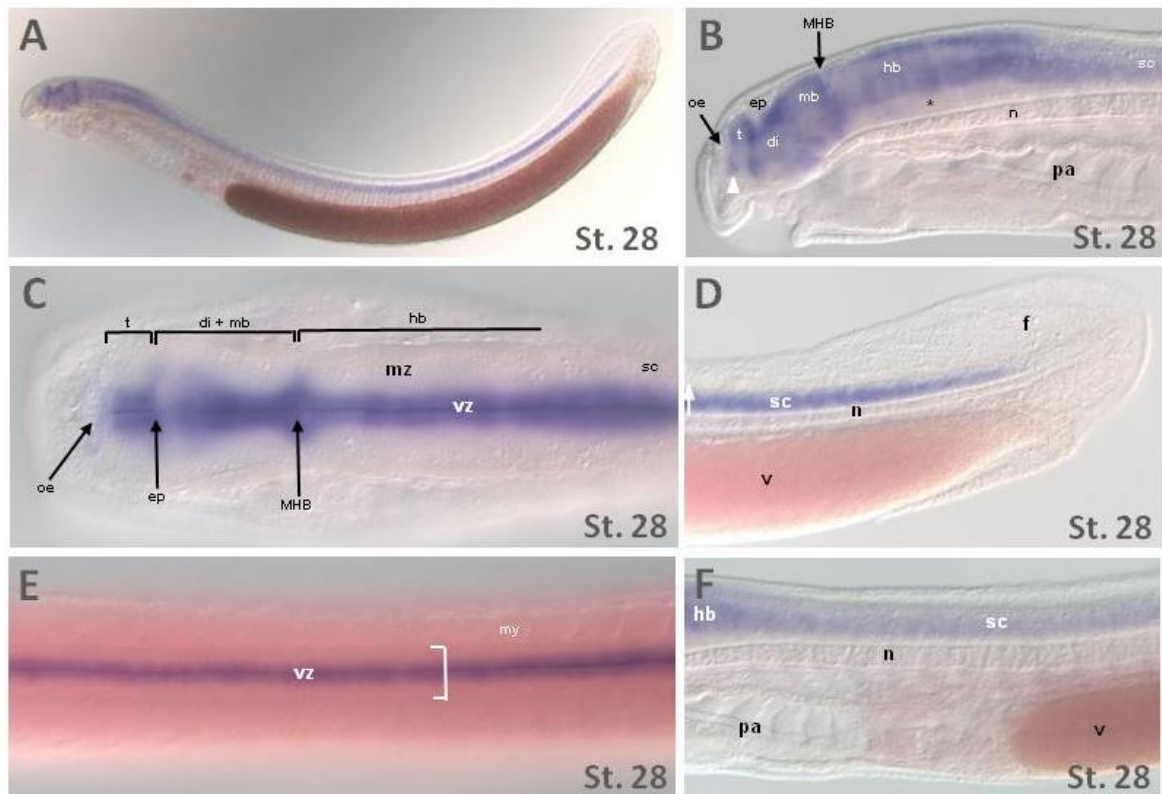


Figure 4-13. Expression of *LpHes-B* in stage 28 *L. planeri* embryos. (A), (B), (D) and (F) are lateral views of the entire embryo, head, tail and the posterior hindbrain/anterior spinal cord region centered at the heart region, respectively. (C) and (E) are dorsal views of the head and trunk region, respectively. Anterior is to the left in all images. (A) *LpHes-B* expression is still restricted to the neural tube, except for a novel expression site in the olfactory epithelium (oe; B, arrowhead), and reaches the end of the spinal cord (D). In the brain, expression is maintained in the same regions as in stages 25 and 26, but expression in the ventral hindbrain has clearly disappeared (asterisk in B). (C) From a dorsal view, expression is clearly seen in the ventricular zone (vz), which is expanded at the MHB and anterior spinal cord levels. (E) *LpHes-B* expression is also maintained in the ventricular zone of the trunk, although expression is much thinner than in previous stages (compare with brackets demarcating the width of the spinal cord). (F) In the spinal cord, expression is seen in the medial part with respect to the dorso-ventral axis. ep: epiphysis, di: diencephalon; f: caudal fin, hb: hindbrain, mz: marginal zone, mb: midbrain, MHB: midbrain-hindbrain boundary, my: myotome, n: notochord, oe: olfactory epithelium, pa: pharyngeal arches, sc: spinal cord, t: telencephalon, vz: ventricular zone, v: vitellum.

4.6. Discussion

4.6.1 Lamprey *Hairy/E(spl)* (*Hes*) Gene Sequence

The identification of lamprey *Hairy/E(spl)* homologs followed the predictions. Based on *Petromyzon marinus* *Hairy/E(spl)* sequences, *PmHes-A/LpHes-A* and *PmHes-B/LpHes-B* contain the predicted bHLH domain as well as the conserved WRPW motif at the C-terminal end (thus supporting their nomenclature as *Hes* genes and not *Hey* or *Dec* genes, which lack the WRPW motif). Both *PmHes-A/LpHes-A* and *PmHes-B/LpHes-B* do not present a short loop region as mouse and *Xenopus* *Hes6* (Figure 4-7), suggesting that, as other vertebrate *Hes* genes, they act as inhibitors of neurogenesis (Bae et al., 2000; Koyano-Nakagawa et al., 2000). However, the region that aligns with the Orange domain of mouse *Hes* proteins is quite short, with no obvious sequence conservation (Figure 4-7). This feature suggests that lamprey *Hes* genes may not have an Orange domain. Given that this domain has been implicated in protein-protein interactions (Kageyama et al., 2007; Van Wayenbergh et al., 2003), in spite of its apparent absence in both lamprey sequences, the presence of a WRPW motif (which binds to Groucho/TLE co-repressors) suggests that lamprey *Hes* proteins are still transcriptional repressors, and that they only display different protein-protein interactions while maintaining their repressor capability.

Phylogenetic analyses also place both lamprey sequences within the *Hairy/E(spl)* / *Hes* clade (although with a somewhat low bootstrap support value of 64), providing more evidence for their classification as genuine *Hes* genes. Both *P. marinus* *Hes* genes appear very closely related in phylogeny (Figure 4-9), suggesting that they are the

product of a recent gene duplication event. This is in agreement with *P. marinus* Ensembl predictions ENSPMAP00000004455 and ENSPMAP00000004457 being closely related (Figure 3-2). The presence of a third *P. marinus* *Hes* gene, Ensembl prediction ENSPMAP00000001555 (*PmHes-C*), which is more distantly related, would be consistent with a whole-genome duplication event before the divergence of lampreys and another genome duplication after lampreys if there was a single ancestral chordate *Hes* gene, *PmHes-A* and *PmHes-B* being the product of an independent gene duplication in the lamprey lineage. However, three *Hes* genes have been found in the *Ciona intestinalis* genome, and further 11-16 *Hes* genes in the *Branchiostoma floridae* genome (Simionato et al., 2007).

While it may seem that there was an ancestral wide repertoire of *Hes* genes in the chordate ancestor based on the gene content only, phylogenetic analyses with several chordate *Hes* sequences (including lamprey sequences; Figure 4-9) reveal that most amphioxus *Hes* genes (*AmphiHairyA-G*; *AmphiHairyH* seems to be more distant) are very closely related, suggesting that they are the product of independent gene duplications within the cephalochordate lineage. A very similar scenario was previously observed, where *AmphiHairyA-H* formed a single clade (Minguillon et al., 2003). Indeed, at least *AmphiHairyA-F* are linked in the amphioxus genome, suggesting that they are the product of tandem gene duplications in the cephalochordate lineage (Minguillon et al., 2003). Of note, it seems that amphioxus *Hairy* genes suffered a rapid subfunctionalisation process shortly after their duplication (Jiménez-Delgado et al., 2006; Minguillon et al., 2003). However, without a comprehensive survey of *Hes* genes in the lamprey genome, both *in silico* and experimentally, the exact evolutionary history of chordate *Hes* genes at the invertebrate-vertebrate transition is still uncertain.

4.6.2 Expression of Lamprey *Hairy/E(spl)* (*Hes*) Genes

Expression of both *LpHes-A* and *LpHes-B* is consistent with a putative role as inhibitors of neurogenesis, as they are both expressed in the ventricular zone of the neural tube, a region of proliferating neural precursors. In particular, *LpHes-B* expression is most similar to mouse *Hes3*, which initially is expressed only in the central nervous system (CNS) (Lobe, 1997). Consistent with this similarity, zebrafish *Hes3* is restricted to the rhombospinal region, although with a gap of expression in the middle and anterior spinal cord except in a few scattered cells (Hans et al., 2004). Furthermore, one of the most similar aspects of mouse *Hes3* and *LpHes-B* is their expression in specific rhombomeres, as both genes are expressed in all rhombomeres except rhombomeres 3 and 5 (Lobe, 1997). Another similarity in the brain is that at early stages both genes are absent from the forebrain, and later in development their expression advances anteriorly towards the forebrain (Lobe, 1997). However, one striking difference between *LpHes-B* and mouse *Hes3* (together with *Hes1*) is that the former is not initially expressed at the MHB. In contrast, *LpHes-B* expression in the brain is most similar to mouse *Hes5* as both are strongly expressed in the forebrain, midbrain and hindbrain but clearly absent from the forebrain-midbrain boundary and the MHB (Hatakeyama et al., 2004). This could –perhaps– be correlated to some extent with *Hes* gene phylogeny. In the Neighbour Joining tree in Figure 4-9, lamprey *Hes* genes are recovered as the sister group of vertebrate *Hes5* homologues, although with a very low bootstrap support value. This clustering might be strengthened by expression of zebrafish *Hes5e* in the entire CNS, although it is also expressed in presomitic mesoderm (Takke et al., 1999). It is not known at present whether there is any evolutionary relation of *LpHes-B* expression in the tailbud with general gnathostome presomitic *Hes* expression. Of note, zebrafish *Hes1b* expression in the brain is unlike

that of lamprey or tetrapod *Hes* genes as it is restricted to the entire forebrain and to subsets of rhombomeres that vary during development of the hindbrain (Pasini et al., 2001).

Interestingly, despite the differences in intensity, expression of both *LpHes-A* and *LpHes-B* in the CNS is nearly identical in distribution and to some extent time frame. For example, at stage 24 both genes are expressed in the ventricular zone of the entire neural tube, and in particular in the brain, both genes are expressed in large “bulges” corresponding to the telencephalon, diencephalon plus midbrain, and rhombencephalon, leaving clear gaps of expression at the telencephalon-diencephalon boundary and MHB. Also, at stage 23, expression of both genes is biased to the ventral region of the neural tube. The overall similarity in the expression patterns of *LpHes-A* and *LpHes-B* in the CNS at stages 23 and 24 is consistent with their very close phylogenetic relationship. In this sense, expression of the third lamprey *Hes* gene, *Hes-C*, is predicted to present more differences with respect to both lamprey *Hes-A* and *Hes-B*.

Many *Hes* genes are also expressed in the spinal cord, particularly in the ventricular zone where neuronal precursors proliferate. In chicken, three *Hes5* (*Hes5-1*, *Hes5-2* and *Hes5-3*) paralogs and one *Hes6* (*Hes6-2*) paralog are expressed in the ventricular zone of the spinal cord (Fior and Henrique, 2005; Vilas-Boas and Henrique, 2010). In rodents, *Hes1*, *Hes3* and *Hes5* are also expressed in the ventricular zone of the spinal cord (Hatakeyama et al., 2004; Lobe, 1997; Takebayashi et al., 1995). In chicken, *Hes5-1*, *Hes5-3* and *Hes6-2* are expressed homogenously along the dorso-ventral axis of the ventricular zone of the spinal cord, except for *Hes5-2* which is more restricted to a ventral domain (Fior and Henrique, 2005). In contrast, expression of rodent *Hes* genes

is not homogenous and instead complementary. Mouse *Hes1* is expressed in a dorsal domain at some distance from the roof plate, whereas *Hes5* is strongly expressed dorsal to *Hes1* and in a large –though weaker– expression domain ventral to *Hes1* covering the remaining of the ventricular zone, thus presenting a dorsal gap of expression where *Hes1* is transcribed (Hatakeyama et al., 2004). On the other hand, *Hes3* is very restricted to a thin dorsal domain of the ventricular zone and to the floor plate of the spinal cord (Lobe, 1997). Both *LpHes-A* and *LpHes-B* are homogeneously expressed in the ventricular zone of the neural tube all along the dorso-ventral axis. In addition, the fact that the loop region in the HLH domain of both *PmHes-A* and *PmHes-B* is similar in length to other Hes proteins, except *Hes6*, is in agreement with the expression of *LpHes-A* and *LpHes-B* in the ventricular zone if we compare to mouse and *Xenopus* (but not chicken) *Hes6* which are expressed mainly in an “intermediate zone” of differentiating neurons (Koyano-Nakagawa et al., 2000). Finally, the available information on *XHes2* expression at early stages of development does not allow an accurate comparison between *LpHes-A* and *LpHes-B* expression at later stages of brain and spinal cord development with vertebrate *Hes2* homologues (Sölter et al., 2006).

In the peripheral nervous system, *LpHes-B* (but not *LpHes-A*) is also similar to mouse *Hes3* as they are both expressed in the olfactory epithelium and cranial placodes/ganglia (Lobe, 1997), though other *Hes* genes, namely *Hes1* and *Hes5*, are also expressed in the olfactory placodes/epithelium (Cau et al., 2000). In particular, mouse *Hes3* is expressed in trigeminal, facial and acoustic ganglia and in the otic vesicle (Lobe, 1997). *Hes* expression in cranial ganglia, however, is not exclusive of *Hes3*. Mouse *Hes1* and zebrafish *Hes5e* have been reported in trigeminal ganglia (Ishibashi et al., 1995; So et al., 2009; Takke et al., 1999). However, *LpHes-B* seems to

be expressed in otic and lateral line placodes/ganglia (Figure 4-9G and H), and so it is only comparable to mouse *Hes3* expression in acoustic/otic ganglia and otic vesicle (Lobe, 1997). In chicken, *Hes5-1*, *Hes5-2* and *Hes5-3* are also expressed in the olfactory placode (Maier and Gunhaga, 2009). Overall, expression of *Hes* homologues in cranial ganglia seems to be an ancient vertebrate feature. Finally, lamprey *Hes* genes were not observed in DRG as rodent *Hes1*, *Hes3*, *Hes5* and *Hes6* (Akazawa et al., 1992; Ishibashi et al., 1995; Lobe, 1997).

4.6.3 *Dec* Genes as a New HLH Gene Subfamily

Dec genes are traditionally included within the *Hairy/E(spl)* (*Hes*) gene subfamily as their bHLH domains are similar in sequence with *Hairy/E(spl)* genes and they also possess an Orange domain. However, there is enough evidence to consider them as a separate HLH gene subfamily. Two paralogues, *Dec1* and *Dec2*, are present in vertebrates. At the sequence level, one of the most important differences of *Dec* proteins with respect to *Hairy/E(spl)* and *Hey* is the lack of the diagnostic motif WRPW or YRPW in *Hairy/E(spl)* or *Hey* proteins, respectively (Fujimoto et al., 2001; Rossner et al., 1997; Shen et al., 1997) which recognises the co-repressor Groucho/TLE to exert transcription inhibition. Instead, *Dec* proteins mediate transcriptional repression through mechanisms involving the recruitment of the histone deacetylase 1 (HDAC1)/Sin3A/NCoR repressor complex as well as through direct physical interaction with the basal transcription machinery (Boudjelal et al., 1997; Garriga-Canut et al., 2001; Sun and Taneja, 2000). Also, the basic region of the bHLH domain contains the proline residue conserved in *Hairy/E(spl)* proteins but two amino acid residues more N-terminal (Boudjelal et al., 1997). Instead, an arginine is present at the conserved

Hairy/E(spl) proline position, preceded from a histidine. *Dec* genes contain other different conserved domains or motives not present in either *Hairy/E(spl)* or *Hey* genes. Vertebrate *Dec1* genes possess four glutamine residues immediately downstream from the bHLH domain (Boudjelal et al., 1997; Yao et al., 2006), while *Dec2* genes possess the same motif except that the third residue is histidine instead of glutamine (Fujimoto et al., 2001). *Dec* genes also have a casein kinase 2 motif (SGSDTDTDS) downstream from the Orange domain and a proline-rich domain near the C-terminal end for *Dec1* (Boudjelal et al., 1997; Yao et al., 2006) and an alanine-glycine-rich region for *Dec2* (Fujimoto et al., 2001). Furthermore, Dec proteins do not bind E- or N-boxes, although they are able to homodimerise (Boudjelal et al., 1997; Garriga-Canut et al., 2001) or heterodimerise with Mash1 and weakly with E12 (Boudjelal et al., 1997).

The patterns and dynamics of expression of *Dec* genes in the nervous system are slightly different from those of *Hairy/E(spl)* or *Hey* genes. Embryonic expression of *Dec* genes has so far been reported in mouse and zebrafish. Mouse *Dec1* is first detected in rhombomere 5 and later extends in the ventricular layer of the brain and spinal cord (Boudjelal et al., 1997). *Dec1* is also expressed in the optic stalk and prospective retinal pigment epithelium, as well as in the olfactory epithelium (Boudjelal et al., 1997; Rossner et al., 1997). Outside the nervous system, prominent expression is observed in prevertebral condensations and later in cartilaginous bone models of the ribs, scapula and limbs (Boudjelal et al., 1997). Zebrafish *Dec1* is expressed in the optic vesicles and later in retinal pigment epithelium and iris and pineal gland. Outside the nervous system zebrafish *Dec1* is expressed in somitic mesoderm.

On the other hand, mouse *Dec2* is initially expressed in a dorsal domain of the telencephalon and diencephalon and later transcripts are observed in the cerebral cortex (Azmi and Taneja, 2002). *Dec2* is also expressed in the optic vesicles and later in neural and pigmented retina and in the lens, as well as in olfactory placodes and later in olfactory and vomeronasal epithelia (Azmi and Taneja, 2002). Cranial sensory ganglia and spinal nerves also express *Dec2* (Azmi and Taneja, 2002). Outside the nervous system, *Dec2* is expressed in several tissues including the first and second branchial arches, heart, liver, thymus, myotomes, prevertebrae, central region of the limb buds and subsequently in interdigital mesenchyme of the developing limbs excluding the digits (Azmi and Taneja, 2002).

Early studies describing the sequence and expression of *Dec* genes have recognised their distinction as a separate subfamily (Azmi and Taneja, 2002; Fujimoto et al., 2001; Rossner et al., 1997), but later works reporting full sets of HLH genes in different vertebrate genomes have consistently included *Dec* genes within the *Hairy/E(spl)* subfamily. Therefore, so far *Dec* genes have not been formally recognised as a distinct HLH gene subfamily when analysing the HLH gene family as a whole. In this regard, broad phylogenetic analysis of lamprey, human and sea urchin HLH genes (Figure 3-1) and a more detailed phylogeny of *Hairy/E(spl)*, *Hey* and *Dec* genes (Figure 4-8), clarify this distinction, as *Dec* genes consistently group as a separate group, with high support value, and distant from either *Hairy/E(spl)* or *Hey* genes. Taken together, based on sequence, expression patterns, mechanisms of repression, mutation phenotypes and phylogenetic analyses, *Dec* genes should be formally erected as a separate HLH gene subfamily.

4.7. Summary of Lamprey *Hairy/E(spl)* Genes

Two genes, *LpHes-A* and *LpHes-B*, belonging to the *Hairy/E(spl)* gene subfamily were experimentally characterised in this study. Sequence analyses of their *P. marinus* orthologues revealed the presence of a well-conserved HLH domain, which contains the diagnostic proline residue in the basic region of the HLH domain and the WRPW motif at the C-terminal end of the sequence. However, both fragments appear to lack an Orange domain. Phylogenetic analyses suggest that both genes are the product of a recent gene duplication event, which is consistent with their very similar expression patterns.

LpHes-A is only expressed in the CNS at the embryonic stages analysed (stages 23-24). Its expression is relatively homogenous and is restricted to the ventricular zone of the neural tube and, in the brain, it progresses from the ventral to the dorsal side. Gaps of expression in the brain seem to highlight embryonic brain segments.

LpHes-B is expressed almost exclusively in the central nervous system, with the exception of cranial ganglia at stage 23, in the tailbud of stage 26 embryos and olfactory epithelium at stage 28. In the neural tube, *LpHes-B* is found in the middle, suggesting the territory of the ventricular zone with expansions at the level of the epiphysis, MHB and posterior hindbrain/anterior spinal cord at stages 26-28. This expression becomes very thin at stage 28 suggesting a reduced ventricular zone. In the brain of stage 22 embryos, however, expression is observed as two lateral stripes with a clear gap in the middle as seen from a dorsal view. At stages 22-24, expression in the hindbrain is

excluded from rhombomeres 3 and 5. Throughout development, expression in the hindbrain generally progresses from the ventral to the dorsal side as *LpHes-A*. Along the anterior-posterior axis, expression first appears in the rhombospinal region, later in the midbrain and finally in the forebrain.

5. *LAMPETRA PLANERI* NEUROGENIN GENES

5.1. Introduction

Neurogenins, catalogued as pro-neural HLH genes, present many different functions during vertebrate development. They are evolutionarily conserved among vertebrates and are almost exclusively expressed in the nervous system in all vertebrates examined to date.

Structure and function of neurogenin genes: At the sequence level, neurogenins are rather simple, possessing only the bHLH domain (Sommer et al., 1996). When neurogenins homodimerise, they do not form a well folded structure, and hence, they do not act efficiently (Aguado-Llera et al., 2010), nor when they dimerise with other tissue-specific neural HLH proteins such as Mash1 in mouse (Gradwohl et al., 1996). Instead, they dimerise with the ubiquitous E12/E47 HLH proteins to bind to E-box promoter sequences of target genes to activate their transcription (Gradwohl et al., 1996).

Developmental roles of Neurogenin genes: Neurogenins have been demonstrated to be powerful inducers of neurogenesis and are necessary for neuronal differentiation. For example, in zebrafish and *Xenopus*, *neurogenin1* mRNA injection causes ectopic neurogenesis in the ectoderm and induces other genes involved in neuronal differentiation such as *NeuroD* (Blader et al., 1997; Ma et al., 1996). Forced expression of mouse *neurogenin1* in chicken premigratory neural crest cells results in neural crest precursors biased to a sensory neuron (DRG) fate (Perez et al., 1999). In addition,

misexpression of mouse *neurogenin1* and *neurogenin2* and chicken *neurogenin1* in migrating chicken neural crest cells induces expression of sensory neuron-specific markers in sympathetic ganglia and glial cells of peripheral nerves (Perez et al., 1999). Furthermore, misexpression of mouse *neurogenin1* and *neurogenin2* and chicken *neurogenin1* in chicken somites is able to induce expression of sensory neuron markers in the dermomyotome (Perez et al., 1999). In the chicken hindbrain, electroporated rat *neurogenin2* promotes cell cycle exit and neuronal differentiation of neural progenitors (Mizuguchi et al., 2001). Indeed, increasing evidence has revealed a role of Ngn2 in driving cell cycle exit of neuronal precursors and independently promoting neuronal differentiation (Ali et al., 2011; Lacomme et al., 2012; Novitch et al., 2001).

Additionally, *in vitro* evidence suggests that both mouse *neurogenin1* and *neurogenin2* might actively inhibit glial (astrocytic) cell differentiation in the CNS independently of promoting neurogenesis (Morrison, 2001; Nieto et al., 2001; Sun et al., 2001). More recently, it has been noted that zebrafish *neurogenin1* presents a similar function in the PNS by inhibiting a glial fate, in this case, myelinating Schwann cells (McGraw et al., 2008). From this, another role of neurogenins that has been distinguished is inhibition of neural crest, as was previously observed in zebrafish where inhibition of *neurogenin1* leads lateral neural plate precursor cells to a neural crest cell fate and their derivatives instead of a Rohon-Beard –primary– sensory neuron fate (Cornell and Eisen, 2002).

Expression of Neurogenin genes in the vertebrate nervous system: Several *neurogenin* homologues have been described in different vertebrate species, which show very similar expression patterns and are almost exclusively expressed in the nervous system. In zebrafish, two *neurogenin* homologues have been identified, *neurogenin1* and *neurogenin3* (Blader et al., 1997; Wang et al., 2001). At neural plate stages (8.5-11 hpf), zebrafish *neurogenin1* is expressed in the three stripes of primary motor, intermediate and sensory neuron precursors and in the trigeminal placodes (Andermann et al., 2002; Blader et al., 1997; Korzh et al., 1998). Later in development (24 hpf), zebrafish *neurogenin1* is expressed in the brain, spinal cord and trigeminal ganglia (Korzh et al., 1998). On the other hand, *neurogenin3* is only expressed in the hypothalamus (Wang et al., 2001).

The genus *Xenopus* has three *neurogenin* genes: *neurogenin1*, *neurogenin2* and *neurogenin3* (Nieber et al., 2009). Expression in *Xenopus laevis* and *Xenopus tropicalis* is quite similar although they also present slight differences in intensity and expression sites between both species. As in zebrafish, expression of *neurogenin1* and *neurogenin2* at open neural plate stages (stages 12-15) is seen in the three parallel stripes of primary neurogenesis (medial, intermediate, lateral) that will give rise to motoneurons, interneurons and Rohon-Beard sensory neurons, respectively (Ma et al., 1996; Nieber et al., 2009). In contrast, *neurogenin3* is expressed only in a longitudinal, bilateral domain close to the midline that lies between the medial and intermediate stripes (Nieber et al., 2009). In the anterior neural plate, *neurogenin1* and *neurogenin2* transcripts are localised in the trigeminal placodes (Nieber et al., 2009). *X. tropicalis neurogenin1* is also expressed in the otic placode, while both *X. tropicalis* and *X. laevis neurogenin2* are expressed in the olfactory placode (Nieber et al., 2009). *neurogenin1* is also

expressed in a horseshoe-shaped domain in the anterior neural plate that gives rise to the ventral midbrain (Nieber et al., 2009).

Later in development at neurula stages (stage 20-22), in both species *neurogenin1* is expressed in the neural tube, trigeminal, otic and olfactory placodes, and in a domain anterior to the MHB (Nieber et al., 2009). On the other hand, *neurogenin2* is found in the midbrain, hindbrain and spinal cord (Nieber et al., 2009). *X. tropicalis neurogenin2* is also expressed in the eye anlage, whereas *X. laevis neurogenin2* is expressed in the trigeminal, otic, olfactory and facial epibranchial placodes, and in the neural tube but not in the eye region (Nieber et al., 2009). *neurogenin3* transcripts are found in the spinal cord, hindbrain and in a domain anterior to the MHB (Nieber et al., 2009). Older embryos (stage 30) show that both *neurogenin3* genes are expressed in a medial and a ventral longitudinal domain along the hindbrain and spinal cord, both domains being localised in the ventricular zone (Nieber et al., 2009).

In chicken, only two *neurogenin* homologues have been identified so far, *neurogenin1* and *neurogenin2*. Expression of both genes has been detected in the spinal cord in a very dynamic pattern. At stages HH14-17, both genes are observed in migrating neural crest cells adjacent to the neural tube (Perez et al., 1999). At the same time, both genes are expressed in the neural tube: at the level of the forelimb, *neurogenin1* is strongly expressed in a dorsal domain and along the periphery of the neural tube except at the ventral-most part; on the other hand, *neurogenin2* expression is split into a dorsal and a much wider ventral domain that covers both the ventricular and marginal zones (Perez et al., 1999). Later in development at stage HH27, *neurogenin1* expression acquires an expression pattern similar to *neurogenin2* in a dorsal

(corresponding to dI2 interneurons) and a wider ventral domain, while *neurogenin2* is now expressed mostly all along the subventricular zone and in a discrete ventral domain within the ventricular zone (Perez et al., 1999; Vilas-Boas and Henrique, 2010). At this stage, *neurogenin2* has completely disappeared in the condensing DRG while *neurogenin1* is expressed as late as stage HH31 in the DRG periphery that consists of non-neuronal cells and/or neuronal precursors (the central region of the DRG consists of differentiating neurons). By this time (stage HH27), neither gene is expressed adjacent to the neural tube in the dorsal region given that neural crest migration is now complete by this stage (Perez et al., 1999). Neither gene is expressed in the differentiated DRG (Aaker et al., 2010).

Chicken *neurogenin1* is expressed in all epibranchial placodes, in the maxillomandibular trigeminal placode and in the vestibuloacoustic portion of the otic vesicle (Begbie et al., 2002; Abu-Elmagd et al., 2001; Furlong and Graham, 2005; Xu et al., 2008). On the other hand, chicken *neurogenin2* is only observed in the ophthalmic trigeminal placode at early stages (Begbie et al., 2002; Furlong and Graham, 2005; Xu et al., 2008), but is later observed in the olfactory epithelium, maxillomandibular trigeminal ganglion, facial ganglion and the neurogenic region of the otic vesicle (Xu et al., 2008). Chicken *neurogenin1* is also expressed in the olfactory placode and epithelium (Abu-Elmagd et al., 2001), while *neurogenin2* is also expressed in the developing retina in retinal ganglion cell precursors (Maier and Gunhaga, 2009).

There are three neurogenin genes identified in mouse, *neurogenin1*, *neurogenin2*, and *neurogenin3*. The expression patterns of mouse neurogenins are transient and restricted to the developing nervous system. In the mouse central nervous system

(CNS), *neurogenin1* and *neurogenin2* are expressed in restricted regions of the brain, particularly in the ventricular zone of the cortical neuroepithelium, diencephalon and hindbrain (Sommer et al., 1996). They are also found in the dorsal thalamus. *neurogenin3*, however, shows an even more restricted expression pattern, as it is mostly limited to the hypothalamus (as with zebrafish *neurogenin3*), where neither *neurogenin1* nor *neurogenin2* are expressed (Sommer et al., 1996). In the murine spinal cord, *neurogenin1* is expressed in a narrow domain below the roof plate and in the ventricular zone of the ventral spinal cord, a region provided with uncommitted progenitors (Ma et al., 1996). *neurogenin2* expression overlaps that of *neurogenin1*, but shows a stronger and broader expression pattern, and its expression is also activated earlier in development than *neurogenin1* (Sommer et al., 1996). On the contrary, *neurogenin3* is restricted to a very narrow domain just dorsal to the floor plate, although it is later activated in a slightly more dorsal domain where it overlaps with *neurogenin1* and *neurogenin2*. Of note, *neurogenin2* and *neurogenin3* are expressed more medially in the ventricular zone than *neurogenin1*.

While in the CNS expression of *neurogenin1* and *neurogenin2* overlaps in many regions, in the PNS their expression is rather complementary, where *neurogenin2* generally precedes *neurogenin1*. Separate expression of *neurogenin1* and *neurogenin2* occurs in the trigeminal and acoustic/otic cranial ganglia, as well as in olfactory receptor neuron (ORN) progenitors of the olfactory epithelium, where only *neurogenin1* is expressed (Cau et al., 1997; Sommer et al., 1996), although there is co-expression of *neurogenin1* and *neurogenin2* in a small ventrocaudal domain of the olfactory epithelium (Cau et al., 1997). On the other hand, *neurogenin2* is expressed in the

epibranchial placodes/ganglia and in the retina, where no *neurogenin1* expression is observed (Gradwohl et al., 1996; Sommer et al., 1996).

neurogenin2 is first expressed in migrating neural crest cells located dorso-lateral to the neural tube, followed by *neurogenin1* in the forming dorsal root sensory ganglia (DRG) (Sommer et al., 1996). However, at subsequent developmental stages, *neurogenin2* expression in the DRG extinguishes before that of *neurogenin1*. Interestingly, *neurogenin2* is expressed in migrating neural crest cells and forming DRG at the thoracolumbar but not at the cervical level, whereas *neurogenin1* is expressed in the forming DRG at both the thoracolumbar and cervical levels, possibly marking two distinct neural crest cell populations at different axial levels. Instead, *neurogenin2* is strongly expressed in the epibranchial placodes, which also contribute neurons to the cranial sensory ganglia. *neurogenin3* does not seem to be expressed in trunk neural crest or its derivatives at any stage. However, neither *neurogenin1* nor *neurogenin2* are expressed in sympathetic or enteric ganglia (Sommer et al., 1996).

5.2. Aims

The aim of this chapter is to characterise the sequence and expression patterns of *Lampetra planeri* genes belonging to the *Neurogenin* gene subfamily based on their identification *in silico* in the *Petromyzon marinus* genome (Chapter 3). This chapter is intended to answer the following questions:

- Does *L. planeri* have *Neurogenin* genes as *P. marinus*?
- Provided the recovery of *L. planeri Neurogenin* genes, where are *L. planeri Neurogenin* genes expressed? Are they found in the nervous system as in gnathostomes?
- In the CNS, are *L. planeri Neurogenin* genes expressed in the ventricular zone of the neural tube as in gnathostomes? Are they expressed in specific compartments of the brain as in gnathostomes? Are they also expressed in longitudinal domains, one dorsal and one ventral, in the spinal cord as in the mouse?
- In the PNS, are *L. planeri Neurogenin* genes expressed in cranial sensory placodes/ganglia as in gnathostomes? Are they also expressed in dorsal root ganglia? Are they expressed in migrating neural crest cells as in gnathostomes?

5.3. Hypothesis

The *Neurogenin* HLH subfamily is widespread in the vertebrate group and homologues of this subfamily are key regulators of neurogenesis in all vertebrate species examined so far. Considering that invertebrate chordates possess one *Neurogenin* homolog, it is predicted that lampreys have at least one neurogenin homolog, more similar in sequence to other vertebrate neurogenins than to their invertebrate chordate relatives. It is also expected that lampreys have more than one

neurogenin homologue, considering the genome-wide duplications that occurred at the origin of vertebrates. It is also predicted that a given lamprey neurogenin homolog is expressed almost exclusively in the nervous system, as occurs with several gnathostome neurogenins. In the same sense, it is predicted that a lamprey neurogenin homolog is expressed in the central nervous system in restricted regions of the brain and along the spinal cord, particularly in the ventricular zone or close to it, where it is also expected to be found in a dorsal and a ventral longitudinal domain. In the peripheral nervous system, it is predicted to be expressed in cranial sensory placodes and/or sensory ganglia, as well as in dorsal root ganglia lateral to the spinal cord.

5.4. Objectives

- To isolate neurogenin homologues from the *Lampetra planeri* genome.
- Perform phylogenetic analyses to examine the relationships of lamprey neurogenin sequences with respect to other deuterostome homologues.
- Study the expression patterns of the isolated sequences by *in situ* hybridisation.
- Compare these expression data with available information from other vertebrates.

5.5. Results

5.5.1. Cloning of *LpNgn*

One *neurogenin* homologue could be isolated from *L. planeri* total RNA. This gene, named *LpNgn*, was isolated by RACE-PCR from a cDNA sequence made from total RNA extracted from stage 24-26 mashed embryos as explained in section 2.2.6. in Methods. This fragment contains the 5'-most region of the transcript and a portion of the bHLH domain, that is, the basic domain, helix 1 and the loop region, but lacks helix 2 and the downstream sequence (Figure 5-1 and 5-2). Compared to a published *neurogenin* homologue from *Lampetra fluviatilis* (*LfNgn1*; Guérin et al. 2009) and the sole *neurogenin*-like gene prediction in the *Petromyzon marinus* genome by the Ensembl genome database, the three sequences are nearly identical at the nucleotide (Figure 5-1) and amino acid level (Figure 5-2).

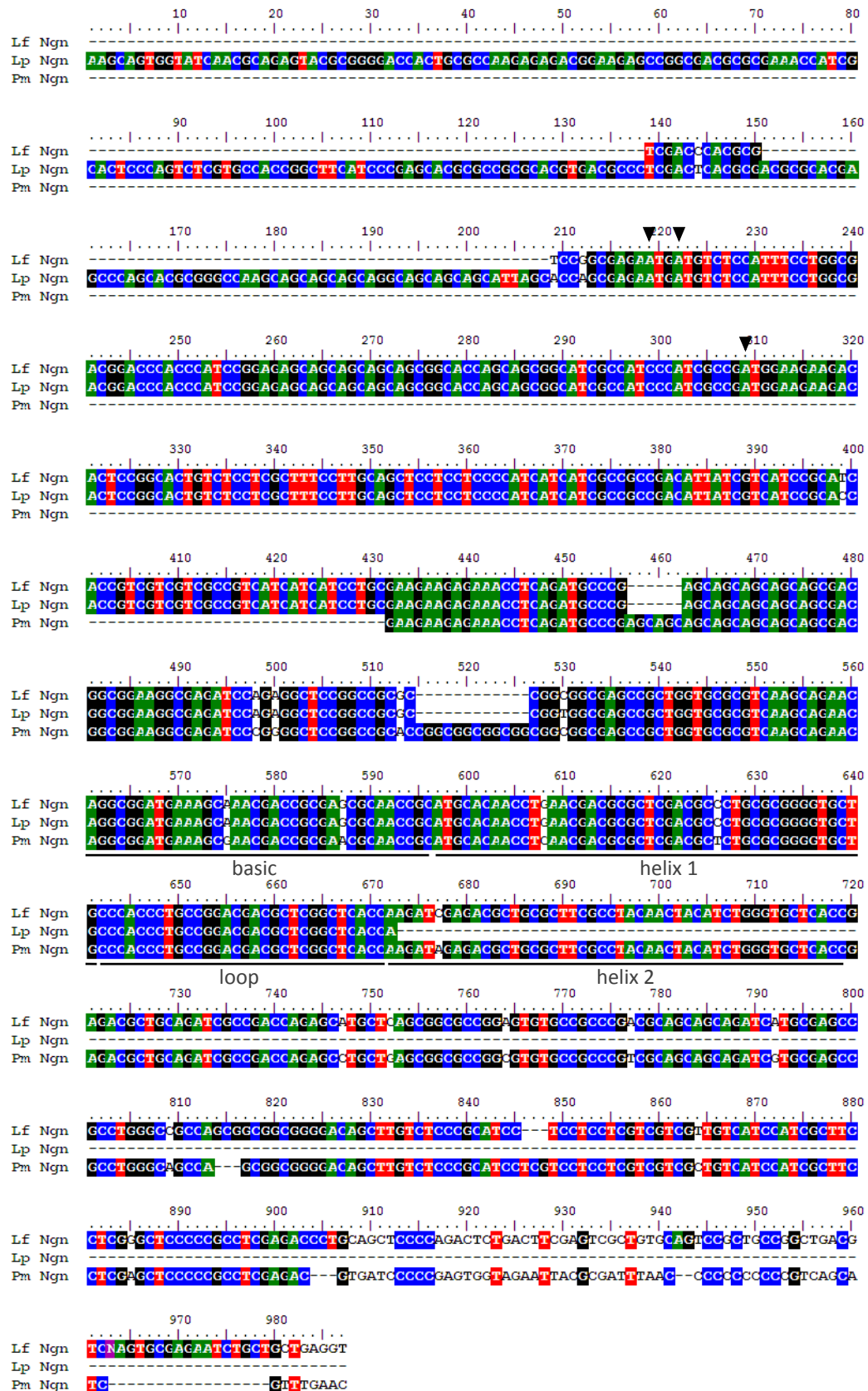


Figure 5-1. Nucleotide alignment of *LpNgn* and *LfNgn* (*LfNgn1*) gene fragments with the predicted *P. marinus* neurogenin homologue (*PmNgn*) in the Ensembl genome database. The three fragments present near-identical sequences where there is available sequence information, except at the 3' end of *LfNgn* and

PmNgn. The bHLH domain is underlined as indicated. Possible methionine start codons are indicated with arrowheads. GenBank accession number for *LfNgn1* is FP243339.1. Ensembl transcript prediction number for *PmNgn* is ENSPMAT00000006843. Basic HLH domain boundaries are according to Ledent and Vervoort (2001).

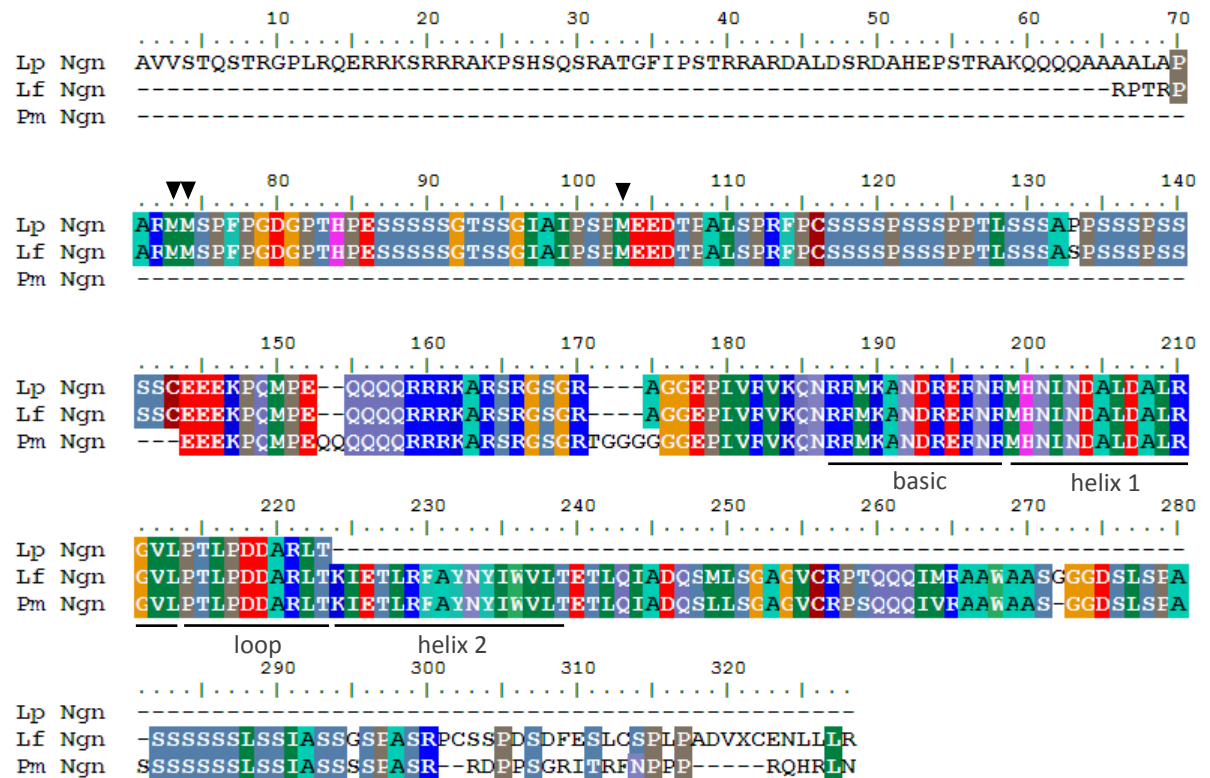


Figure 5-2. Alignment of predicted amino acid sequences of *LpNgn*, *LfNgn* (*LfNgn1*) and *PmNgn* fragments. The bHLH domain is underlined as indicated. Possible methionine start codons are indicated with arrowheads. Basic HLH domain boundaries are according to Ledent and Vervoort (2001).

When compared to sequences from other vertebrates, it can be seen that the three lamprey sequences have a well conserved bHLH domain (Figure 5-3).

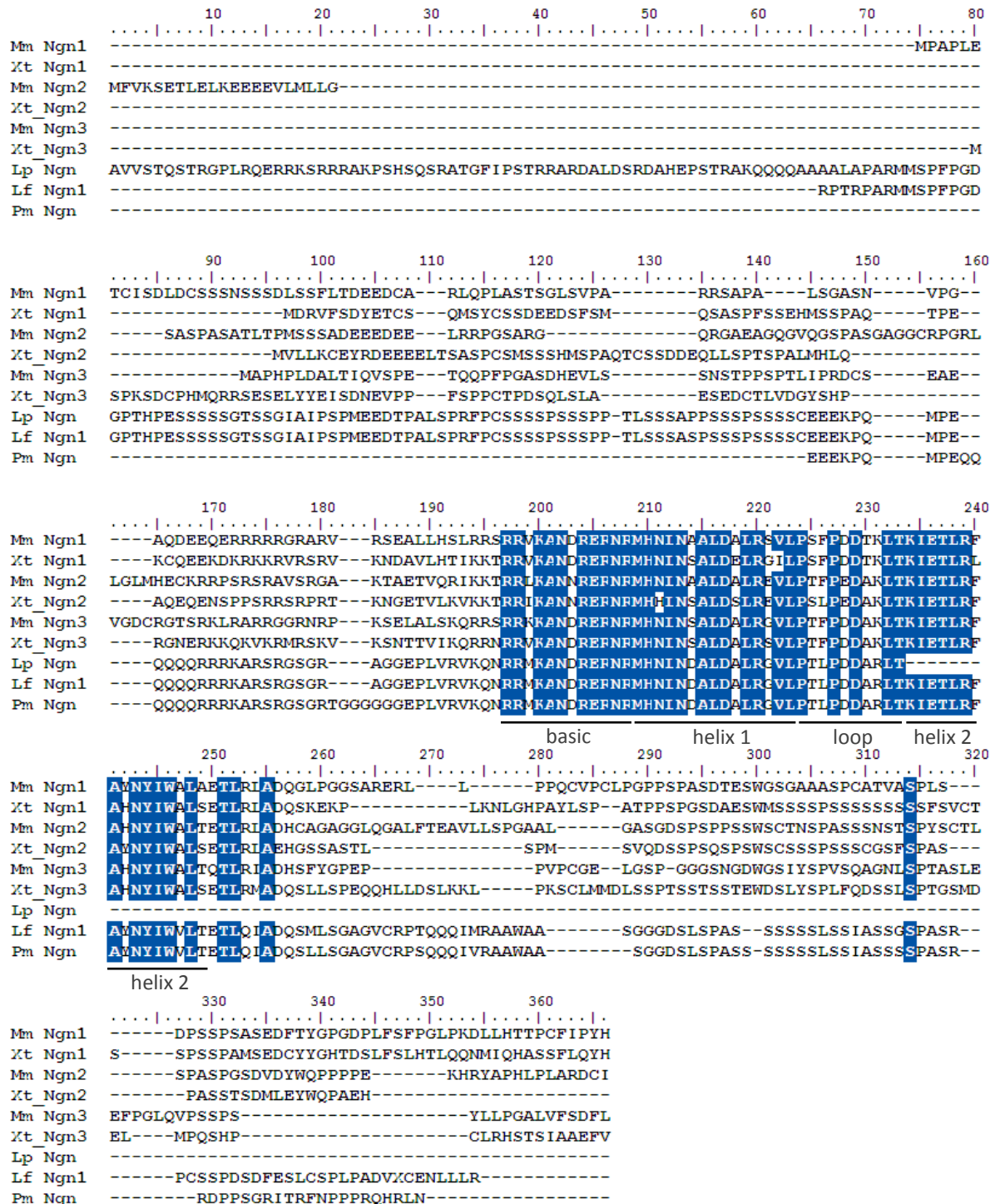


Figure 5-3. Alignment of predicted amino acid sequences of *Lampetra planeri* (Lp), *Lampetra fluviatilis* (Lf) and *Petromyzon marinus* (Pm) neurogenin gene fragments with neurogenin1-3 amino acid sequences from *Xenopus tropicalis* (Xt) and *Mus musculus* (Mm). Shadows highlight conserved residues in at least six of the eight compared sequences. Basic HLH domain boundaries are according to Ledent and Vervoort (2001).

5.5.2. Phylogenetic Analysis of *LpNgn*

To determine the phylogenetic relationships of *LpNgn*, *LfNgn* and *PmNgn*, *neurogenin* homologues of representative chordate species were used to construct a phylogenetic tree (Table 5-1). Sequences were first aligned using MAFFT software version 6.864b as described in Methods (Figure 5-4). A Neighbour joining tree was constructed as described in Methods.

Species	Gene name	Accession number	Species	Gene name	Accession number
<i>Homo sapiens</i>	<i>HsNgn1</i>	NP_006152.2	<i>Homo sapiens</i>	<i>HsNeuroD1</i>	NP_002491.2
<i>Homo sapiens</i>	<i>HsNgn2</i>	NP_076924.1	<i>Homo sapiens</i>	<i>HsNeuroD2</i>	NP_006151.3
<i>Homo sapiens</i>	<i>HsNgn3</i>	NP_066279.2	<i>Homo sapiens</i>	<i>HsNeuroD4</i>	NP_067014.2
<i>Mus musculus</i>	<i>MmNgn1</i>	NP_035026.1	<i>Homo sapiens</i>	<i>HsNeuroD6</i>	NP_073565.2
<i>Mus musculus</i>	<i>MmNgn2</i>	NP_033848.1	<i>Mus musculus</i>	<i>MmNeuroD1</i>	NP_035024.1
<i>Mus musculus</i>	<i>MmNgn3</i>	NP_033849.3	<i>Mus musculus</i>	<i>MmNeuroD2</i>	NP_035025.3
<i>Gallus gallus</i>	<i>GgNgn1</i>	NP_990214.1	<i>Mus musculus</i>	<i>MmNeuroD4</i>	NP_031527.1
<i>Gallus gallus</i>	<i>GgNgn2</i>	NP_990127.1	<i>Mus musculus</i>	<i>MmNeuroD6</i>	NP_033847.1
<i>Xenopus laevis</i>	<i>XlNgn1</i>	DC114135.1			
<i>Xenopus laevis</i>	<i>XlNgn2</i>	NP_001081804.1	<i>Homo sapiens</i>	<i>HsATOH1</i>	NP_005163.1
<i>Xenopus laevis</i>	<i>XlNgn3</i>	NP_001128257.1	<i>Homo sapiens</i>	<i>HsATOH7</i>	NP_660161.1
<i>Danio rerio</i>	<i>DrNgn1</i>	NP_571116.1	<i>Mus musculus</i>	<i>HsATOH1</i>	NP_031526.1
<i>Danio rerio</i>	<i>DrNgn3</i>	NP_571890.1	<i>Mus musculus</i>	<i>HsATOH7</i>	NP_058560.1
<i>Lampetra fluviatilis</i>	<i>LfNgn1</i>	FP243339.1			
<i>Lampetra planeri</i>	<i>LpNgn</i>	This study			
<i>Petromyzon marinus</i>	<i>PmNgn</i>	ENSPMAP00000006813			
<i>Ciona intestinalis</i>	<i>CiNgn</i>	NP_001071999.1			
<i>Branchiostoma lanceolatum</i>	<i>BlNgn</i>	ACE79717.1			
<i>Branchiostoma floridae</i>	<i>BfNgn</i>	AAF81766.1			

Table 5-1. List of chordate *neurogenin* sequences used to construct the phylogenetic tree shown in Figure 5-4. GenBank (or Ensembl for *P. marinus*) accession numbers are to the right of each gene.



Figure 5-4. MAFFT alignment of *neurogenin* sequences (shown in Table 5-1) used to construct the phylogenetic tree in Figure 5-5.

The three lamprey sequences group together with a 99 bootstrap support value, *LfNgn* and *PmNgn* being the closest among them with a 76 bootstrap support value (Figure 5-5). The three lamprey sequences are contained inside a clade defined by a bootstrap support value of 99 (colored in red in Figure 5-5) that includes all other chordate *neurogenins*. However, they do not cluster with a particular vertebrate *neurogenin* subgroup (i.e. *neurogenin* 1-3) with a significant bootstrap support value. In addition, this phylogenetic tree includes vertebrate members of the closely related *NeuroD* gene subfamily, which forms a distinct clade defined by a bootstrap support value of 99 (colored in blue in Figure 5-5). The *Atonal* HLH subfamily was used as an outgroup (colored in green in Figure 5-5). All together, this analysis provides strong support for the classification of *LpNgn*, *LfNgn* and *PmNgn* as true *neurogenin* homologues.

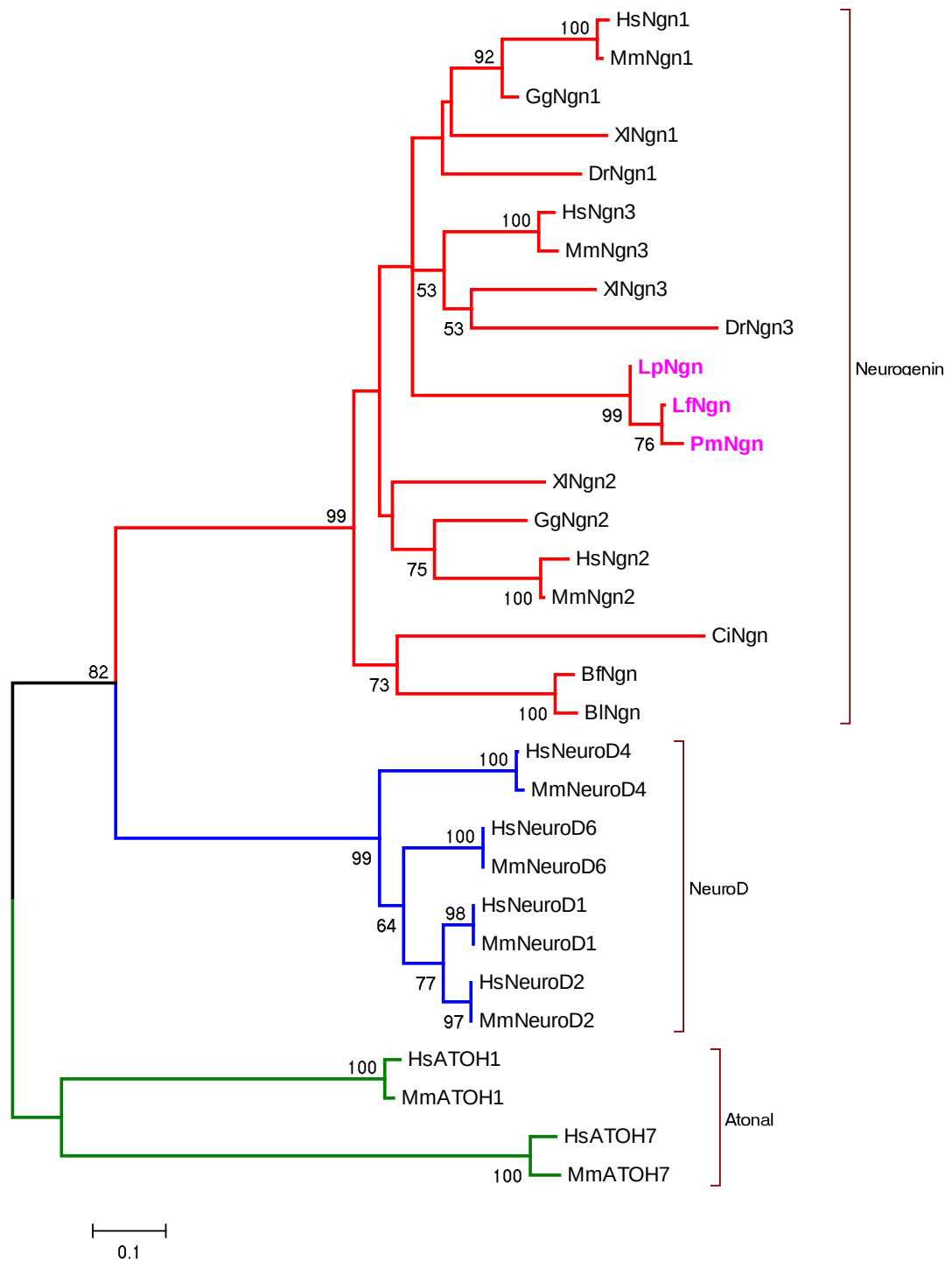


Figure 5-5. NJ phylogenetic tree using the alignment shown in Figure 5-4. Tree based on the Jones-Taylor-Thornton (JTT) matrix amino acid substitution model, 1000 bootstrap replicates and gamma distribution. Only bootstrap support values above 50 are shown. GenBank accession numbers for each sequence are given in Table 5-1. Scale bar, amino acid substitutions per position. Note how *LpNgn*, *LfNgn* and *PmNgn* form a single clade within chordate *neurogenins*.

5.5.3. *LpNgn* Expression in Developing *L. planeri* Embryos

Expression of *LpNgn* was analysed from Tahara (1988) stage 21 to stage 28. At stage 21, expression of *LpNgn* is seen in the developing spinal cord, in a discrete domain of the forebrain, and in the presumptive nasohypophyseal plate (NHP) (Figure 5-6A). From a dorsal view, expression in the spinal cord at this stage is seen as two lateral stripes that gradually meet towards the posterior with a punctuated pattern (Figure 5-6B). At stage 22, expression remains in the same areas, but in the spinal cord a dorsal and a ventral longitudinal expression domain are discerned (Figure 5-6C). At stage 23, expression in these domains increases running along the spinal cord (Figure 5-6D). In the brain, expression considerably increases in the developing diencephalon and midbrain (Figure 5-6D). Interestingly, expression domains in the brain are generally biased to the ventral half. Expression also persists in the NHP and transcripts are now seen in glossopharyngeal, vagal and lateral line placodes (Figure 5-6D). Cranial placodes and ganglia at this and subsequent stages were identified following descriptions of (Damas, 1944; Kuratani et al., 1997; Modrell et al., 2014). Anatomical internal references to aid in their identification were pharyngeal arches, otic vesicle, MHB, epiphysis and the transition of the notochord to the precordial plate.

At stage 24, expression considerably intensifies in these placodes and it is observed in what seems to correspond to the facial placode/ganglion and in trigeminal ganglia (Figure 5-6E-H). In the brain, expression increases even more in the diencephalon in a domain delimited by the zona limitans intrathalamica (Figure 5-6E-G, white asterisks) and a novel expression site appears in a more dorsal domain of the diencephalon (Figure 5-6E-G, solid black arrowheads). Expression in the midbrain persists and it is also seen

in the ventral hindbrain, with some expression expanded dorsally at the level of rhombomere 4-5 (Figure 5-6E and G). Staining is now seen in the telencephalon. Notably, expression in the diencephalon and midbrain progressively extends dorsally into the dorsal half of the brain. Mild expression is also observed in the epiphysis (Figure 5-6G). Expression in the spinal cord is still observed as two longitudinal (one dorsal and one ventral) expression domains along its length (Figure 5-6E).

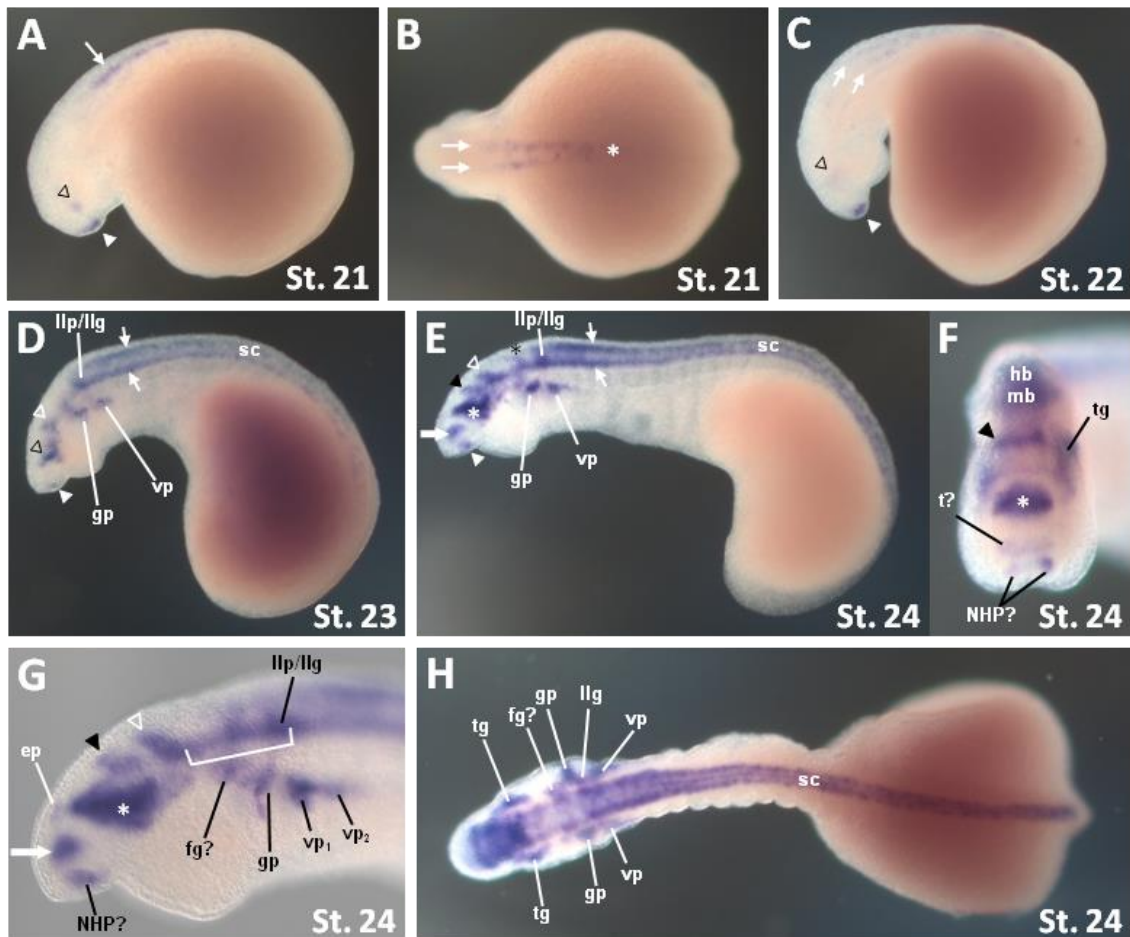


Figure 5-6. Expression of *LpNgn* at stages 21-24 in *L. planeri* embryos. (A, C, D, E and G) are lateral views. (B and H) are dorsal views. In all images anterior is to the left except in (F) which is a frontal view of the embryo shown in (G). (A) At stage 21, *LpNgn* is expressed in the nascent spinal cord (arrow), in the prospective nasohypophyseal plate (NHP; solid white arrowhead) and in a discrete domain in the presumptive diencephalon (open black arrowhead). (B) From a dorsal view, expression in the spinal cord is observed as two lateral stripes (arrows) with a punctuated pattern, and gradually meet posteriorly (asterisk). (C) At stage 22, expression in the spinal cord is now seen as a dorsal and a ventral longitudinal expression domain (arrows). Expression persists in prospective NHP (solid white arrowhead) and presumptive diencephalon (open black arrowhead). (D) At stage 23, expression appears in

glossopharyngeal (gp) and vagal (vp) placodes, lateral line placode/ganglion (llp/llg) and midbrain (open white arrowhead). Expression in the diencephalon (open black arrowhead) and spinal cord (arrows) considerably increases, while that in the NHP (solid white arrowhead) seems reduced when compared with previous stages. (E-H) At stage 24, expression persists in the midbrain (open white arrowheads), in the region of the diencephalon demarcated by the zona limitans intrathalamica (white asterisks), and it is now observed in the telencephalon (large arrow), in an additional domain of the diencephalon (solid black arrowheads), epiphysis (ep), ventral hindbrain (white brackets), in a dorsally extending domain at the level of rhombomere 4-5 (black asterisk), trigeminal ganglia (tg), in the nascent second vagal placode (vp₂) and, possibly, in facial ganglia (fg). Cranial placodes were identified following descriptions of (Damas, 1944; Kuratani et al., 1997; Modrell et al., 2014) using the pharyngeal arches, otic vesicle, MHB, epiphysis and the transition of the notochord to the precordal plate as anatomical internal references. Abbreviations: hb, hindbrain; mb, midbrain; sc, spinal cord; t, telencephalon; vp₁₋₂, vagal placode 1-2.

At stage 25, expression notably expands in the telencephalon, covering mostly the dorsal half and the anterior aspect of the ventral half (Figure 5-7A and B, white arrows). Diencephalic expression remains strong (Figure 5-7A and B, white asterisk) and the midbrain now presents distinct dorsal and ventral expression domains (Figure 5-7A and B, open white arrowheads). The hindbrain elongates and two patches of expression on the dorsal side are observed, one at the posterior hindbrain and another at the level of rhombomeres 3-5 (Figure 5-7B, black asterisks). Expression posterior to the big patch continues along the spinal cord as a longitudinal dorsal line. A ventral longitudinal line of expression is observed running continuously from the MHB towards the spinal cord. Therefore, two longitudinal expression domains, one dorsal and one ventral, are observed in the rhombospinal region of the neural tube.

An interesting feature of stage 25 embryos is the presence of tiny dots of staining in the middle of each pharyngeal arch with respect to the dorso-ventral axis (Figure 5-7B). This staining is only observed in stage 25 embryos. Weak expression is also observed in the developing epibranchial ganglia, but it is stronger in the posterior-most branchial region where they are most likely still at the placodal state (Figure 5-

7B). From a dorsal view, expression is observed in the middle of the neural tube, that is, in the ventricular –and possibly subventricular– part of the neural tube. Interestingly, expression in the spinal cord posterior to the big patch of expression along the spinal cord is observed as small spaced clusters (Figure 5-7C), in comparison to older embryos that show a slightly more homogenous pattern (Figure 5-7I). Expression also appears in the developing dorsal root ganglia for the first time, in comparison to stage 24 embryos which do not show expression in these ganglia.

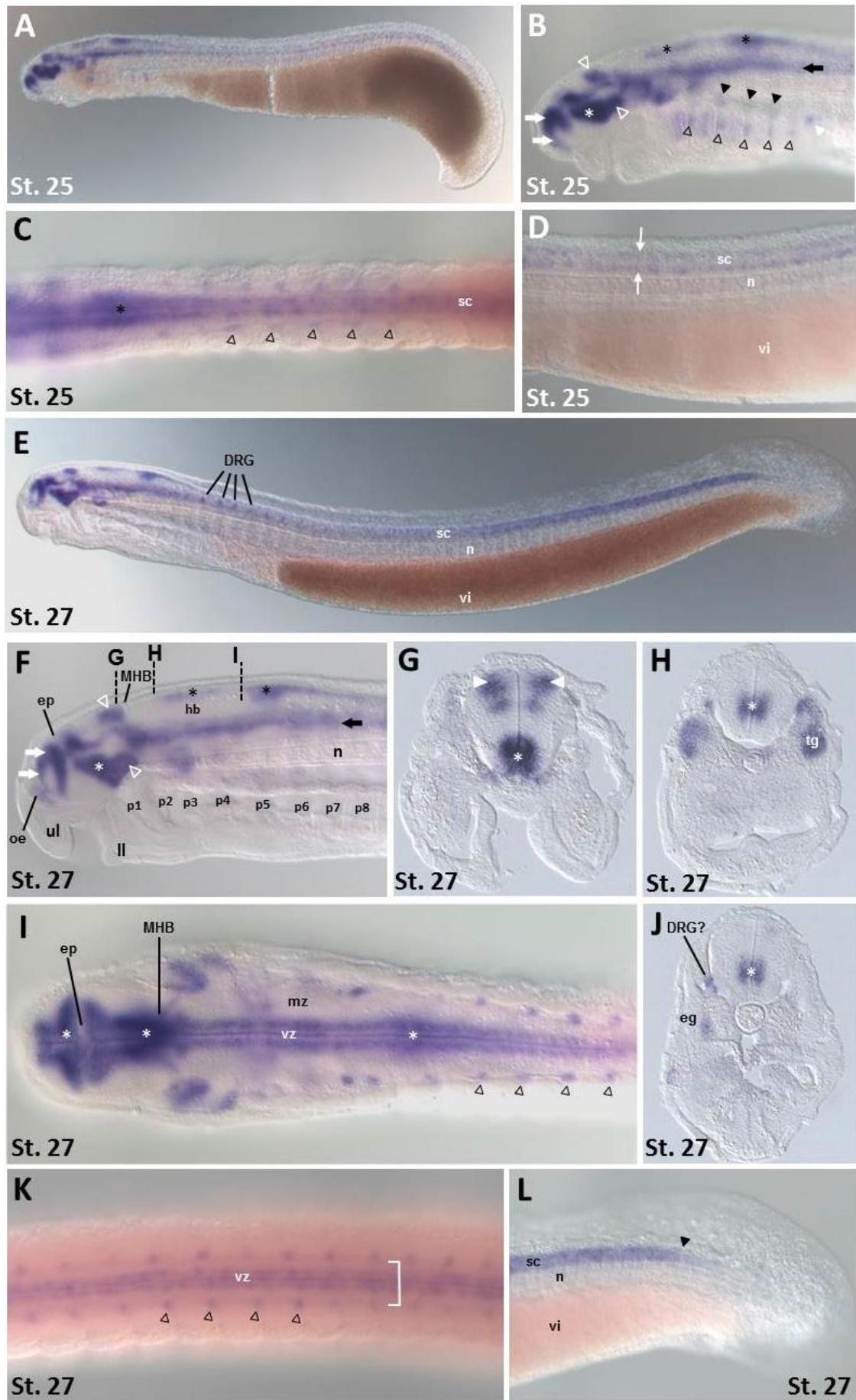


Figure 5-7. Expression of *LpNgn* at stages 25 and 27 in *L. planeri* embryos. (A, B, D-F and L) are lateral views. (C, I and K) are dorsal views. Anterior is to the left in all lateral and dorsal views. (G, H and J) are cross-sections, dorsal to the top. (A-D) At stage 25, *LpNgn* is expressed in a dorsal and a ventral domain along the spinal cord (A and D, arrows). *LpNgn* is also expressed in all epibranchial placodes/ganglia (B, black solid arrowheads), showing strongest expression at the placodal stage (B, white solid arrowhead; compare with Figure 5-6D, E and G). Note expression in each pharyngeal arch in a medial position with respect to the dorsal-ventral axis of each arch (B, open arrowheads). Transcripts are also observed in the telencephalon (B, white arrows), diencephalon (B, white asterisk), midbrain (B, open white arrowheads) and hindbrain, which presents two dorsal patches of expression (B, black asterisks) and a continuous ventral domain running from the MHB towards the spinal cord (B, black arrow). From a dorsal view (C), expression can be observed in the middle part of the neural tube which corresponds to the ventricular zone, with a high expression level in the posterior hindbrain (asterisk). Expression in the spinal cord shows a punctuated pattern within the ventricular zone. At this stage, expression starts to be seen in the developing DRG (C, open black arrowheads). (E-L) At stage 27, expression in the brain maintains the same pattern as stage 25 embryos, although expression is more resolved as the brain differentiates (F, labelling as in B). In the spinal cord the strongest expression is seen at the posterior, gradually diminishing towards the anterior (E); however, at the posterior expression stops somewhat abruptly a short distance from the end of the spinal cord (L, solid black arrowhead). DRG are better appreciated at this stage; following the anterior to posterior wave of neurogenesis in the spinal cord, *LpNgn* is strongly expressed at the anterior-most DRG and gradually diminishes towards the posterior (E, I and K, black open arrowheads). Dorsal views show that *LpNgn* is expressed in the ventricular zone of the neural tube (I and K; bracket in K marks the width of the spinal cord), and in the brain expression expands at the level of the epiphysis, MHB and posterior hindbrain (I, white asterisks). Cross sections at this stage show ventricular expression in the ventral neural tube (G, H and J, white asterisks) and probable subventricular expression in dorsal midbrain (G, solid white arrowheads). Expression is also observed in trigeminal ganglia (H) and epibranchial ganglia (J). Approximate plane of sections are indicated in (F). Abbreviations: DRG, dorsal root ganglia; eg, epibranchial ganglion; ep, epiphysis; hb, hindbrain; ll, lower lip; MHB, midbrain-hindbrain boundary; mz, marginal zone; n, notochord; oe, olfactory epithelium; p1-8, pharyngeal pouch 1-8; sc, spinal cord; tg, trigeminal ganglion; ul, upper lip; vi, vitellum; vz, ventricular zone.

At stage 27, the embryonic brain compartments are quite defined (Figure 5-7F). Expression in the telencephalon maintains a dorsal expression domain, but it also extends ventrally along its rostral-most and posterior-most aspects, leaving a gap in the middle. In the diencephalon, the epiphysis is slightly stained but stronger staining is observed ventrally. The hypothalamus is left unstained as well as the pretectum. In the midbrain, the dorsal expression domain is observed between the periphery and the ventricular zone, whereas the ventral domain is restricted to the ventricular zone. Expression in the hindbrain is generally maintained with respect to stage 25 embryos. From a dorsal view, expression is maintained in the ventricular zone all along the neural tube, but it expands at the level of the epiphysis, MHB and posterior hindbrain

(Figure 5-7I). Towards the tail, expression does not reach the very end of the spinal cord (Figure 5-7L).

In the peripheral nervous system, transcripts are observed in the olfactory epithelium, trigeminal and epibranchial ganglia (Figure 5-7F, H and J). Expression in the dorsal root ganglia is maintained along the spinal cord and intensifies at the anterior as the ganglia mature (Figure 5-7E, I and K). Both in stage 25 and stage 27 embryos, the dorsal root ganglia are observed in a zigzag pattern (Figure 5-7C, I, K).

Cross sections of a stage 27 embryo reveal the dynamic expression pattern of *LpNgn* along the spinal cord (Figure 5-8). At the posterior end, expression is seen covering the entire extent of the dorsoventral and mediolateral axes, with stronger staining observed dorsally (Figure 5-8A, B). Towards the anterior, expression gradually becomes restricted to the ventricular zone preserving stronger expression on the dorsal side (Figure 5-8C, D). In the middle of the body, expression diminishes but is still observed within the ventricular zone (Figure 5-8E). More anteriorly, expression becomes stronger again (Figure 5-8F) and it starts to separate into two domains, one dorsal and one ventral, still in the ventricular zone (Figure 5-8G, H).

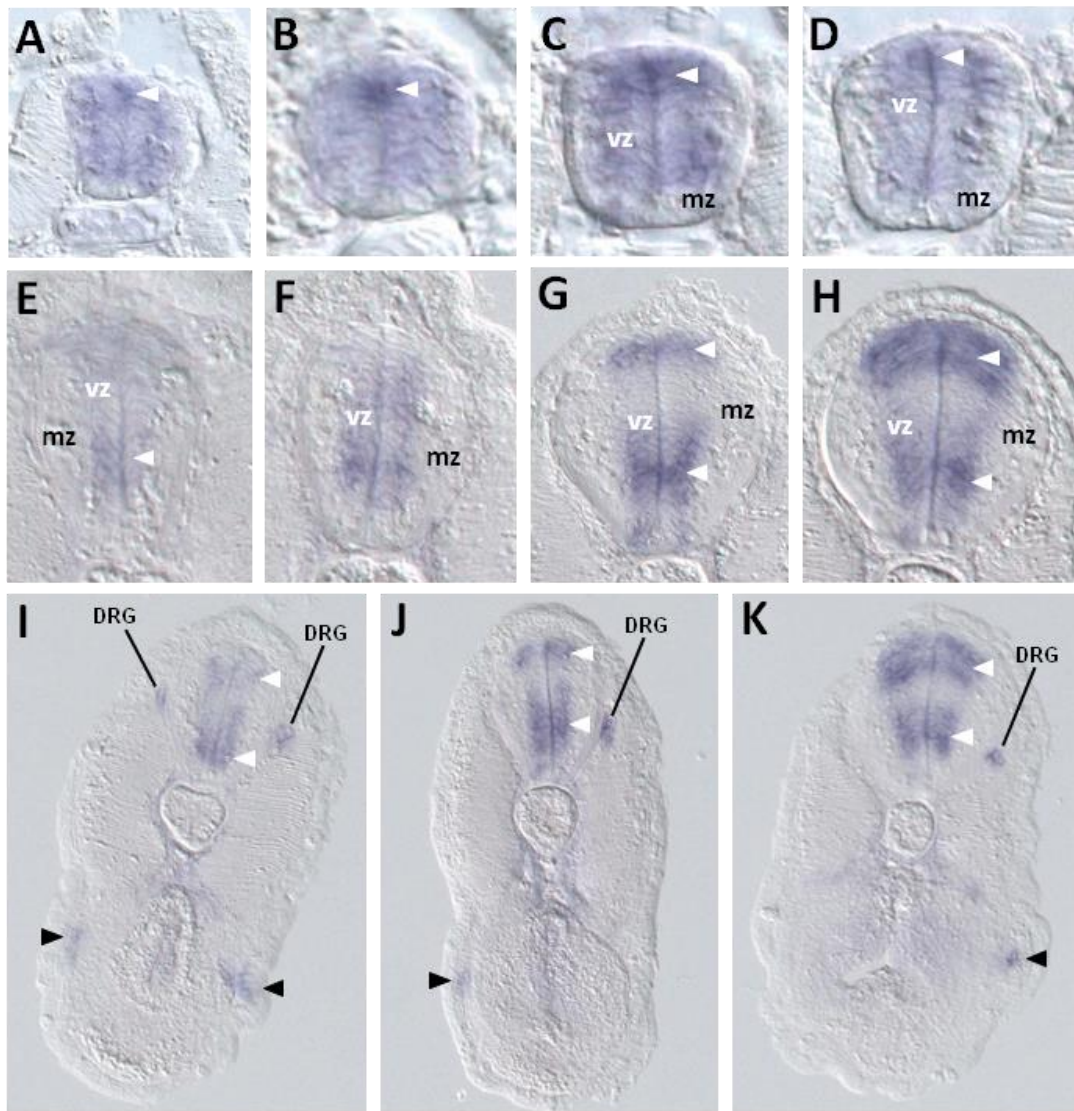


Figure 5-8. Expression of *LpNgn* in the spinal cord of stage 27 *L. planeri* embryos. (A-H) Serial cryosections at different levels of the spinal cord presented from posterior to anterior. In all images dorsal is at the top. At the posterior spinal cord, *LpNgn* expression covers the entire dorso-ventral and medio-lateral extent of the spinal cord (A, B). Subsequent sections towards the anterior show how expression gradually becomes restricted medially, a region that corresponds to the ventricular zone (C, D). Additionally, stronger expression is observed in a small dorsal domain (A-D, white arrowheads). In the middle part of the embryo, expression becomes considerably weaker and is mostly observed in a ventral domain in the ventricular zone (E, white arrowhead). More anteriorly, expression becomes gradually stronger (F) and splits into a dorsal and a ventral domain of staining in the ventricular zone (G, H, white arrowheads). (I-K) Serial cryosections at the level of the hindbrain and anterior spinal cord. Apart from the dorsal and ventral expression domains in the ventricular zone of the anterior neural tube (white arrowheads), *LpNgn* is also expressed in DRG and putative branchial ganglia (black arrowheads). Abbreviations: DRG, dorsal root ganglia; mz, marginal zone; vz, ventricular zone.

Cross sections of stage 27 embryos also reveal staining in the developing dorsal root ganglia (Figure 5-8I-K). Given that these ganglia are placed in a zigzag pattern (Figure 5-7I, K), quite often only one ganglion is observed in cross sections. Staining is observed next to the neural tube in a mid-ventral position with respect to the dorsoventral axis. It is clear that expression of *LpNgn* is seen in the developing ganglion itself but not dorsally adjacent to the neural tube where ventrally migrating neural crest cells that form DRG are expected to be present. These sections also show expression lateral to the pharynx, possibly corresponding to –novel– branchial ganglia inferred from *LpNgn* expression in pharyngeal arches of stage 25 embryos (Figure 5-7B).

At stage 28, expression is notably downregulated, although it is generally maintained in the same expression sites as in stage 27 embryos (Figure 5-9). Most of the spinal cord is devoid of staining except at the posterior, where it remains somewhat strong (Figure 5-9A, C, F). Of note, expression at the posterior spinal cord also ends somewhat abruptly before the end of the spinal cord as in stage 27 embryos but with a dotted pattern (Figure 5-7L). On the other hand, in the head, weak expression still persists in the ventricular zone in the same brain compartments as in stage 27 embryos, as well as in the trigeminal ganglia (Figure 5-9B, D). In the body, expression in the DRG is severely reduced, if not absent (Figure 5-9E). Finally, expression in the olfactory epithelium and epiphysis seems to have disappeared (Figure 5-9B).

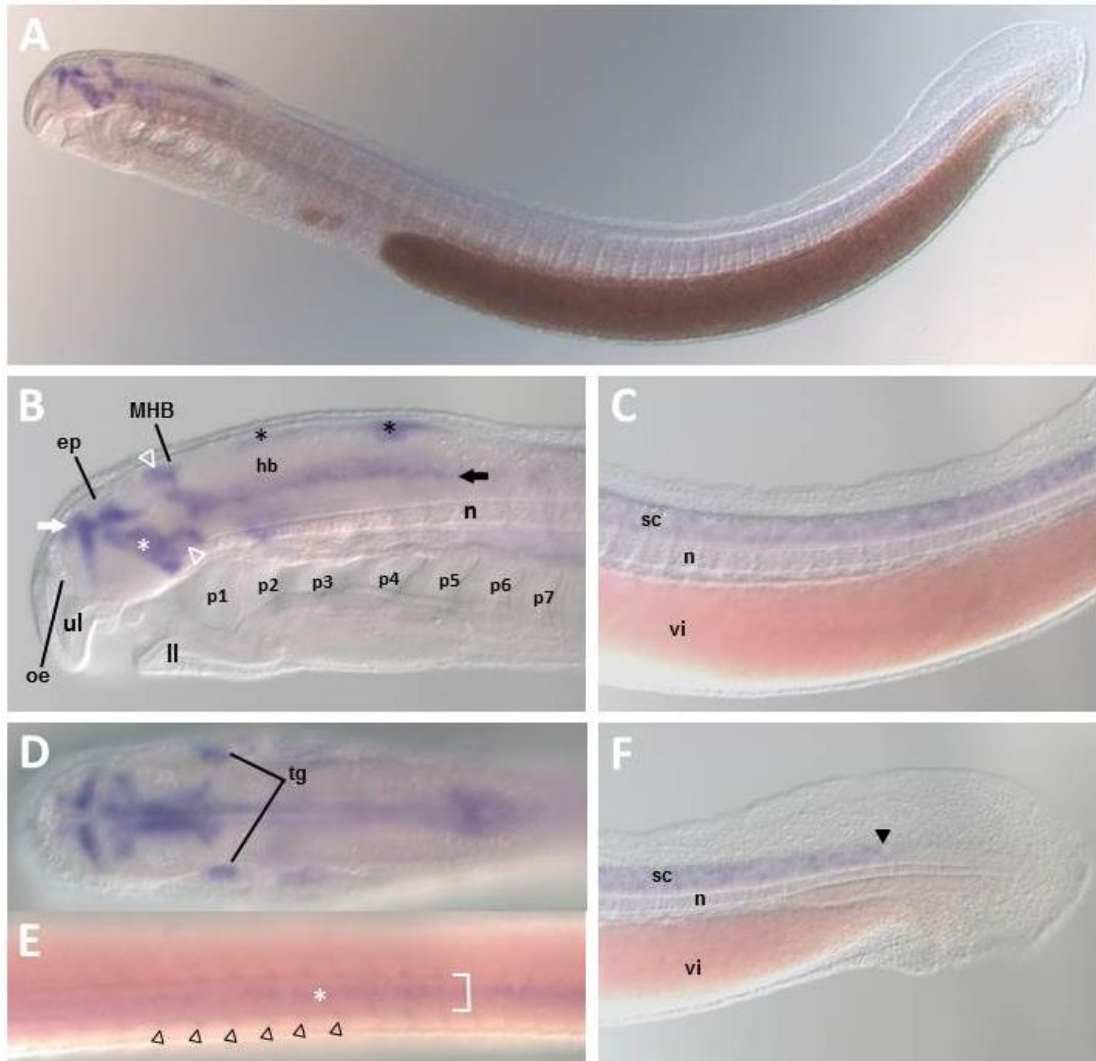


Figure 5-9. Expression of *LpNgn* at stage 28 in *L. planeri* embryos. (A-C, F) are lateral views; (D and E) are dorsal views. In all images anterior is to the left. Overall, *LpNgn* expression is maintained in the same sites but is notably downregulated with respect to earlier stages as can be seen in the brain (B, D). Note how expression in the spinal cord is virtually absent except at the posterior (A, C, F). Also, expression in the DRG is very weak, if not absent (E, open black arrowheads), and expression in the olfactory epithelium has disappeared (B). Labelling in B as in Figure 5-7B and F. Labelling in E as in Figure 5-7K). Abbreviations: ep, epiphysis; hb, hindbrain; ll, lower lip; MHB, midbrain-hindbrain boundary; n, notochord; oe, olfactory epithelium; p1-7, pharyngeal pouch 1-7; sc, spinal cord; tg, trigeminal ganglia; ul, upper lip; vi, vitellum.

5.6. Discussion

5.6.1. *LpNgn* Expression in the Central Nervous System

Brain: Expression of *LpNgn* in the brain will be compared mostly with mouse expression data, given that there is extensive information on mouse *neurogenin* genes but also because, anatomically, the lamprey brain is best mapped to the mouse brain as seen in sagittal sections from published information.

TELENCEPHALON: *LpNgn* expression in the telencephalon starts at stage 24 in a dorsal domain (Figure 5-6E). This expression continues through stage 28 with a concomitant expansion of expression along the anterior and posterior aspects of the telencephalon, leaving an unstained region in the middle (Figure 5-7B, F; Figure 5-9B). In the mouse telencephalon, *neurogenin1* and *neurogenin2* are expressed dorsally in the pallium in the proliferative ventricular zone and are absent from the subpallium where *Mash1* is expressed in a complementary fashion (Fode et al., 2000; Gradwohl et al., 1996; Sommer et al., 1996). Expression studies have determined that the lamprey telencephalon is also subdivided into pallium and subpallium (Murakami et al., 2005; Sugahara et al., 2011). Expression of *LpNgn* is observed dorsally as in mouse in the region corresponding to the pallium. However, its expression reaches more ventral areas in the subpallium, in close apposition to the olfactory epithelium where *LpNgn* is also expressed. Pallium-restricted *neurogenin* expression is also observed in zebrafish *neurogenin1* and *Xenopus neurogenin2* (Mueller and Wullimann, 2003; Wullimann et al., 2005). Therefore, while *LpNgn* shares a dorsal (pallial) expression with other vertebrates, it differs in that it is also found in ventral (subpallial) regions.

DIENCEPHALON: Expression of *LpNgn* demarcates territories in the diencephalon that are consistent with the prosomeric model to some extent. *LpNgn* demarcates regions that could correspond to prosomere 1 (P1) and P2 in the diencephalon. P3 is not easily demarcated when compared to mouse *neurogenin2* (Osório et al., 2010), in particular considering the region delimited by the zona limitans intrathalamica (Figure 5-7F, white asterisk). Therefore, expression of *LpNgn* in the diencephalon suggests the absence of P3 at the embryonic stages analysed (see Figure 9-1 in Discussion). However, both a P3 and a prethalamus (that corresponds to the alar part of P3) have been described in the adult lamprey brain with some difficulty, although they are comparatively thinner with respect to P1 and P2 (Pombal and Puelles, 1999; Pombal et al., 2009). In this updated lamprey brain model, *LpNgn* expression also highlights the secondary prosencephalon, which is composed of the hypothalamus and telencephalon (Figure 9-1 in Discussion, dark blue region). These forebrain subdivisions are more consistent with the lamprey prosomeric model described by Murakami and colleagues (Murakami et al., 2001, 2002; see also Villar-Cheda et al., 2005) than that described by Osório and colleagues (Osório et al., 2005), particularly considering the position of the epiphysis dorsal to P2 (dorsal thalamus). Also, *LpNgn* is expressed in the epiphysis as *Xenopus neurogenin1* and zebrafish *neurogenin1*, which has an important role in pineal neurogenesis in conjunction with *ash1a* (Cau and Wilson, 2003; Nieber et al., 2009). *LpNgn* is not expressed in the hypothalamus, where vertebrate *neurogenin* paralogy group 3 is localised (Nieber et al., 2009; Sommer et al., 1996; Wang et al., 2001). No *LpNgn* expression was observed in retinal cells as occurs with zebrafish *neurogenin1*, *Xenopus neurogenin2*, chicken *neurogenin1* and *neurogenin2* and mouse *neurogenin2* (Matter-Sadzinski et al., 2001; Mueller and Wullimann, 2003; Nieber et al., 2009; Perron et al., 1999; Sommer et al., 1996; Yan et al., 2009).

MIDBRAIN: In the midbrain, *LpNgn* shows two dorsal and a ventral expression domain. The dorsal domains could correspond to the tectum (midbrain roof), which is not stained completely, whereas the ventral domain could correspond to the entire tegmentum (ventral midbrain) as seen with mouse *neurogenin2* (Osório et al., 2010). In *Xenopus*, *neurogenin1* (but not *neurogenin2*) is expressed in the midbrain (Nieber et al., 2009), but in *Xenopus* larvae, *neurogenin2* is reported to be expressed in the tectum and tegmentum (Wullimann et al., 2005). In zebrafish, *neurogenin1* is also found in the tectum and tegmentum (Mueller and Wullimann, 2003). *LpNgn* expression in the tegmentum is contiguous with expression in the basal part of P1 and P2 as occurs with mouse *neurogenin2*.

HINDBRAIN: In the hindbrain, *LpNgn* is expressed in two longitudinal domains, one dorsal and one ventral. The ventral domain is continuous through all rhombomeres and continues into the spinal cord. The dorsal domain, however, is in two patches, one at the level of r3-5, and another, much stronger patch at the hindbrain-spinal cord transition (possibly corresponding to r6-7) that continues into the spinal cord. A strong patch of proliferating cells has been observed at this level with PCNA immunostaining (Meléndez-Ferro et al., 2003). *LpNgn* expression in the hindbrain is most similar to mouse *neurogenin2* (Gradwohl et al., 1996; Osório et al., 2010), which also presents two longitudinal expression domains that project into the spinal cord. In *Xenopus* embryos, *neurogenin1-3* genes are all expressed in the hindbrain, but their expression is generally found more ventrally (Nieber et al., 2009), whereas in *Xenopus* larvae *neurogenin2* is also expressed more dorsally (Bellefroid et al., 1996; Wullimann et al., 2005). In zebrafish, *neurogenin1* is expressed in a more segmental pattern marking each rhombomere (Amoyel et al., 2005; Korzh et al., 1998). *neurogenin* genes are also

expressed in the cerebellum (Zordan et al., 2008) and are important players during its development (Florio et al., 2012). However, a dorsal *LpNgn* expression domain at the level of r1-2 suggestive of cerebellar expression was not observed at any of the stages examined.

Spinal cord: Expression of *LpNgn* in the spinal cord is also very similar to gnathostomes. At the posterior-most region of the lamprey spinal cord, *LpNgn* expression extends into the entire mediolateral and dorsal-ventral axes of the neural tube as seen in cross-sections. Given the wave of neuronal differentiation in the spinal cord from anterior to posterior, this expression pattern is consistent with the role of *neurogenin* genes in the specification of neuronal precursors which are located at the posterior spinal cord as it elongates. As neurons differentiate in the anterior spinal cord and the ventricular zone narrows in this region, *LpNgn* expression becomes restricted to the ventricular zone. Additionally, serial sections of the spinal cord show that *LpNgn* expression is gradually split –in a posterior to anterior gradient– into a dorsal and a ventral domain within the ventricular zone. These expression patterns are most similar to mouse and chicken *neurogenin2* (Gradwohl et al., 1996; Sommer et al., 1996; Perez et al., 1999).

5.6.2. *LpNgn* Expression in the Peripheral Nervous System

Dorsal Root Ganglia: A prominent *LpNgn* expression site is seen in the DRG. Expression in the DRG starts at stage 25 (Figure 5-6C) to stage 27 (Figure 5-6I and K; compare with Figure 5-5H). Lamprey DRG are found in intersomitic/intermyotome spaces in a zigzag distribution (Hirata et al., 1997; Nakao and Ishizawa, 1987), in

contrast to gnathostomes in which each dorsal root ganglion is found in register with each somite along the entire rostrocaudal axis (Raible and Ungos, 2006). At stage 28 (Figure 5-8E), the DRG barely express *LpNgn*. To our knowledge, this is the first molecular marker found for DRG in lampreys. Other vertebrate *neurogenin* homologues are also expressed in DRG. Mouse *neurogenin2* is expressed in early migrating neural crest cells that will form DRG and when DRG start to condense (Gradwohl et al., 1996; Sommer et al., 1996; Ma et al., 1999). On the other hand, mouse *neurogenin1* is expressed only after DRG have begun their condensation (Sommer et al., 1996; Ma et al., 1999). In chicken, both *neurogenin1* and *neurogenin2* are expressed in migrating crest cells adjacent to the neural tube (Perez et al., 1999). Besides their expression, vertebrate neurogenins are essential for the development of DRG. In mouse, disruption of *neurogenin1* and *neurogenin2* function leads to a complete absence of DRG (Ma et al., 1999). Similarly, in zebrafish, *neurogenin1* function disrupted by antisense morpholino oligonucleotides leads to a complete loss of DRG in addition to cranial ganglia (Andermann et al., 2002; Cornell and Eisen, 2002). In chicken, forced expression of *neurogenin1* directs premigratory neural crest cells to a sensory (dorsal root ganglia) rather than an autonomic (sympathetic ganglia) fate (Perez et al., 1999). Similarly, in zebrafish, *neurogenin1* regulates the progression of neural crest precursors of dorsal root ganglia into sensory neurons at the expense of glial cells. In mouse, *neurogenin1* and *neurogenin2* mutants lack cranial and spinal sensory ganglia (Fode et al., 1998; Ma et al., 1998, 1999). Therefore, in different vertebrate species *neurogenins* are predominant factors that control sensory neurogenesis.

Placodes

OLFACTORY PLACODE: Expression of *LpNgn* in the anterior-most part of the head starting from stage 21 is revealing as it prefigures the future olfactory placode, which is not morphologically discerned until stage 23 in *Lampetra japonica* embryos (Kuratani et al., 2001; Uchida et al., 2003) (according to FishBase, the accepted name for *L. japonica* is *Lethenteron camtschaticum*; *Lethenteron* is a subgenus of *Lampetra*; junior synonyms include *Lethenteron japonicum* and *Lampetra japonica*). This feature makes *LpNgn* the first-known molecular marker in the anterior ectoderm of the prospective olfactory placode from –at least– stage 21. *LpNgn* is later restricted to the olfactory derivative of the NHP and is not expressed in the adenohypophysis. Expression of *neurogenins* in olfactory placodes and olfactory epithelium is found in *Xenopus*, chicken and mouse embryos (Cau et al., 2002; Ma et al., 1997; Maier and Gunhaga, 2009; Nieber et al., 2009; Schlosser and Northcutt, 2000; Sommer et al., 1996).

TRIGEMINAL PLACODES: *LpNgn* is also expressed in a trigeminal placode anlage, although it is hard to distinguish whether it is found in the ophthalmic or maxillomandibular placode, or whether it is a continuous trigeminal placode area. The ophthalmic and maxillomandibular placodes have been recognised morphologically as independent placodes (Damas, 1944; von Kupffer, 1895), but *LpNgn* seems not to distinguish between them. Interestingly, at stage 27, *LpNgn* in the ophthalmic ganglion is observed in the posterior half and in the maxillomandibular ganglion posterolaterally. Since *LpNgn* expression was not observed in neural crest cells or their derivatives, and instead found *LpNgn* expression in trigeminal placodes, it is likely that the partial expression of *LpNgn* in both the ophthalmic and maxillomandibular ganglia reflects the extent of placode-derived neurons, while the unstained parts of both ganglia

represent the neural crest-derived population (a neural crest contribution to the maxillomandibular ganglion was observed by Modrell et al., 2014).

OTIC PLACODE: Comparison of expression of all vertebrate *neurogenin* homologues in cranial sensory placodes (Table 5-2) reveals the absence of *LpNgn* expression in the otic placode and/or its derivatives. The vertebrate otic placode forms the otic vesicle, which generates sensory neurons of the inner ear and mechanosensory hair cells. Cells also delaminate from the otic vesicle to form neurons of the statoacoustic ganglion. The otic placode has been recognised morphologically in *Lampetra fluviatilis* at around stage 22 dorsally at the level of the second pharyngeal pouch (Damas, 1944). The lamprey otic placode invaginates to form an otic vesicle which produces hair cells, and cells also delaminate from this vesicle to form the statoacoustic ganglion (Hammond and Whitfield, 2006). Zebrafish *neurogenin1* is expressed in the otic vesicle (Andermann et al., 2002), while *Xenopus neurogenin1* and *neurogenin2* are expressed in the otic placodes (Nieber et al., 2009) and chicken *neurogenin1* is expressed in the otic vesicle (Begbie et al., 2002; Abu-Elmagd et al., 2001). Mouse *neurogenin1* is expressed in the otic placode, otic vesicle and inner ear, where it is involved in generating neural precursors and maintaining an uncommitted progenitor cell population for sensory epithelial cells (Friedman et al., 2005; Ma et al., 1998, 2000; Matei et al., 2005; Raft et al., 2007). Therefore, a lamprey *neurogenin* homologue is predicted to be present in the otic placode and/or statoacoustic ganglia, considering that in jawed vertebrates at least one *neurogenin* gene is expressed in all placode-derived cranial sensory neurons. This might thus suggest the existence of a second lamprey *neurogenin* gene expressed in otic placodes and/or its derivatives.

LATERAL LINE PLACODES: Expression of *LpNgn* in the posterior lateral line ganglion parallels that of zebrafish *neurogenin1* and *Xenopus neurogenin1* and *neurogenin2* in lateral line placodes (Andermann et al., 2002; Nieber et al., 2009; Schlosser and Northcutt, 2000). The posterior lateral line ganglion has been identified in lamprey embryos (Kuratani et al., 1997; Modrell et al., 2014; Takio et al., 2007). However, lamprey embryos and adults have three lateral line ganglia: an anterior and a posterior lateral line ganglion, and an intracapsular ganglion (Koyama et al., 1990; Modrell et al., 2014; Ronan and Northcutt, 1987). This suggests the presence of at least another (anterior) lateral line placode in lamprey embryos. Given that the facial and anterior lateral line ganglion form a single complex as they develop (Modrell et al., 2014), it is possible that the facial and a putative anterior lateral line placode also form a single complex marked by *LpNgn* in stage 24 embryos.

EPIBRANCHIAL PLACODES: *LpNgn* is expressed in epibranchial placodes as in other vertebrates. Mouse *neurogenin2* is expressed in the three epibranchial placodes (Gradwohl et al., 1996). Chicken *neurogenin1* is expressed in all epibranchial placodes and its expression starts to diminish as cells migrate inside to differentiate into neurons that will form the epibranchial ganglia (Abu-Elmagd et al., 2001). Compared to lamprey, it is interesting to see how expression of *LpNgn* is seen stronger at the posterior-most epibranchial placode/ganglion region. Considering the position and strength of staining, it seems that *LpNgn* is expressed more strongly at the placodal stage and its expression starts to diminish as placodal cells start to delaminate to form epibranchial ganglia as occurs with chicken *neurogenin1*.

NOVEL (BRANCHIAL) PLACODES: One of the most intriguing aspects of *LpNgn* expression is its localisation in pharyngeal arches. At stage 25, expression can be observed dorsal to each pharyngeal arch, which corresponds to epibranchial ganglia. However, expression at the middle of each pharyngeal arch with respect to the dorsal-ventral axis is quite intriguing. Considering that *neurogenins* are expressed mainly in the nervous system, particularly in the sensory neuron lineage, this expression pattern suggests the presence of a sensory neuron lineage. Given that this expression site is in the ectoderm, *LpNgn* expression might reveal another set of placodes, in particular sensory placodes. Damas (1944) described the presence of small thickened territories at the middle of each pharyngeal arch, which are at the level of the putative placodes identified by *LpNgn* expression. However, Damas (1944) suggested that these ectodermal thickenings corresponded to a mesodermal lineage. The expression of *LpNgn* strongly suggests that these regions are ectodermal placodes (as opposed to mesoderm-related structures), and that these placodes likely give rise to sensory neurons. Experiments such as lineage tracing might follow this observation and establish whether those –putative– placodes indeed give rise to neurons, particularly sensory neurons. One possibility could be that they generate additional neurons for the larval and/or adult lateral line system dorsal to pharyngeal arches such as the suprabranchial neuromasts (Gelman et al., 2008), which seem to be the closest to the position of the embryonic “branchial placodes”, possibly substituting for the lack of several lateral line placodes as in zebrafish and *Xenopus*. To our knowledge, no placodes in that position have been described in any other vertebrate species. No *LpNgn* expression ventral to pharyngeal pouches suggestive of hypobranchial placodes/ganglia was observed as occurs with *Xenopus neurogenin2* (Schlosser and Northcutt, 2000).

Comparison of expression of all vertebrate *neurogenin* homologues in cranial placodes (summarised in Table 5-2) reveals an apparent absence of *LpNgn* expression in otic placodes/ganglia, where is predicted to be present considering that vertebrate *neurogenin* genes are well-known to be expressed (all together) in all placode-derived cranial sensory neurons. Therefore, in this context, I predict the existence of a second lamprey *neurogenin* homologue expressed in otic placodes/ganglia.

Species	Ngn homologues	Olf	Trigeminal	Profundal	Otic	Lateral Line				Epibranchial			Branchial	Hypobranchial	References
						aLL		mLL	pLL	Facial	Glosso.	Vagal			
						ad	av								
Lamprey	LpNgn LpNgn-B?	LpNgn	LpNgn			LpNgn				LpNgn	LpNgn	LpNgn	LpNgn		This study
Zebrafish	ngn1 ngn3	ngn1? -	ngn1 -	N/A? N/A?	ngn1 -	ngn1 -	ngn1 -	ngn1 -	ngn1 -	ngn1 -	ngn1 -	ngn1 -			Andermann et al., 2002 ¹ Wang et al., 2001 ²
Xenopus	ngn1 ngn2 ngn3	ngn1 ngn2 -	ngn1 ngn2 -	ngn1 ngn2 -	ngn1 ngn2 -	ngn1 ?		ngn1 ?	ngn1 ?	ngn1 ngn2	ngn1 ngn2	ngn1 ngn2		? ngn2 ?	Nieber et al., 2009 ³ Nieber et al., 2009 ³ Nieber et al., 2009
Chicken	ngn1 ngn2	ngn1 -	ngn1 -	- ngn2	ngn1 -					ngn1 -	ngn1 -	ngn1 -			Abu-Elmagd et al., 2001
Mouse	ngn1 ngn2 ngn3	ngn1 ngn2 -	ngn1 ngn2 -		ngn1 - -					ngn1 ngn2 -	ngn1 ngn2 -	ngn1 ngn2 -			Ma et al., 1998 Gradwohl et al., 1996

Table 5-2. Comparison of vertebrate neurogenin homologues in all cranial sensory placodes. All gnathostome *neurogenin* homologues –together– are expressed in all sensory cranial placodes/ganglia. *LpNgn* is expressed in olfactory, trigeminal, epibranchial, lateral line, and branchial placodes/ganglia. From this, the existence of a second lamprey *neurogenin* homologue is predicted to be expressed in otic placodes/ganglia. Dash indicates lack of structures in that particular species.

5.7. Summary of Lamprey *Neurogenin* Genes

LpNgn clusters with *LfNgn1* and *PmNgn* forming a single clade, but they do not group with a specific vertebrate (gnathostome) *Ngn* paralogy group. As in gnathostomes, *LpNgn* is expressed in specific regions of the brain corroborating a regionalised lamprey brain and adding information to the prosomeric model of the embryonic lamprey brain. *LpNgn* is also expressed in cranial sensory placodes and their derivatives including olfactory, trigeminal, epibranchial and lateral line. A notable *LpNgn* expression site is in the DRG. In addition to expression in epibranchial placodes/ganglia, *LpNgn* is also expressed in other patches of ectodermal cells of the pharyngeal arches, suggesting the presence of novel cranial sensory placodes. As with other vertebrate Ngns, *LpNgn* is not expressed in the autonomic nervous system.

6. *LAMPETRA PLANERI* COE GENES

6.1. Introduction

Structure and function of COE genes: COE genes (also known as *Ebf* genes in some species) constitute an atypical subfamily of HLH transcription factors conserved across the Metazoa, which are mainly characterised by their lack of the basic region next to the HLH domain, and instead possess a unique DNA-binding domain (DBD) upstream of the HLH domain specific of this subfamily (Dubois and Vincent, 2001; Liberg et al., 2002). Inside the COE DBD there is also an atypical Zn-finger domain with a consensus sequence H-X₃-C-X₂-C-X₅-C that is important for DNA binding (Fields et al., 2008; Hagman et al., 1995) (Figure 6-1). Between the DBD and HLH domains, an IPT/TIG (Immunoglobulin-like, Plexins, Transcription-factors/Transcription factor Immunoglobulin) domain is also present and has been suggested to be involved in protein-protein interactions, dimerization, and even DNA binding (Siponen et al., 2010; Treiber et al., 2010a). Immediately downstream of the IPT/TIG domain, a HLH domain is present with which COE proteins form homo- or heterodimers with other family members to bind DNA and activate transcription of target genes (Daburon et al., 2008; Dubois and Vincent, 2001; Hagman et al., 1993; Liberg et al., 2002; Pozzoli et al., 2001; Travis et al., 1993; Wang et al., 1997, 2002) (Figure 6-2). Following the HLH domain, a C-terminal transactivation region is present with no conserved sequence (Hagman et al., 1995). COE proteins recognise variations of the palindromic sequence ATTCCCNNGGGAAT (Travis et al., 1993; Wang and Reed, 1993).

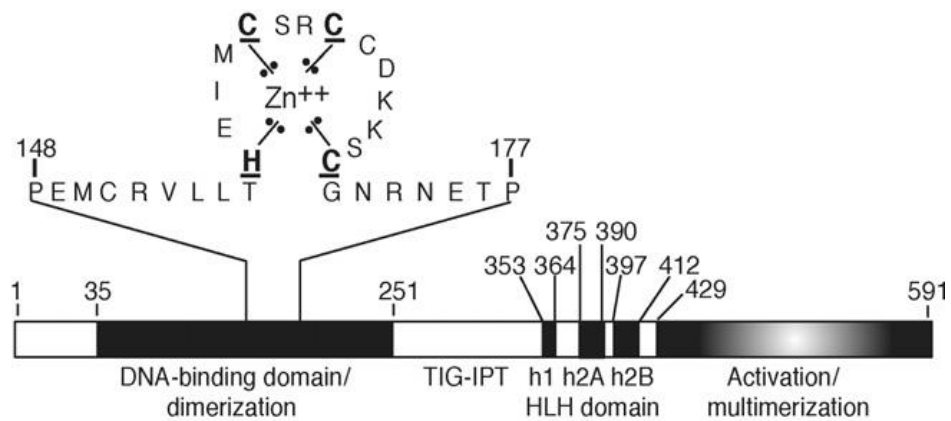


Figure 6-1. Modular structure of COE proteins. Depicted is mouse COE1/Ebf1 protein, whose structure is representative of all metazoan COE proteins. It contains a unique DNA-binding domain (DBD) that includes an atypical Zn-finger domain (expanded above). Next to the DBD is the IPT/TIG (TIG-IPT) domain followed by an atypical HLH domain composed of three α -helices (h1, h2A and h2B, also named H1, H2d and H2a, respectively). An activation/multimerization region is located at the C-terminal end with no conserved sequence. Taken from Fields et al. (2008).

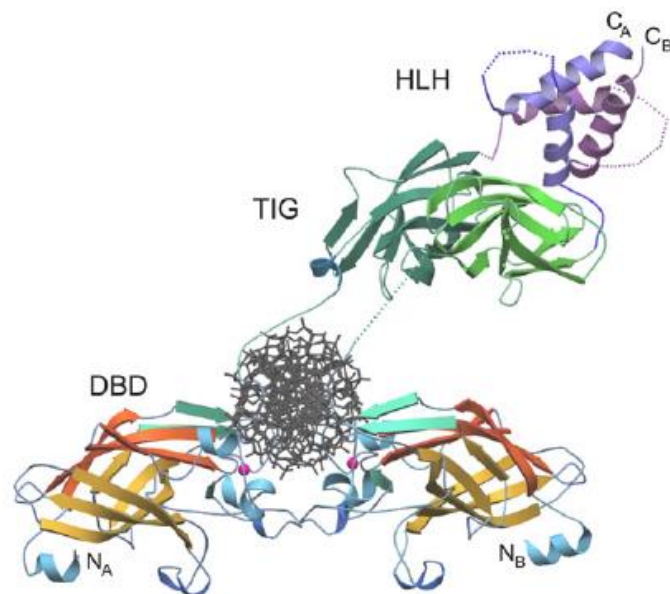


Figure 6-2. Crystal structure of dimeric COE1/Ebf1 bound to DNA. DNA-binding domains (DBD) are β strands colored in yellow and orange. Zinc ions in the Zn-finger domain (light blue) are represented by pink spheres. IPT/TIG domains are shown in light and dark green, and HLH domains in dark blue and purple. Only two helices of the HLH domain are represented in this model. Unstructured loops are indicated as dashed lines. DNA is shown in dark gray. Taken from (Treiber et al., 2010b).

One of the most interesting aspects of COE proteins is that only vertebrates possess three helices in the HLH domain, while all invertebrates studied so far (including tunicates and cephalochordates) possess only two helices in the HLH domain (Jackson et al., 2010; Mazet et al., 2004; Pang et al., 2004) (Figure 6-3). This is additionally correlated with gene content as all invertebrates studied so far, including tunicates and cephalochordates, possess only one *COE* gene copy whereas in jawed vertebrates four paralogy groups have been identified (Jackson et al., 2010; Mazet et al., 2004; Pang et al., 2004).

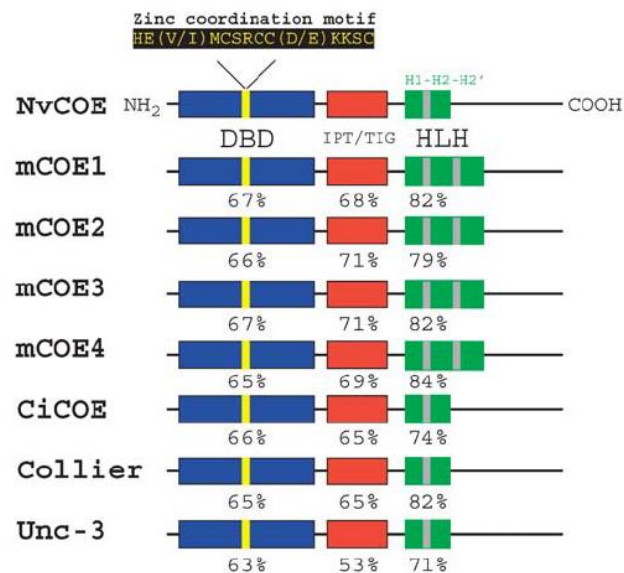


Figure 6-3. Modular structure of COE/Ebf transcription factors. Conserved domains of COE/Ebf genes across the Metazoa include: DBD (in blue), a Zn-coordination motif (yellow) contained within the DBD, an IPT/TIG domain (red), and the HLH domain (green and grey). Note the presence of a duplicated helix 2 (H2') in mouse COE proteins (mCOE1-4) representing vertebrates, whereas in invertebrates, where only one *COE* gene has been identified (here represented by *Nematostella vectensis* NvCOE, *Drosophila melanogaster* Collier, *Caenorhabditis elegans* Unc-3 and *Ciona intestinalis* CiCOE), only two helices are found. Taken from Pang et al. (2004).

Daburon et al. (2008) reported the identification of two lamprey *COE* genes, *PmCOE-A* and *PmCOE-B*, in the genome of *Petromyzon marinus*, and also reported

LfCOE-B EST sequences from the river lamprey *Lampetra fluviatilis*. They further showed that *PmCOE-B* and one splice variant of *LfCOE-B* have the duplicated H2 and the first H2 (H2d) is differentially spliced, rendering transcripts with two or three helices (Daburon et al., 2008) (Figure 6-4).

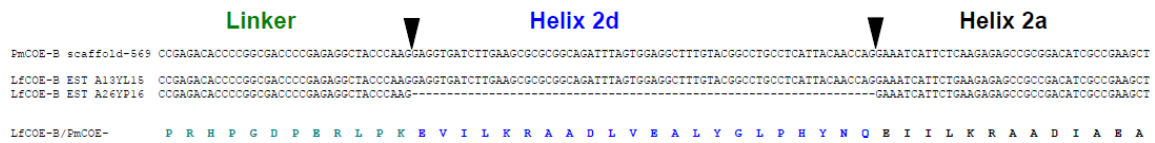


Figure 6-4. Structure of *P. marinus* and *L. fluviatilis* H2 domain regions of *COE-B* transcripts. The top line shows the *PmCOE-B* mRNA sequence predicted from the genomic scaffold 569. Arrowheads indicate the positions of introns. Lines in the middle correspond to partial *L. fluviatilis* *LfCOE-B* EST sequences. One EST fragment lacks the duplicated H2d whereas the other EST fragment contains both H2d and H2a. The bottom line corresponds to the deduced amino acid sequences of *LfCOE-B* and *PmCOE-B* together given that they are identical. Taken from Daburon et al. (2008).

Expression of COE genes in the vertebrate nervous system: In zebrafish, up to five *COE* genes have been identified (Wang et al., 2009), but only the expression of *Zcoe2/Ebf2* and *Ebf3* has been analysed. *Zcoe2/Ebf2* is expressed in migrating neural crest cells, olfactory and trigeminal placodes, and in the olfactory bulb, diencephalon, midbrain, MHB, hindbrain and spinal cord (Bally-Cuif et al., 1998; Li et al., 2010). On the other hand, zebrafish *Ebf3* is expressed in the olfactory bulb, diencephalon and MHB (Li et al., 2010). Earlier in development at neural plate stages, *Zcoe2* is expressed in the anterior part of the neural plate in the three stripes of primary neurogenesis (motor-, inter- and sensory neurons) (Bally-Cuif et al., 1998). Expression of zebrafish *Ebf3* has not been analysed at neural plate stages.

In *Xenopus*, *Xcoe2* is also expressed in the three stripes of primary neuron precursors (motor-, inter- and sensory neurons), trigeminal placodes/ganglia and anterior neural crest cells at early stages (Dubois et al., 1998). After neural tube closure, *Xcoe2* is expressed transiently in differentiating neurons in the marginal zone of the spinal cord, in neural crest cells, mesencephalon, rhombencephalon, and in the olfactory placodes and cranial ganglia (Dubois et al., 1998). Another *COE/Ebf* gene in *Xenopus*, *Xebf3*, has a similar expression profile as *Xcoe2* but it differs from the latter in that it is also expressed in the retina and in the otic vesicle (Pozzoli et al., 2001).

Chicken *COE1-3* homologs have also been cloned (Garcia-Dominguez et al., 2003; Mella et al., 2004; Nieminen et al., 2000), although their expression patterns in the nervous system have not been analysed in much detail. Chicken *Ebf1* is expressed in the pretectum in the diencephalon (Merchán et al., 2011), and both chicken *Ebf1* and *Ebf3* are expressed in the hindbrain in all rhombomeres, although expression in rhombomeres 3 and 5 is delayed (Garcia-Dominguez et al., 2003). In the spinal cord, they are expressed in the marginal zone along the dorso-ventral axis and they have a wider expression domain ventrally in a region corresponding to motor neurons (Garcia-Dominguez et al., 2003).

In mice, *Ebf1*, *Ebf2* and *Ebf3* are expressed in both central (CNS) and peripheral (PNS) nervous systems, particularly in post mitotic neurons (Davis and Reed, 1996; Garel et al., 1997a). In the brain, *Ebf1*, *Ebf2* and *Ebf3* are expressed in partially overlapping domains of the telencephalon, diencephalon, midbrain and hindbrain (Davis and Reed, 1996; Garel et al., 1997a). Expression in the hindbrain first appears in

rhombomeres 2 and 4 and later extends into all other rhombomeres and inter-rhombomeric boundaries (Garel et al., 1997a). *Ebf1* and *Ebf3* expression in the hindbrain is restricted to the ventricular zone, whereas *Ebf2* hindbrain expression is restricted to the subventricular zone (Garel et al., 1997a). In the cerebellum, mouse *COE/Ebf* genes are expressed in the cortical and nuclear transitory zones and cerebellar nuclei (Croci, 2006; Garel et al., 1997a), and later in Purkinje cells, external granular layer and deep cerebellar nuclei, but absent from the molecular layer (Croci, 2006; Wang et al., 1997). In the spinal cord, *Ebf1* and *Ebf3* are first expressed ventrolaterally in a region corresponding to motor neurons, and later their expression covers the entire marginal zone (Garel et al., 1997a). *Ebf2*, on the other hand, is weakly expressed in the inner-most part of the same ventrolateral motor neuron region in the basal plate, and later its expression expands into the alar plate in a thin layer of early post-mitotic neurons conforming the subventricular zone (Garel et al., 1997a). *Ebf1* and *Ebf3* (but not *Ebf2*) are also expressed in the developing retina (Davis and Reed, 1996; Wang et al., 1997).

Altogether, mouse *COE/Ebf* genes are expressed along the entire neural tube in the subventricular (*Ebf2*) and marginal zones (*Ebf1* and *Ebf3*) coincident with neurogenesis. Their expression is generally activated first in the basal plate, and later expanded into the alar plate, with a concomitant downregulation in the basal plate. This pattern reflects a pattern of neurogenesis from the basal to the alar plate (i.e. from ventral to dorsal).

In the PNS, mouse *COE/Ebf* genes are expressed in the olfactory epithelium, vomeronasal organ, trigeminal and glossopharyngeal cranial ganglia and dorsal root ganglia (DRG) at different levels (Corradi et al., 2003; Davis and Reed, 1996; Malgaretti et al., 1997; Wang et al., 1997). *Ebf1* expression has also been observed in the developing inner ear (Wang et al., 1997) and *Ebf2* expression has been identified in Schwann cells of sciatic nerves (Corradi et al., 2003). Expression of a fourth *Ebf* gene in mouse, *COE4/Ebf4*, has so far been determined by in situ hybridisation only in the olfactory epithelium (absent in the vomeronasal organ), although transcripts were also detected in adult tissues such as olfactory epithelium, olfactory bulb, cerebellum, eye, lung, spleen and kidney by RT-PCR (Wang et al., 2002).

Finally, in elasmobranchs, three *COE/Ebf*-like genes have been identified from the spleen and brain of *Raja eglanteria*: skate *EBF-1*, *EBF-B* and *EBF-C*, although their expression patterns have not been determined yet (Anderson et al., 2004).

Developmental roles of COE genes in vertebrates: *COE* genes have been shown to induce neuronal differentiation and regulate migration. Mouse *Ebf1* is able to induce the exit from the cell cycle in neuroblasts located in the ventricular zone of the chicken spinal cord, and it promotes their migration towards the marginal zone with concomitant neuronal differentiation (Garcia-Dominguez et al., 2003). Also, mouse *Ebf1* regulates early neuronal differentiation in the embryonic striatum and their migration from the subventricular zone to the marginal zone (Garel et al., 1999). *Ebf1* also regulates migration of facial branchiomotor neurons (Garel et al., 2000) and the early topography of axons that travel from the thalamus to the neocortex (Garel et al., 2002). On the other

hand, mouse *Ebf2* regulates migration and survival of Purkinje cells in the cerebellum (Crocì et al., 2006), migration of gonadotropin-releasing hormone (GnRH)-synthesising neurons from the vomeronasal organ to the hypothalamus, and myelination in peripheral nerves (Corradi et al., 2003). Furthermore, *Ebf2* and *Ebf3* regulate migration of Cajal-Retzius cells in the cerebral cortex (Chiara et al., 2012). In addition, disruption of *Ebf2* and *Ebf3* causes a defect in the projection of olfactory axons to the olfactory bulb, although it does not affect the olfactory epithelium (Wang, 2004).

In *Xenopus*, injection of *XCoe2* mRNA causes ectopic neurogenesis both in the neural plate and epidermis; conversely, its inactivation reduces neuronal differentiation (Dubois et al., 1998). Similarly, ectopic expression of *Xebf3* converts ectodermal cells into neurons (Pozzoli et al., 2001). Furthermore, overexpression of *XCoe2* and *XEbf3* upregulates a number of genes involved in multiple aspects of neuronal differentiation such as axon growth and stability, synaptic plasticity and maintenance of the membrane potential (Green and Vetter, 2011). This indicates that *Xenopus* COE genes participate in an extensive transcriptional network that regulates neuronal development.

6.2. Aims

The aim of this chapter is to characterise the sequence and expression patterns of *Lampetra planeri* genes belonging to the *COE* gene subfamily based on their identification *in silico* in the *Petromyzon marinus* genome (Chapter 3). This chapter is intended to answer the following questions:

- Does *L. planeri* have *COE* genes as *P. marinus*?
- Provided the recovery of *L. planeri COE* genes, do they have all conserved domains in their sequence as all other metazoan *COE* genes? Is their HLH domain similar to gnathostome or invertebrate *COE* HLH domains with respect to the duplication of the second helix? Do *L. planeri COE* genes present splicing in the HLH domain as occurs in *P. marinus COE-B*?
- Are *L. planeri COE* genes expressed in the nervous system as in gnathostomes? If so, which organs/structures express them? Are their expression patterns also correlated to neuronal differentiation as gnathostome *COE* genes?

6.3. Hypothesis

COE genes are found throughout the animal kingdom and *P. marinus* has at least two *COE* genes as explained in Chapter 3. This is consistent with a genome-wide duplication before the divergence of lampreys/cyclostomes. The question is whether *L. planeri* also has at least two *COE* genes as in *P. marinus*. Apart from the number of genes, it would also be interesting to know whether –predicted– lamprey *COE* genes have a single Helix 2 or two Helix 2 domains as occurs in jawed vertebrates. Therefore,

the study of lamprey *COE* genes will undoubtedly shed light into the evolution of this gene family at the invertebrate-vertebrate transition.

In addition to gene number and sequence aspects, *L. planeri* *COE* gene(s), if present, are predicted to be expressed in restricted regions of the brain, in the spinal cord, pharyngeal arches, as well as in the olfactory placode/epithelium and trigeminal placode/ganglia as in gnathostomes. In the spinal cord, they are predicted to be expressed in the marginal zone in accordance with their role in neuronal differentiation in gnathostomes. Their expression in the spinal cord is also predicted to be biased towards the ventral part of the spinal cord where motor neurons are present in gnathostomes.

6.4. Objectives

- To isolate homologous sequences of gnathostome *COE* genes from the *L. planeri* genome.
- Study the HLH domain of lamprey *COE* genes to identify the number of helices (in particular helix 2) present in those sequences.
- Perform phylogenetic analyses to study the phylogenetic relationships of the isolated sequences with respect to other deuterostome *COE* genes.
- Characterise the expression patterns of the isolated sequences in *L. planeri* embryos at different developmental stages by *in situ* hybridisation.
- Combine both expression and phylogenetic data to identify differences and similarities of expression patterns in different vertebrates mapped into a phylogenetic context in order to identify evolutionary trends of this HLH gene subfamily in vertebrates.

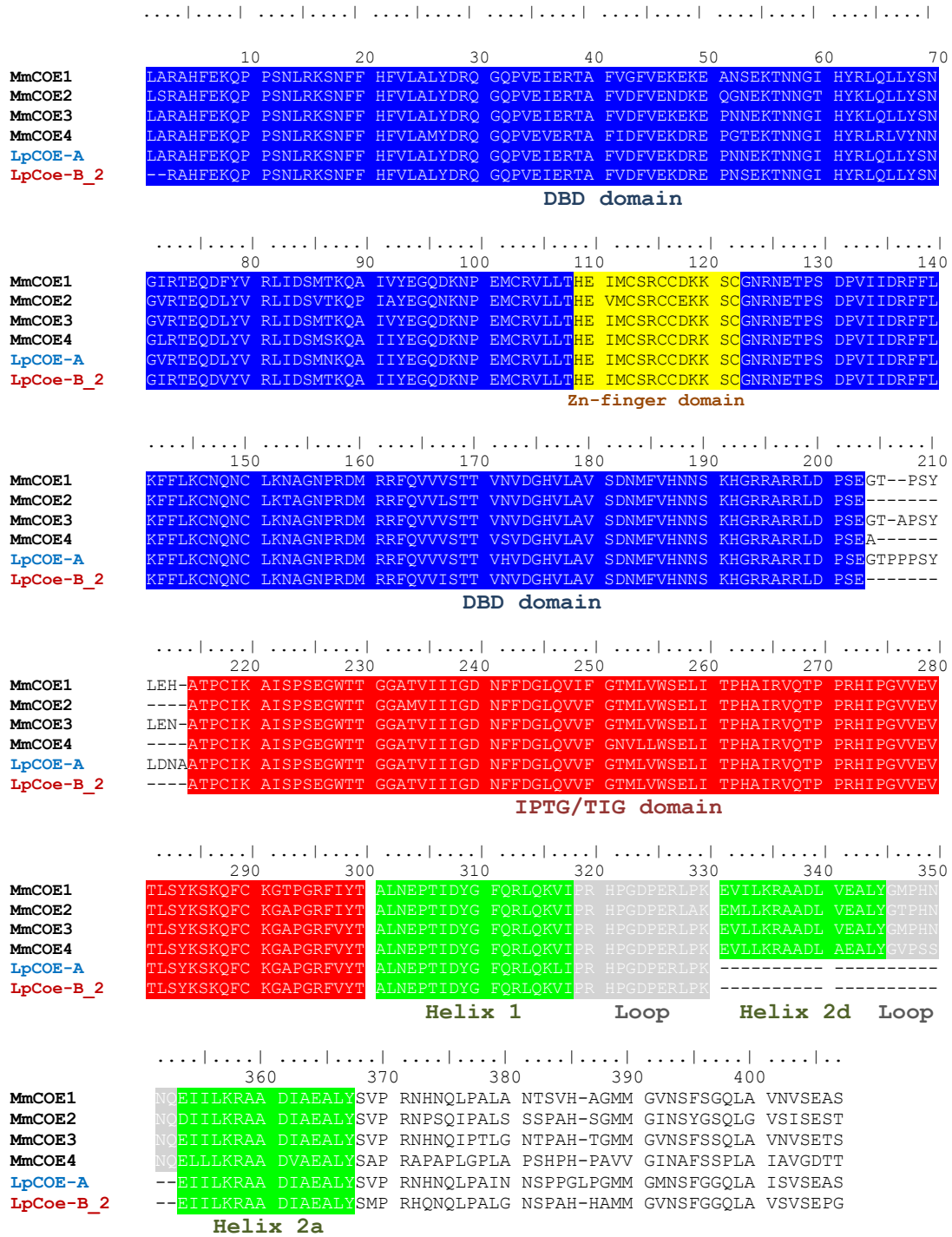


Figure 6-6. Alignment of predicted amino acid sequences of the two *L. planeri* Coe fragments, *LpCoe-A* and *LpCoe-B* isoform 2 (*LpCoe-B_2*, explained below), with the four mouse COE proteins COE1-4. Both lamprey fragments possess highly conserved DBD, IPT/TIG and HLH domains.

Interestingly, both cDNA fragments present only one helix2 (H2) as seen in all invertebrates examined to date, in contrast to other vertebrate COE proteins which have two H2 domains (Figure 6-7).

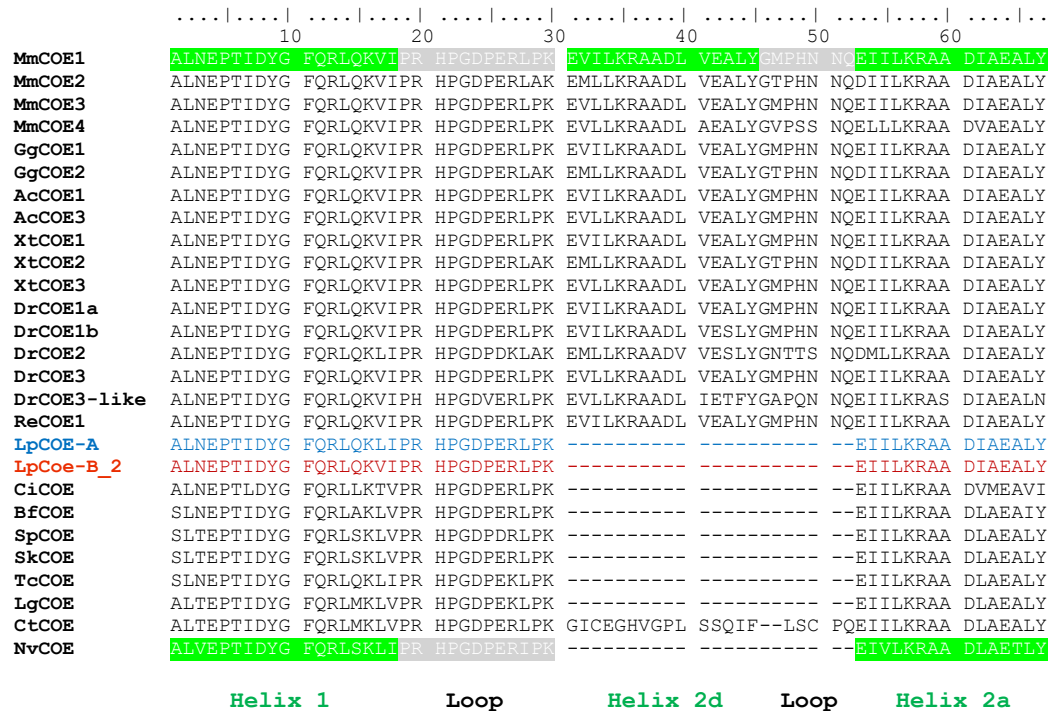


Figure 6-7. Alignment of *LpCoe-A* and *LpCoe-B_2* predicted amino acid sequences with various metazoan COE proteins. Both genes show strong sequence conservation in the HLH domain with other metazoan COE proteins. Notably, both *LpCoe-A* and *LpCoe-B_2* cDNAs lack H2d and the second loop region.

However, careful examination of the contigs containing the HLH domain of both genes showed that, on one hand, the contig containing *LpCoe-A* presents –as expected– H1 and H2a. On the other hand, the contig containing *LpCoe-B* presents H1 and H2a, but in addition, a second H2 (H2d) is found between H1 and H2a (Figure 6-8).

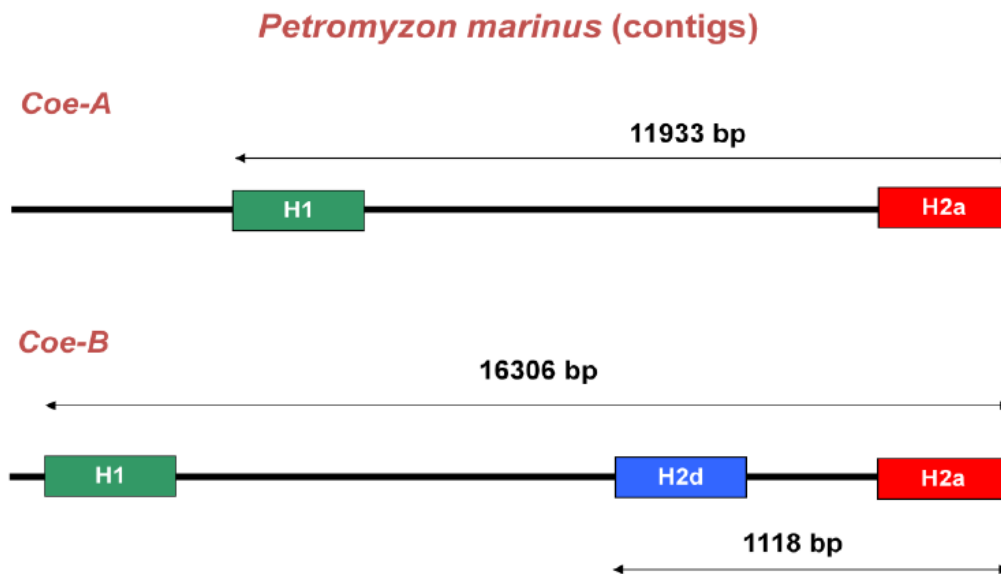


Figure 6-8. Genomic structure of the HLH region of *Petromyzon marinus* *Coe-A* and *Coe-B*. *PmCoe-A* has H1 and H2a, whereas *PmCoe-B* has H1, H2d and H2a. Approximate distances between exons containing each domain are indicated in base pairs (bp).

Given the presence of three helices in the *PmCoe-B* locus (Figure 6-8), different clones of *LpCoe-B* were screened for the possible presence of additional splice variants that could contain both H2 helices (H2d and H2a), that is, three helices in total. In this way, three different splice variants were identified (Figure 6-9). When compared together, the initially isolated sequence containing H1 and H2a, *LpCoe-B_1* (splice variant 1) was shown to have a deletion in the IPT/TIG domain that disrupts the reading frame (Figure 6-9, red arrowheads; Figure 6-10A). Another splice variant, *LpCoe-B_2* (splice variant 2), also containing H1 and H2a, did not present this deletion in the IPT/TIG domain and thus the reading frame is not disrupted in this variant (Figure 6-9). Finally, as predicted, a third splice variant, *LpCoe-B_3* (splice variant 3) containing both H2 helices (H2d and H2a, that is, three helices in total) was retrieved (Figure 6-9, green arrowheads; Figure 6-10B). This variant does not have the IPT/TIG deletion. However, splice variant *LpCoe-B_3* also contains an insert right in the middle of the Zn-

finger domain that is included within the DBD domain (Figure 6-9, orange arrowheads).

This insertion disrupts the reading frame (Figure 6-10C).

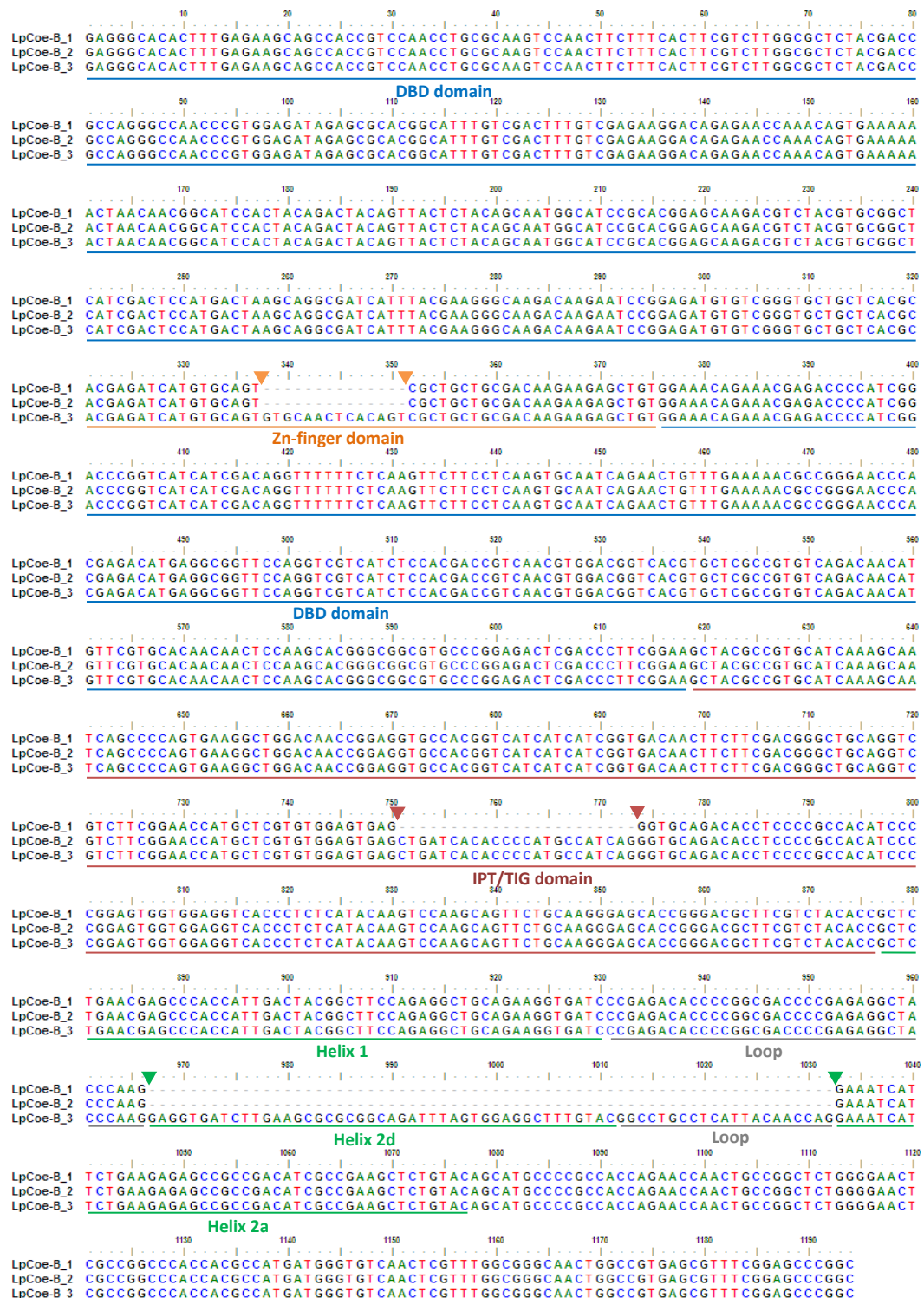


Figure 6-9. Nucleotide sequence of the three *LpCoe-B* splice variants identified in the present study. Splice variant *LpCoe-B_1* possesses only two helices, H1 and H2a, in the HLH domain and it presents a deletion in the IPT/TIG domain (marked by red arrowheads) that disrupts the reading frame (cf. Figure 6-10A). Splice variant *LpCoe-B_2* also contains H1 and H2a helices only but does not present the deletion in the IPT/TIG domain. Splice variant *LpCoe-B_3* presents three helices, namely, H1, H2d (marked by green arrowheads) and H2a (cf. Figure 6-10B). However, this splice variant additionally presents an insertion in the Zn-finger domain (marked by orange arrowheads) that disrupts the reading frame (cf. Figure 6-10C).

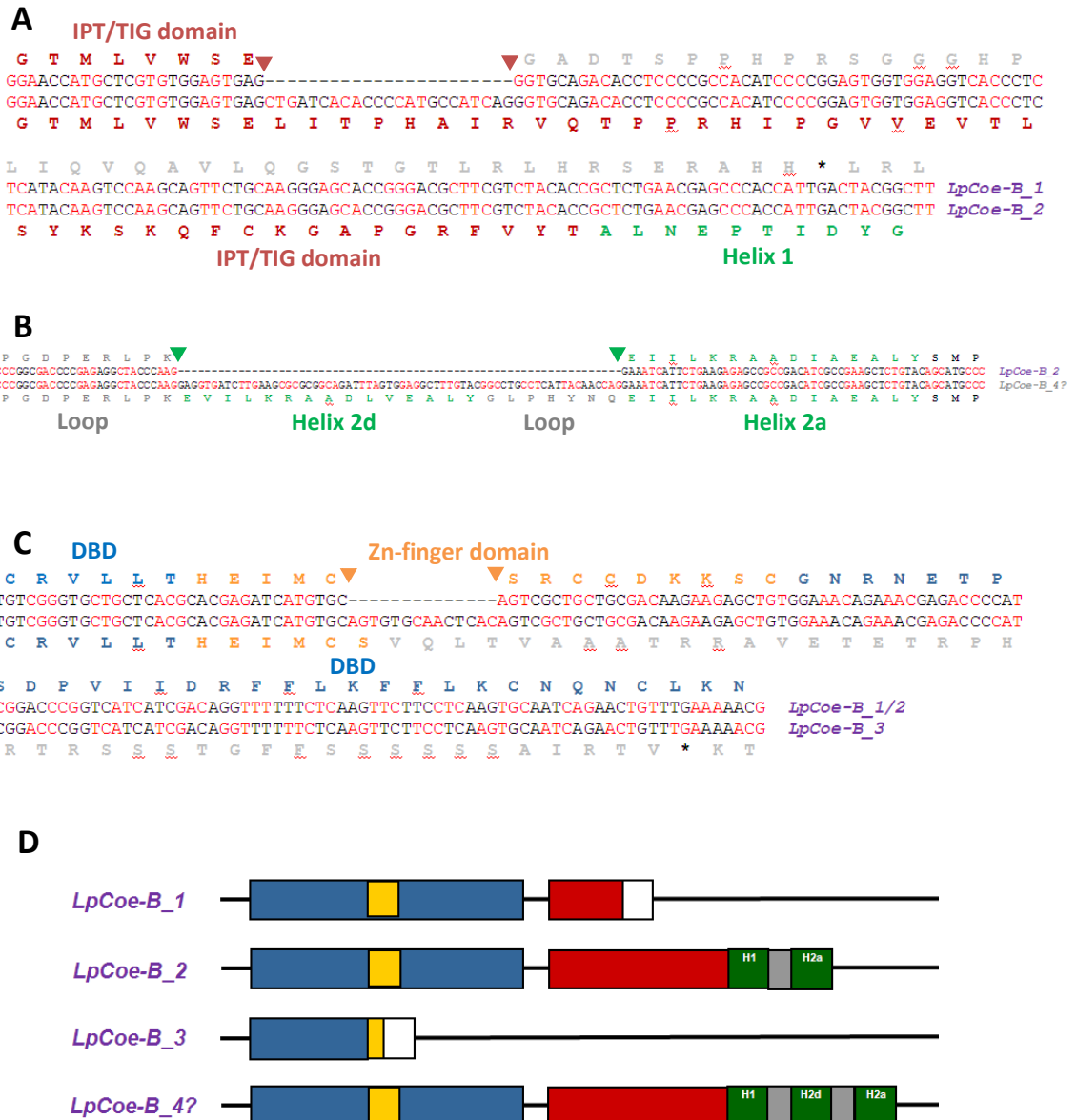


Figure 6-10. Regions within the isolated *LpCoe-B* splice variants surrounding the different splicing sites. (A) Splicing site within the IPT/TIG domain exemplified by splice variants *LpCoe-B_1* and *LpCoe-B_2* (cf. Figure 6-9). Note how the reading-frame in splice variant *LpCoe-B_1* is disrupted by a deletion in this domain (marked by red arrowheads). (B) Splicing site in the HLH domain exemplified by splice variants *LpCoe-B_2* and a predicted *LpCoe-B_4* splice variant (also observed in splice variant *LpCoe-B_3*; cf.

Figure 6-10D). Insertion of Helix 2d and a loop domain (marked by green arrowheads) does not disrupt the reading frame in this region, rendering a potential HLHLH domain. (C) Splicing site in the Zn-finger domain of splice variant *LpCoe-B_3* (marked by orange arrowheads) that also contains a potential HLHLH domain (cf. Figure 6-9). This insertion, however, disrupts the reading-frame within the Zn-finger domain and, therefore, a HLHLH domain would not be produced by this splice variant. (D) Schematic representation of peptides predicted to be produced by the three *LpCoe-B* splice variants identified in the present study. Splice variant *LpCoe-B_1* would contain a normal DBD, with a Zn-finger within it, and only –roughly– the first half of the IPT/TIG domain. Splice variant *LpCoe-B_2* would contain an intact DBD (with a Zn-finger), the IPT/TIG domain, and a HLH domain composed of Helix 1 and Helix 2a, rendering a predicted fully-functional protein. Splice variant *LpCoe-B_3* would contain only –roughly– the first half of a disrupted DBD domain that stops after several amino acids where a Zn-finger domain is expected to be located. In addition, a fourth splice variant, *LpCoe-B_4* is predicted to exist that does not contain the insertion within the Zn-finger domain, it does not present the deletion within the IPT/TIG domain, and also possesses a HLHLH domain composed of Helix 1, Helix 2d and Helix 2a, rendering a predicted fully-functional protein with three helices.

Both *LpCoe-A* and *LpCoe-B* fragments were aligned with the two *Petromyzon marinus* Ensembl gene predictions discussed in Chapter 3 that grouped with other *COE* genes to determine whether they correspond to respective *LpCoe-A* and *LpCoe-B* orthologs. The alignment with *LpCoe-A* and *P. marinus* transcript prediction ENSPMAT00000003606 (corresponding to amino acid prediction ENSPMAP00000003591) at the nucleotide level showed near-identical sequence similarity in the aligned region, strongly suggesting that they are gene orthologs (Figure 6-11). This prediction, however, does not contain the first half of the DBD and the Zn-finger domain as seen when aligned at the –predicted– amino acid level (Figure 6-12). Additionally, this prediction contains a stretch between the loop region and Helix 2a, where a second Helix 2, corresponding to Helix 2d in *LpCoe-B_3*, could potentially be found. However, this stretch does not present any sequence similarity to Helix 2d in *LpCoe-B_3* (Figure 6-12, black arrowheads).

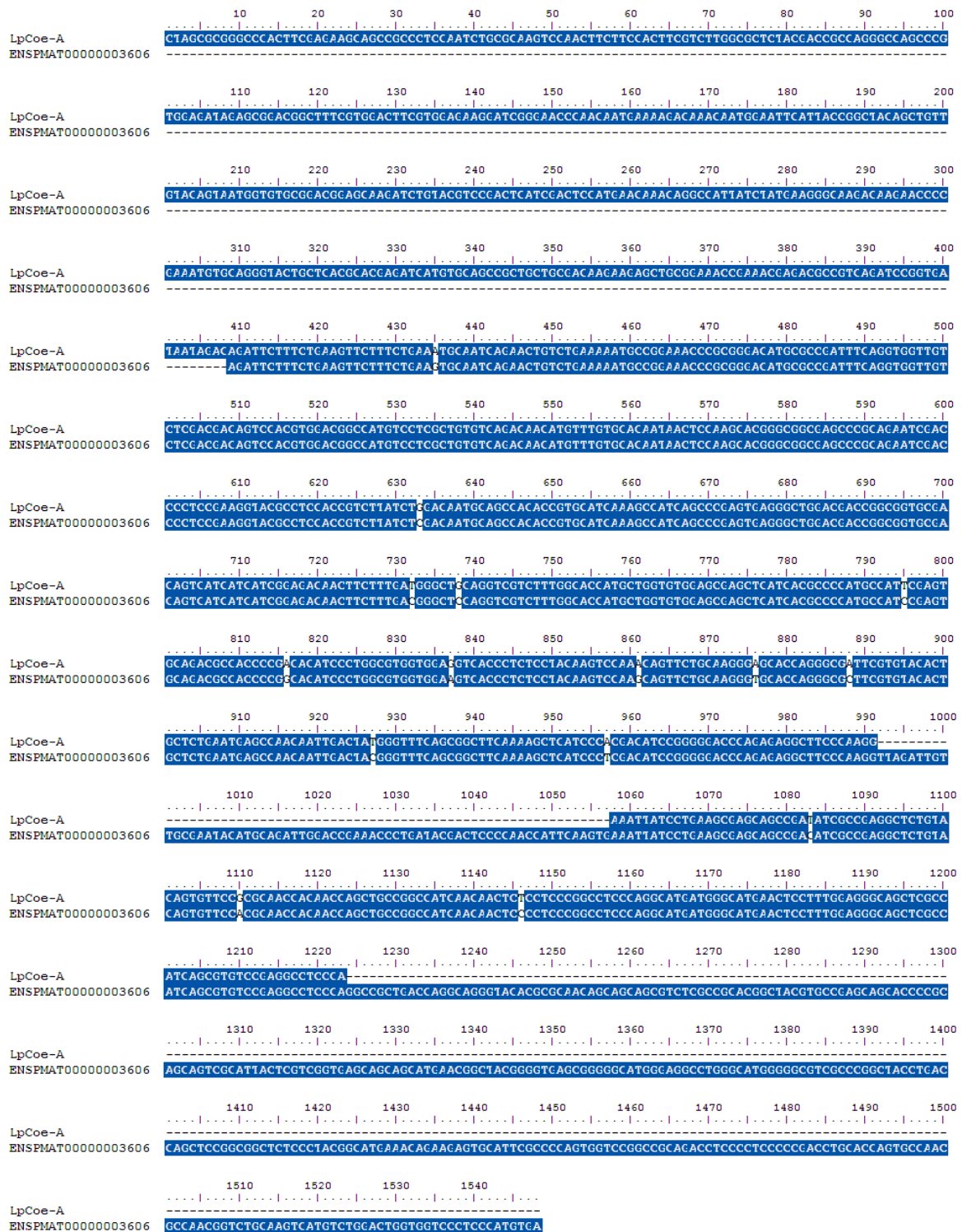


Figure 6-11. Nucleotide alignment of *LpCoe-A* and *Petromyzon marinus* Ensembl transcript prediction ENSPMAT00000003606. Sequences show near-identical similarity in the aligned region.

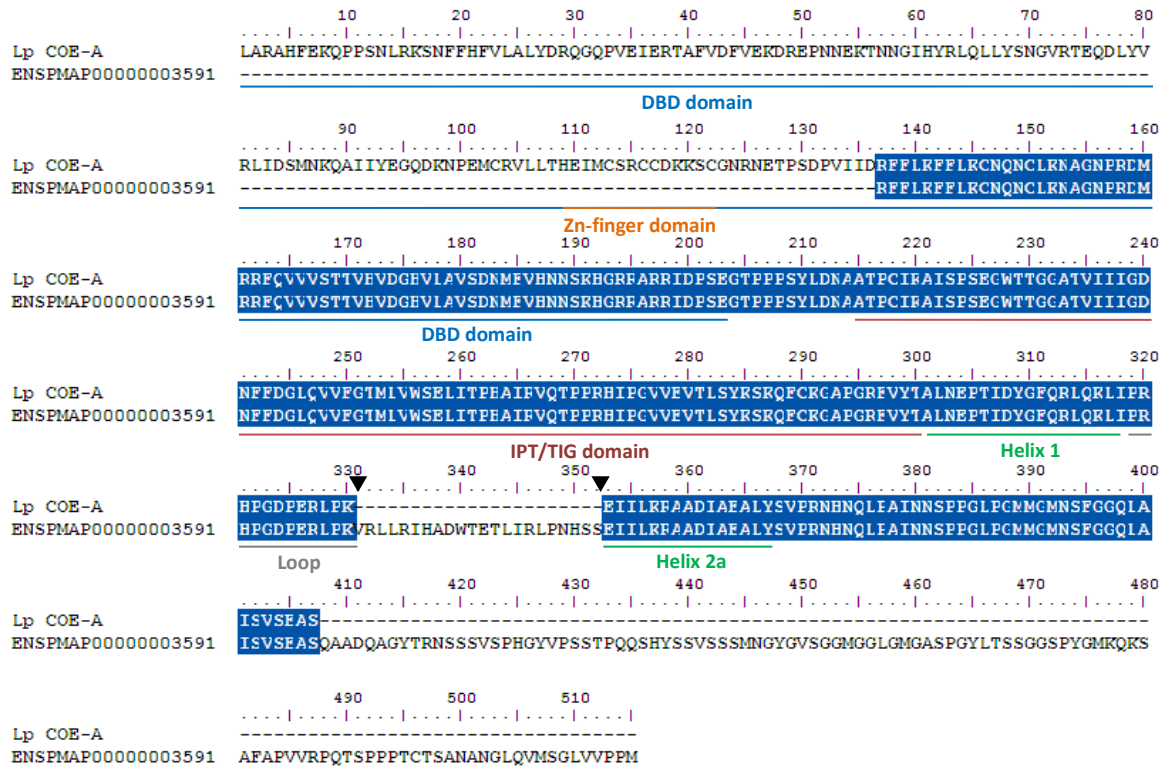


Figure 6-12. Predicted amino acid alignment of *LpCoe-A* and *Petromyzon marinus* Ensembl protein prediction ENSPMAP00000003591. The Ensembl prediction contains all conserved domains except the first half of the DBD domain and the Zn-finger domain. All predicted amino acids in the aligned region are identical. However, note how the Ensembl prediction includes a stretch (delimited by arrowheads) between the loop region and Helix 2a, in a region where another Helix 2, corresponding to Helix 2d in *LpCoe-B_3*, could potentially be found. This stretch, however, does not show any sequence similarity to Helix 2d in *LpCoe-B_3* (cf. Figure 6-9B).

For *LpCoe-B*, isoform *LpCoe-B_3* was used as it is the only isolated splice variant with sequence information for the three helices in the HLH domain, even though the insertion in the Zn-finger domain of this splice variant disrupts the reading frame (see Figures 6-9 and 6-10C). However, as this analysis is only to study gene orthology, this insertion was removed to obtain a predicted (undisrupted) reading frame containing all conserved domains (i.e. the DBD with an intact Zn-finger domain, the IPT/TIG domain, and the HLH domain). Alignment of *LpCoe-B_3* with *P. marinus* transcript prediction ENSPMAT00000004655 (corresponding to amino acid prediction

ENSPMAP00000004636) at the nucleotide level revealed near-identity in most of the aligned region; however, two aligned regions presented very different sequences (Figure 6-13, arrowheads). When compared at the amino acid level, these two regions correspond to a small part of the DBD and to about the second half of the IPT/TIG domain (Figure 6-14, arrowheads).

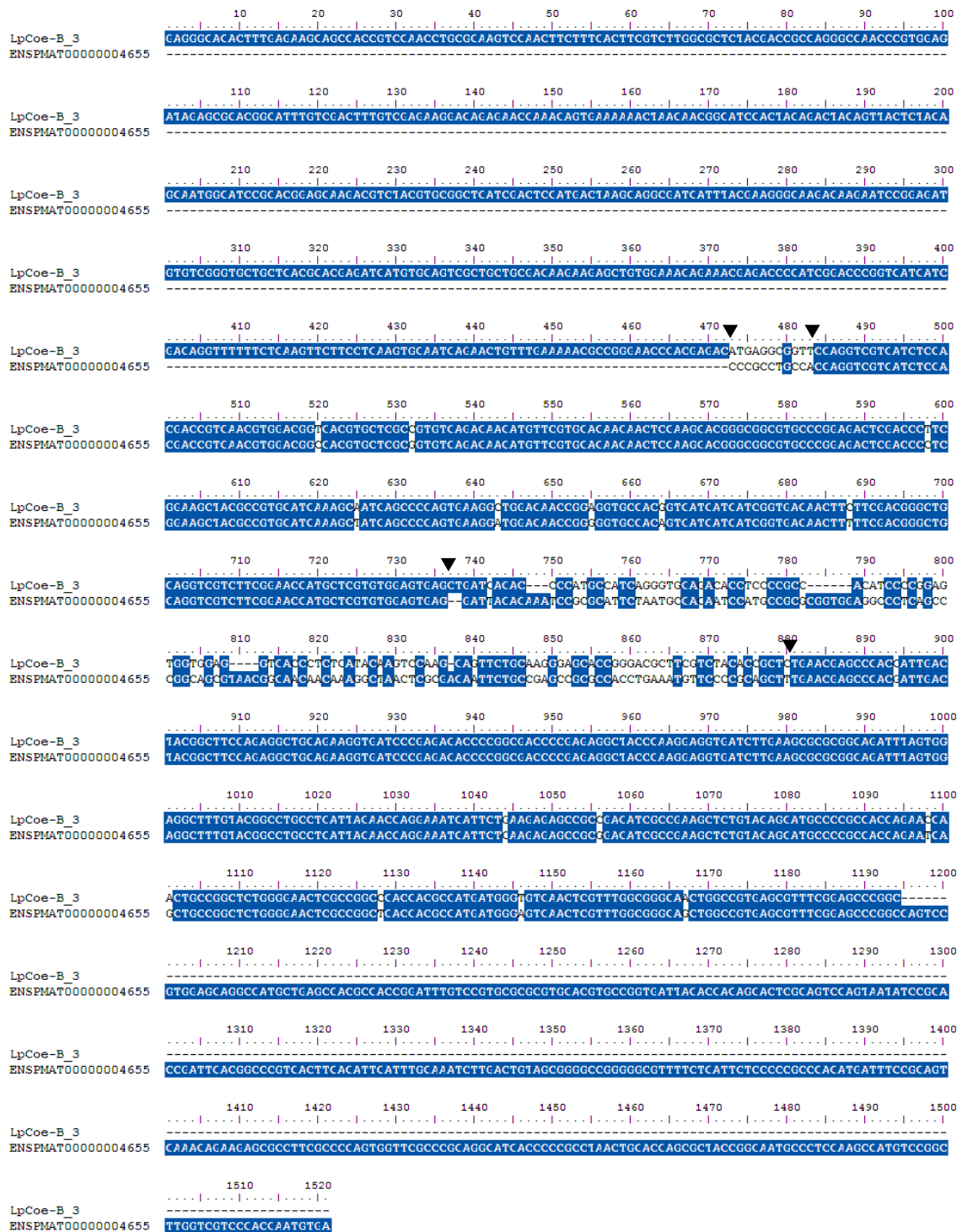


Figure 6-13. Nucleotide alignment of *LpCoe-B_3* (with the insertion in the Zn-finger domain removed; cf. Figures 6-8 and 6-9C) and *Petromyzon marinus* Ensembl transcript prediction ENSPMAT00000004655. Sequences show near-identity in most of the aligned region except in the two regions delimited by arrowheads.

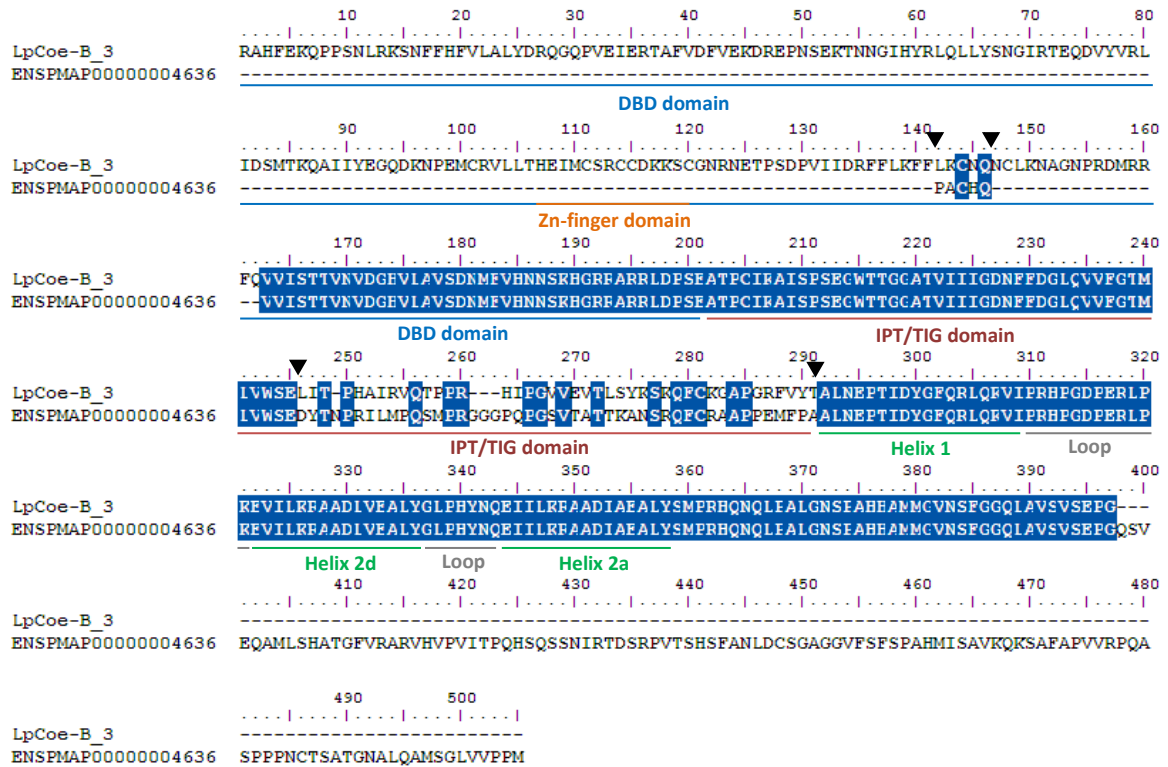


Figure 6-14. Predicted amino acid alignment of *LpCoe-B_3* (with the insertion in the Zn-finger domain removed) and *Petromyzon marinus* Ensembl protein prediction ENSPMAP00000004636. The Ensembl prediction contains only a small portion of the DBD, excluding the Zn-finger domain, the IPT/TIG domain, and the HLHLH domain. However, a small region in the DBD and about the second half of the IPT/TIG domain (delimited by arrowheads) in the Ensembl prediction present very different sequences with respect to *LpCoe-B_3*. Otherwise, predicted amino acids of both sequences in the aligned region are identical.

All together, considering the alignments of *LpCoe-A* and *LpCoe-B_3* with corresponding *Petromyzon marinus* Ensembl gene predictions, particularly at the nucleotide level, and that gene predictions could be inaccurate, it is very likely that Ensembl transcript prediction ENSPMAT00000003606 reveals a gene orthologous to *LpCoe-A*, whereas Ensembl transcript prediction ENSPMAT00000004655 reveals a gene orthologous to *LpCoe-B*.

After the identification of different *LpCoe-B* splice variants, I looked particularly at the temporal distribution of the HLH domain variants, that is, at what point in development splice variants containing either two or three helices are expressed. For this purpose, RT-PCR was performed using total RNA extracted from embryos at four different time periods: from stage 17 to 20, from stage 21 to 23, from stage 24 to 26 and from stage 28 to 30. Primers were designed within each helix of both *LpCoe-A* and *LpCoe-B* as shown below:

For *LpCoe-A*:

- Sense primer in Helix 1 (H1): 5'-TCTGAATGAGCCAACAATTGACTAT-3'
- Antisense primer in Helix 2a (H2a): 5'-TGGTTGCGCGGAACACTGTACAGAGCC-3'

For *LpCoe-B*:

- Sense primer in Helix 1 (H1): 5'-TCTGAACGAGCCCACCATTGACTAC-3'
- Antisense primer in Helix 2d (H2d): 5'-TGGTTGTAATGAGGCAGGCCGT-3'
- Antisense primer in Helix 2a (H2a): 5'-TGGTGGCGGGGCATGCTGTACAGAGCT-3'

Three combinations of primers were used for PCR: for *LpCoe-A*, only H1 + H2a were used, whereas for *LpCoe-B*, either H1 + H2d or H1 + H2a were used. In this way, based on the splice variants described above (i.e. splice variants *LpCoe-B_1-3*), splice variants containing two or three helices can be discriminated by size.

All primer combinations were first tested for PCR, cloning and sequencing, to corroborate that they amplify corresponding *LpCoe-A* and *LpCoe-B* HLH domains using cDNA derived from stage 24-26 embryos. As with this pilot PCR, two bands of approximately 150 bp and 230 bp were observed in all time frames analysed using primer combination H1 + H2a for *LpCoe-B* (Figure 6-14A, blue and red arrow, respectively), which recognises transcripts containing sequence information for only two (*LpCoe-B_1* and/or *LpCoe-B_2*) or three (*LpCoe-B_3*) helices, respectively, regardless of whether these transcripts encode functional or truncated proteins. Using primer combination H1 + H2d for *LpCoe-B*, a band of approximately 150 bp was identified, accordingly corresponding to Helix 1 and Helix 2d (Figure 6-14A, blue arrow). This band, however, was seen in all time frames analysed except from stages 17-20.

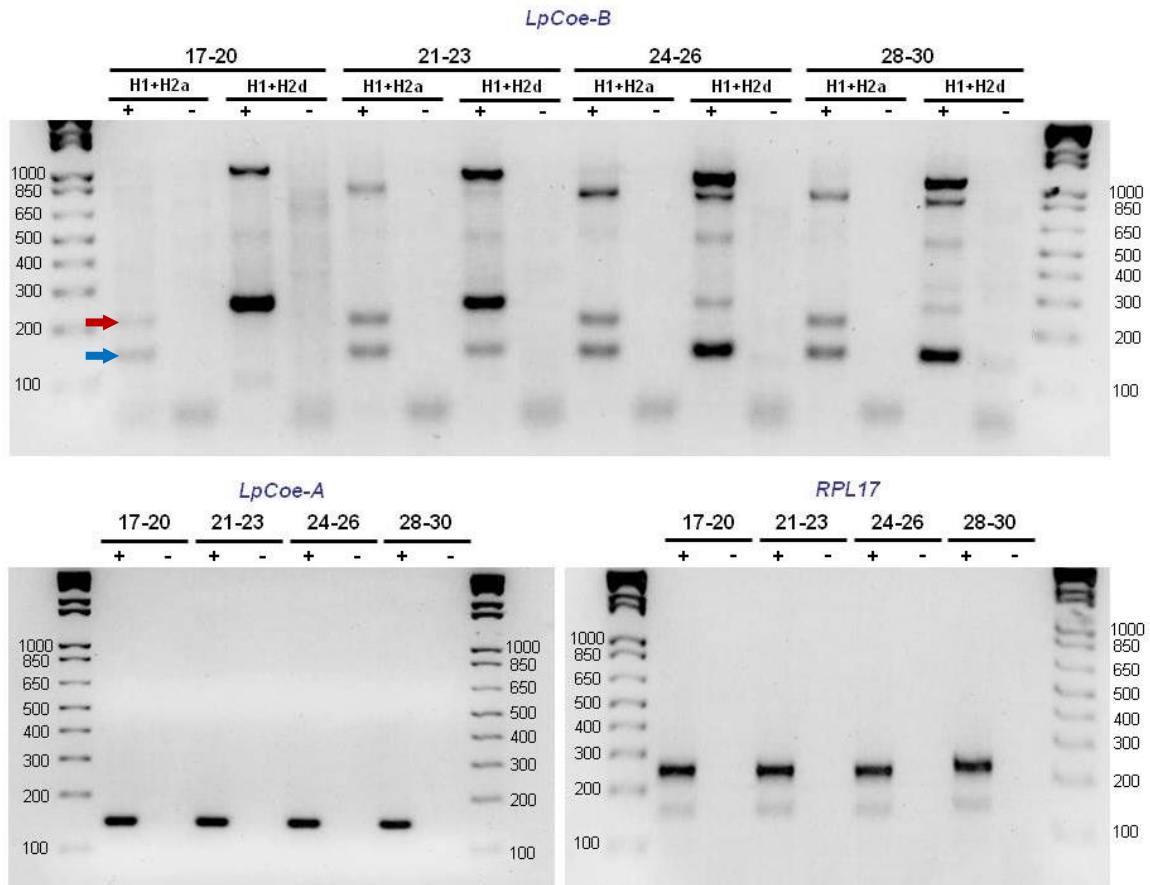


Figure 6-14. Expression of the HLH domains of both *LpCoe-A* and *LpCoe-B* at different stages of development. (A) Expression of *LpCoe-B* at 4 intervals of development as indicated. Bands of the expected sizes according to each primer combination are seen. (B) Expression of *LpCoe-A* at the same time intervals. One single band of the corresponding size according to the pair of primers is observed. Note that other larger bands are also detected. (C) RPL17 was used as a control to test the integrity of the cDNA that was used in these experiments at the respective time periods.

Although this analysis is only semi-quantitative, a general increase in the amount of each splice variant is observed after stage 20. In addition, several longer bands were also amplified in this particular experiment with both primer combinations. These bands, however, were not observed in the pilot experiment to test the specificity of all primer combinations. Due to time restrictions, the identity of these bands could not be determined.

On the other hand, as expected, primer combination H1 + H2a for *LpCoe-A* yielded only one band of approximately 150 bp corresponding to Helix 1 and Helix 2a (Figure 6-14B). This band was observed in all time periods examined. Finally, to test the integrity of cDNA synthesised for these experiments, an additional PCR was performed using primers for lamprey ribosomal protein L17 (RPL17), a protein that forms part of the exit channel of the large ribosomal subunit (Treiber, Mandel et al. 2010) (Meng, Zhang et al. 2010) (sense primer: 5'-CACATCCGCAAGGCAACCAAGTA-3'; antisense primer: 5'-TGGGGGGCCTTGTTACCTG-3'). One strong band of approximately 260 bp was observed in all time periods, confirmed to be RPL17 by sequencing. This result shows that all samples are comparable.

6.5.2. Phylogenetic Analysis of *LpCoe-A* and *LpCoe-B*

After the identification of *Lampetra planeri* COE gene homologs, a Neighbour Joining tree was constructed as described in section 2.1.3. in Methods using the alignment shown in Figure 6-15 with COE amino acid sequences from representative metazoan species (Table 6-1). For *LpCoe-B*, splice variant *LpCoe-B_2* was used as it is the only splice variant which does not have a disrupted frame in the region that was cloned which spans all conserved domains. The NJ tree places *LpCoe-A* as the sister group of mouse *COE4*, but with very low support (bootstrap value of 36), whereas *LpCoe-B_2* is placed as the sister group of the *COE1-4* clade containing *LpCoe-A*, but again, with very low support (bootstrap support value of 38) (Figure 6-16). Support values are generally very low in this tree, likely due to the very well conserved domains, a region that was used to construct this tree. However, these conserved domains precisely indicate that both cloned fragments are true COE genes.

Species	Gene name	Accession number	Species	Gene name	Accession number
<i>Homo sapiens</i>	<i>HsCOE1</i>	NP_076870.1	<i>Raja eglanteria</i>	<i>ReCOE1</i>	AAL86576.1
<i>Homo sapiens</i>	<i>HsCOE2</i>	NP_073150.2	<i>Lampetra planeri</i>	<i>LpCoe-A</i>	This study
<i>Homo sapiens</i>	<i>HsCOE3</i>	NP_001005463.1	<i>Lampetra planeri</i>	<i>LpCoe-B</i>	This study
<i>Homo sapiens</i>	<i>HsCOE4</i>	NP_001103984.1			
<i>Mus musculus</i>	<i>MmCOE1</i>	NP_031923.1	<i>Ciona intestinalis</i>	<i>CiCOE</i>	NP_001072030.1
<i>Mus musculus</i>	<i>MmCOE2</i>	NP_034225.1			
<i>Mus musculus</i>	<i>MmCOE3</i>	NP_001106886.1	<i>Branchiostoma floridae</i>	<i>BfCOE</i>	CAE45569.1
<i>Mus musculus</i>	<i>MmCOE4</i>	NP_001103983.1			
			<i>Strongylocentrotus purpuratus</i>	<i>SpCOE</i>	XP_782395.3
<i>Gallus gallus</i>	<i>GgCOE1</i>	NP_990083.1	<i>Saccoglossus kowalevskii</i>	<i>SkCOE</i>	NP_001158446.1
<i>Gallus gallus</i>	<i>GgCOE2</i>	NP_001158008.1			
<i>Gallus gallus</i>	<i>GgCOE3(ex)</i>	XP_421824.3	<i>Drosophila melanogaster</i>	<i>DmCOE</i>	NP_524813.2
<i>Anolis carolinensis</i>	<i>AcCOE1</i>	XP_003223510.1	<i>Tribolium castaneum</i>	<i>TcCOE</i>	XP_973686.2
<i>Anolis carolinensis</i>	<i>AcCOE3</i>	ENSACAP00000008592			
<i>Xenopus laevis</i>	<i>XlCoe2</i>	NP_001079147.1	<i>Caenorhabditis elegans</i>	<i>CeCOE</i>	NP_510453.1
<i>Xenopus laevis</i>	<i>XlCoe3</i>	NP_001083801.1			
			<i>Lottia gigantea</i>	<i>LgCOE</i>	jgi Lotgi1 225788
<i>Xenopus tropicalis</i>	<i>XtCOE1</i>	XP_002939654.1	<i>Capitella teleta</i>	<i>CtCOE</i>	ADQ39029.1
<i>Xenopus tropicalis</i>	<i>XtCOE2</i>	NP_989200.1			
<i>Xenopus tropicalis</i>	<i>XtCOE3</i>	XP_002932694.1	<i>Helobdella robusta</i>	<i>HrCOE</i>	jgi Helro1 112029
<i>Danio rerio</i>	<i>DrCOE1a</i>	XP_693680.3	<i>Nematostella vectensis</i>	<i>NvCOE</i>	AAS48917.1
<i>Danio rerio</i>	<i>DrCOE1b</i>	XP_685990.5			
<i>Danio rerio</i>	<i>DrCOE2</i>	NP_571493.1			
<i>Danio rerio</i>	<i>DrCOE3</i>	NP_001074071.1	<i>Amphimedon queenslandica</i>	<i>AqCOE</i>	geneid_v1.2_predicted_protein_1 688_AA
<i>Danio rerio</i>	<i>DrCOE4</i>	XP_684384.1			

Table 6-1. List of metazoan COE sequences used to construct the phylogenetic tree shown in Figure 6-15. GenBank or Joint Genome Institute (JGI) accession numbers are to the right of each gene.

	10	20	30	40	50	60	70	80	90
MmCOE1	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EANSEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
MmCOE2	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVENDK	---EQNEKTINGHYKLQLLYSNGVTEODLYVRLIDSVTK							
MmCOE3	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EPNNEKTINGHYKLQLLYSNGVTEODLYVRLIDSMTK							
MmCOE4	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDR	---EPGTEKTINGHYRLRLVYNGVTEODLYVRLIDSMK							
GgCOE1	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EANSEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
GgCOE2	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVENDK	---EQNEKTINGHYKLQLLYSNGVTEODLYVRLIDSVTK							
AcCOE1	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EANSEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
AcCOE3	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EPNNEKTINGHYKLQLLYSNGVTEODLYVRLIDSMTK							
XtCOE1	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---ETSSEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
XtCOE2	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVENEK	---ELSTEKTINGHYKLQLLYSNGVTEODLYVRLIDSVTK							
XtCOE3	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EPNSEKTINGHYKLQLLYSNGVTEODLYVRLIDSMTK							
DrCOE1a	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---ESTGEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
DrCOE1b	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---ETTGEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
DrCOE2	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---TGEKTINGHYKLQLLYSNGVTEODLYVRLIDSVTK							
DrCOE3	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EPNSEKTINGHYKLQLLYSNGVTEODLYVRLIDSMTK							
DrCOE3-like	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EPVSEKTINGHYKLQLLYSNGVTEODLYVRLIDSVTK							
ReCOE1	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---ELSGEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
LpCOE-A	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EPNNEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
LpCoe-B_2	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EPNSEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
CiCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVENET	---ETAGEKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
BfCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EPDNRKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
SpCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EPDATKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
SkCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EPDNQKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
TcCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EADGQKTINGHYRLQLLYSNGVTEODLYVRLIDSVTK							
LgCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKDK	---EHEGQKTINGHYRLQLLYSNGVTEODLYVRLIDSVTK							
CtCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---EPEG-KTINGHYRLQLLYSNGVTEODLYVRLIDSVTK							
NvCOE	RAHSEKQPPNLRKSNFFBEVIALDRGQGVIEDRTAFVGFVEKEK	---LPEGRKTINGHYRLQLLYSNGVTEODLYVRLIDSMTK							
	100	110	120	130	140	150	160	170	180
MmCOE1	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
MmCOE2	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
MmCOE3	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
MmCOE4	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
GgCOE1	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
GgCOE2	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
AcCOE1	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
AcCOE3	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
XtCOE1	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
XtCOE2	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
XtCOE3	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
DrCOE1a	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
DrCOE1b	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
DrCOE2	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
DrCOE3	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
DrCOE3-like	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
ReCOE1	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
LpCOE-A	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
LpCoe-B_2	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
CiCOE	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
BfCOE	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
SpCOE	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
SkCOE	QPIAYEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
TcCOE	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
LgCOE	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
CtCOE	QAIIVLEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
NvCOE	QIVPVEGQDNFEMORVLLTHEIMCSRCCDRSCCN	NETPSDFVIIDRFLKGFPLRCNCNCLK	ACNPRDMRFECVWSTTVMDGHVL						
	190	200	210	220	230	240	250	260	270
MmCOE1	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
MmCOE2	AVSINMVFVNSRBGRRAR-RDPSE	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
MmCOE3	AVSINMVFVNSRBGRRAR-RDPSEGT-APSYL	-----EN--ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
MmCOE4	AVSINMVFVNSRBGRRAR-RDPSEA	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
GgCOE1	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
GgCOE2	AVSINMVFVNSRBGRRAR-RDPSE	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
AcCOE1	AVSINMVFVNSRBGRRAR-RDPSEA	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
AcCOE3	AVSINMVFVNSRBGRRAR-RDPSEGT-APSYL	-----ENEVATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
XtCOE1	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--AATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
XtCOE2	AVSINMVFVNSRBGRRAR-RDPSE	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
XtCOE3	AVSINMVFVNSRBGRRAR-RDPSEGT-APSYL	-----EN--VATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
DrCOE1a	AVSINMVFVNSRBGRRAR-RDPSEA	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
DrCOE1b	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--AATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
DrCOE2	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--AATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
DrCOE3	AVSINMVFVNSRBGRRAR-RDPSEA	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
DrCOE3-like	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--AATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
ReCOE1	AVSINMVFVNSRBGRRAR-RDPSEGT--PSYL	-----EH--AATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
LpCOE-A	AVSINMVFVNSRBGRRAR-RDPSEGT-PPSYL	-----DN--AATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
LpCoe-B_2	AVSINMVFVNSRBGRRAR-RDPSE	-----ATPCIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						
CiCOE	AVSINMVFVNSRBGRRAR-RDPSE	-----ASPTIFAI	SPSECWTGGATVIIIGDIFFDGLOWFGTMLVWS						

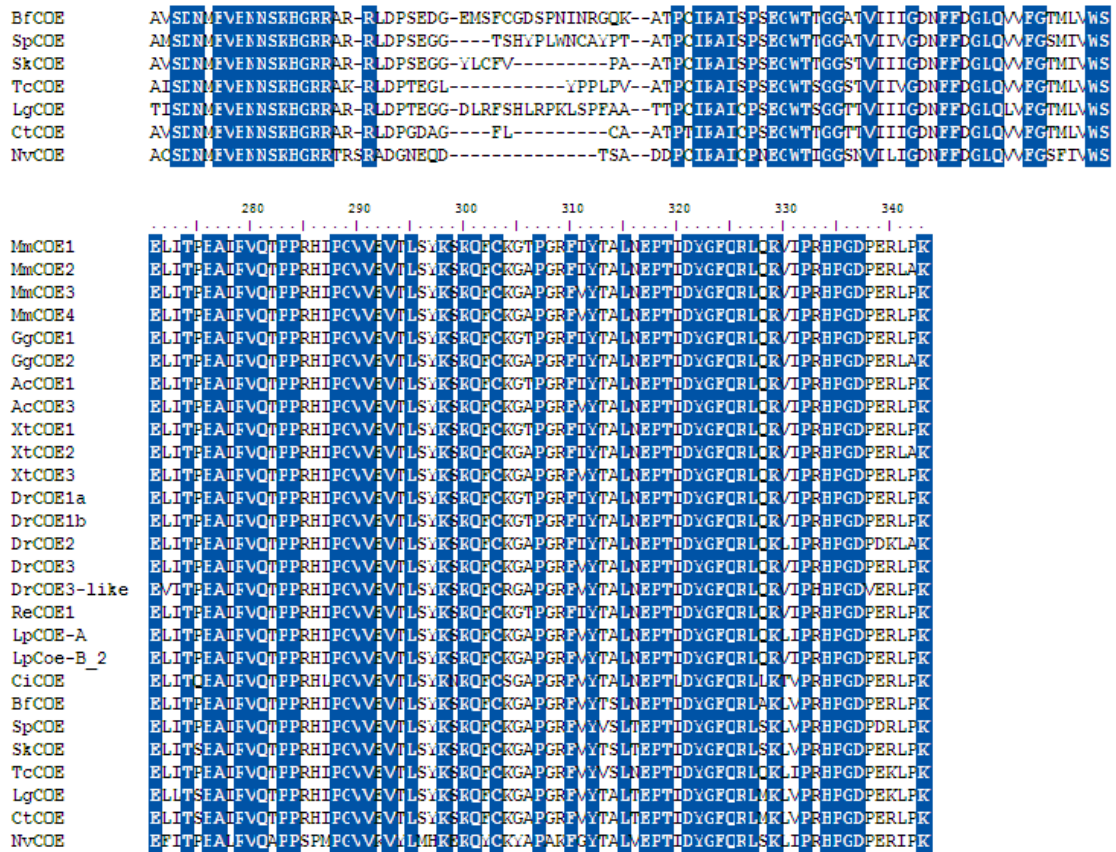


Figure 6-15. MAFFT alignment of COE sequences (shown in Table 6-1) used to construct the phylogenetic tree in Figure 6-16.

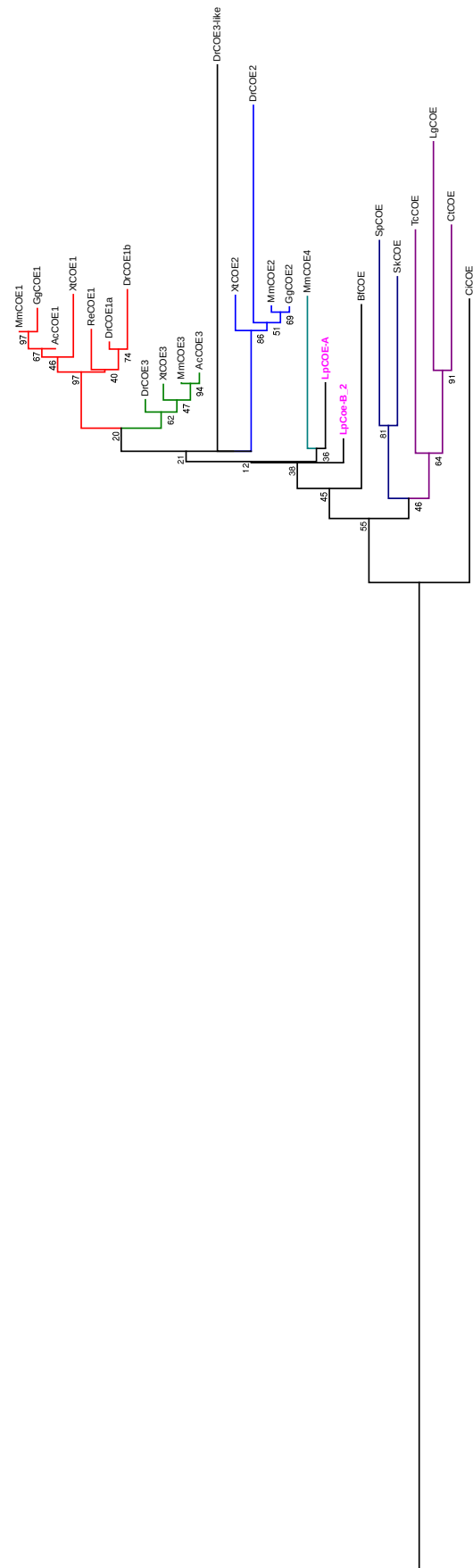


Figure 6-16. NJ consensus tree using the alignment shown in Figure 6-15 showing the phylogenetic relationships of *LpCoe-A* and *LpCoe-B_2* with respect to representative metazoan COE sequences. The tree was built with the JTT amino acid substitution model and gamma distribution. Bootstrap support values after 1000 replications. Scale bar, amino acid substitutions per position. Species abbreviations: Ac, *Anolis carolinensis*; Bf, *Branchiostoma floridae*; Ci, *Ciona intestinalis*; Ct, *Capitella teleta*; Dr, *Danio rerio*; Gg, *Gallus gallus*; Lg, *Lotia gigantea*; Mm, *Mus musculus*; Nv, *Nematostella vectensis*; Re, *Raja eglanteria*; Sk, *Saccoglossus kowalevskii*; Sp, *Strongylocentrotus purpuratus*; Tc, *Tribolium castaneum*; Xt, *Xenopus tropicalis*.

6.5.3. *LpCoe-A* Expression in Developing *L. planeri* Embryos

Expression of *LpCoe-A* was analysed by digoxigenin-labelled *in situ* hybridisation from Tahara (1988) stages 21 to 28. At stage 21, expression is seen in the developing mandibular arch as well as in a domain under the notochord (Figure 6-17A). At stage 22, expression appears in the nascent second and third pharyngeal arches (Figure 6-17B). At stage 23, expression dramatically increases and now appears in the spinal cord and ventral regions of the midbrain and hindbrain (possibly also of the diencephalon), in the trigeminal, facial and lateral line placodes, as well as in the developing fourth pharyngeal arch and in an extending domain below the notochord (Figure 6-17C and D). At this stage, faint expression starts to be seen in the NHP (Figure 6-17D). Cranial placodes and ganglia at this and subsequent stages were identified following descriptions of (Damas, 1944; Kuratani et al., 1997; Modrell et al., 2014). Anatomical internal references to aid in their identification were pharyngeal arches, otic vesicle, MHB, epiphysis and the transition of the notochord to the precordial plate.

At stage 24, *LpCoe-A* expression in both the midbrain and hindbrain extends dorsally, leaving a gap of expression at the level of the MHB (Figure 6-17G). Expression extends posteriorly along the growing spinal cord and also appears in the

developing fifth pharyngeal arch (Figure 6-17E and G). Expression also increases in the NHP and cranial placodes/ganglia, with new expression appearing in the ophthalmic placode (Figure 6-17G). From a dorsal view, expression is seen as two lateral stripes all along the neural tube, leaving the medial region unstained which corresponds to the ventricular zone of the neural tube (Figure 6-17F). Accordingly, expression is found in the marginal zone of the neural tube. At the level of the vitellum, expression in the spinal cord is not continuous, and instead is seen as intermittent patches of expression as observed from a dorsal view (Figure 6-17H).

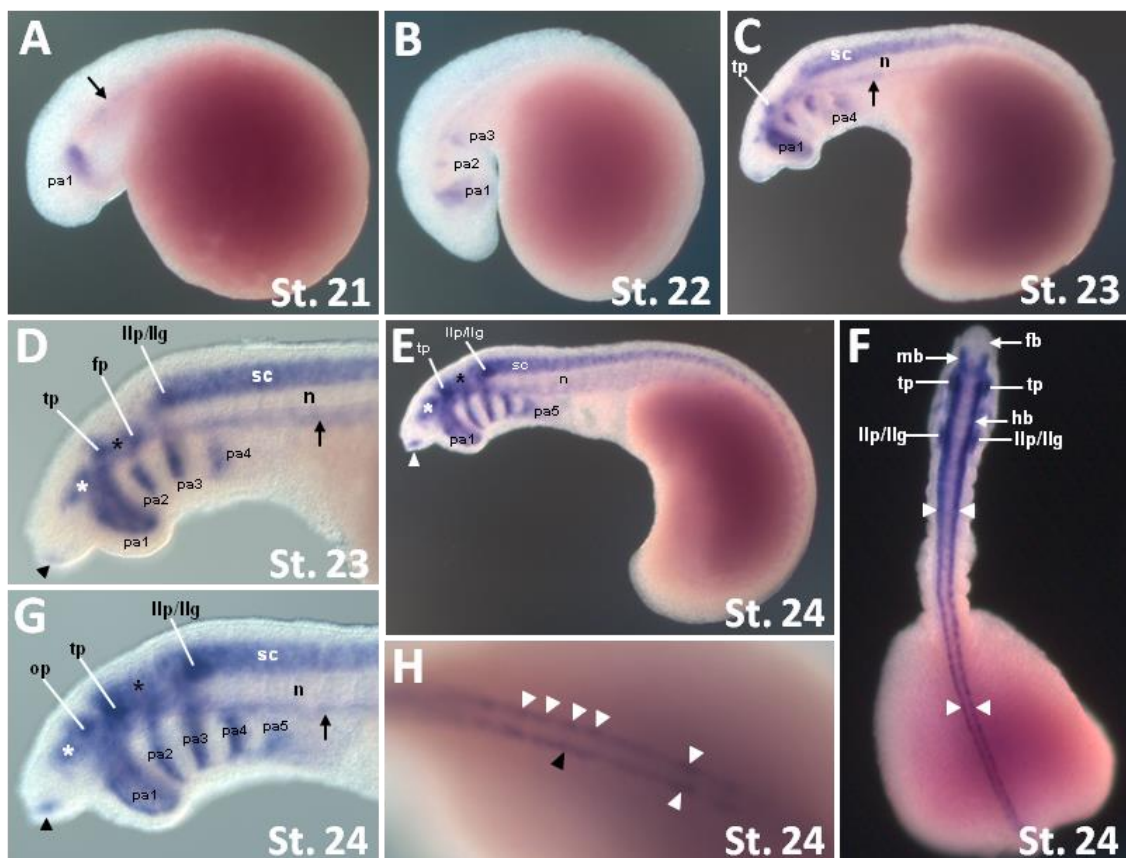


Figure 6-17. *LpCoe-A* expression during *Lampetra planeri* development at stages 21-24. (A-E, G) are lateral views. (F, H) are dorsal views. In all photographs anterior is to the left except in (F) where anterior is to the top. (A) At stage 21, expression is seen in the mandibular arch (pa1) and below the notochord (arrow). (B) At stage 22, expression appears in the second and third pharyngeal arches (pa2, pa3). (C, D) At stage 23, expression appears in the spinal cord (sc), ventral midbrain (white asterisk) and hindbrain (black asterisk), in the trigeminal (tp), facial (fp), and lateral line placode/ganglia (llp/lpg), nasohypophyseal plate (black arrowhead), as well as in the fourth pharyngeal arch (pa4). Expression

below the notochord (n) extends posteriorly (arrow). (E-H) At stage 24, expression extends dorsally in the midbrain (white asterisk) and hindbrain (black asterisk) (E, G). Expression increases in the NHP and cranial placodes/ganglia, and now appears in the ophthalmic placode (op) (F, G). Transcripts are also observed in the nascent fifth pharyngeal arch (G). In the spinal cord, expression progresses posteriorly (E). From a dorsal view, expression is seen as two lateral stripes along the neural tube (paired white arrowheads) (F). In the posterior half of the spinal cord, at the level of the vitellum, transcripts are observed as separate clusters (white arrowheads), or as a relatively continuous expression domain (black arrowhead) (H). Cranial placodes were identified following descriptions of (Damas, 1944; Kuratani et al., 1997; Modrell et al., 2014) using the pharyngeal arches, otic vesicle, MHB, epiphysis and the transition of the notochord to the precordal plate as anatomical internal references. Abbreviations: fb, forebrain; hb, hindbrain; mb, midbrain.

At stage 25, expression expands in most of the CNS, running the entire length of the spinal cord (Figure 6-18A). In the brain, expression covers most of the diencephalon (except its ventral-most region possibly corresponding to the hypothalamus), midbrain and hindbrain, and transcripts are now present in the telencephalon (Figure 6-18C). Faint expression is also seen in the epiphysis. In the hindbrain, a big patch of expression at the level of rhombomeres 3-5 can be seen on the dorsal side (Figure 6-18C, asterisk). Notably, the MHB remains unstained (Figure 6-18C). Outside the brain, expression increases in the forming olfactory epithelium and trigeminal ganglia (Figure 6-18C). Transcripts are now seen in the eight pharyngeal arches forming seven pharyngeal pouches and in the upper lip (Figure 6-18C). From a ventral view, expression is distinguished in the mesoderm of each pharyngeal arch (Figure 6-18F). From a dorsal view, staining is seen as two lateral stripes along the neural tube as in stage 24 embryos (Figure 6-18G), and, at the level of the vitellum, transcripts are still observed as separate clusters or as small continuous stretches (Figure 6-18J, white and black arrowheads, respectively). Expression under the notochord also progresses posteriorly as the embryo elongates (Figure 6-18A, K, arrows).

Cross sections of stage 25 embryos reveal strong staining in most of the neural tube, but with a thin, unstained domain in the middle (Figure 6-18D, E, H, I, L, M). This pattern is consistent with dorsal views showing expression along the neural tube as two lateral stripes of staining (Figure 6-18G and J). Expression on the lateral sides of the neural tube thus delineate the ventricular zone, conversely marking the marginal (perhaps also subventricular) zone of the brain and spinal cord where differentiated (differentiating for the subventricular zone if present) neurons are located. At the level of the otic vesicle, there is another gap of expression in the middle of the marginal zone with respect to the dorso-ventral axis (Figure 6-18H, white asterisks). Interestingly, there is also a wide area in the dorsal-most part of the brain unstained (Figure 6-18D, E, H, I, black asterisks), whereas in the spinal cord it is mostly seen in the ventral-most part (Figure 6-18B and M, black asterisks). Apart from the brain, expression is mostly seen in mesoderm of the pharyngeal arches, clearly excluding the ectoderm and endoderm (Figure 6-18D, E, H, I). Expression is also seen in mesenchyme surrounding the otic vesicles (Figure 6-18H, white arrows). Given the strong staining it is difficult to unambiguously distinguish internal structures in this area, but expression medial to the otic vesicles could correspond to the intracapsular-vestibuloacoustic ganglion complex (Figure 6-18H, white arrowheads). In the body, transcripts are found in mesenchyme ventro-lateral to the notochord and in an extended domain just above the vitellum (Figure 6-18L, black and white arrows, respectively).

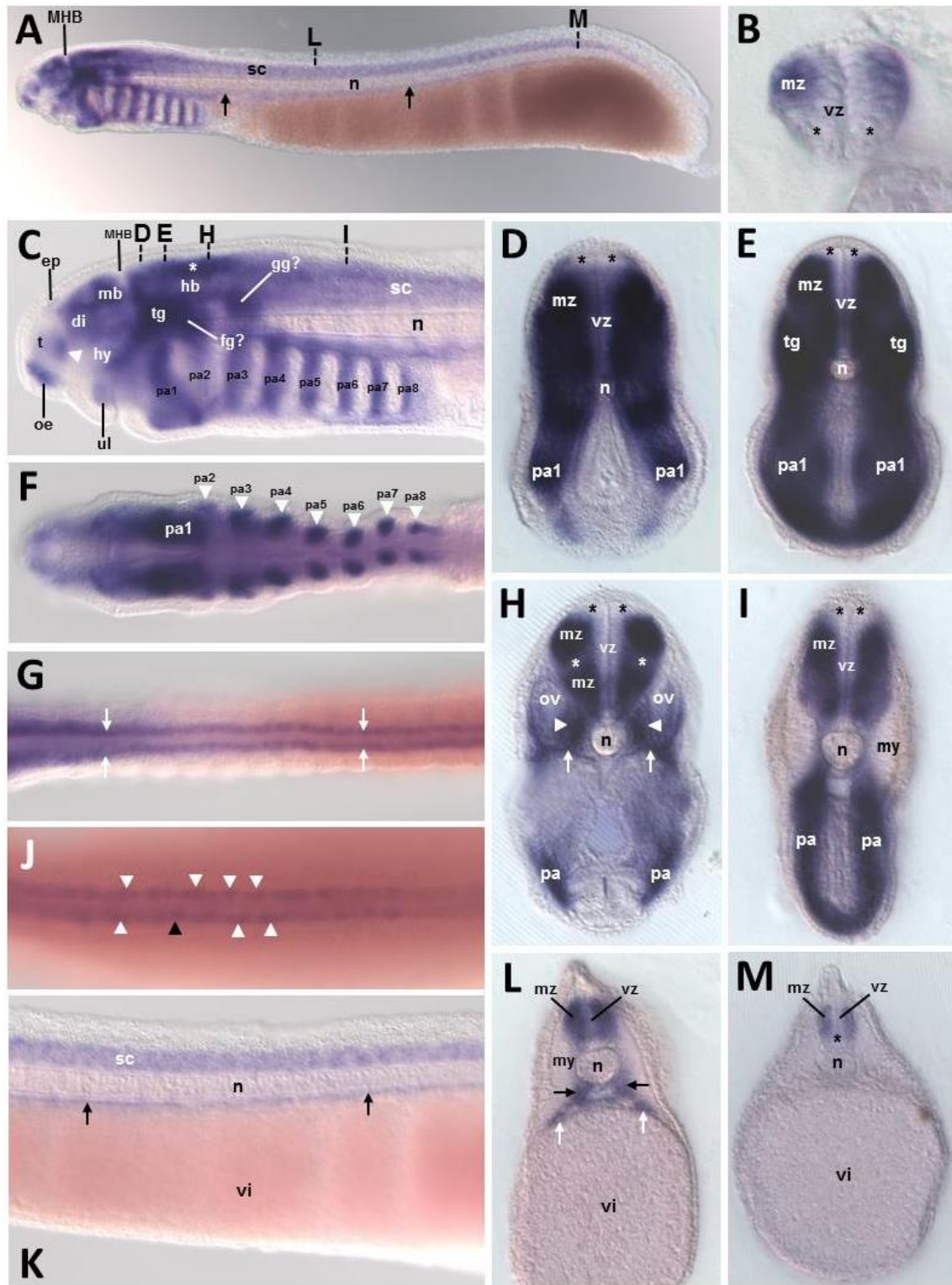


Figure 6-18. *LpCoe-A* expression during *Lampetra planeri* development at stage 25. (A), (C) and (K) are lateral views. (F) is a ventral view of the pharyngeal region. (G) and (J) are dorsal views. Anterior to the left in all these images. (B, D, E, H, I, L and M) are cross-sections at the indicated levels in (A) and (C), dorsal side to the top. Expression dramatically increases in the CNS and transcripts are now seen in the telencephalon (A, C, white arrowhead). A strong patch of staining is observed in the hindbrain (C, white asterisk). Transcripts are observed throughout the spinal cord. From a ventral view, expression is seen in the middle of each pharyngeal arch, likely corresponding to neural crest and/or mesoderm (F). From a dorsal view, two parallel stripes of expression can be seen in the CNS (G, white arrows; J, arrowheads),

which correspond to the marginal zone (possibly also the subventricular zone). At the level of the vitellum transcripts are observed as separate clusters (J, white arrowheads) or as small continuous stretches (J, black arrowhead). Clear staining is also observed in a thin domain under the notochord running along the length of the body (A and K, black arrows). Cross sections reveal strong staining on the sides of the neural tube corresponding to the marginal zone and possibly subventricular zone (B, D, E, H, I, L, M). Transcripts are also observed in trigeminal ganglia (E) and mesenchyme of pharyngeal arches corresponding to mesoderm and/or neural crest (D, E, H, I). (C) Staining is also seen surrounding the otic vesicle (H, white arrows), and some of this expression may correspond to the intracapsular-vestibuloacoustic ganglion complex (H, white arrowheads). In the body, transcripts are found in mesenchyme ventro-lateral to the notochord and in an extended domain just above the vitellum (L, black and white arrows, respectively; compare with K, black arrows). This expression is not observed more posteriorly (M). Abbreviations: di, diencephalon; ep, epiphysis; fg, facial ganglion; gg, glossopharyngeal ganglion; hb, hindbrain; hy, hypothalamus; mb, midbrain; MHB, midbrain-hindbrain boundary; my, myotome; mz, marginal zone; n, notochord; oe, olfactory epithelium; ov, otic vesicle; pa1-8, pharyngeal arch 1-8; sc, spinal cord; t, telencephalon; tg, trigeminal ganglion; ul, upper lip; vi, vitellum; vz, ventricular zone.

At stage 26, *LpCoe-A* expression diminishes both in the brain and spinal cord with respect to stage 25 embryos in all regions (Figure 6-19A-C, E and F). In the head, reduction of expression is notably seen in the ventral midbrain, dorsal diencephalon and olfactory epithelium (Figure 6-19B). A marked expression patch is maintained in the hindbrain (Figure 6-19B, white asterisk). Staining is seen in a forming ninth pharyngeal arch depicting eight pharyngeal pouches (Figure 6-19B). From a dorsal view, it is clearly seen that staining is restricted to the lateral side of the brain and spinal cord delineating the ventricular zone in the middle all along the neural tube (Figure 6-19D and F). In the brain, the ventricular zone expands at the level of the epiphysis, the MHB, and posterior hindbrain (Figure 6-19D, white asterisk). Strong staining is also seen medial to the otic vesicle (which is not stained) that may correspond to the intracapsular-vestibuloacoustic ganglion complex (Figure 6-19D). Additional strong expression is seen in trigeminal ganglia (Figure 6-19D). Notably, expression under the notochord disappears (Figure 6-19C and E, white arrows). At stage 27, expression

diminishes even more with respect to stage 26, becoming more restricted, although it maintains the same pattern in general (Figure 6-19G).

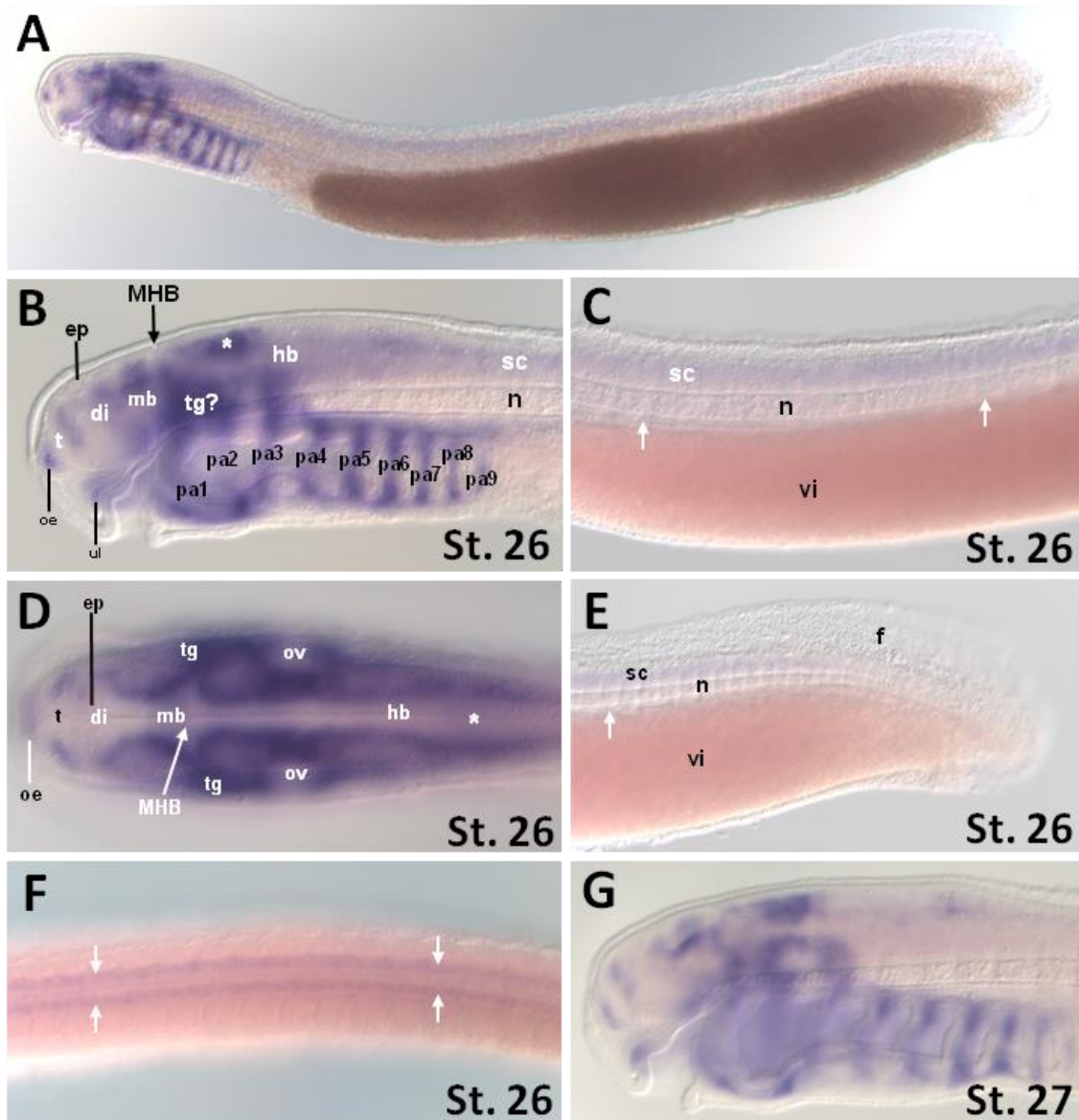


Figure 6-19. *LpCoe-A* expression during *Lampetra planeri* development at stages 26 and 27. (A), (B), (C), (E) and (G) are lateral views. (D) and (F) are dorsal views. Anterior to the left in all images. At stage 26, *LpCoe-A* shows a diminished, more restricted expression in the head and spinal cord when compared to stage 25 (A-C, E, F), although it preserves the same pattern in general (asterisk in B marks a strong patch of staining in the hindbrain). Dorsal views show a conserved expression pattern as two lateral stripes of staining in the neural tube, marking the marginal zone (D, F). Strong staining medial to the otic vesicles may correspond to the intracapsular-vestibuloacoustic ganglion complex (D). Expression in the spinal cord is notably weaker with respect to stage 25 (compare Figure 6-18G and J with Figure 6-19F, white arrows) and expression under the notochord disappears (C and E, white arrows). At stage 27, expression is even more reduced than stage 26 embryos although it preserves the same general expression

patterns (G). Abbreviations: di, diencephalon; ep, epiphysis; f, fin; hb, hindbrain; mb, midbrain; MHB, midbrain-hindbrain boundary; n, notochord; oe, olfactory epithelium; ov, otic vesicle; pa1-9, pharyngeal arch 1-9; sc, spinal cord; t, telencephalon; tg, trigeminal ganglion; ul, upper lip; vi, vitellum;.

At stage 28, expression is slightly more reduced in the head, but with a finer and clearer expression pattern in the same regions as in previous stages (Figure 6-20A). Expression in the brain is also more restricted than at stages 26 and 27, particularly in the forebrain (Figure 6-20B, C, E). Notably, expression in the spinal cord is virtually absent (Figure 6-20A, D, F). Expression in the olfactory epithelium and pharyngeal arches is maintained (Figure 6-20C).

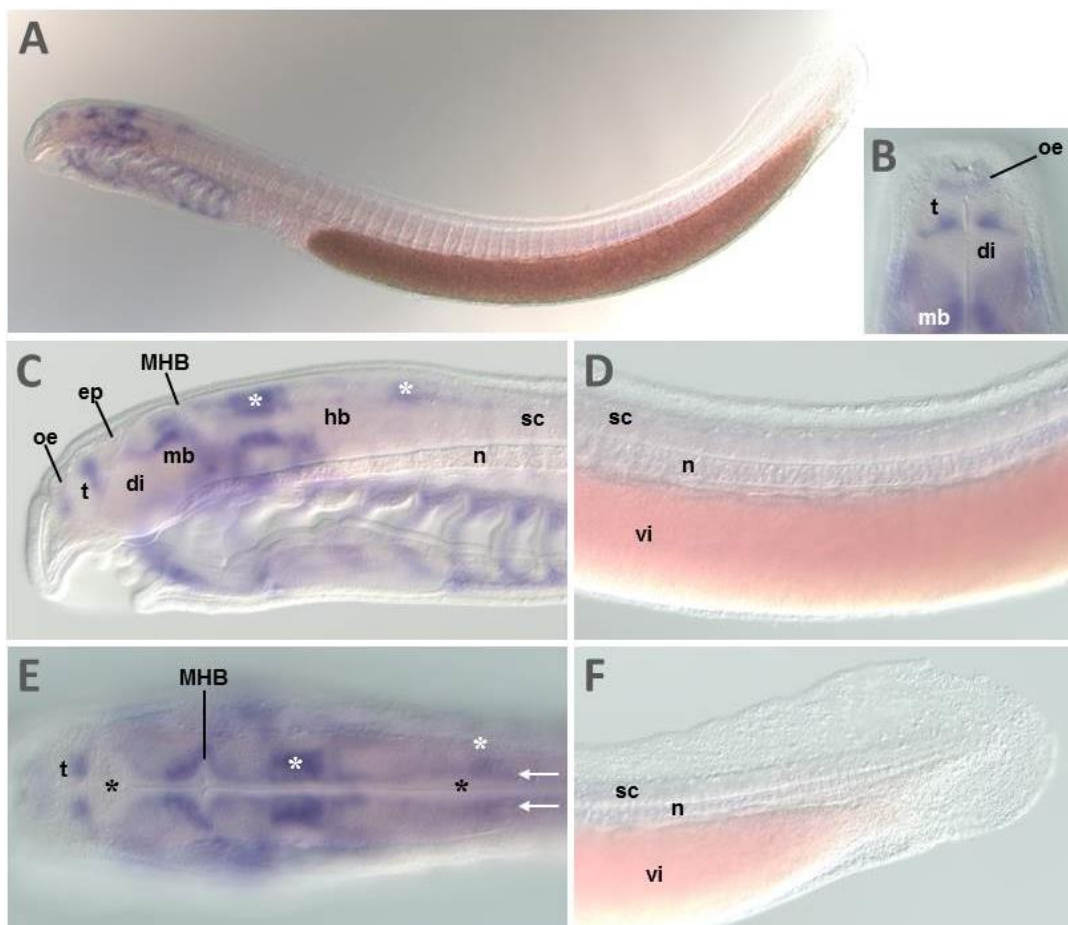


Figure 6-20. *LpCoe-A* expression during *Lampetra planeri* development at stage 28. (A), (C), (D) and (F) are lateral views. (B) and (E) are dorsal views. In all images anterior is to the left except in (B) in which anterior is to the top. At this stage expression decreases even more with respect to stage 27 and it becomes more restricted in the head, although generally conserving the same expression sites (A, C). In the hindbrain, two main patches of expression are observed dorsally (white asterisks in C and E). The ventricular zone in the brain has considerably narrowed but expansions are still observed at the level of the epiphysis, MHB, and posterior hindbrain (E, black asterisks). In the spinal cord, expression is virtually absent (D, F). Expression in the telencephalon and diencephalon is observed posteriorly (B). Abbreviations: di, diencephalon; ep, epiphysis; hb, hindbrain; mb, midbrain; MHB, midbrain-hindbrain boundary; n, notochord; oe, olfactory epithelium; sc, spinal cord; t, telencephalon; vitellum.

6.5.4. *LpCoe-B* Expression in Developing *L. planeri* Embryos

LpCoe-B (*LpCoe-B_1*) expression was analysed from stage 21 to stage 28 embryos. At stage 21, mild expression is observed in the anterior spinal cord as distinct, widely-spaced dots forming two lateral stripes as seen from a dorsal view (Figure 6-21A and B). Later at stage 22, this expression pattern increases in intensity and extends posteriorly along the growing spinal cord (Figure 6-21C and D). Expression is still seen as widely-spaced dots forming two parallel stripes. At this stage expression is also observed in the trigeminal placodes and presumptive NHP (Figure 6-21C, black and white arrowheads, respectively). Cranial placodes and ganglia at this and subsequent stages were identified following descriptions of (Damas, 1944; Kuratani et al., 1997; Modrell et al., 2014). Anatomical internal references to aid in their identification were pharyngeal arches, otic vesicle, MHB, epiphysis and the transition of the notochord to the precordal plate. At stage 23, expression is maintained in the forming maxillomandibular ganglion and it now appears in the developing facial and glossopharyngeal ganglia (Figure 6-21E). At this stage, expression persists in the spinal cord as two lateral stripes with a more homogenous (continuous) pattern with respect to earlier stages that present a dotted (more separated) pattern.

At stage 24 *LpCoe-B* is upregulated in the brain, particularly in the midbrain, hindbrain, and in a small region in the diencephalon (Figure 6-21F, G). Notably, expression is not seen in the telencephalon at this stage. Expression intensifies in the NHP, ophthalmic, maxillomandibular and facial ganglia, the latter possibly corresponding to the facial-anterior lateral line ganglion complex (Figure 6-21F, G). Transcripts are now seen in what seems to correspond to the posterior lateral line placode/ganglion (Figure 6-21G). Expression in the spinal cord is considerably upregulated with respect to earlier stages (Figure 6-21G), preserving a pattern of two parallel lateral stripes of expression as seen from a dorsal view (Figure 6-21H). In addition, all expression in the CNS presents an overall bias towards the ventral half of the neural tube (Figure 6-21G). In slightly older embryos, however, expression in the hindbrain and spinal cord starts to be restricted dorsally (Figure 6-21F). Outside the nervous system, expression is also seen in mesenchyme of the first pharyngeal arch and in an extending domain between the notochord and the pharynx (Figure 6-21F, G).

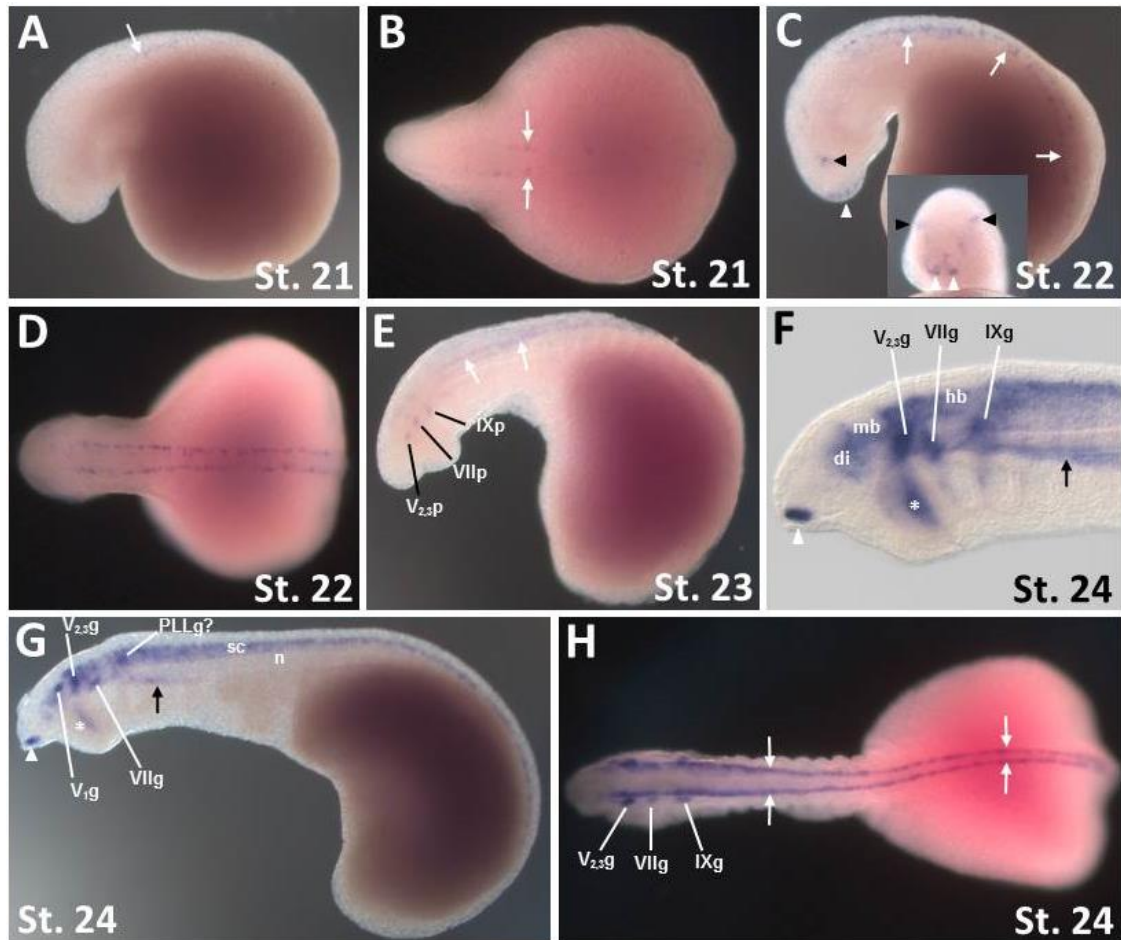


Figure 6-21. *LpCoe-B* expression during *Lampetra planeri* development at stages 21-24. (A, C, E, F and G) are lateral views. (B, D and H) are dorsal views. In all images anterior is to the left. (A, B) At stage 21, *LpCoe-B* is observed in the anterior spinal cord as two lateral stripes (arrows). (C, D) At stage 22, expression extends posteriorly along the growing spinal cord (C, arrows), which is seen as two parallel stripes from a dorsal view (D). Transcripts are also observed in the trigeminal placodes (black arrowheads) and the NHP anlage (white arrowheads). Inset, frontal view of the embryo shown in (C). (E) At stage 23, expression is maintained in the spinal cord (arrows) and trigeminal placodes, and it now appears in the developing facial and glossopharyngeal placodes. (F-H) At stage 24 *LpCoe-B* appears in the hindbrain, midbrain and in a small region in the diencephalon (F). Expression considerably intensifies in the ophthalmic, maxillomandibular, facial and glossopharyngeal ganglia, as well as in the NHP (white arrowheads) and spinal cord (F, G). Expression in the facial ganglion may comprise the facial-anterior lateral line ganglion complex. Transcripts are also observed in what seems to correspond to the posterior lateral line ganglion (G). Outside the nervous system, strong expression is seen in mesenchyme of the first pharyngeal arch (asterisks) and in an extending domain between the pharynx and notochord (black arrows). Expression in the neural tube is localised to the lateral sides as in previous stages (H, white arrows). Cranial placodes were identified following descriptions of (Damas, 1944; Kuratani et al., 1997; Modrell et al., 2014) using the pharyngeal arches, otic vesicle, MHB, epiphysis and the transition of the notochord to the precordial plate as anatomical internal references. Abbreviations: di, diencephalon; hb, hindbrain; mb, midbrain; n, notochord; PLLg, posterior lateral line ganglion; sc, spinal cord; V₁g, ophthalmic ganglion; V_{2,3}g, maxillomandibular ganglion; VIIg, facial ganglion; IXg, glossopharyngeal ganglion; V_{2,3}p, maxillomandibular placode; VIIp, facial placode; IXp, glossopharyngeal placode.

At stage 25, *LpCoe-B* is strongly expressed in the midbrain and hindbrain and posterior-most diencephalon, as well as in ophthalmic, maxillomandibular, facial and glossopharyngeal ganglia and in the forming olfactory epithelium (Figure 6-22A-C). In the hindbrain, expression starts to segregate to the anterior- and posterior-most regions, leaving a gap of mild expression in the middle (Figure 6-22B, C). Notably, *LpCoe-B* is not expressed in the telencephalon at this stage. Expression in the spinal cord becomes very restricted to the dorsal side (Figure 6-22D, white arrow) and reaches the end of the spinal cord where it covers the entire dorsal-ventral axis (Figure 6-22E, white arrow). Expression is also seen in the anterior-most five pharyngeal arches, and most prominently in the mandibular arch (Figure 6-22B, C). Like *LpCoe-A*, an expression domain below the notochord is also appreciated, although it does not go as far to the posterior (Figure 6-22B-D, black arrows). From a dorsal view, staining is also restricted to the lateral sides of the neural tube as with *LpCoe-A*, delimiting the ventricular zone (Figure 6-22F).

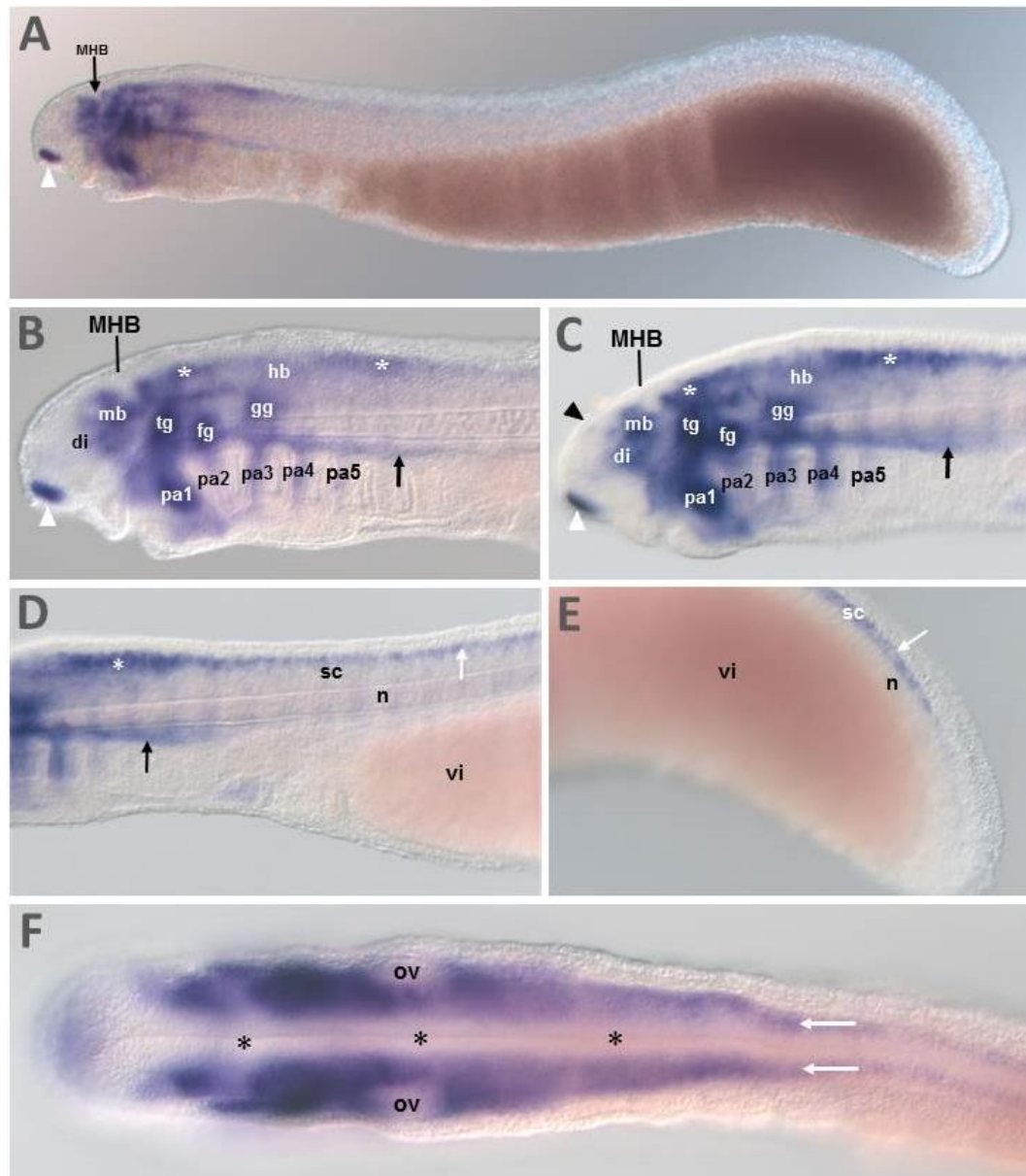


Figure 6-22. *LpCoe-B* expression during *Lampetra planeri* development at stage 25. (A-E) Lateral views. (F) Dorsal view. Anterior is to the left in all images. (A) *LpCoe-B* expression is localised in the head and spinal cord. (B) In the brain, transcripts are observed in most of the midbrain, in restricted domains of the hindbrain mostly towards the dorsal side, and possibly in the posterior-most diencephalon, but transcripts are not observed in the telencephalon. Note how there is a gap of expression at the posterior-most side of the brain. Two main patches of expression are seen dorsally in the anterior and posterior hindbrain (white asterisks in B and C). In the PNS transcripts are seen in the developing olfactory epithelium (white arrowheads in A-C) and in the trigeminal maxillomandibular, facial and glossopharyngeal ganglia. Outside the nervous system expression is seen in the first five pharyngeal arches and in an extending domain just above the pharynx (black arrows in B-D). (C) In a slightly older stage 25 embryo, expression in the head starts to increase, especially in cranial ganglia. (D) In the spinal cord, expression is restricted to the dorsal side (white arrow) and is continuous with the big patch of expression in the posterior hindbrain (white asterisk). (E) Expression in the spinal cord reaches the posterior end covering the entire dorso-ventral axis (white arrow). (F) From a dorsal view, signals are detected at the lateral margins of the neural tube corresponding to the marginal zone (white arrows), leaving an unstained region in the middle which corresponds to the ventricular zone (black asterisks). Abbreviations: di, diencephalon; fg, facial ganglion; gg, glossopharyngeal ganglion; hb, hindbrain; mb, midbrain; MHB, midbrain-hindbrain

boundary; n, notochord; ov, otic vesicle; pa1-5, pharyngeal arch 1-5; sc, spinal cord; tg, trigeminal maxillomandibular ganglion; vi, vitellum.

At stage 26, expression in the head dramatically increases in the head, particularly at the level of the mandibular arch, and expression in the spinal cord is preserved (Figure 6-23A). Expression sites in the head are more defined, and transcripts are now detected in the telencephalon and epiphysis (Figure 6-23C). In the midbrain, two large expression domains, one dorsal and one ventral, are separated by an unstained region (Figure 6-23C). In the hindbrain, a big patch of staining is seen on the dorsal side in rhombomeres 3-5. Similarly, another patch of staining is seen at the posterior hindbrain (Figure 6-23C). Expression in the olfactory epithelium remains strong and signals are now detected in the upper and lower lips (Figure 6-23C). Eight pharyngeal arches are now stained, all along their dorsoventral axis covered by a dorsal and a ventral expression domain (Figure 6-23C). The spinal cord maintains its dorsal expression domain that reaches the end of the tail (Figure 6-23D and L). From a dorsal view, staining is still observed at the lateral margins of the neural tube as two lateral stripes, demarcating the ventricular zone in the middle (Figure 6-23B and H). As with *LpCoe-A*, in the head, the two bands of *LpCoe-B* expression are not straight, showing expansions at the level of the MHB and posterior hindbrain (Figure 6-23B).

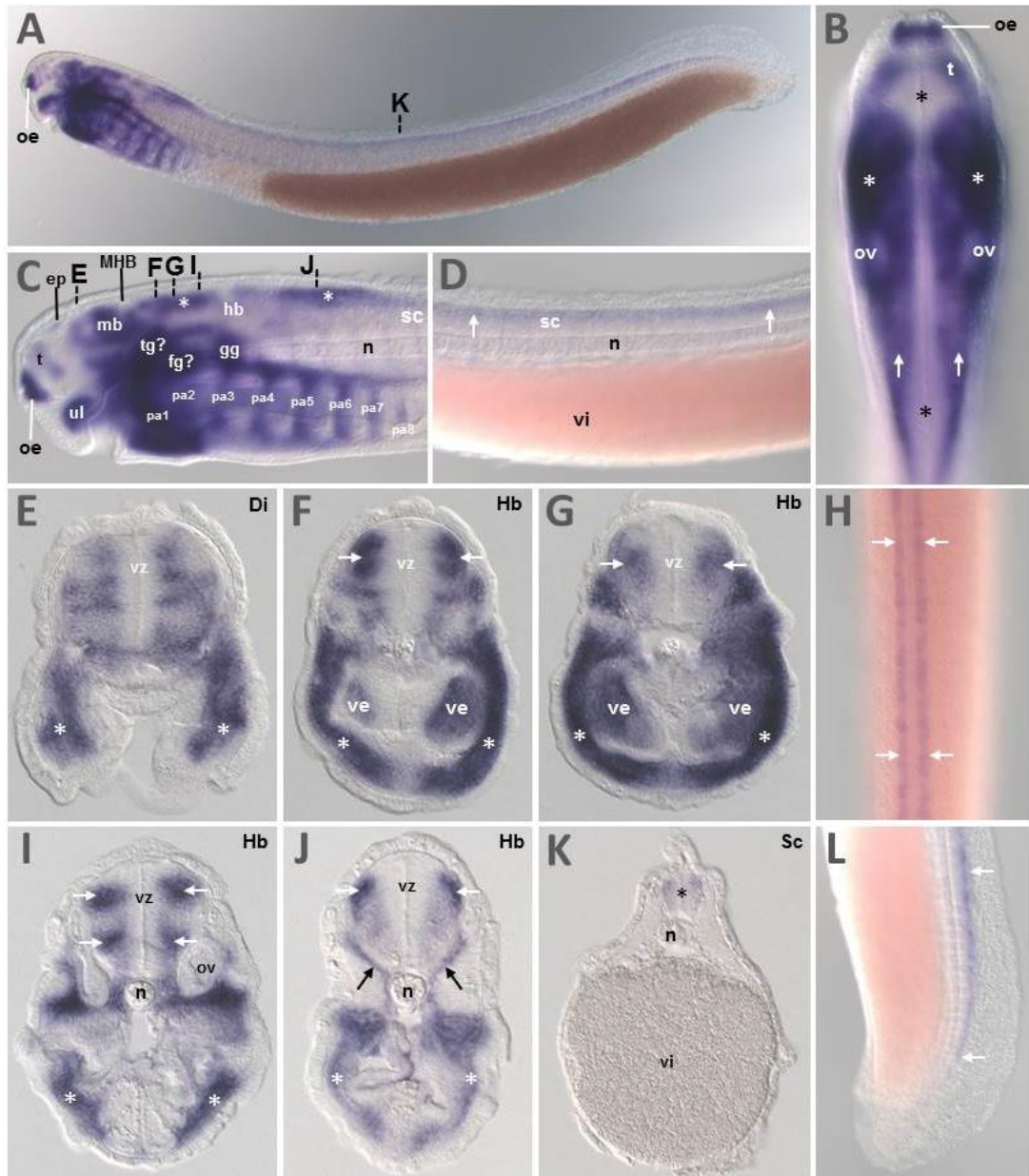


Figure 6-23. *LpCoe-B* expression during *Lampetra planeri* development at stage 26. (A), (C), (D) and (L) Lateral views. (B) and (H) Dorsal view, anterior to the top. In (A), (C) and (D), anterior is to the left. In (L) anterior is to the top. (E-G) and (I-K) are cross sections at the indicated levels in (A) and (C), dorsal to the top. (A) *LpCoe-B* expression dramatically increases in the head and is maintained in the spinal cord. (B) From a dorsal view, expression is observed on the lateral sides of the neural tube, corresponding to the marginal zone (white arrows) and in turn delimiting the extent of the ventricular zone, which presents expansions at the level of the epiphysis and posterior hindbrain (black asterisks). Very strong expression at the level of the trigeminal/facial ganglia precludes identifying structures with certainty (white asterisks). (C) In the brain, transcripts are observed in the telencephalon, posterior diencephalon, in dorsal and ventral midbrain, and in restricted domains of the hindbrain. Two patches of expression are seen dorsally in the anterior and posterior hindbrain (white asterisks). Transcripts are also seen in the olfactory epithelium, in the maxillomandibular trigeminal, facial and glossopharyngeal ganglia, in the upper lip and pharyngeal arches. (D) In the spinal cord, expression is clearly restricted to the dorsal side (white arrows). From a dorsal view, expression is seen at the lateral margins of the spinal cord corresponding to the marginal zone (H, white arrows), leaving a space in the middle which corresponds to the ventricular zone. This expression is continuous with brain expression. Expression in the neural tube continues to the end of

the spinal cord (L, white arrows). (E-G, I, J) Cross sections at different levels of the head reveal a dynamic expression in the marginal zone of different brain compartments (indicated at the upper right corner) both in intensity and distribution along the dorso-ventral axis. Expression is also observed in the mesodermal and/or neural crest component of pharyngeal arches (white asterisks), and is clearly excluded from the ectoderm and endoderm. Abbreviations: Di, diencephalon; ep, epiphysis; fg, facial ganglion; gg, glossopharyngeal ganglion; hb, hindbrain; mb, midbrain; MHB, midbrain-hindbrain boundary; n, notochord; oe, olfactory epithelium; ov, otic vesicle; pa1-8, pharyngeal arch 1-8; sc, spinal cord; t, telencephalon; tg, trigeminal ganglion; ul, upper lip; ve, vellum; vi, vitellum; ventricular zone.

At stage 27, expression clearly diminishes, in particular in the spinal cord where it is virtually absent (Figure 6-24A, F and G). Expression in the diencephalon also disappears, leaving only the telencephalon and midbrain stained (Figure 6-24B). The two big patches of staining in the dorsal hindbrain persist although at a much lower level. Expression at the level of the mandibular arch and upper lip is also maintained at a lower level. The pharyngeal arches show only mild staining, the ventral expression vanishes and the dorsal domain remains somewhat strong (Figure 6-24B). From a dorsal view, expression in the brain is fairly similar to stage 26 embryos, demarcating the ventricular zone, although staining is weaker (Figure 6-24D). At stage 28, expression diminishes even more in the head but the same expression sites are maintained (Figure 6-24C and E).

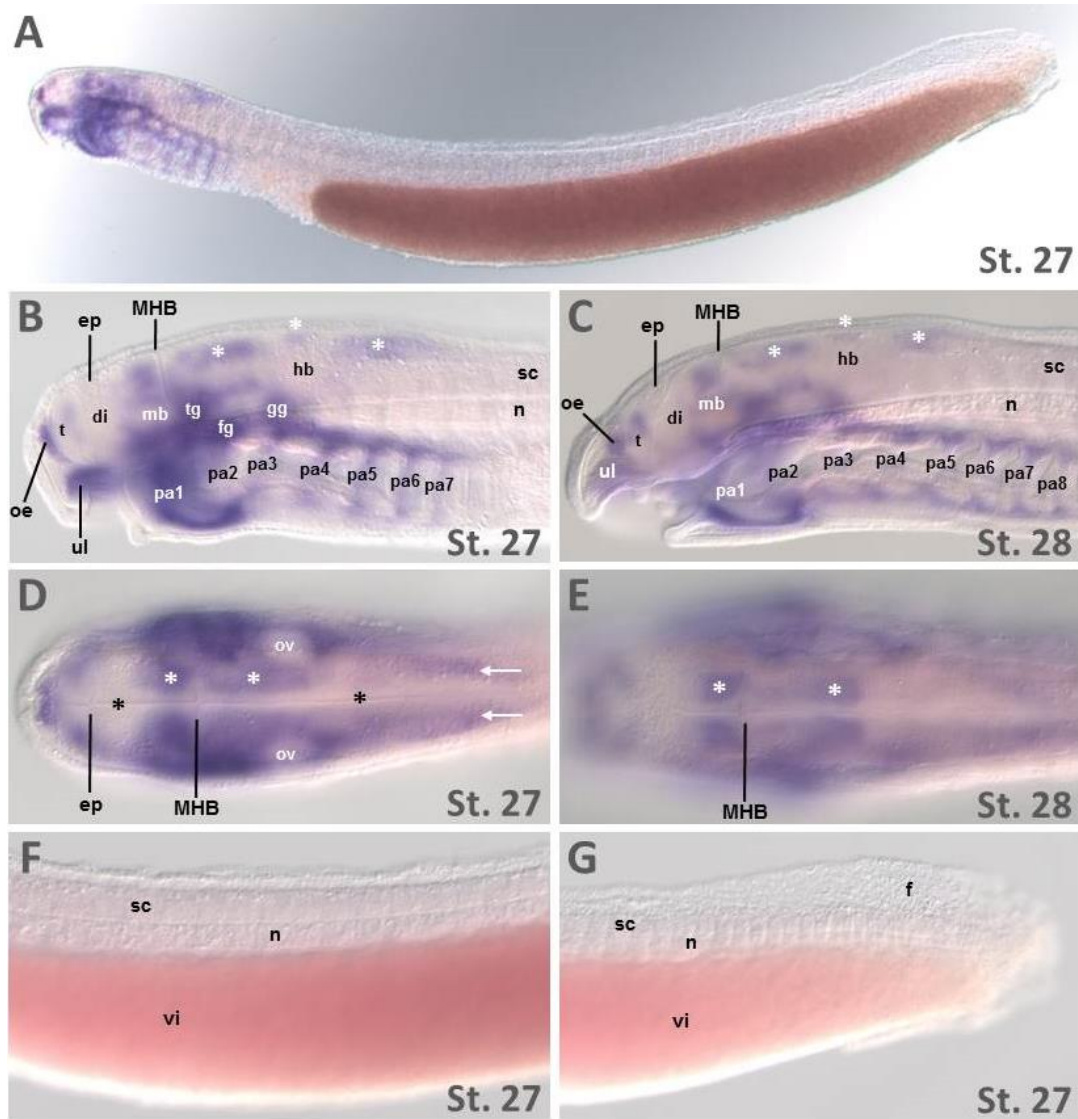


Figure 6-24. *LpCoe-B* expression during *Lampetra planeri* development at stage 27 and 28. (A-C) and (F, G) are lateral views. (D) and (E) are dorsal views. Anterior is to the left in all images. Stages are as indicated. At stage 27, expression in the spinal cord disappears and that in the head diminishes (A), (F), (G). In the brain, expression is almost absent in the diencephalon while other expression sites are maintained with respect to stage 26 embryos although at a much lower level (B). Three distinct patches of expression are observed in the dorsal hindbrain (B, white asterisks). Dorsal views show expression in the midbrain and hindbrain (D, white asterisks and white arrows) demarcating the ventricular zone in the middle; the ventricular zone expands at the level of the epiphysis, MHB and posterior hindbrain (black asterisks). Similar expression sites are observed at stage 28 (C, E) although with weaker signals as compared to previous stages. Labels in (C) and (E) as in (B) and (D), respectively. Abbreviations: di, diencephalon; ep, epiphysis; f, fin; fg, facial ganglion; gg, glossopharyngeal ganglion; hb, hindbrain; mb, midbrain; MHB, midbrain-hindbrain boundary; n, notochord; oe, olfactory epithelium; ov, otic vesicle; pa1-8, pharyngeal arch 1-8; sc, spinal cord; t, telencephalon; tg, trigeminal ganglion; ul, upper lip; vi, vitellum.

6.5.5. *COE* Genes in Vertebrate Lymphocytes

Another evolutionary interesting aspect of *COE* genes comes from the expression of mouse *COE1* early in B-lymphocyte development, where it has an important role in the development of this cell lineage by inducing progenitor cells to differentiate into pro-B cells (Hagman and Lukin, 2005; Lukin et al., 2008). Mouse *COE1* activates genes necessary for immunoglobulin V(D)J domains rearrangements (*RAG-1* and *RAG-2*), for B-cell lineage commitment (*CD19*), for the assembly of B-cell receptors ($\lambda 5$ and *VpreB1*), and for immunoglobulin subunits (*mb-1* and *B29*) (Hagman and Lukin, 2005; Lukin et al., 2008) (Figure 6-25). *COE1* participates in a global network of transcription factors including E2A and Foxo1 that regulates B cell fate (Lin et al., 2010). In mice deficient of *COE1*, B-lymphocytes arrest at a stage resembling early B-lymphocyte progenitors, whereas T-cell development proceeds normally (Lin and Grosschedl, 1995). Conversely, forced expression of *COE1* restricts lymphopoiesis to the B cell lineage at the expense of other lymphoid-derived lineages such as NK or dendritic cells (Zhang et al., 2003). Interestingly, skate (*Raja eglanteria*) *COE* genes are expressed in the spleen, a site of B-lymphocyte production (Anderson et al., 2004).

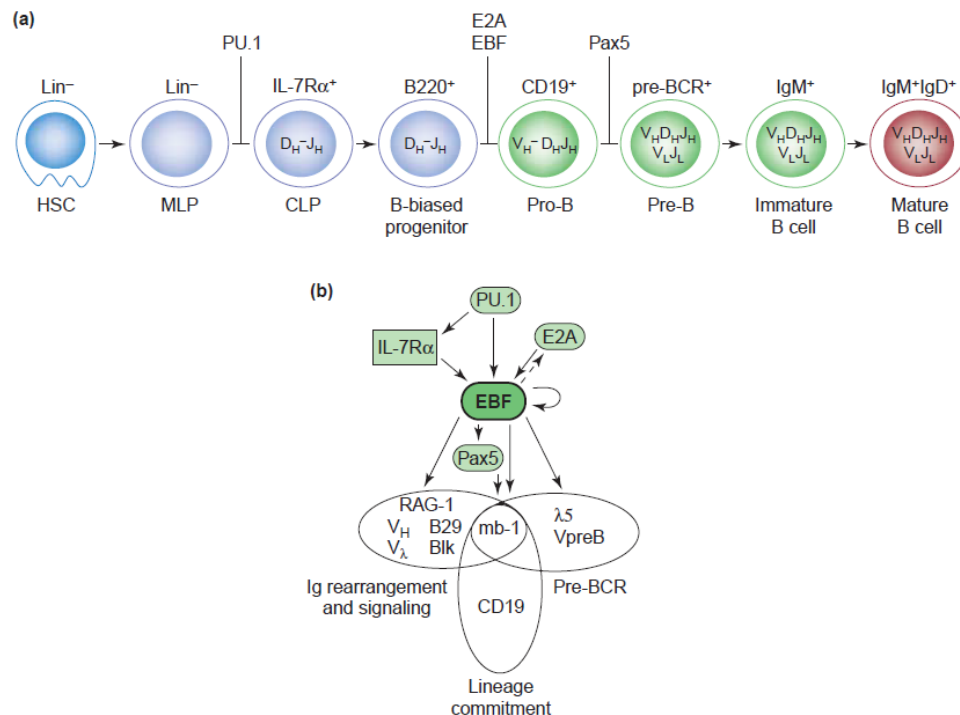


Figure 6-25. B-cell development and the roles of EBF/COE1. (a) Progressive stages of B-cell lymphopoiesis are shown with (i) cell designations indicated below, (ii) the status of V(D)J recombination indicated within each cell type and (iii) characteristic cell surface markers indicated above each cell. The approximate points at which B-cell lymphopoiesis is arrested in PU.1^{-/-}, E2A^{-/-}, EBF^{-/-} and Pax5^{-/-} mice are indicated above the cells. Cells lacking EBF or E2A closely resemble B-biased progenitor cells but lack all Ig gene rearrangements. (b) Central role of EBF in B-cell development. Regulatory circuits in developing B cells are shown to highlight the functions of EBF. Transcription of the *ebf1* gene is a function of the transcription factors PU.1 and E2A and is reinforced by IL-7R signaling. EBF might autoregulate its own expression. EBF regulates genes, including *pax5*, *RAG-1*, $\lambda 5$ and *VpreB1*, *B29*, B-lymphoid tyrosine kinase (*blk*) and *CD19*. Many EBF targets have been confirmed as targets of E2A, which functions in synergy with EBF. A role for EBF in the regulation of E2A expression in the B-cell compartment (broken line) is suggested by the reduction of E2A in the absence of EBF. Pax5, in turn, regulates *mb-1* and *CD19* synergistically with EBF. Pax5 regulates many other genes required for the B-lineage programme and commitment. Abbreviations: Lin⁻, lineage-specific markers. Taken from Hagman and Lukin (2005).

Lymphocyte-like cells have been isolated and characterised in lampreys (Mayer et al., 2002a) (Figure 6-26, left), and they seem to express genes that are homologous to mammalian genes involved in the immunological response (Uinuk-ool et al., 2002). Surveys from larval *Petromyzon marinus* lymphocyte-like cells isolated from the hematopoietic organs and blood have revealed homologs of genes involved in vertebrate lymphocyte development, differentiation, proliferation, signal transduction, activation,

chemotaxis and innate immune responses (Kuroda et al., 2003; Mayer et al., 2002a, 2002b; Pancer et al., 2004a; Sato et al., 2003; Shintani et al., 2000; Uinuk-ool et al., 2002). Importantly, some lamprey lymphocyte-like cells express T cell receptor (TCR) and CD4 coreceptor (Pancer et al., 2004a; Yu et al., 2009).

Although no immunoglobulin or MHC genes have been identified in the lamprey lymphocyte-like cells (Mayer et al., 2002a), they contain a novel type of antigen receptors, the variable lymphocyte receptors (VLRs), composed of highly diverse leucine-rich repeats (LRR) and that also present somatic recombination akin to immunoglobulins (Alder et al., 2005; Cooper and Alder, 2006; Pancer and Cooper, 2006; Pancer et al., 2004b). These receptors are expressed in a monoclonal fashion (Nagawa et al., 2007; Pancer et al., 2004b). However, no *RAG1* or *RAG2* recombination activation genes, used for *TCR* and *BCR* gene assembly, have been found in lampreys, and instead *VLR* gene assembly might occur by gene conversion (Alder et al., 2005; Guo et al., 2009; Rogozin et al., 2007) or copy choice (Nagawa et al., 2007).

Three *VLRs* (*VLRA*, *VLRB* and *VLRC*) have been described both in lampreys and hagfish (Alder et al., 2005; Das et al., 2013; Kasamatsu et al., 2010; Kim et al., 2006; Li et al., 2013; Pancer et al., 2004b, 2005). Lamprey *VLRA* is expressed in T cell-like lymphocytes, whereas *VLRB* are expressed by B cell-like lymphocytes, and the immune functions of lamprey T cell-like and B cell-like lymphocytes also resemble those of mammalian T and B cells, respectively (Alder et al., 2008; Guo et al., 2009). Remarkably, lamprey B cell-like cells expressing specific *VLRB* molecules undergo lymphoblastoid transformation upon antigen stimulation, expand clonally, and

differentiate into plasmacytes that secrete *VLRB* molecules in a manner analogous to the secretion of immunoglobulins by mammalian B cells (Alder et al., 2008; Kasamatsu et al., 2010). On the other hand, T cell-like lymphocytes also proliferate in response to antigen stimulation, but do not differentiate into antibody-secreting cells and instead maintain cell surface expression of their receptors while increasing the expression of proinflammatory cytokines, macrophage migration inhibitory factor (MIF) and interleukin-17 (IL-17). Yet a different lymphocyte population specifically expresses *VLRC*, and its characterisation has led to the hypothesis that *VLRA*⁺ lymphocytes are equivalent to mammalian *TCRαβ*⁺ T lymphocytes, whereas *VLRC*⁺ lymphocytes are equivalent to mammalian *TCRγδ*⁺ T lymphocytes (Hirano et al., 2013; Kasamatsu et al., 2010). The primordial vertebrate lymphocyte repertoire might thus have consisted of three different lymphocyte lineages, although both lamprey *VLRA*⁺ and *VLRC*⁺ lymphocytes are predicted to come from a single precursor as mammalian *TCRαβ*⁺ and *TCRγδ*⁺ T lymphocytes (Hirano et al., 2013). Therefore, it appears that lamprey lymphocyte-like cells have the same immunological functions as mammalian lymphocytes, although with a different molecular machinery.

The lamprey lymphocyte-like cells are mainly present in the gills, opisthonephros, typhlosole and blood of the ammocoetes (Amemiya et al., 2007; Fujii, 1982; Mayer et al., 2002a; Tanaka et al., 1981). The typhlosole is considered the major hematopoietic tissue in larvae (Alder et al., 2005), and there is also abundant expression of *VLRB* (thus marking B cell-like lymphocytes) in the typhlosole (Bajoghli et al., 2011), making it a good candidate to search for *COE* gene expression considering that *COE1/Ebf1* is expressed in mammalian B lymphocytes. The typhlosole is a single longitudinal in-fold

of the ventral part of the intestine where haematopoiesis occurs in a cellular nest-like structure (Tanaka et al., 1981) (Figure 6-26, middle, right). In the middle of the typhlosole, the typhlosolar artery runs along its entire length, which provides the blood sinuses that contain mature and immature blood cells (Amemiya et al., 2007; Ardavín and Zapata, 1987). In this region, haematopoietic cells of the erythroid, lymphoid, monocytic and granulocytic lineages differentiate (Ardavín and Zapata, 1987; Fujii, 1982). During metamorphosis, the typhlosole loses its lympho-hemopoietic capacity and is mainly composed of connective tissue (Amemiya et al., 2007; Ardavín and Zapata, 1987).

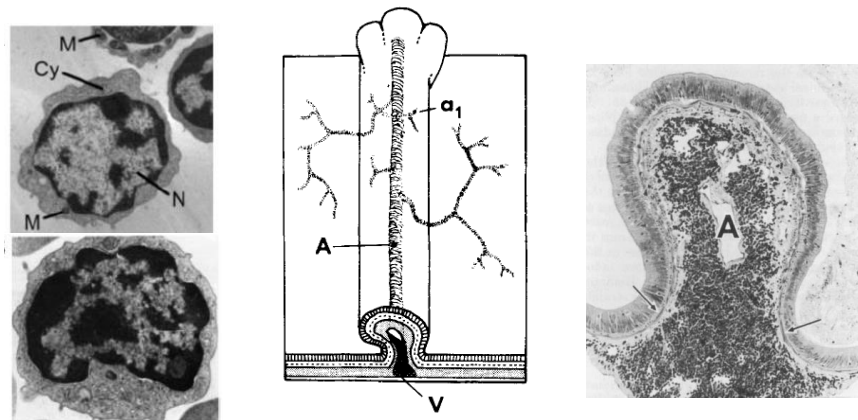


Figure 6-26. Morphology of the lymphocyte-like cells and their production in the typhlosole. Left panel: transmission electron microscopy of lamprey lymphocyte-like cells (top) and mouse intestinal epithelial lymphocytes (bottom). Both cell types display the similar morphology of having large nuclei surrounded by a thin layer of cytoplasm. Abbreviations: M, mitochondria; Cy, cytoplasm; N, nucleus Middle panel: schematic of typhlosole anatomy in the interior of the ammocoete intestine. The typhlosole constitutes a single longitudinal in-fold of the intestine occupied by a lymphohemopoietic tissue, irrigated by a system of enlarged blood sinuses arising from the typhlosolar artery that runs along the interior of the typhlosole. Note that the typhlosole becomes lobulated towards the anterior region (top). Abbreviations: A, typhlosolar artery; V, hematopoietic area. a_1 , arteriole. Right panel: cross section of the intestine showing the developing blood cells occupying most of the space of the typhlosole. The typhlosolar artery can be seen in the middle (A). Left panel, taken from Mayer *et al.*, 2002; middle and right panels, taken from Tanaka *et al.*, 1981.

In order to test whether the isolated *Lampetra planeri* COE fragments are expressed in lymphocyte-like cells, RT-PCR was performed with total RNA from ammocoete intestine using the same pair of primers employed to isolate each *L. planeri* COE fragment (i.e. *LpCoe-A* and *LpCoe-B*). Animals were anaesthetised in 0.1% tricaine methane sulfonate in river water. Once anaesthetised, animals were decapitated with a clean scalpel, the intestine dissected and placed in Trizol reagent. The lamprey gene *Spi*, a member of the purine box 1 (*PU.1*)/spleen focus-forming virus integration B (*Spi-B*) gene family (Shintani et al., 2000) and known to be expressed in lymphocyte-like cells, was used as a positive control. Also, cDNA derived from stage 24-26 embryos was used as a positive control for expression of *LpCoe-A* and *LpCoe-B*.

PmSpi expression in the intestine RNA could be detected, however, no expression of either *LpCoe-A* or *LpCoe-B* was observed in the ammocoete intestine cDNA pool, except for a small fragment that needs further identification, but its small size which does not correspond to the expected size of the two *L. planeri* COE genes suggests that it is an artefact (Figure 6-27).

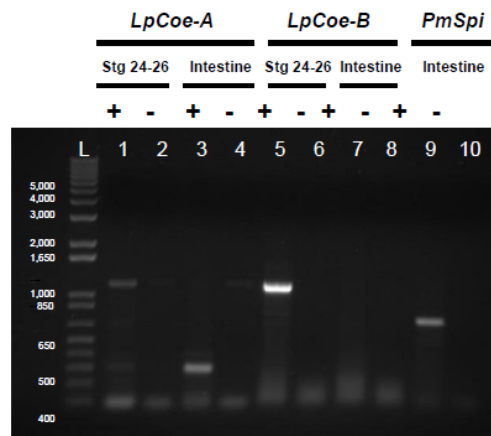


Figure 6-27. Analysis of *Lampetra planeri* *COE* genes expression in the ammocoete intestine. RT-PCR of *LpCoe-A* and *LpCoe-B* with total RNA isolated from ammocoete intestine. Neither *LpCoe-A* nor *LpCoe-B* are detected in the ammocoete intestine in these experiments, whereas the lymphocyte positive control gene *PmSpi* is detected at the expected size (721 bp). cDNA derived from stage 24-26 embryos was used as a positive control for *LpCoe-A* and *LpCoe-B* expression. Negative control was no cDNA added, indicated with a (-). The first lane on each gel corresponds to the ladder (L), molecular weights in base-pairs.

These results suggest that neither of the two lamprey *COE* genes is expressed in the intestine/typhlosole (and hence in the lymphocyte-like cells in this structure), or that their transcript levels are very low to be detected with this method. However, other sites of lymphopoiesis such as branchial arches, intertubular spaces of the larval opisthonephros, fat cells of the dorsal nephric fold and the supraneural body (also known as dorsal fat body or protovertebral arch) or blood (Amemiya et al., 2007; Kilarski and Plytycz, 1981; Zapata et al., 1981) remain to be examined for *COE/EBF* gene expression.

6.6. Discussion

6.6.1. *LpCoe-A* and *LpCoe-B* Sequence and Gene Structure

The phylogenetic position of *LpCoe-A* and *LpCoe-B* is difficult to establish with certainty with respect to other vertebrate *COE/Ebf* genes, which very likely is due to the well conserved amino acid sequences employed in the phylogenetic analyses. Longer sequences containing regions outside the conserved domains are certainly necessary to clarify these issues. However, considering the conserved amino acid sequences and domains, it can be concluded with certainty that both *LpCoe-A* and *LpCoe-B* are true members of the *COE/Ebf* gene subfamily.

The predicted amino acid sequences of both *LpCoe-A* and *LpCoe-B* obtained from this study are very conserved not only with other vertebrate *COE/Ebf* genes but also with other metazoan sequences in basal groups such as cnidarians. As mentioned before, vertebrate *COE/Ebf* genes contain a duplicated helix 2 which, in principle, suggests different and novel interactions with other proteins with respect to invertebrate *COE/Ebf* proteins that may be specific to vertebrates. Furthermore, the existence of alternative splicing of the HLH domain at least in lamprey *COE-B* genes suggests particular interactions of lamprey *COE-B* proteins that may be different from other vertebrate *COE/Ebf* proteins. Daburon and colleagues (Daburon et al., 2008) recently observed the same situation, and further identified two different EST sequences of the orthologous gene in *Lampetra fluviatilis* (Figure 6-4). One of the EST sequences equally lacks H2d, but the other sequence presents both H2 domains (H2a for ancestral and H2d for duplicated), showing that at least the duplicated H2d can be differentially

spliced out. Particularly interesting are the splice variants *LpCoe-B_1* and *LpCoe-B_3*. *LpCoe-B_1* has a truncated IPT/TIG domain (approximately half of it), but its DBD is left intact. This raises the possibility that, having an intact DBD, the *LpCoe-B_1* protein could still bind to DNA. It is uncertain whether *LpCoe-B_1* could form functional homo- or heterodimers with distinct properties, but it could also form non-functional dimers, in which case it would compete with intact COE dimers and act as an inhibitor. It would be interesting to perform binding assays with proteins synthesised from the different lamprey COE splice variants to assess these possibilities, both in terms of their dimer formation and DNA binding, and also in terms of their ability to induce (or repress) transcription. None of the lamprey COE splice variants have been reported in any other COE gene in metazoans.

Another COE gene reported to have splice variants, but different from lamprey COE-Bs, is mouse COE4. Several splice variants were described for mouse COE4 that disrupt the reading frame from the HLH domain towards the C-terminal region (Wang et al., 2002). Interestingly, the 5'-most splice site introduces an in-frame stop codon between H2d and H2a, leaving a predicted protein containing a HLH domain with H1 and H2d only. This raises the intriguing possibility for the existence of an additional lamprey COE-B splice variant containing H1 and H2d only. Wang and colleagues (2002) also suggested that perhaps any two helices from the vertebrate COE HLHLH domain would be sufficient for a functional COE protein (Wang et al., 2002). Finally, with the present study, there is no evidence for alternative splicing in the *LpCoe-A* HLH domain.

6.6.2. *LpCoe-A* and *LpCoe-B* Expression in the Central Nervous System

Brain: The expression of *LpCoe-A* and *LpCoe-B* in the brain is in agreement with the role of *COE/Ebf* genes in neuronal differentiation. Expression of *COE/Ebf* genes in the brain has been consistently observed in other vertebrates in specific compartments:

TELENCEPHALON: In the telencephalon, both lamprey *COE* genes are expressed in a small domain slightly biased to the dorsal half of the telencephalon (Figure 6-19B). However, a clear division of the pallium and subpallium is not discerned by either gene, although it is more likely that both genes are expressed in pallial regions. This would be in sharp contrast with mouse *COE1* which is expressed in the lateral and medial ganglionic eminences which are part of the subpallium, although mouse *COE* expression is observed in the neocortex (absent in lampreys) which is a pallium derivative (Garel et al., 1997a, 1999). Lamprey *COE* expression, however, would be more consistent with zebrafish *COE* genes as they are expressed in the olfactory bulb which is a pallial derivative (Bally-Cuif et al., 1998; Li et al., 2010).

DIENCEPHALON: In the diencephalon, *LpCoe-A* is initially expressed relatively homogeneously in the region ventral to the epiphysis and next to the midbrain along the dorsoventral axis, and it is gradually downregulated in most of these regions except in a very small stripe next to the midbrain at stage 28. On the other hand, *LpCoe-B* is mostly expressed next to the midbrain. The region ventral to the epiphysis in the alar plate likely corresponds to the gnathostome dorsal thalamus, whereas that next to the midbrain corresponds to the pretectum. Mouse *COE1-3* are expressed in the pretectum, whereas only *COE1* and *COE3* are expressed in the dorsal thalamus (Garel et al.,

1997a) as well as zebrafish *Ebf2* and *Ebf3* (Li et al., 2010). In the pretectum, mouse *COE* genes present a gap of expression between the alar and basal plates, similar to that observed for *LpCoe-B*. In the hypothalamus, is not very clear whether lamprey *COE* genes are expressed in this region, perhaps only in a very small domain next to the telencephalon. Hypothalamic expression has been observed for mouse *COE3* (Garel et al., 1997a). Chicken *COE1* is also expressed in the pineal gland. No lamprey *COE* expression was observed in the retina as occurs with *Xenopus Xebf3* and mouse *COE1* and *COE3* (Garel et al., 1997a; Pozzoli et al., 2001; Wang et al., 1997).

MIDBRAIN: In the midbrain, early *LpCoe-A* expression was observed along the dorsoventral axis, but later it is restricted dorsally, whereas *LpCoe-B* expression is initially biased dorsally, but later appears ventrally, and subsequently downregulated ventrally preserving its dorsal expression domain. In mice, initial midbrain *COE* expression is observed in the alar plate in the marginal zone, but later expanded into the basal plate with a gap at the level of the longitudinal limit between the alar and basal plates (Garel et al., 1997a). Later, mouse *COE* expression is downregulated in the basal plate but maintained in the tectum. Midbrain expression of mouse *COE* genes is most similar to that of *LpCoe-B* than *LpCoe-A*, but dorsal midbrain expression of both lamprey genes at late stages likely marks the extent of the tectum. Specific expression of either lamprey *COE* gene in the MHB was not detected as with zebrafish *Ebf2* and *Ebf3* (Li et al., 2010).

HINDBRAIN: An intriguing aspect in the hindbrain is that both *LpCoe-A* and *LpCoe-B* expression starts to be restricted to rhombomeres 2-4 (3-5?) from stage 26 onwards. A similar restriction occurs in chicken, where *Ebf1* and *Ebf3* are expressed in all

rhombomeres except rhombomeres 3 and 5 at stages HH13-HH14 (Garcia-Dominguez et al., 2003). However, such a distinction of rhombomeres 3 and 5 is not distinguished with *LpCoe-A* and *LpCoe-B*. Restriction of *LpCoe-A* and *LpCoe-B* expression is seen at late stages as compared to stages 21-24 where rhombomeres 3 and 5 are already marked by lamprey *Krox20* (Jimenez-Guri and Pujades, 2011). Comparing with *LpHes-B*, which distinguishes rhombomeres 3-5 at stages 23-24 (see Figure 7-7), *LpCoe-A* and *LpCoe-B* expression in this region at late stages suggests a site of prolonged neurogenesis in the lamprey hindbrain.

Early expression of *LpCoe-A* and *LpCoe-B* is relatively homogenous, but at later stages becomes restricted dorsally to a strong patch at the level of r3-5 and a smaller patch in the posterior hindbrain. A similar restriction of *COE* genes in the hindbrain (though not identical) occurs in chicken, where *Ebf1/COE1* and *Ebf3/COE3* are expressed in all rhombomeres except rhombomeres 3 and 5 at stages HH13-HH14. In mouse, *COE1* is initially expressed in the motor columns of r2 and r4, later extending to other rhombomeres and rhombomeric boundaries (Garel et al., 1997a). Also in zebrafish, *Zco2* is expressed early mainly in r2 and r4 (Bally-Cuif et al., 1998). *LpCoe-A* and *LpCoe-B* expression in the two patches at late stages suggests a site of prolonged neurogenesis in the lamprey hindbrain, but no expression of either lamprey gene reminiscent of rhombomeric boundaries was observed. It is also difficult to ascertain whether restricted lamprey expression corresponds to hindbrain nuclei such as the facial motor nucleus, the superior olive or vestibular nuclei where mouse *COE* genes are expressed (Garel et al., 1997a, 2000). Also, no lamprey *COE* expression in the ventricular zone was observed as occurs with mouse *COE* genes in the anterior hindbrain (Garel et al., 1997a). As mentioned before, *COE* genes are expressed in the

cerebellum, and expression of *LpCoe-A* and *LpCoe-B* was observed at the level of r1-r2, but the presence of a cerebellum in lampreys is still debatable.

Spinal cord: In the spinal cord, both *LpCoe-A* and *LpCoe-B* are observed as two parallel stripes of expression as seen from a dorsal view. Similar stripes of expression in the CNS have also been observed in zebrafish (Bally-Cuif et al., 1998). Such expression in *Lampetra planeri* marks the marginal zone of the spinal cord as inferred from expression of *COE/Ebf* genes in postmitotic neurons in the spinal cord of mouse, chicken and *Xenopus* (Dubois et al., 1998; Garcia-Dominguez et al., 2003; Garel et al., 1997a). In particular, mouse *Ebf1* and *Ebf3* are expressed mainly in the motor column (their expression also reaches the alar plate) at early stages, while *Ebf2* expression surrounds the motor column (Garel et al., 1997a). Also, mouse *Ebf2* is expressed specifically in the subventricular zone of differentiating neuroblasts at later stages (Garel et al., 1997a). Neither lamprey *COE* gene was observed in a region reminiscent of a subventricular zone as mouse *Ebf2*. *COE/Ebf* gene expression in the spinal cord is observed first in the basal plate and later in the alar plate (Davis and Reed, 1996; Garel et al., 1997a; Wang et al., 1997). For both lamprey *COE* genes, there was an overall permanence of expression of *LpCoe-A* along the dorsoventral axis and of *LpCoe-B* on the dorsal side at the stages analysed. Interestingly, *Xenopus Xebf3* expression has also been reported in the dorsal marginal zone (Pozzoli et al., 2001). Lamprey *COE* expression in the spinal cord is thus generally conserved with other vertebrates in the marginal zone of differentiated neurons.

6.6.3. *LpCoe-A* and *LpCoe-B* Expression in the Peripheral Nervous System

In the peripheral nervous system, both lamprey *COE* genes were observed in the olfactory placode/epithelium as in zebrafish, *Xenopus* and mouse (Bally-Cuif et al., 1998; Pozzoli et al., 2001; Wang et al., 1997, 2002). Expression of *LpCoe-A* and *LpCoe-B* in the trigeminal ganglia is also conserved with zebrafish, *Xenopus* and mouse *COE* expression in the trigeminal placodes/ganglia (Bally-Cuif et al., 1998; Dubois et al., 1998; Pozzoli et al., 2001; Wang et al., 1997). *LpCoe-A* expression was also observed associated with the otic vesicle. Expression of *COE* genes in the otic vesicle/inner ear has been reported with *Xenopus Xebf3* and mouse *COE1* (Pozzoli et al., 2001; Wang et al., 1997). Both lamprey *COE* genes are also expressed in facial and glossopharyngeal ganglia, but so far only mouse *COE1* has been observed to be expressed in glossopharyngeal ganglia among epibranchial placodes/ganglia (Wang et al., 1997). Both lamprey *COE* genes also seem to be expressed in the posterior lateral line placode/ganglion at stages 23-24, and no other vertebrate *COE* gene has been reported in these structures. Overall, neural expression of both *LpCoe-A* and *LpCoe-B* is similar to that in other vertebrates, and lamprey *COE* genes are expressed in cranial sensory placodes/ganglia as in gnathostomes. No expression of either gene was observed in DRG, a major expression site for mouse *COE* genes (Davis and Reed, 1996; Wang et al., 1997). No expression in migrating cranial neural crest cells was observed either as in zebrafish and *Xenopus* (Bally-Cuif et al., 1998; Dubois et al., 1998; Pozzoli et al., 2001).

6.6.4. *LpCoe-A* and *LpCoe-B* Expression in non-Neural Tissues

Expression in the branchial region has previously been observed in *Xenopus*. *Xcoe2* is expressed in the three anterior-most pharyngeal arches (Dubois et al., 1998). The authors report expression in the neural crest component giving rise to pharyngeal cartilage. However, their expression data do not allow an unambiguous argument for a neural crest-only component. Expression in the first three pharyngeal arches is also observed with *Xebf3* (Pozzoli et al., 2001). In zebrafish, it is difficult to ascertain with the available information whether *Zcoe2* is expressed or not in pharyngeal arches, as *Zcoe2* expression has been observed in migrating cranial neural crest cells but it has not been analysed in the more differentiated pharynx (Bally-Cuif et al., 1998).

Comparison with *LpCoe-A* and *LpCoe-B* expression, particularly by optical horizontal sections for *LpCoe-A* (Figure 6-18F), suggests that lamprey *COE* expression in pharyngeal arches extends to both neural crest and mesodermal components. However, to determine with certainty whether lamprey *COE* expression is in pharyngeal neural crest and/or mesoderm, it is necessary to double-stain lamprey *COE* genes in horizontal sections with markers such as *AP-2* (Meulemans and Bronner-Fraser, 2002) or *RAR2* (Jandzik et al., 2014) which stain pharyngeal neural crest cells that surround pharyngeal mesoderm.

6.6.5. *COE/EBF* Expression in Lymphocytes

One of the most intriguing questions regarding lamprey *COE* genes is whether they are also expressed in lymphocyte-like cells. Given the resemblance of B cell-like

lymphocytes to mammalian B lymphocytes (Guo et al., 2009), it is predicted that lamprey *COE* genes are expressed in these cells. As described in section 6.5.5., a preliminary survey of lamprey *COE* gene transcription in the intestine (and by extension in the typhlosole) did not show any indication of transcription of neither *LpCoe-A* nor *LpCoe-B*. This result was significant as cells expressing *VLRB* (i.e. B cell-like lymphocytes) are abundant in the typhlosole (Bajoghli et al., 2011). However, there are other sites of hematopoiesis such as the opisthonephros, supraneural body, branchial arches and blood (Amemiya et al., 2007; Kilarski and Plytycz, 1981; Zapata et al., 1981) where lymphocyte-like cells could express one or both *COE* genes.

In spite of the presumed absence of lamprey *COE* expression in the typhlosole (section 6.5.5.), there might be an indication of lymphocyte-related *COE* expression in pharyngeal arches. Both *LpCoe-A* and *LpCoe-B* are strongly expressed in pharyngeal arches in regions corresponding to mesoderm and/or neural crest. A recent study suggested that the larval pharyngeal arches might house thymus-like structures called “thymoids” located at the tips of gill filaments and in secondary lamellae of ammocoetes (Bajoghli et al., 2011). However, “thymoids” express *VLRA*, thus marking T cell-like lymphocytes instead of B cell-like cells (Bajoghli et al., 2011). Nevertheless, while expression of *VLRB* was not detected in “thymoids”, it was found in the vasculature of the pharyngeal basket and not in pharyngeal epithelium (Bajoghli et al., 2009, 2011). Furthermore, *VLRB* expression in the pharyngeal basket and oral tentacles has been observed as early as 45 days-old ammocoetes (Amemiya et al., 2007). This opens the possibility of lamprey *COE* expression in the pharyngeal region.

The above results thus suggest that expression of *LpCoe-A* and *LpCoe-B* could also be indicative of lamprey lymphocytes. So far the typhlosole and the nephric fold (including the larval opisthonephros and the associated adipose tissue) are considered the earliest blood-forming tissues in ammocoetes and the main hematopoietic organs at this stage of development (Amemiya et al., 2007). It is not certain at present the exact stage at which the earliest ever hematopoietic (and, in particular, lymphocyte-producing) tissues are found in developing lamprey embryos, as all studies on lamprey lymphocyte-like cells are conducted in larvae and adults. Therefore, studies on lamprey lymphocytes should also explore earlier developmental stages to clarify these issues.

Finally, data from protochordates are also suggestive of lymphocyte-related *COE* gene expression in pharyngeal arches. Lymphocyte-like cells have been described in amphioxus gills, and expression of a number of lymphocyte-related homologs, including *AmphiCoe*, is observed in the same region (Huang et al., 2007). *AmphiCoe* was also observed in the intestine (Huang et al., 2007), reminiscent of typhlosole lymphocyte-like cells in lampreys, and possible *COE* expression in lamprey lymphocyte-like cells. Also, the amphioxus lymphocyte-like cells increase in size upon antigen stimulation (Huang et al., 2007). Furthermore, *CiCOE* is expressed in pharyngeal mesoderm and has an important role in its specification (Stolfi et al., 2010). Overall, there seems to be a developmental and evolutionary correlation between lymphocyte-like cells present in branchial arches of –at least– amphioxus and vertebrates, and *COE* gene expression. Whether *AmphiCoe* and *LpCoe-A* and/or *LpCoe-B* expression in branchial arches indeed corresponds to lymphocyte-like cells awaits further investigation.

6.7. Summary of Lamprey COE Genes

Lampetra planeri has at least two COE genes, here named *LpCoe-A* and *LpCoe-B*, which are orthologous to *Petromyzon marinus* *PmCOE-A* and *PmCOE-B*. Both *L. planeri* genes possess all conserved domains of the COE subfamily, namely, the DBD with an atypical Zn-finger, the IPT/TIG domain and the HLH domain. However, phylogenetic analyses do not show any specific grouping of either lamprey gene with gnathostome COE1-4 paralogy groups. Lamprey COE-A genes seem to possess only two helices in the HLH domain, H1 and H2a. On the other hand, lamprey COE-B genes possess three helices (H1, H2d and H2a) in the genome. However, H2d can be differentially spliced, rendering transcripts with two or three helices. Alternative splicing of *LpCoe-B* was also observed in the Zn-finger and in the IPT/TIG domain.

Both *L. planeri* COE genes are expressed in the nervous system: they are expressed in restricted regions of the brain and at the lateral periphery of the spinal cord which corresponds to the marginal zone of differentiated neurons, in agreement with the role of gnathostome COE genes in neuronal differentiation. *LpCoe-A* and *LpCoe-B* are also expressed in cranial sensory placodes and ganglia. Outside the nervous system, both genes are expressed in pharyngeal arches in regions corresponding to mesoderm and/or neural crest. Additionally *LpCoe-A* is expressed in a characteristic longitudinal domain below the notochord. So far there is no clear evidence of expression of either gene in hematopoietic regions containing lymphocytes.

7. *LAMPETRA PLANERI* ASC GENES

7.1. Introduction

Structure and function of ASC genes: Similar to *Neurogenin* genes, ASC genes are relatively simple at the sequence level possessing only the bHLH domain, although they also present a small region at the C-terminal end with conserved sequence within a larger acidic region (Turner and Weintraub, 1994). Vertebrate ASC proteins form homo- and heterodimers with the ubiquitous E12 bHLH protein to bind to E-boxes and activate transcription of target genes (Ferreiro et al., 1992; Johnson et al., 1992; Meierhans et al., 1995; Zimmerman et al., 1993). Their DNA binding ability can also be inhibited by partnering with Emc proteins which lack the DNA-binding basic domain (Ferreiro et al., 1992). Upon binding to DNA, the bHLH domain of murine Mash-1 suffers a conformational change from a mainly unfolded to a largely α -helical form (Meierhans et al., 1995).

Expression of ASC genes in the vertebrate nervous system: A total of five ASC paralogues, ASC1-5 (*Mash1-5*), have been described in rodents (Johnson et al., 1990), and the most extensively studied ASC gene is *Mash1*, whose expression is restricted to the nervous system and for which there is more expression data available, thus serving as a reference in the study of vertebrate ASC gene expression. In the CNS, vertebrate ASC genes are expressed in neuronal precursors but their expression is extinguished upon overt neuronal differentiation. Thus, expression of ASC genes in the CNS is

generally biased towards the ventricular zone of neural progenitors, and excluded from differentiated neurons in the marginal zone (Lo et al., 1991).

Expression of *Mash1* in the entire nervous system is generally complementary to that of the neurogenins. In the brain, *Mash1* shows a restricted expression pattern in specific compartments, with a general complementarity with *neurogenin1* and *neurogenin2* gene expression (Figure 7-1). For example, in the telencephalon, *Mash1* is expressed in the presumptive basal ganglia while *neurogenin1* and *neurogenin2* are expressed in the presumptive cortex, and in the olfactory lobes *neurogenin1* is expressed in the dorsal part while *Mash1* is expressed ventrally (Gradwohl et al., 1996; Ma et al., 1997; Osório et al., 2010; Porteus et al., 1994; Sommer et al., 1996). In the diencephalon, *Mash1* is expressed in the hypothalamus and thalamus except for a small region between those two domains where *neurogenin1* is expressed, and similar complementarity is observed in the optic vesicles (Ma et al., 1997). In the hindbrain, *Mash1* is expressed in the medulla while *neurogenin1* is expressed in the pons (Ma et al., 1997). *Mash1* is also expressed in the retina (Ahmad, 1995; Jasoni and Reh, 1996).

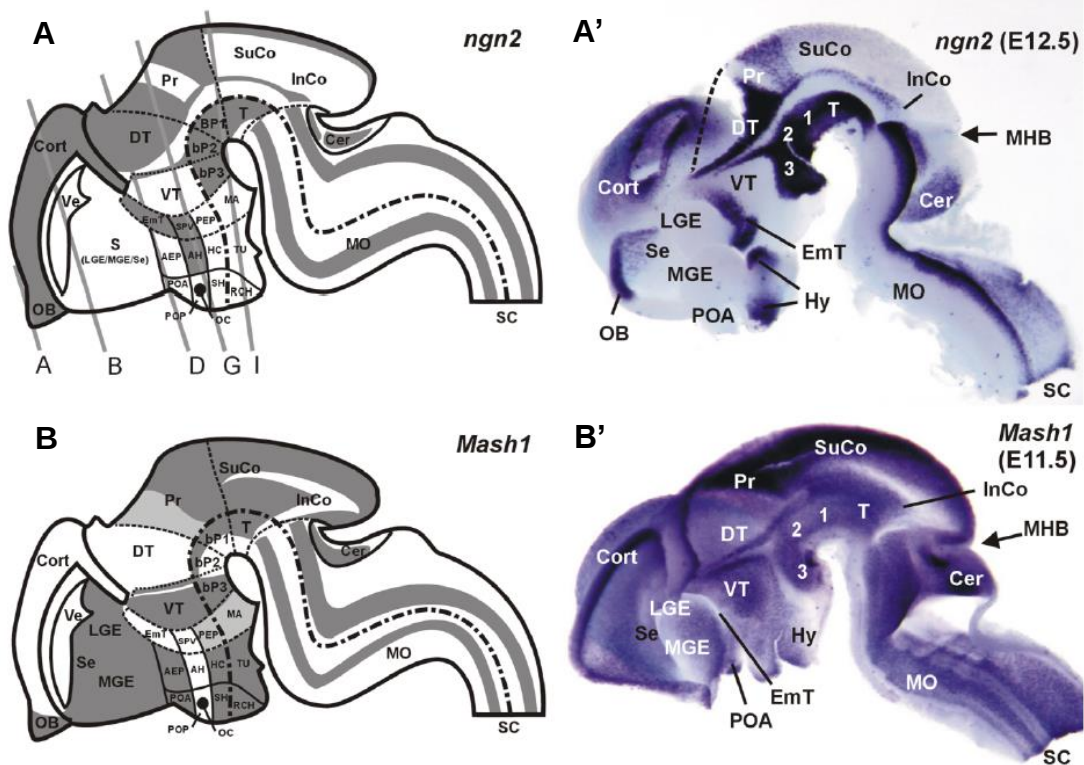


Figure 7-1. Expression of *Mash1* and *Ngn2* in mouse brain showing general complementarity of expression. Medial views of wholemount mouse brains (A' and B') and their respective schematic representations (A and B). (A and A') Expression of *neurogenin2* (*ngn2*) at E12.5. (B and B') Expression of *Mash1* at E11.5. *Mash1* is expressed in different compartments of the brain and its expression is mostly complementary to that of *neurogenin2* particularly in the forebrain. Abbreviations: AEP, anterior entopeduncular area; AH, anterior hypothalamus; bP1, basal plate of prosomere 1; bP2, basal plate of prosomere 2; bP3, basal plate of prosomere 3; Cer, cerebellum; Cort, cortex; DT, dorsal thalamus (thalamus); E, epiphysis; EmT, eminentia thalami; Hy, hypothalamus; HC, hypothalamic cell cord; InCo, inferior colliculus; LGE, lateral ganglionic eminence; MA, mammillary hypothalamus; MGE, medial ganglionic eminence; MHB, midbrain-hindbrain boundary; MO, medulla oblongata (rhombencephalon without cerebellum); OB, olfactory bulb; oc, optic chiasma; PEP, posterior entopeduncular area; POA, anterior preoptic area; POP, posterior preoptic area; Pr, pretectum; RCH, retrochiasmatic hypothalamus; S, subpallium; SC, spinal cord; Se, septum; SH, supra-chiasmatic area; SPV, supraoptic/paraventricular area; SuCo, superior colliculus; T, midbrain tegmentum; TU, tuberal hypothalamus; Ve, telencephalic brain ventricle; VT, ventral thalamus (prethalamus). Thick dotted line marks the alar-basal boundary. Taken from (Osório et al., 2010).

In the spinal cord, *Mash1* is initially expressed in a ventral domain adjacent to the floor plate on each side (Guillemot and Joyner, 1993; Lo et al., 1991). Later in development it is expressed in a wide dorsal domain and a smaller, medial domain with respect to the dorsal-ventral axis (Guillemot and Joyner, 1993). This expression is found in the ventricular zone and excluded from the marginal zone. As in the brain, expression of *Mash1* in the spinal cord is complementary to that of *neurogenin1* and to some extent

to *neurogenin2* at later stages. *neurogenin1* and *neurogenin2* are expressed in a small dorsal domain of the spinal cord and in a wider ventral domain, whereas *Mash1* is expressed between these two domains in the ventricular zone (Gowan et al., 2001; Gradwohl et al., 1996; Ma et al., 1997; Sommer et al., 1996).

In the peripheral nervous system, *Mash1* is expressed in some cranial placodes/ganglia but most prominently in neurons of the three major autonomic divisions, namely sympathetic, parasympathetic and enteric lineages. For example, *Mash1* is expressed in the olfactory placode and in olfactory receptor neuron (ORN) progenitors of the olfactory epithelium, as well as in enteric ganglia lining the gut, neurons of the sympathetic lineage (sympathetic ganglia, sympatho-adrenal neurons and coeliacum), and neurons of the parasympathetic lineage (sphenopalatine ganglia and ganglia along the lung and tongue) (Blaugrund et al., 1996; Cau et al., 1997; Gordon et al., 1995; Guillemot and Joyner, 1993; Lo et al., 1991; Ma et al., 1997). Rodent ASC genes are not expressed in neural crest derivatives such as DRG, Schwann cells or melanocytes (Lo et al., 1991; Ma et al., 1997).

There are, however, some regions where *Mash1* and *neurogenin1* are co-expressed such as the olfactory epithelium, midbrain and ventral spinal cord at early stages, although their expression patterns in these domains is not identical but rather sequential (Ma et al., 1997). Finally, *Mash2* is expressed only in the extraembryonic trophoblast (Alders et al., 1997; Guillemot and Joyner, 1993; Guillemot et al., 1994; Tanaka et al., 1997).

ASC genes have been described in other vertebrates. Chicken *CASH-1* is expressed in the subpallium of the telencephalon with mild expression reaching into the pallium (Frowein et al., 2002). In the diencephalon, *CASH-1* is expressed in the ventral thalamus seen as arcuate stripes of expression contiguous with the tegmentum, and surrounding the *zona limitans intrathalamica* (Sanders et al., 2002). *CASH-1* is also expressed in the retina (Jasoni et al., 1994). In the midbrain, *CASH-1* is expressed in the tegmentum as arcuate stripes of expression contiguous with the ventral thalamus stripes, and broadly in the tectum leaving a gap between both midbrain regions (Sanders et al., 2002). In the hindbrain, it is expressed in dorsal and ventral domains (Sanders et al., 2002). In the spinal cord, *CASH-1* is expressed in a narrow dorsal domain close to the roof plate just ventral to *neurogenin1* expression, and *CASH-1* is also expressed in a wider ventral domain of the spinal cord (Chesnutt et al., 2004; Ernsberger et al., 1995). In the PNS, *CASH-1* expression is observed in the olfactory placode/epithelium, trigeminal ganglia and neural crest cells that will form the DRG (Chesnutt et al., 2004; Ernsberger et al., 1995; Maier and Gunhaga, 2009). Also, as in rodents, *CASH-1* is observed in the autonomic nervous system in developing sympathetic ganglia, in the Remak ganglion (parasympathetic division), preaortic ganglia and enteric neurons (Ernsberger et al., 1995; Groves et al., 1995; Howard et al., 2000; Jasoni et al., 1994; McPherson et al., 2000).

In *Xenopus*, *XASH1* is expressed in the ventral telencephalon, diencephalon, midbrain and hindbrain (Ferreiro et al., 1992). *XASH1* is also expressed in the developing eye vesicles and later in the sensory retina, and transiently in the pineal gland (Ferreiro et al., 1992). Expression of *XASH1* in the brain is found in the ventricular zone of proliferating neuroblasts (Ferreiro et al., 1992). *XASH1* is not

expressed in the spinal cord or in neural crest derivatives (Ferreiro et al., 1992). On the other hand, *XASH3* is expressed early in the anterior-most forebrain, in the MHB region, in the prospective hindbrain, and in the presumptive spinal cord region seen as two bilateral longitudinal stripes (Turner and Weintraub, 1994; Zimmerman et al., 1993). Expression in the closed neural tube is seen in the anterior-most forebrain, sulcus limitans along the neural tube, and adjacent to the floor plate in the hindbrain (Turner and Weintraub, 1994; Zimmerman et al., 1993). *XASH3* is also expressed in putative neural crest cells, in the developing eye vesicle and later in the retina (Turner and Weintraub, 1994; Zimmerman et al., 1993).

In zebrafish, expression of two ASC genes, *Zash-1a* and *Zash-1b*, has been reported. *Zash-1a* is expressed in restricted compartments of the embryonic brain as *Mash1*. In the telencephalon, *Zash-1a* is expressed in the subpallium and in the most caudal pallial ventricular zone (Wullimann and Mueller, 2002), although mild expression also extends into pallial territory (Allende and Weinberg, 1994). In the diencephalon, *Zash-1a* is expressed in the ventral thalamus, hypothalamus, preoptic region, dorsal thalamus, epiphysis and in the neural retina (Allende and Weinberg, 1994; Wullimann and Mueller, 2002). In the midbrain it is strongly expressed in the tegmentum and tectum (Allende and Weinberg, 1994). In the hindbrain it is expressed dorsally at the level of the alar-basal plate boundary, first covering the entire mediolateral axis in a segmented pattern, but later only on the lateral side of each rhombomere (thus seen as a pair of lateral stripes of expression) and additionally in ventral domains of rhombomeres 1 and 2 (Allende and Weinberg, 1994; Amoyel et al., 2005). *Zash-1a* is also expressed near the floor plate of the presumptive cerebellum (Allende and Weinberg, 1994). In the spinal cord, *Zash-1a* is expressed in a large ventral domain abutting the floor plate and

in occasional cells in the dorsal-most region of the spinal cord (Allende and Weinberg, 1994). *Zash-1a* is also expressed in the adenohypophyseal placode and later in the adenohypophysis (Allende and Weinberg, 1994; Pogoda et al., 2006) and possibly in (unspecified) cranial ganglia at the level of the anterior hindbrain (Allende and Weinberg, 1994). In the autonomic nervous system, *Zash-1a* is expressed in sympathetic ganglia and enteric neurons (Lucas et al., 2006).

Zash-1b, on the other hand, is expressed in a generally complementary fashion to *Zash-1a*. *Zash-1b* is expressed in the ventricular zone of the brain (Mueller and Wullimann, 2003). In the telencephalon, *Zash-1b* is expressed dorsally (Allende and Weinberg, 1994). In the diencephalon, *Zash-1b* is expressed ventrally including the hypothalamus and later it covers the dorsal thalamus (Allende and Weinberg, 1994; Mueller and Wullimann, 2003). *Zash-1b* is also expressed in the inner nuclear layer of the eye which is constituted by postmitotic cells but not in proliferating peripheral retinal edge cells (Mueller and Wullimann, 2003). In the midbrain, *Zash-1b* is expressed in the tegmentum and tectum (Allende and Weinberg, 1994). In the hindbrain, *Zash-1b* is first expressed dorsally at high levels in presumptive rhombomeres 2 and 4 and later expression extends to rhombomere boundaries (Allende and Weinberg, 1994; Amoyel et al., 2005). *Zash-1b* is also initially expressed in smaller ventral domains of the hindbrain and cerebellum and subsequently broad expression is observed in the entire basal plate (Allende and Weinberg, 1994). In the spinal cord, *Zash-1b* is expressed more dorsally than *Zash-1a* (Allende and Weinberg, 1994). *Zash-1b* is not expressed in the epiphysis, olfactory epithelium or the adenohypophysis (Mueller and Wullimann, 2003; Pogoda et al., 2006).

Developmental roles of ASC genes in vertebrates: In agreement with its expression pattern, *Mash1* participates in a cascade of neural HLH gene expression and ORN formation in the olfactory epithelium, as *Mash1* mutant embryos fail to express *neurogenin1* and *neuroD* in the ORN lineage and ORN progenitors remain undifferentiated (Cau et al., 1997). Targeted mutation of *Mash1* affects the formation of ORNs and their progenitors, as well as of sympathetic and parasympathetic (sphenopalatine, submandibular, otic and ciliary) ganglia and enteric (myenteric neurons of the esophagus, but not the myenteric plexus of the stomach and intestine) neurons, but not sensory neurons, chromaffin cells of the adrenal medulla or sympathetic glia (Blaugrund et al., 1996; Guillemot et al., 1993; Hirsch et al., 1998; Lo et al., 1998; Sommer et al., 1995). The brain and spinal cord in these mutants seemed not to be affected overall (Guillemot et al., 1993). However, later studies showed that *Mash1* mutants present a severe loss of neuronal progenitors in the subventricular zone of the medial ganglionic eminence and lack the locus coeruleus and subcoeruleus and noradrenergic cells in the hindbrain, and probably the mesencephalic nucleus of the trigeminal nerve and a small population of spinal cord cells located laterally just dorsal to the sulcus limitans (Hirsch et al., 1998).

In *Xenopus*, injection of *XASH3* RNA in embryos at the 2-cell stage converts ectodermal cells into neural tissue causing the lateral expansion of the CNS along its entire length (Turner and Weintraub, 1994). These effects depend on *XASH3* binding to DNA (Turner and Weintraub, 1994). In addition, in *XASH3*-injected embryos presumptive cranial neural crest cells that will form cartilage and pharyngeal arch mesenchyme adopt a neural fate (Turner and Weintraub, 1994).

7.2. Aims

The aim of this chapter is to characterise the sequence and expression patterns of *Lampetra planeri* genes belonging to the ASC gene subfamily based on their identification *in silico* in the *Petromyzon marinus* genome (Chapter 3). This chapter is intended to answer the following questions:

- Does *L. planeri* have ASC genes as *P. marinus*?
- Provided the recovery of *L. planeri* ASC genes, where are *L. planeri* ASC genes expressed? Are they found in the nervous system as in gnathostomes?
- In the CNS, are *L. planeri* ASC genes expressed in the ventricular zone of the neural tube as in gnathostomes? Are they expressed in specific compartments of the brain as in gnathostomes? Do they show complementarity in their expression patterns both in the brain and spinal cord with respect to *Neurogenins*?
- In the PNS, are *L. planeri* ASC genes expressed in sympathetic or other autonomic ganglia as *Mash1* in mouse?

7.3. Hypothesis

The ASC HLH subfamily is widespread in the vertebrate group and homologues of this subfamily are key regulators of neurogenesis in all vertebrate species examined so far. Considering that invertebrate chordates possess one ASC homolog, it is predicted that *Lampetra planeri* has at least one ASC homolog, more similar in sequence to other vertebrate ASC genes than to their invertebrate chordate relatives. Based on the expression patterns of gnathostome ASC genes explained in section 7.1., *L. planeri* ASC

homologs are predicted to be expressed in the central nervous system in restricted regions of the brain, in a general complementary pattern with *LpNgn*, and in the eye anlage and/or retina. In the spinal cord, lamprey ASC genes are predicted to be expressed in the ventricular zone in a narrow dorsal domain and a wider ventral domain. In the peripheral nervous system, lamprey ASC genes are predicted to be expressed mainly in ganglia of the autonomic nervous system, in particular in enteric neurons. However, expression in some cranial sensory placodes/ ganglia such as olfactory placode (NHP), olfactory epithelium, trigeminal placode/ganglia could also occur based on gnathostome ASC expression.

7.4. Objectives

- To isolate homologous sequences of gnathostome *ASCL* genes from the *L. planeri* genome.
- Perform phylogenetic analyses to study the phylogenetic relationships of the isolated sequences with respect to other Deuterostome *ASCL* genes.
- Characterise the expression patterns of the isolated sequences in *L. planeri* embryos at different developmental stages by *in situ* hybridisation.
- Combine both expression and phylogenetic data to identify differences and similarities of expression patterns in different vertebrates mapped into a phylogenetic context in order to identify evolutionary trends of this HLH gene subfamily in vertebrates.

7.5. Results

7.5.1. Cloning of *LpAsh*

A fragment of 472 bp was cloned from *Lampetra planeri* genomic DNA. Based on other vertebrate *ASCL* homologues, which are contained in a single exon, the coding sequence was inferred from a predicted open reading frame in the genomic sequence obtained from PreEnsembl. Primers were designed within this region. The resultant fragment presents near-identical sequence similarity with Ensembl prediction ENSPMAT00000008966 at the nucleotide level within its first 273 bp, which covers the entire bHLH domain (Figure 7-2). The predicted amino acid sequences show high sequence conservation within the bHLH domain with respect to several other deuterostome sequences (Figure 7-3). Outside the bHLH domain, however, sequences are very divergent.

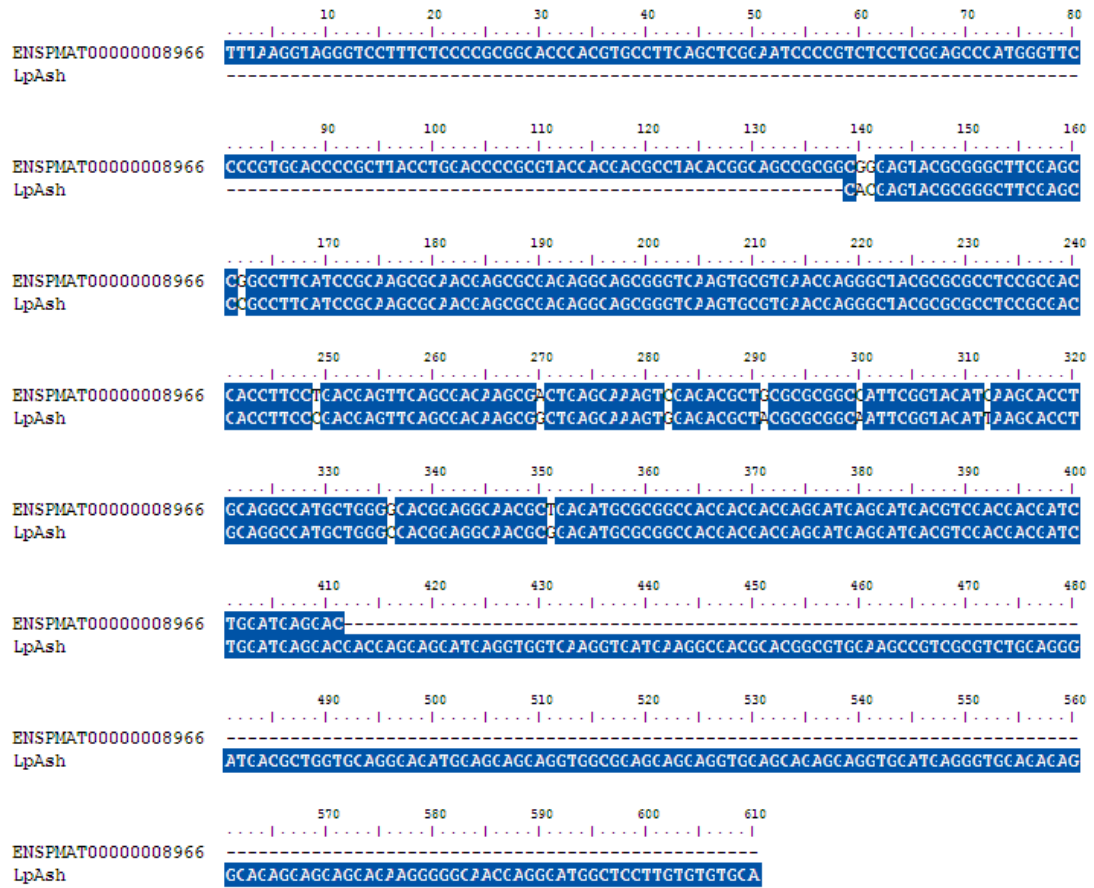


Figure 7-2. Nucleotide alignment of the Ensemble gene prediction ENSPMAT00000008966 (which recognises an *ASCL*-like gene) with *LpAsh*, a cloned fragment from *Lampetra planeri* that has sequence similarity to gnathostome *ASCL* genes. Although both sequences align only in a small portion, they present a near-identical sequence in that region, suggesting that they are orthologous genes.

7.5.2. Phylogenetic Analysis of *LpAsh*

A NJ tree was constructed with the alignment shown in Figure 7-3. Both *LpAsh* and the Ensembl prediction ENSPMAT00000008966 cluster within the ASCLb group, which contains *ASCL3-5*, with high support (Figure 7-4). However, both sequences do not cluster with any particular group (i.e. *ASCL3-5*). Therefore, the only conclusion from this tree is that both sequences belong to the ASCLb group.

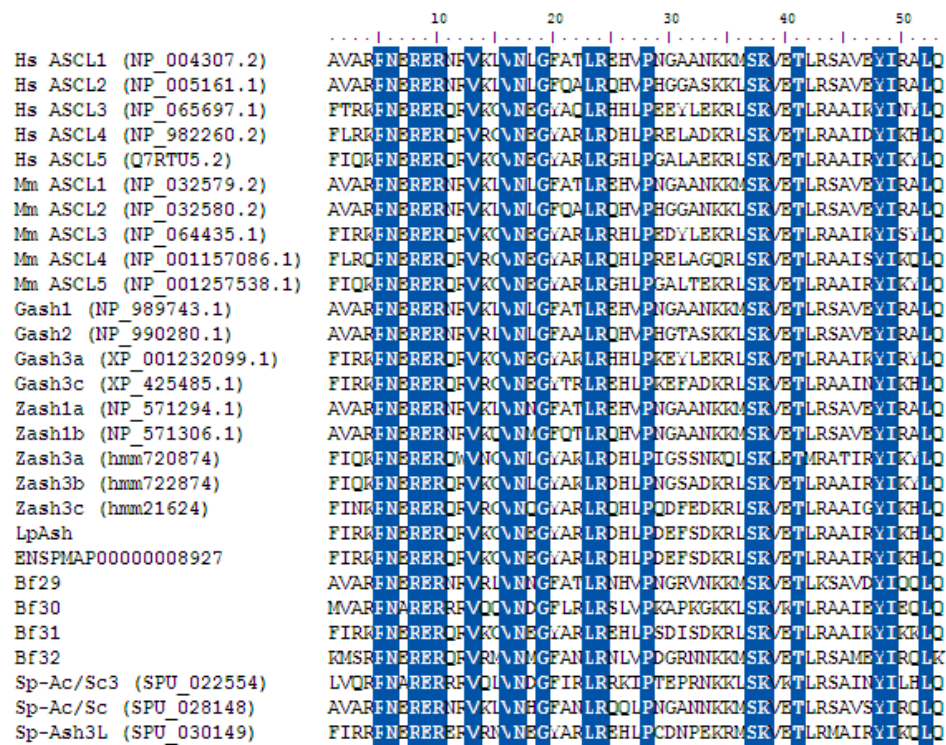


Figure 7-3. Amino acid alignment of the bHLH domain of several deuterostome ASCL genes including the Ensemble amino acid prediction ENSPMAP00000008927 (a predicted amino acid sequence from the transcript model ENSPMAT00000008966) and *LpAsh*. Amino acids conserved in all sequences are shaded. Abbreviations: Hs: *Homo sapiens*, Mm: *Mus musculus*, Gash: *Gallus gallus* ASCL genes, Zash: *Danio rerio* ASCL genes, Bf: *Branchiostoma floridae*, Sp: *Strongylocentrotus purpuratus*.

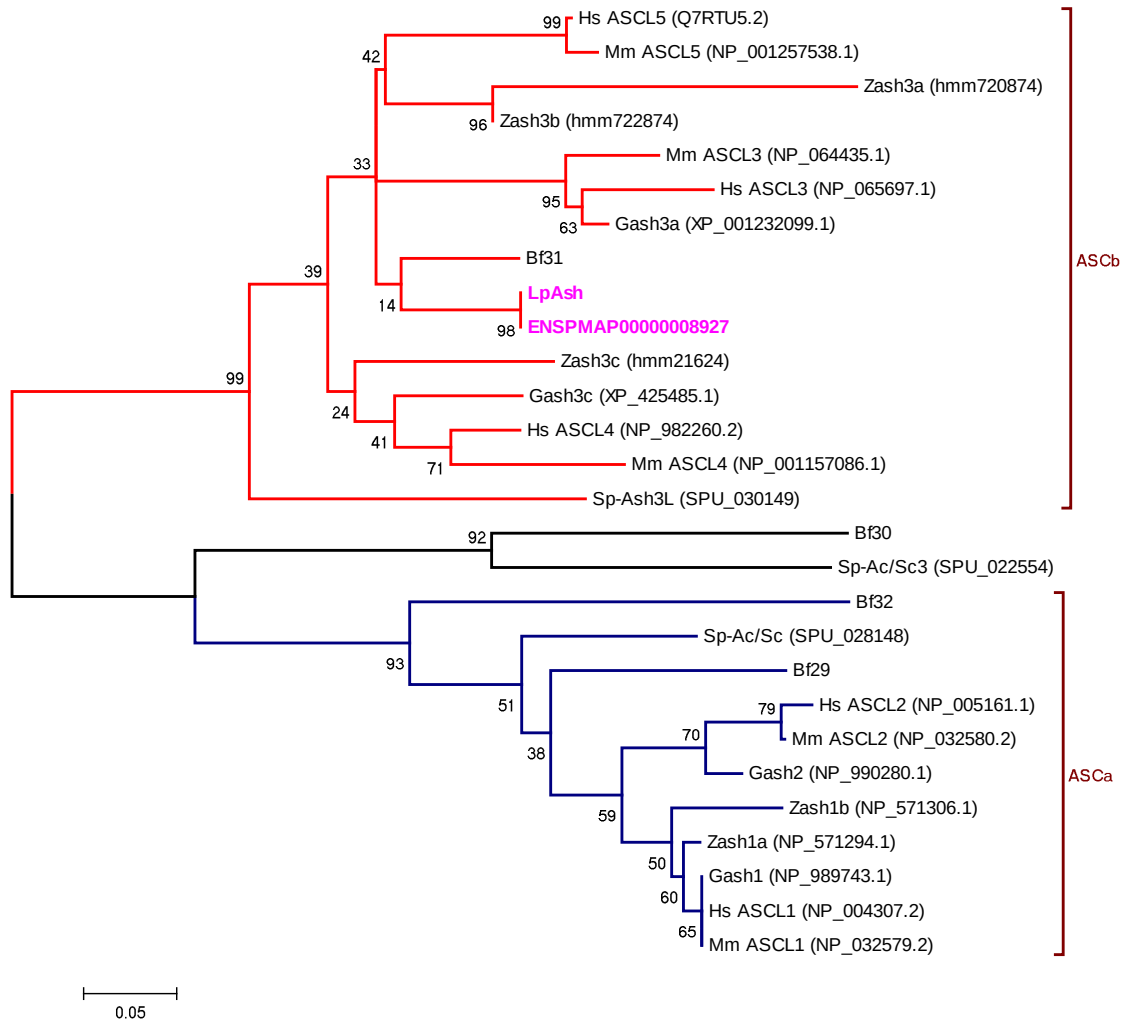


Figure 7-4. Unrooted Neighbour joining tree using the alignment in Figure 7-3 based on the Jones-Taylor-Thornton (JTT) amino acid substitution model. Numbers at nodes are bootstrap support values after 1000 replicates. Both Ensembl prediction ENSPMAP0000008927 and *LpAsh* cluster within the ASCb group (branches in red) supported by a bootstrap value of 99, suggesting that they are true ASC homologs. Abbreviations: Hs: *Homo sapiens*, Mm: *Mus musculus*, Gash: *Gallus gallus* ASC genes, Zash: *Danio rerio* ASC genes, Bf: *Branchiostoma floridae*, Sp: *Strongylocentrotus purpuratus*. Scale bar: amino acid substitutions per position.

7.5.3. *LpAsh* Expression in Developing *L. planeri* Embryos

Expression of *LpAsh* was analysed by digoxigenin-labelled *in situ* hybridisation from Tahara (1988) stages 21 to 28. At stage 21, expression is observed ventral to the spinal cord (Figure 7-5A), and from a dorsal view is seen as two parallel lateral stripes of expression (Figure 7-5B). At stage 22, these expression domains disappear, and

instead *LpAsh* is upregulated in the entire brain in an homogenous pattern (Figure 7-5C). Later at stage 23, expression is downregulated in the brain (although this could be due to less staining exposure time) except in the ventral midbrain and hindbrain, and transcripts are now observed in the anterior spinal cord (Figure 7-5D). At stage 24, expression is upregulated in the entire brain in an homogenous pattern, and is observed much stronger than at stage 22 (Figure 7-5E). Expression is also observed in the ventral spinal cord (Figure 7-5E, F). At this stage expression appears in the mandibular arch and in a very discrete domain of the developing endostyle (Figure 7-5E). From a dorsal view transcripts are seen in the middle part of the neural tube in a domain corresponding to the ventricular zone and possibly extending into the marginal zone (Figure 7-5G). At stage 25 expression is completely downregulated in the spinal cord but persists in the brain although at a slightly lower level with respect to stage 24 (Figure 7-5H-J). Expression also persists in the mandibular arch and is now observed in the upper and lower lips (Figure 7-5I). Expression in the endostyle increases as the endostyle develops (Figure 7-5I, J).

At stage 26 expression persists in the brain and is upregulated in the posterior spinal cord and tailbud (Figure 7-6A, B, G). The upper and lower lips are strongly stained and expression is even stronger in the mandibular arch and endostyle (Figure 7-6B). Given that riboprobes are trapped inside the pharynx in sense probe-control embryos, it is not possible to discern whether other pharyngeal arches are specifically stained (Figure 7-5K and Figure 7-6D).

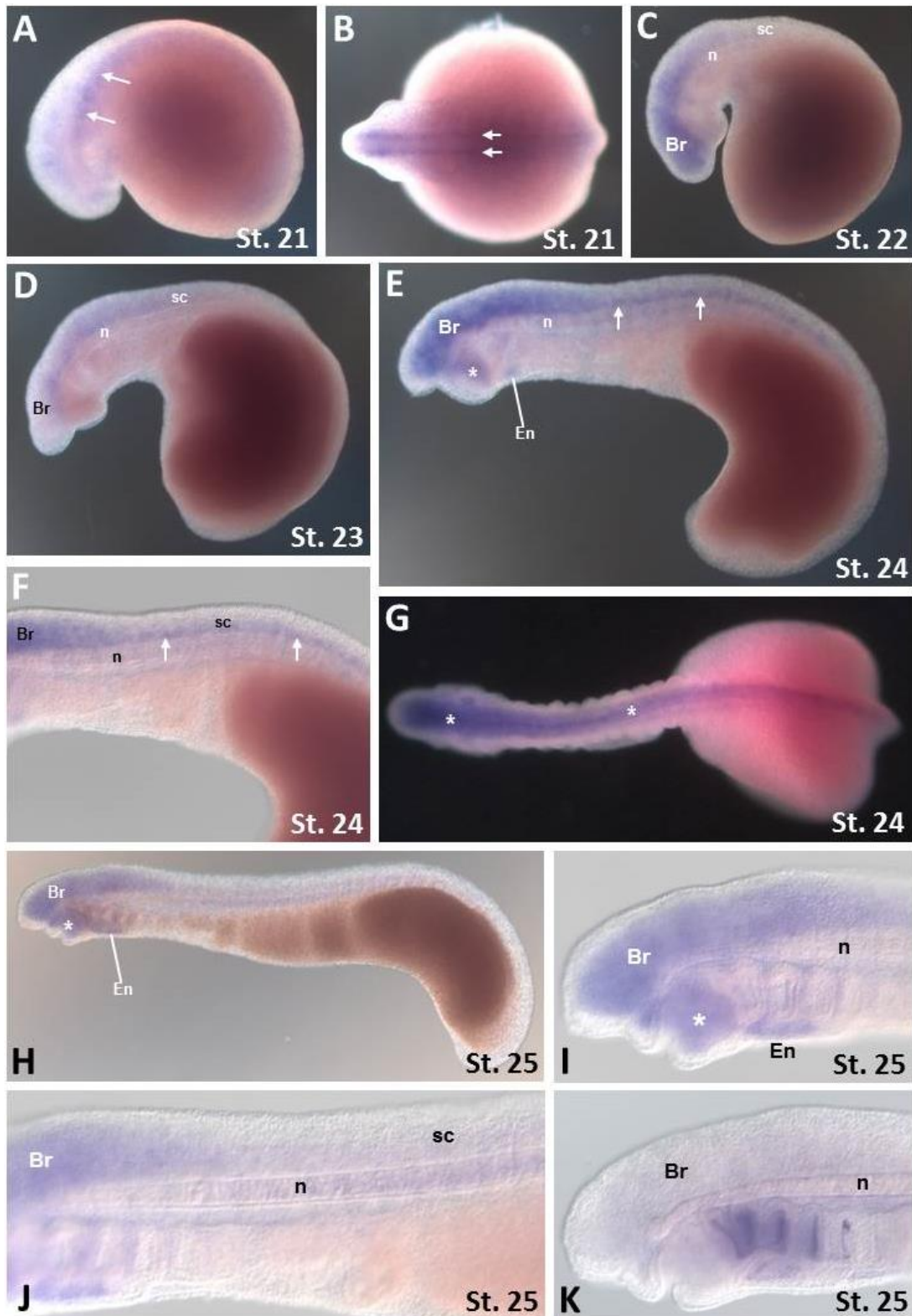


Figure 7-5. Expression of *LpAsh* in developing *Lampetra planeri* embryos at stages 21 to 25. All images are lateral views except (B) and (G) which are dorsal views. Anterior is to the left in all images. (A) At stage 21, expression is observed ventral to the spinal cord (white arrows). However, in a dorsal view, it is seen as two lateral stripes likely found in the spinal cord (B, white arrows). (C) At stage 22 expression dramatically increases in a homogenous pattern covering the entire brain. (D) Stage 23 embryo showing weaker expression in the brain and spinal cord due to less time of staining. (E) Stage 24 embryo showing strong expression throughout the brain and ventral spinal cord (white arrows in E and F), first pharyngeal

arch (white asterisk) and in the nascent endostyle. (G) From a dorsal view, expression is seen in the middle part of the spinal cord (white asterisks), possibly delimiting the extent of the ventricular zone or also including part of the marginal zone. (H-J) At stage 25, expression decreases in the spinal cord and is virtually absent, but is maintained somewhat strong in the brain, first pharyngeal arch (white asterisks) and endostyle. (K) Control stage 25 embryo with *LpAsh* sense probe showing trapped probe inside the pharynx but no staining in the embryo.

At stage 28, expression considerably diminishes in the brain (Figure 7-6C, F) and, in the spinal cord, very mild –if any– expression remains covering the entire medio-lateral extent of the spinal cord as seen from dorsal view (Figure 7-6E). In contrast, strong expression remains in the endostyle and mandibular arch (Figure 7-6C, F), where transcripts are mainly confined medially (Figure 7-6I). Expression in the tailbud and posterior spinal cord completely disappears (Figure 7-6H).

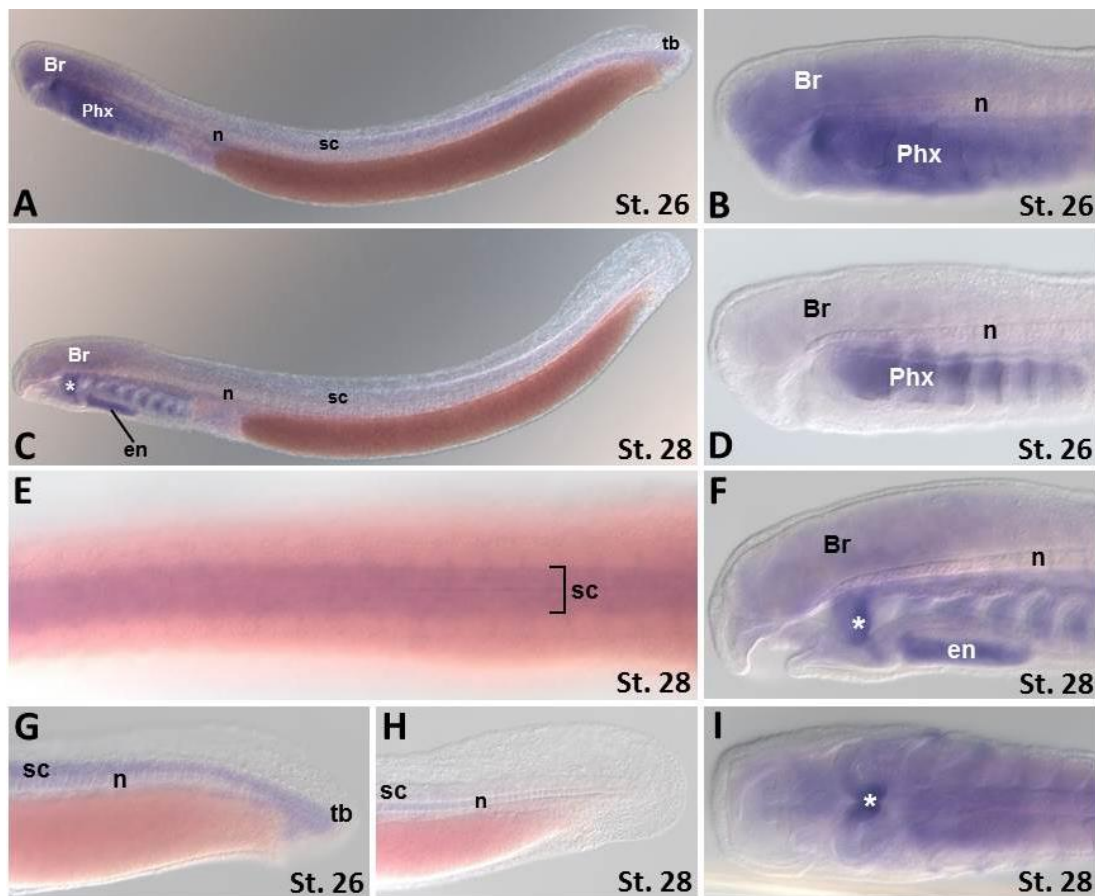


Figure 7-6. Expression of *LpAsh* in developing *Lampetra planeri* embryos at stages 26 and 28. All images are lateral views except (E) which is a dorsal view of the trunk region and (I) which is a ventral view of the head. Anterior is to the left in all images. (A, B) At stage 26, expression is observed in the brain, pharyngeal region, endostyle, posterior spinal cord and tailbud. (D) Control stage 26 embryo with *LpAsh* sense probe showing trapped probe in the pharynx. (C) At stage 28, expression diminishes in the neural tube but increases in the first pharyngeal arch (white asterisk) and endostyle (F). (E) Dorsal view of a stage 28 embryo shows mild expression in the spinal cord that seems to cover the entire medio-lateral extent (bracket). (G, H) Expression in the posterior spinal cord and tailbud is relatively strong at stage 26 (G) but vanishes at stage 28 (H). (I) Expression in the first pharyngeal arch is strongest in the middle (white asterisk). Expression of *LpAsh* does not reveal any structures similar to gnathostome sympathetic ganglia. Abbreviations: Br, brain; en, endostyle; n, notochord; Phx; pharynx; sc, spinal cord; tb, tailbud.

7.6. Discussion

7.6.1. *LpAsh* Expression in the Central Nervous System

Perhaps one of the most conspicuous differences in expression of *LpAsh* with respect to other vertebrate ASC genes is the general homogenous expression in the brain. As mentioned in section 7.1., gnathostome ASC genes are expressed in specific compartments of the developing brain and, at least in rodents, their expression is generally complementary to that of *neurogenins*. *LpAsh*, however, is expressed in a homogenous pattern in the entire brain (Figures 7-5 and 7-6), in sharp contrast to murine *Mash1* and *neurogenin2* (Figure 7-1). Therefore, the general complementarity of expression between *LpAsh* and *LpNgn* (Figures 5-6 and 5-7), as expected from data from rodents, is not observed. *LpAsh* does not seem to be expressed in the epiphysis as gnathostome ASC genes. Furthermore, *LpAsh* expression is comparable to zebrafish *Zash-1b* as it is expressed mostly in the ventricular zone at postembryonic stages (Mueller and Wullmann, 2003).

Another lamprey ASC gene, *Petromyzon marinus Ascl1*, has been cloned and its expression analysed. *P. marinus Ascl1* is expressed in the pituitary gland, lens, anterior lip mesoderm, trigeminal ganglia and possibly in the notochord (Häming et al., 2011). Although *P. marinus Ascl1* shows a somewhat more restricted expression in the brain with respect to *L. planeri LpAsh*, its expression pattern is certainly quite different from murine *Mash1*. The grouping of *LpAsh* within the *ASCb* subfamily and of *Ascl1* within the *ASCa* subfamily is in agreement with their different expression patterns in the brain.

In the spinal cord, particularly at stage 24 (Figure, 7-5E, F), *LpAsh* is expressed in the ventral spinal cord, possibly corresponding to the ventral domain of expression of *LpNgn* in the spinal cord (Figure 5-5E). However, because of time restrictions, sections of embryos stained for *LpAsh* could not be obtained and therefore it was not possible to describe the distribution of *LpAsh* transcripts with more accuracy. Double staining with *LpAsh* and *LpNgn* will be necessary to determine whether both genes are co-expressed in the ventral spinal cord or present complementary expression domains. From dorsal views (Figure 7-5G and 7-6E), *LpAsh* expression is observed in the ventricular zone but its expression extends more to the periphery of the spinal cord with respect to *LpNgn* or *LpHes-B*, suggesting that *LpAsh* might also be expressed in the marginal zone and/or subventricular zone.

7.6.2. *LpAsh* Expression in the Peripheral Nervous System

Like expression in the CNS, *LpAsh* expression in the lamprey PNS is quite different from that in gnathostomes. As explained in section 7.1., in the PNS, ASC genes are

mainly expressed in ganglia of the autonomic nervous system, but they are excluded from sensory neurons/ganglia. The autonomic nervous system in lampreys is not well understood. Lampreys have been reported to lack the paired chain of interconnected, paravertebral sympathetic ganglia observed in gnathostomes (Nicol, 1952), but they have neural crest-derived enteric ganglia (Häming et al., 2011).

Overall, there seems to be an evolutionary difference in expression of ASC genes between agnathans and gnathostomes in the nervous system, and more conspicuously in the brain and PNS. This may reflect differences both in anatomy and protein function or binding site recognition of ASC proteins between these two vertebrate groups.

7.6.3. *LpAsh* Expression in non-Neural Tissues

One of the strongest expression sites of *LpAsh* outside the nervous system is in the endostyle. The endostyle has been well-recognised in lampreys as homologous to the vertebrate thyroid gland (Kluge et al., 2005). The ammocoete endostyle is known to differentiate into a thyroid gland in the adult after metamorphosis, and it is specifically marked by expression of transcription factor *TTF-1* (i.e. *Nkx-2.1*), which is detected in the endostyle anlage as early as stage 21 (Ogasawara et al., 2001). The expression and role of ASC homologues in the endostyle/thymus is, however, less clear. *Petromyzon marinus* *Ascl1* is not detected in the endostyle at embryonic stages (Häming et al., 2011). *LpAsh* is first detected in the endostyle at stage 24 (Figure 7-5E), much later than *TTF-1*. This suggests that *LpAsh* is not involved in the specification of the lamprey endostyle and that it might be downstream of *TTF-1* action. In gnathostomes, it is known that human *Hash1* is expressed in thyroid tissues and that it has an important

implication in thyroid cancer (Ball et al., 1993; Chen et al., 1997, 2005; Sippel et al., 2003). Therefore, there is a precedent of at least one *ASC* homolog expressed in the thyroid, and expression of *LpAsh* in the endostyle thus adds more information to their evolutionary homology and further suggests that *ASC* expression in endostyle/thyroid tissues is conserved in vertebrates.

Finally, two other *LpAsh* expression sites outside the nervous system are the first pharyngeal arch and tailbud, although so far no analysis on *ASC* genes in these structures has been reported in other vertebrates.

7.7. Summary of Lamprey *ASC* Genes

LpAsh has high sequence similarity to other vertebrate *ASC* genes and clusters with the *ASCb* group within the *ASC* subfamily together with its *Petromyzon marinus* predicted ortholog. *LpAsh* is expressed in the entire CNS. In the brain, a prominent difference with respect to gnathostome *ASC* genes is that *LpAsh* is expressed in the entire brain in a homogenous pattern, whereas gnathostome *ASC* genes are expressed in specific compartments. Expression in the brain was expected to be complementary to that of *LpNgn*. In the spinal cord, strongest expression is observed at stage 24 where it is restricted to the ventral side. In the PNS, no expression was observed in cranial ganglia or DRG. Notably, *LpAsh* does not seem to be expressed in any structure reminiscent of sympathetic or other autonomic ganglia (such as enteric), supporting increasing evidence that lampreys do not have a chain of sympathetic ganglia. Also, another

lamprey ASC gene, *Petromyzon marinus Ascl1*, is expressed in small domains of the brain and no expression is observed in structures reminiscent of autonomic ganglia, similar to *LpAsh*. *P. marinus Ascl1* clusters with the *ASCa* group within the ASC subfamily, in agreement with its different expression pattern with respect to *LpAsh*. *LpAsh* is also strongly expressed in the endostyle and mandibular arch since they start to form.

8. REGULATION OF NEURAL HLH GENES BY THE NOTCH SIGNALLING PATHWAY IN THE LAMPREY SPINAL CORD

8.1. Introduction

This chapter is dedicated to the study of the regulation of lamprey neural HLH gene expression by the Notch signalling pathway. A general description of the Notch signalling pathway will be given, with a special reference to an enzymatic complex, the γ -secretase complex, which is the direct target of a Notch pathway inhibitor, DAPT, which was used in this study as explained below. It is beyond the scope of this project to describe in detail each component and/or protein domains within the Notch signalling pathway. For comprehensive descriptions of this pathway please consult –for example– (Bray, 2006; Kopan and Ilagan, 2009) and references therein.

8.1.1. The Notch Signalling Pathway

The Notch signalling pathway is one of the most important pathways that function during vertebrate development, and one of its main roles is regulation of neurogenesis. This pathway relies on cell-cell contacts. The vertebrate Notch protein is a single-pass transmembrane receptor that has a large extracellular part composed of 29-36 tandem epidermal growth factor (EGF)-like repeats, which bind to calcium ions thereby affecting the conformation and affinity of Notch to ligands (Bray, 2006; Kopan and Ilagan, 2009) (Figure 8-1). Following the EGF-like repeats towards the membrane, the extracellular part of the Notch receptor also possesses three cysteine-rich LIN repeats

and a heterodimerisation domain which conform the negative regulatory region (NRR) that prevents receptor activation in the absence of ligands (Kopan and Ilagan, 2009). On the other hand, the intracellular part of the Notch molecule contains a RAM domain, six ankyrin repeats (ANK domain) and a carboxyl-terminal PEST domain that possesses degradation signals. The Notch protein also possesses nuclear localisation signals between the RAM and ANK domains, and between the ANK and PEST domains (Bray, 2006; Kopan and Ilagan, 2009).

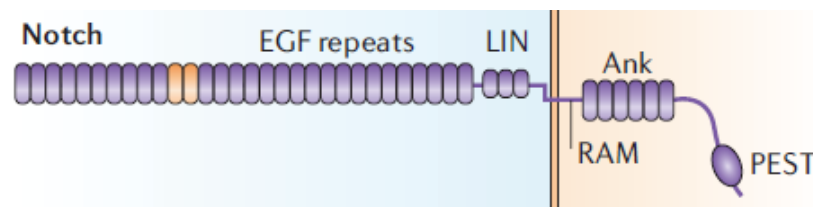


Figure 8-1. Domain structure of the Notch molecule. The Notch extracellular domain contains 29-36 EGF repeats, three cystein-rich LIN repeats and the NRR (not shown). EGF repeats 11 and 12 (orange) in the extracellular part are essential for ligand binding. The intracellular region contains the RAM, ANK and PEST domains, as well as nuclear localisation signals (not shown). Taken from (Bray, 2006).

Notch ligands include Delta and Serrate (known as Jagged in mammals), which are also transmembrane proteins. These ligands use their large extracellular domains that also consist of EGF-like repeats to bind Notch receptors (Bray, 2006; Kopan and Ilagan, 2009). Upon ligand binding, the Notch molecule suffers two proteolytic cleavages (Figure 8-2). The first cleavage is catalysed by ADAM-family metalloproteases some 12 amino acids before the transmembrane domain in the extracellular part of the Notch molecule (Kopan and Ilagan, 2009). This process is known as “ectodomain shedding”. For a detailed description of the ADAM protein family see, for example, (Klein and Bischoff, 2011) and references therein. Once the Notch extracellular domain is shed, the second cleavage is carried out by the intramembrane γ -secretase enzyme complex

within the lipid bilayer in the Notch transmembrane domain (Bray, 2006; Kopan and Ilagan, 2009). After these two rounds of cleavages, the Notch intracellular domain (NICD) is released into the cytosol; accordingly, the NICD contains the RAM, ANK and PEST domains, together with the nuclear localisation signals. The NICD then translocates to the nucleus and interacts with the DNA-binding protein CSL through the RAM and ANK domains to release it from binding to co-repressors of transcription (Fischer and Gessler, 2007). Later, this complex recruits the coactivator Mastermind to promote transcription of target genes (Bray, 2006; Fischer and Gessler, 2007).

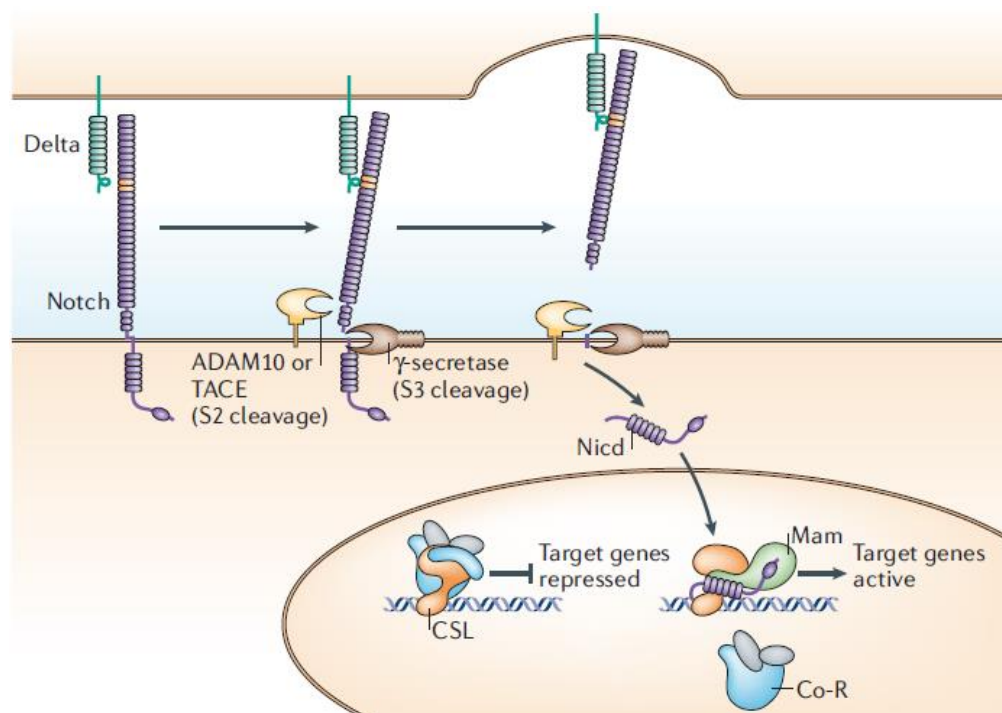


Figure 8-2. Overview of the Notch signalling pathway. Upon Delta ligand binding from an adjacent cell, the Notch molecule suffers two proteolytic cleavages. The first cleavage (S2 cleavage) is carried out by ADAM10 or ADAM17 (TACE; TNF- α -converting enzyme) metalloproteases. Later, the second cleavage (S3 cleavage) is carried out by the γ -secretase complex. After these two rounds of cleavages, the NICD (Nid) is released and it translocates to the nucleus to interact with CSL to release it from binding to co-repressors (Co-R). This complex then recruits the co-activator Mastermind (Mam) to promote transcription of target genes. Taken from (Bray, 2006).

8.1.2. Notch Expression in the Vertebrate Spinal Cord

Mouse *Notch1*, *Notch2* and *Notch3* and chicken *Notch1* are all expressed in the ventricular zone of the spinal cord (Fior and Henrique, 2005; Vilas-Boas and Henrique, 2010; Yang et al., 2006), whereas zebrafish *Notch1a*, *Notch1b* and *Notch5* are also expressed in neural precursors but *Notch1a* is additionally found in differentiating neurons of the spinal cord (Appel et al., 2001; Kim et al., 2008a). *Petromyzon marinus* *Notch* is also expressed in the ventricular zone of the neural tube (Guérin et al., 2009) (Dr. Tatjana Sauka-Spengler, personal communication; Figures 8-3 and 8-4).

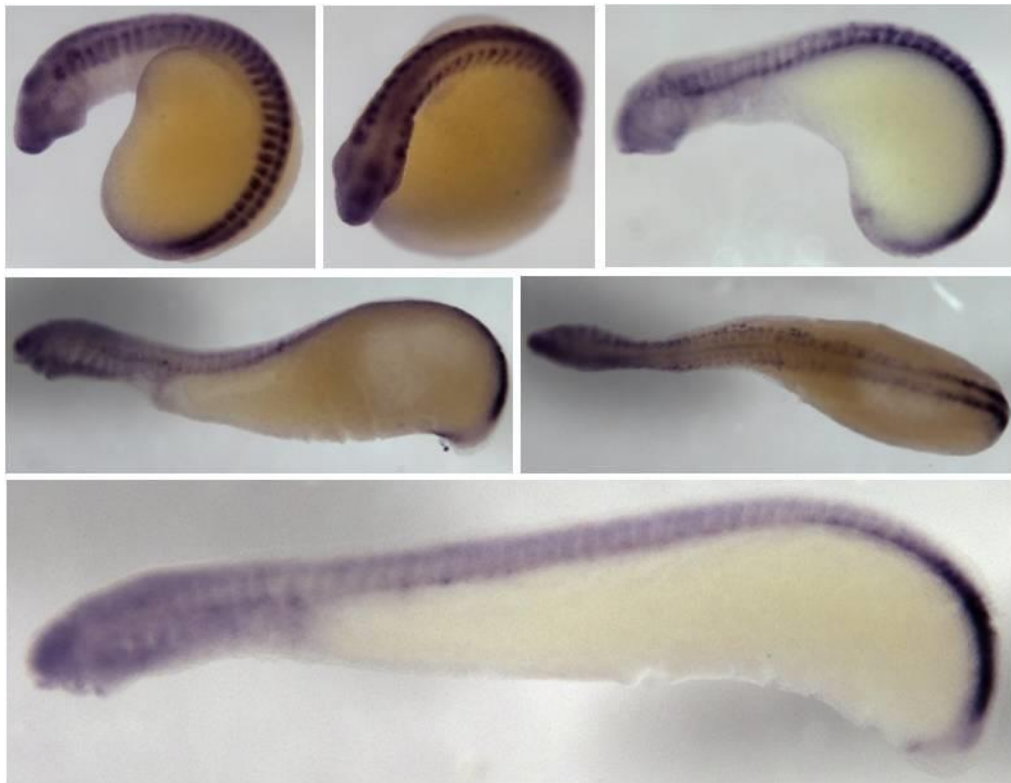


Figure 8-3. Expression of *Petromyzon marinus* *Notch* at different stages of development. All images are kind courtesy of Dr. Tatjana Sauka-Spengler.

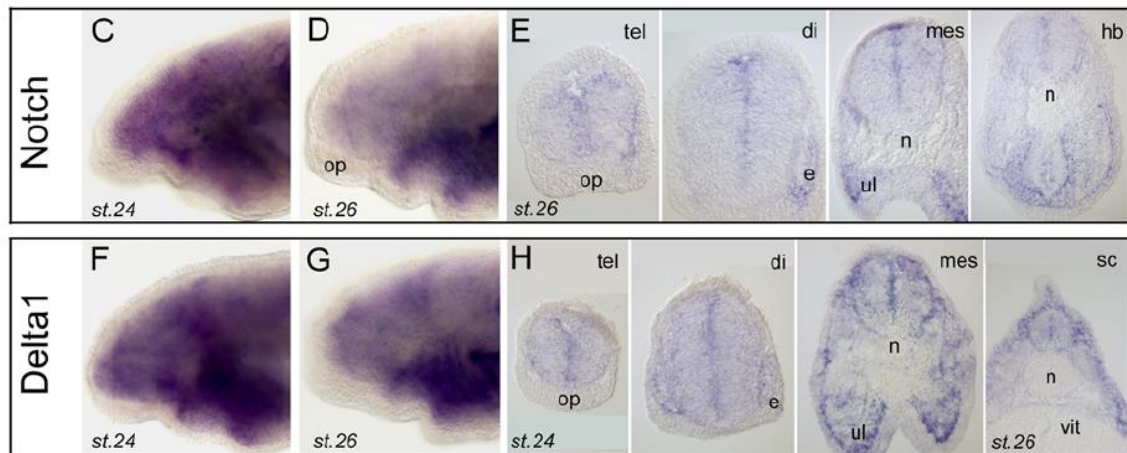


Figure 8-4. Expression of *Petromyzon marinus* *Notch* and its ligand *Delta1* at different stages of development. Taken from Guérin et al. (2009).

8.1.3. The γ -Secretase Complex

The γ -secretase complex is composed of one molecule each of presenilin 1 or 2 (PS1 or PS2), nicastrin, presenilin enhancer-2 (Pen-2) and anterior pharynx defective 1a or 1b (Aph1a or Aph1b) hydrophobic proteins (Sato et al., 2007) (Figure 8-5). This complex is able to cut a wide range of single-pass membrane receptors such as Notch (Beel and Sanders, 2008). Nicastrin is a transmembrane glycoprotein whose ectodomain has been suggested to bind the new N-terminal end of 1st-shedded Notch molecule and recruit it into the γ -secretase complex (Dries et al., 2009; Sato et al., 2007). However, this view has been challenged by a number of papers which suggest that instead of being a substrate receptor, Nicastrin plays a role in stabilising the γ -secretase complex assembly and is dispensable for γ -secretase activity (Chavez-Gutierrez et al., 2008; Futai et al., 2009; Shah et al., 2005). Therefore, a certain debate still remains concerning the exact role of Nicastrin within the γ -secretase complex. The exact sequence of events for the γ -secretase complex assembly still remains to be elucidated; however, it seems that this

complex becomes finally active only until the four elements are properly assembled (Zhao et al., 2010).

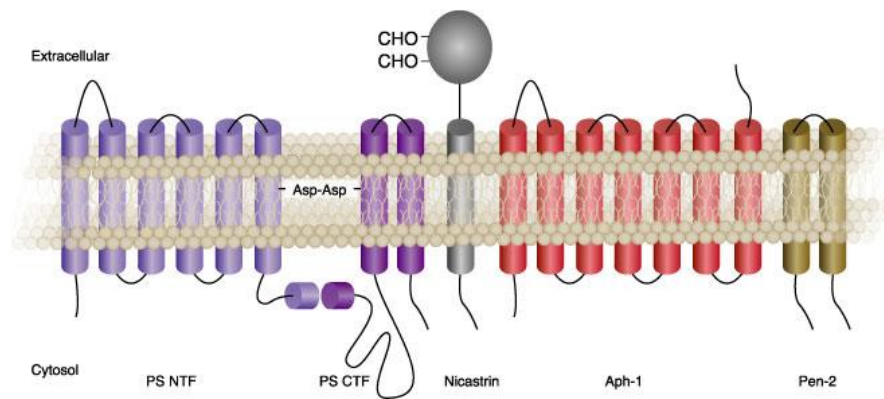


Figure 8-5. The γ -secretase complex. The active γ -secretase complex is composed of one molecule each of Presenilin (PS), Nicastrin, Aph-1 and Pen-2. Upon assembly of the four components, PS suffers endoproteolysis into an N-terminal fragment (NTF) and a C-terminal fragment (CTF). The active site of the complex is constituted by two aspartate residues (Asp) in the PS transmembrane domains. Nicastrin also becomes glycosylated (CHO) to form a mature complex. Taken from (Wolfe, 2003).

This complex hydrolyses peptide bonds within the transmembrane domain of the Notch molecule (Kopan and Ilagan, 2009). Hydrolysis is carried out by presenilin, which spans the membrane eight times (De Strooper and Annaert, 2001). Presenilin suffers endoproteolysis yielding an N-terminal fragment (NTF) and a C-terminal fragment (CTF) which remain associated forming a heterodimer (Podlisny et al., 1997). Two transmembrane aspartates within the Presenilin molecule are essential for Presenilin endoproteolysis and form the active site (Kimberly et al., 2000; Wolfe et al., 1999; Zhao et al., 2010). That state constitutes the functional form of Presenilin (Wolfe, 2003). It has also been suggested that Presenilin CTF is also involved in substrate docking for intramembrane proteolysis (Fraering et al., 2005).

8.1.4. The Notch Pathway Inhibitor DAPT

Compound (N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester), DAPT (Figure 8-6), is a dipeptidic γ -secretase-specific inhibitor. It was isolated from an initial pool of 25,000 compounds that was screened to identify compounds that significantly lowered production of amyloid β -peptide ($A\beta$) by human kidney 293 cells (HEK293) (Dovey et al., 2001). $A\beta$ is a peptide produced by activity of the γ -secretase complex and constitutes an important factor in Alzheimer's disease (Fortini, 2002). DAPT action does not seem to affect physiological cellular processes such as protein synthesis or general secretion (Dovey et al., 2001), indicating that it does not cause cellular toxicity. DAPT was later shown to affect the Notch pathway by inhibiting NICD production by HEK293 cells and altering Notch-regulated somitogenesis and neurogenesis in live zebrafish embryos (Geling et al., 2002). A number of studies have also made use of DAPT in zebrafish embryos to study the regulation of nervous system development by the Notch signalling pathway, most of them with a DAPT concentration of 100 μ M (Bae et al., 2005; Filippi et al., 2005; Geling et al., 2004; Kim et al., 2008a, 2008b; Latimer et al., 2005; Webb et al., 2011).

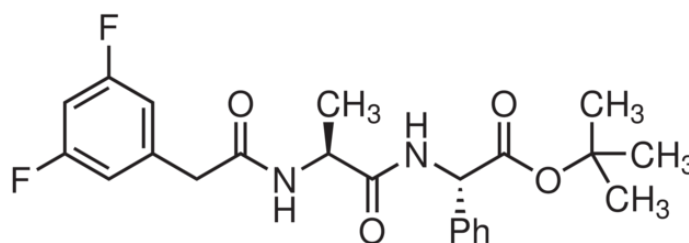


Figure 8-6. Molecular structure of the DAPT molecule.

DAPT binds to the C-terminal region of human Presenilin 1 CTF that contains TM7 and TM8 (residues 372-467) and this binding site seems to be different from either the Presenilin 1 catalytic site or the substrate binding site (Morohashi et al., 2006). The authors suggested that DAPT may inhibit γ -secretase activity by blocking substrate movement or the substrate unwinding process when passing from the binding site to the catalytic site by –directly or indirectly– altering the conformation of γ -secretase. However, these experiments and model need to be confirmed independently.

In this study, I took advantage of the specificity of DAPT on inhibiting the Notch pathway to study its effects on the expression of the previously cloned genes, namely, *LpHes-B*, *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B*. For this purpose, the spinal cord was chosen as a model system given that: 1) The Notch pathway is known to regulate the development of the spinal cord in gnathostomes by maintaining a proliferative state of neuroblasts at the posterior end of the neural tube (Akai et al., 2005); 2) As mentioned before, several vertebrate *Notch* homologues are expressed in the ventricular zone of the spinal cord; 3) A number of HLH gene subfamilies (including *Hairy/E(spl)*, *ASC*, *Neurogenin* and *COE* gene subfamilies) are expressed in the gnathostome spinal cord, with very dynamic expression patterns both in the anterior-posterior and dorsal-ventral axes throughout development; 4) All five lamprey HLH genes cloned in the present study are expressed in the spinal cord; 5) The spinal cord is a tissue that can be uniformly treated in embryos grown in media containing substances such as DAPT. This approach is based on studies in zebrafish that have used DAPT treatments to study axial structures such as the spinal cord and somites (Geling et al., 2002; Kim et al.,

2008a; Latimer et al., 2005). Moreover, some of these studies have analysed expression of neural HLH genes in the spinal cord.

8.2. Aims

The aim of this chapter is to analyse the sensitivity of expression of *LpHes-B*, *LpNgn*, *LpCoe-A*, *LpCoe-B* and *LpAsh* in the spinal cord of developing *Lampetra planeri* embryos to the Notch signalling pathway upon its pharmacological inhibition. This chapter is intended to answer the following questions:

- Is the expression of *L. planeri* *LpHes-B*, *LpNgn*, *LpCoe-A*, *LpCoe-B* and *LpAsh* sensitive to Notch signalling?
- Is their sensitiveness to Notch signalling similar to that of their gnathostome homologues?
- Is there any correlation between the sensitiveness of expression to Notch signalling with respect to the timing of activation of *LpHes-B*, *LpNgn*, *LpCoe-A*, *LpCoe-B* and *LpAsh* along the neuronal differentiation programme?

8.3. Hypothesis

Considering the importance of the Notch signalling pathway in neurogenesis and its conservation among different vertebrate groups, both in expression patterns and regulation on neural HLH gene's expression, in particular in the spinal cord, it is predicted that the Notch pathway is present, and acts in a similar manner in the lamprey spinal cord as in gnathostomes. *Notch* genes are expressed in the neural plate of

different gnathostomes as the spinal cord elongates (Akai et al., 2005; Beck and Slack, 1998; Fior and Henrique, 2005). Given that lamprey *notch* is expressed in the neural tube during development (Guérin et al., 2009)(Dr. Tatjana Sauka-Spengler –personal communication), it is predicted that Notch activity regulates the expression of lamprey neural HLH genes during spinal cord development in the same way as in gnathostomes. DAPT is therefore predicted to inhibit the lamprey Notch pathway, observed particularly in the spinal cord. Given the homology of the cloned lamprey neural HLH genes with particular gnathostome HLH gene subfamilies, it is predicted that *LpHes-B* will be downregulated in the spinal cord upon DAPT treatment, whereas expression of *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B* will instead be upregulated after this treatment.

8.4. Objectives

- Treat developing *L. planeri* embryos with 100 μ M DAPT from stages 24 to 26.
- Analyse expression of *LpHes-B*, *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B* in the spinal cord region by *in situ* hybridisation in control and DAPT-treated embryos.
- Compare expression data from DAPT-treated *L. planeri* embryos with information from gnathostomes.

8.5. Results

As a first step, I set out to look for lamprey *presenilin* homologues considering that Presenilin proteins of the γ -secretase complex are the direct targets of DAPT molecules. Blast searches against the Ensembl genome database version 7.0 for *Petromyzon marinus* using mouse and human Presenilin-1 and Presenilin-2 amino acid sequences rendered one predicted lamprey sequence with high sequence similarity to mouse and human Presenilin proteins (Figure 8-7). This sequence possesses the C-terminal part that corresponds to CTF where DAPT binds in human and mouse proteins. Therefore, this sequence represents a good gene candidate that encodes a lamprey Presenilin protein where DAPT is predicted to bind. No further sequences with high similarity were obtained from Blast searches.

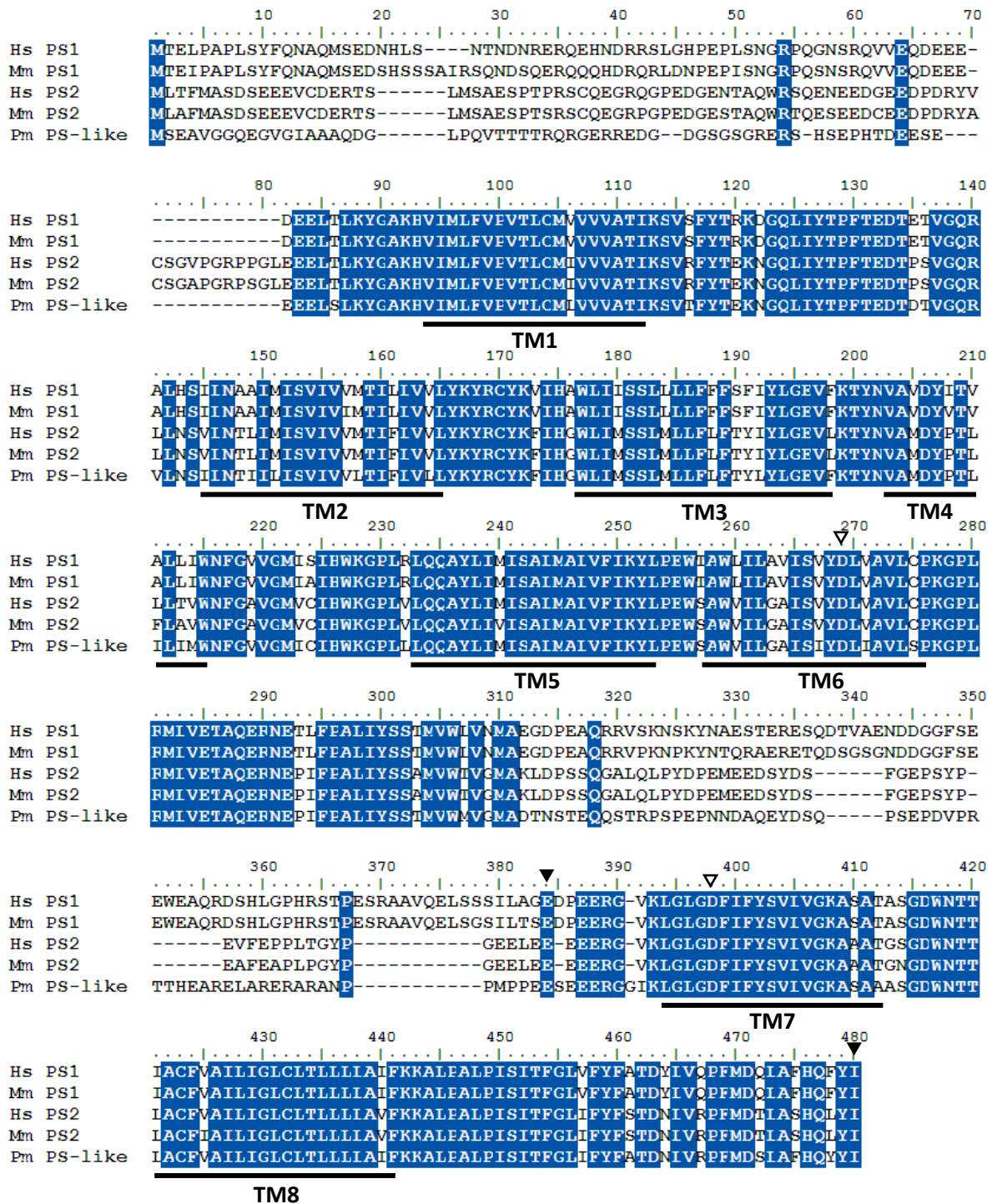


Figure 8-7. Alignment between human and mouse Presenilin-1 and Presenilin-2 amino acid sequences with lamprey *Presenilin-like* predicted amino acid sequence. Note the highly conserved regions covering the transmembrane domains (TM) 1-8, and in particular at the C-terminal part of the sequences corresponding to amino acids 372-467 (delimited by arrowheads) of human Presenilin-1 where DAPT is predicted to bind according to Morohashi et al. (2006). Conserved residues among all five sequences are shaded. Conserved aspartate residues (D) in TM6 and TM7 essential for the endoproteolytic cleavage are indicated with open arrowheads. Transmembrane domains are according to Groth et al. (2002).

Lamprey embryos were treated either with 100 μ M DAPT in river water or an equivalent concentration of DMSO in river water from stage 24 to stage 26 as explained in section 2.2.16. in Methods. Approximately 120 embryos were dechorionated and treated for each condition. Three batches of experiments were performed, rendering over 700 treated embryos in total. Expression of all cloned genes was analysed both in DAPT-treated and in DMSO-control embryos of the same batch. At the gene expression level, all DAPT treated embryos analysed displayed the same change in phenotype with respect to DMSO-control embryos.

The first, most obvious morphological phenotype that could be observed was that embryos treated with DAPT were consistently (over 90% of embryos analysed, $n=25$) shorter in length and bent backwards (compare A and B in Figures 8-8, 8-9, 8-10, 8-11, 8-12), suggesting that Notch inhibition affects normal growth and extension of the body. Some embryos (less than 10%, $n=25$), however, were not bent backwards or shorter in length (hence looking morphologically normal). On the other hand, the head was morphologically unaffected overall (compare C and F in Figure 8-8, 8-9, 8-10; and C and D in Figure 8-11 and 8-12). Any obvious internal anatomical differences in DAPT-treated embryos were observed such as expansion or reduction of the brain, spinal cord, notochord or adjacent somites as compared to control embryos, which displayed a near-identical phenotype as normal, untreated embryos. The gross phenotype of DAPT-treated embryos (i.e. backwards bending and shorter length) constitutes a good indicator of an effective DAPT inhibition irrespective of gene expression, which is very useful in cases where DAPT treatments would not affect expression of a particular gene at all.

The first gene to be analysed was *LpHes-B*. As explained in section 4.1. in Chapter 4, some *Hes* genes are direct effectors –and upregulated by– the Notch signalling pathway in gnathostomes. Therefore, changes in their expression can indicate whether DAPT affects the Notch signalling pathway in a specific manner. Indeed, expression of *LpHes-B* was completely abolished in the spinal cord upon DAPT treatment (Figure 8-8A, B, D, E, G and H). Interestingly, expression of *LpHes-B* was only mildly downregulated in the brain, where it maintains a very similar expression pattern when compared to control embryos treated with DMSO only (Figure 8-8C and F). Inhibition of *LpHes-B* also seemed to reach the tail bud (Figure 8-8G and H); however, given that not all normal embryos express *LpHes-B* in the tail bud, it is not certain whether DAPT has a real effect on this region.

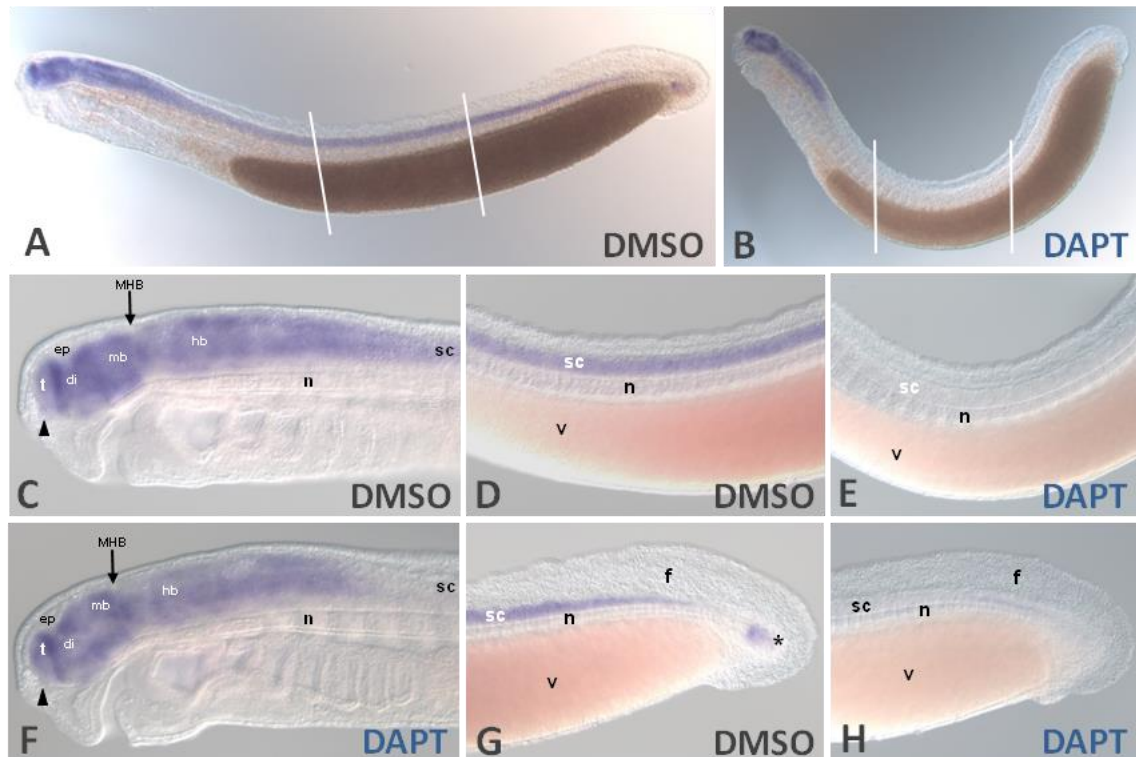


Figure 8-8. Expression of *LpHes-B* in treated embryos from stage 24 to stage 26. (A), (C), (D) and (G) are DMSO-control embryos. (B), (E), (F) and (H) are DAPT-treated embryos. (D) and (E) are photographs of corresponding regions in control-DMSO and DAPT-treated embryos as indicated by lines in (A) and (B), respectively. (G) and (H) are photographs of the tail region in control-DMSO and DAPT-

treated embryos, respectively. In all images anterior is to the left. *LpHes-B* is expressed in the entire central nervous system and in the tail bud (asterisk in G) in DMSO-control embryos as occurs in normal embryos (A, C, D, G; cf. Figure 4-12). Upon DAPT treatment, expression in the spinal cord is completely abolished (B, E) together with –possibly– expression in the tail bud (H). Note how expression in the brain is maintained overall (C, F). Abbreviations: ep: epiphysis, di: diencephalon; f: caudal fin, hb: hindbrain, mb: midbrain, MHB: midbrain-hindbrain boundary, n: notochord, sc: spinal cord, t: telencephalon, v: vitellum.

Expression of *LpAsh* was only mildly downregulated in the brain and spinal cord as compared to DMSO-control embryos (Figure 8-9A-F). In the posterior spinal cord, however, downregulation was more noticeable with respect to control embryos (Figure 8-9G and H).

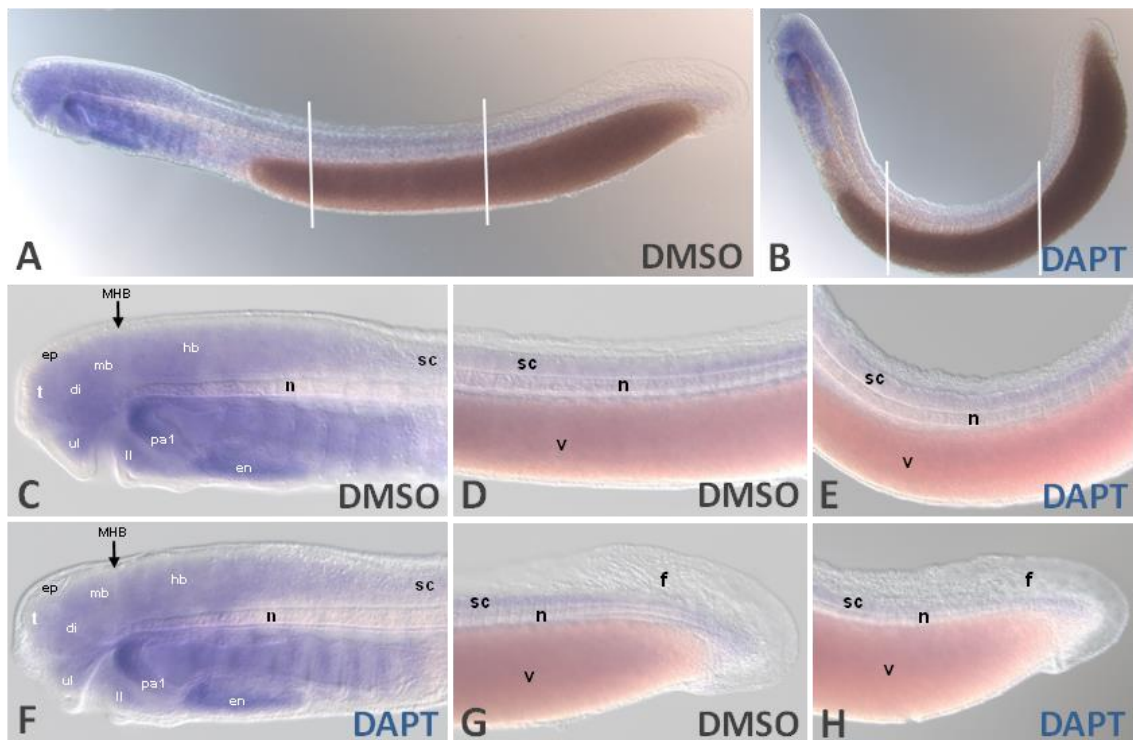


Figure 8-9. Expression of *LpAsh* in treated embryos from stage 24 to stage 26. (A), (C), (D) and (G) are DMSO-control embryos. (B), (E), (F) and (H) are DAPT-treated embryos. (D) and (E) are photographs of corresponding regions in control-DMSO and DAPT-treated embryos as indicated by lines in (A) and (B), respectively. (G) and (H) are photographs of the tail region in control-DMSO and DAPT-treated embryos, respectively. In all images anterior is to the left. In the nervous system, *LpAsh* is expressed mainly in the brain and faintly in the spinal cord in DMSO-control embryos as in normal embryos (A, C, D, G; cf.

Figure 7-6). Upon DAPT treatment, expression in the brain and spinal cord is only slightly downregulated (B, E, F, H), notably at the posterior (H). Abbreviations: en: endostyle, ep: epiphysis, di: diencephalon; f: caudal fin, hb: hindbrain, ll: lower lip, mb: midbrain, MHB: midbrain-hindbrain boundary, n: notochord, pa1: pharyngeal arch 1, sc: spinal cord, t: telencephalon, ul: upper lip, v: vitellum.

Expression of *LpNgn* is slightly upregulated in the hindbrain in DAPT-treated embryos (Figure 8-10A-C and F). In the spinal cord however, expression is drastically modified from a nearly homogenous pattern to discrete clusters of putative neuroblasts (Figure 8-10A-K). From a normal expression pattern as two bands in the anterior spinal cord, one dorsal and one ventral, expression of *LpNgn* in DAPT-treated embryos changes to clusters from this region all along the spinal cord (Figure 8-10C-K). However, the intensity of staining is weak at the anterior and progressively becomes stronger towards the posterior.

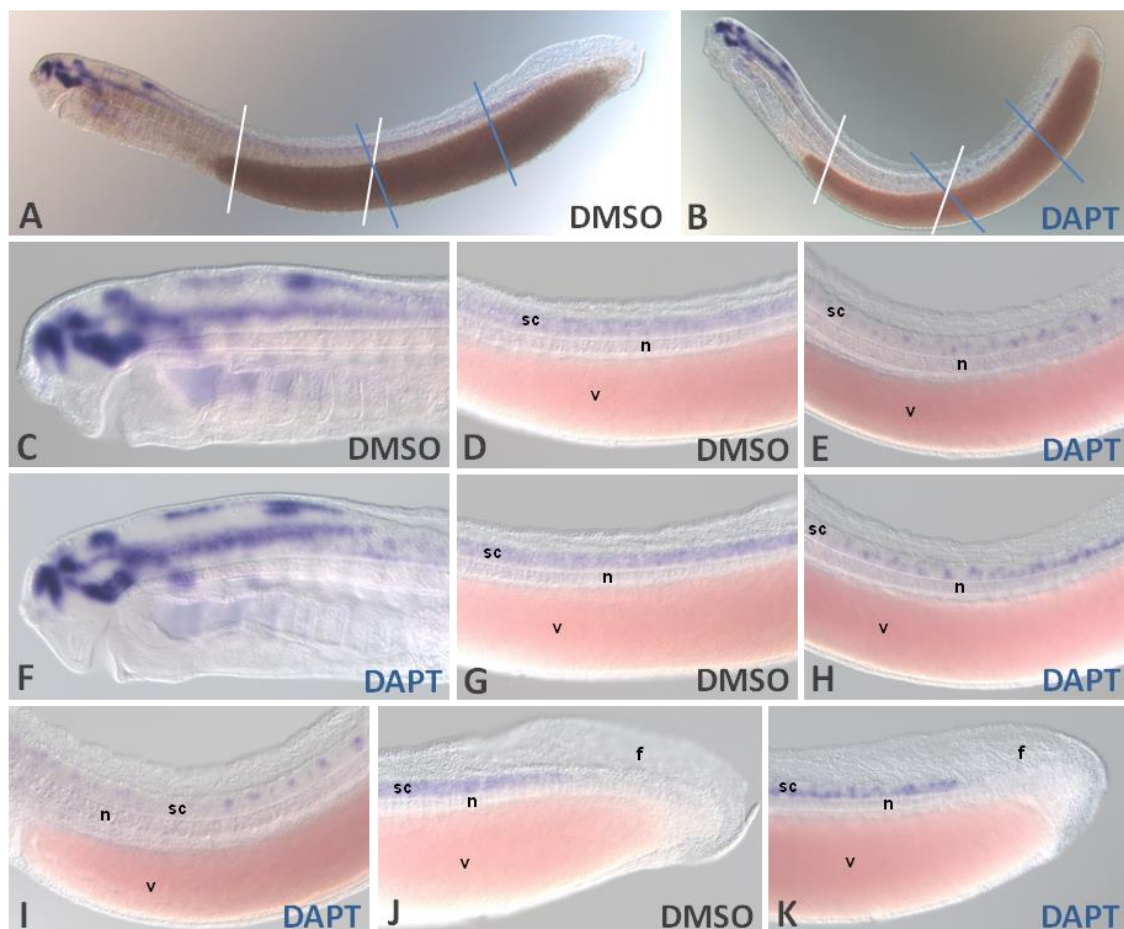


Figure 8-10. Expression of *LpNgn* in treated embryos. (A), (C), (D), (G) and (J) are control-DMSO embryos. (B), (E), (F), (H), (I) and (K) are DAPT-treated embryos. (D) and (E) are photographs of corresponding regions in control-DMSO and DAPT-treated embryos as indicated by white lines in (A) and (B), respectively. (G) and (H) are photographs of corresponding regions in control-DMSO and DAPT-treated embryos as indicated by blue lines in (A) and (B), respectively. (I) is a photograph of a different embryo; the region of the trunk shown in this photograph is slightly anterior to that delimited by white lines in (A) and (B). In all images anterior is to the left. *LpNgn* is expressed in the brain and spinal cord in control-DMSO embryos as in normal embryos (A, C, D, G; cf. Figure 5-7). Upon DAPT treatment, expression in the brain is slightly upregulated (F), particularly in the hindbrain, while maintaining the same expression pattern. In the spinal cord, however, *LpNgn* expression changes to clusters of putative neuroblasts all along the spinal cord (B, E, H, I). These clusters present an irregular organisation. Additionally, they are progressively more densely packed towards the posterior and terminate abruptly just before the end of the spinal cord (K). Two transition zones of expression in the spinal cord of DAPT-treated embryos are observed (E and H) as explained in the text below. Abbreviations: f: caudal fin, n: notochord, sc: spinal cord, v: vitellum.

Two transition zones of expression in DAPT-treated embryos are observed within the spinal cord. At the anterior clusters are sparse, weakly stained, and they gradually increase in number, density and strength of staining at the middle of the body (Figure 8-10E). These clusters are found at different levels along the dorso-ventral axis. From this region, clusters gradually change to an even more compact pattern localised in the ventral half of the spinal cord (Figure 8-10H). Corresponding regions of these transition zones in DMSO control embryos show a homogenous expression pattern both in the anterior-posterior and dorso-ventral axes (Figure 8-10D and G, respectively). In one embryo, the first transition zone presents widely-spaced clusters, and clusters in the anterior part are only weakly stained (Figure 8-10I). On the other hand, clusters in the posterior part of this transition zone are strongly stained. Interestingly, both parts are separated by a large region with no staining (Figure 8-10I).

Towards the tail, expression of *LpNgn* does not reach the very end of the developing spinal cord both in control DMSO- and DAPT-treated embryos (Figure 8-10J and K). However, expression in DAPT-treated embryos stops abruptly as strongly-stained

clusters densely packed that reach the dorsal part of the spinal cord (Figure 8-10K). In contrast, expression in control-DMSO embryos is homogenous and stops gradually (Figure 8-10J).

LpCoe-A expression is clearly upregulated in the spinal cord in DAPT-treated embryos (Figure 8-11B and F) with respect to DMSO-control embryos (Figure 8-11A and E), whereas in the head expression seems unaffected (Figure 8-11C and D). Towards the tail, expression in the spinal cord is also much stronger in DAPT-treated than in control-DMSO embryos, and expression in both treatments fades gradually towards the tip (Figure 8-11G and H), in contrast to *LpNgn* whose expression stops abruptly in DAPT-treated embryos (cf. Figure 8-10K). From a dorsal view, expression of *LpCoe-A* in the spinal cord of control-DMSO embryos is seen as two parallel lateral stripes (Figure 8-11I and K), whereas expression in DAPT-treated embryos is broken from this pattern and seen with a somewhat homogenous distribution (Figure 8-11J and L). However, expression in DAPT-treated embryos gradually acquires a two-stripe pattern towards the posterior end of the spinal cord (Figure 8-11J), showing that the DAPT treatment affects expression of *LpCoe-A* mostly in the anterior part of the spinal cord as it elongates.

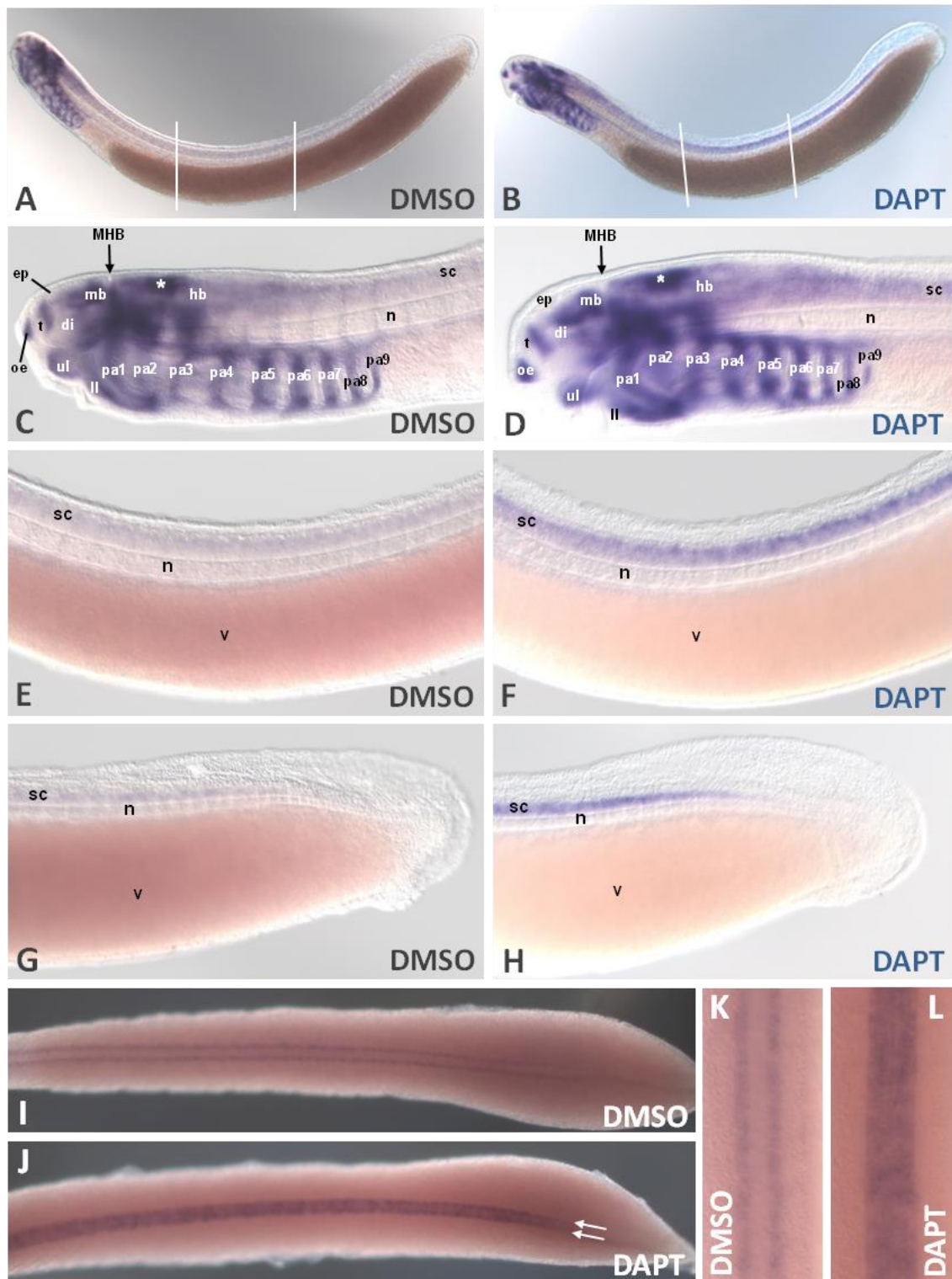


Figure 8-11. Expression of *LpCoe-A* in treated embryos. (A), (C), (E), (G), (I) and (K) are DMSO-control embryos. (B), (D), (F), (H), (J) and (L) are DAPT-treated embryos. (E) and (F) are photographs of corresponding regions in DMSO-control and DAPT-treated embryos as indicated by lines in (A) and (B), respectively. In all images anterior is to the left except in (K) and (L) in which anterior is to the top. In the nervous system, *LpCoe-A* is expressed in the brain and faintly in the spinal cord in DMSO-control embryos as in normal embryos (A, C, E, G; cf. Figure 6-19). Upon DAPT treatment, expression in the brain is generally upregulated and in the spinal cord it is considerably increased as compared to DMSO-control embryos (B, D, F, H). From a dorsal view, it can be observed that *LpCoe-A* is expressed in the spinal cord as two lateral stripes (I and K), whereas in DAPT-treated embryos expression is seen

relatively homogenous and expression expands to the middle of the spinal cord (J and L). However, in the newly-forming spinal cord at the posterior end, expression is seen as two lateral bands as in normal and DMSO-control embryos (J, arrows). Abbreviations: ep: epiphysis, di: diencephalon; f: caudal fin, hb: hindbrain; ll, lower lip; mb: midbrain, MHB: midbrain-hindbrain boundary, n: notochord; oe, olfactory epithelium; pa1-9, pharyngeal arch 1-9; sc: spinal cord, t: telencephalon; ul, upper lip; v: vitellum.

LpCoe-B responds to DAPT in a similar manner as *LpCoe-A*. *LpCoe-B* expression is upregulated in the head as well as in the spinal cord in DAPT-treated embryos with respect to DMSO-control embryos (Figure 8-12A-F), where from a very faint signal localised at the dorsal part of the spinal cord in control embryos, it becomes stronger in DAPT-treated embryos preserving the same dorsal position (Figure 8-12E and F). Towards the tail, from a faint signal biased to the dorsal part of the spinal cord in DMSO-control embryos, DAPT-treated embryos uniformly express *LpCoe-B* in the dorso-ventral axis reaching the end of the developing spinal cord (Figure 8-12G and H). From a dorsal view, similar to *LpCoe-A*, *LpCoe-B* expression in the spinal cord is observed as two parallel lateral stripes in DMSO-control embryos, whereas DAPT-treated embryos express *LpCoe-B* in a more homogenous pattern that covers the entire medio-lateral axis (Figure 8-12I-L). However, *LpCoe-B* expression is seen as slightly bigger clusters than *LpCoe-A*, generally arranged in a zigzag pattern with diffuse staining between them (compare Figures 8-11L and 8-12L). In contrast, *LpCoe-A* expression is slightly stronger and observed as only weakly discernible clusters with similar strong staining between them as compared to *LpCoe-B* (Figure 8-11L). Clusters of both genes are localised at the periphery of the spinal cord, in a position where both *LpCoe-A* and *LpCoe-B* are expressed in DMSO-control embryos (Figure 8-11I-L; Figure 8-12I-L).

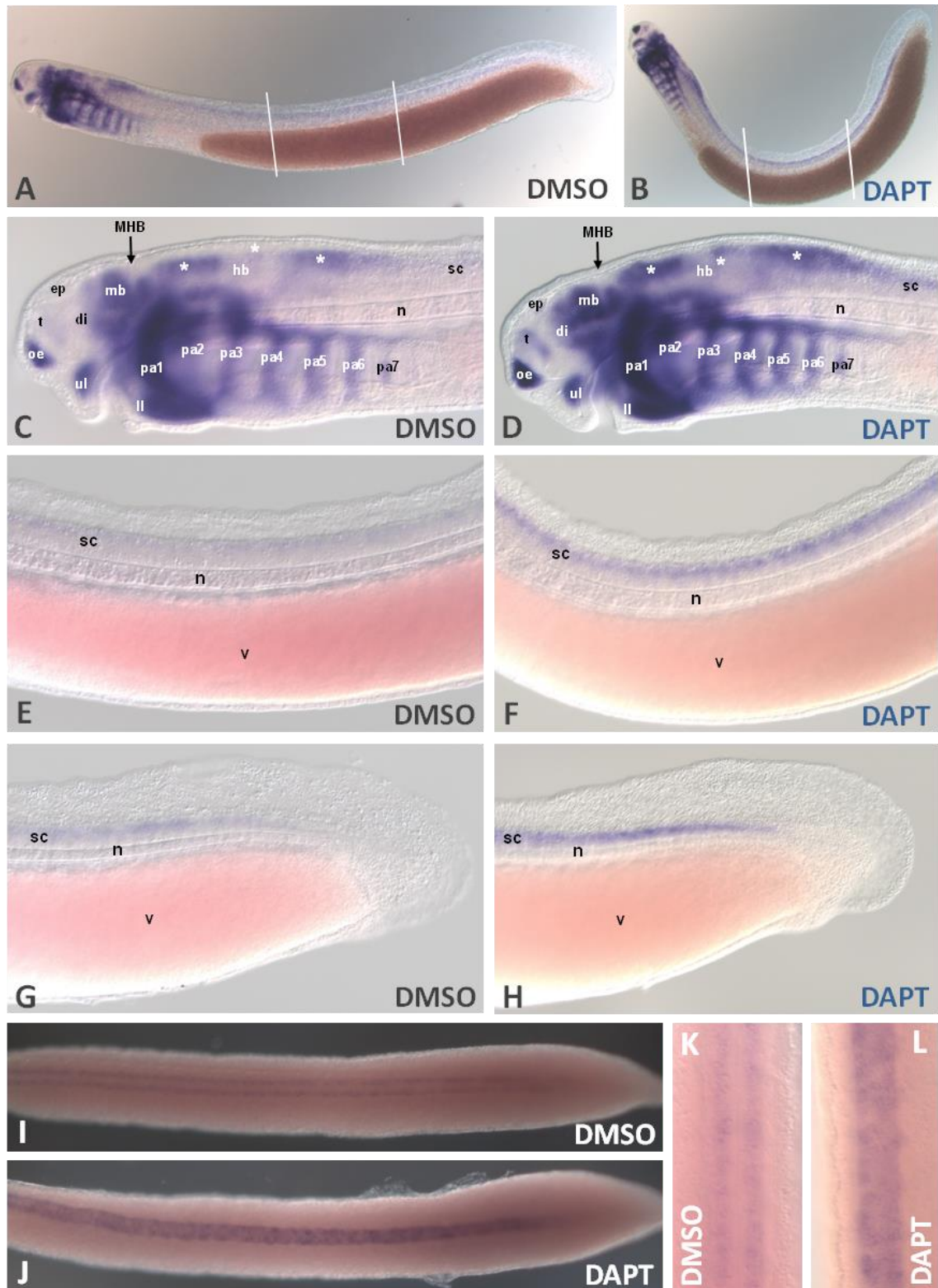


Figure 8-12. Expression of *LpCoe-B* in treated embryos. (A), (C), (E), (G), (I) and (K) are DMSO-control embryos. (B), (D), (F), (H), (J) and (L) are DAPT-treated embryos. (E) and (F) are photographs of corresponding regions in DMSO-control and DAPT-treated embryos as indicated by lines in (A) and (B), respectively. In all images anterior is to the left except in (K) and (L) in which anterior is to the top. In the nervous system, *LpCoe-B* is expressed in the brain and faintly in the spinal cord dorsally in DMSO-control embryos as in normal embryos (A, C, E, G; cf. Figure 6-23). After DAPT treatment, expression in the brain is generally upregulated and in the spinal cord it is considerably increased maintaining the same dorsal position as compared to DMSO-control embryos (B, D, F, H). From a dorsal view, it can be

observed that *LpCoe-B*, similar to *LpCoe-A*, is expressed in the spinal cord as two lateral stripes in DMSO-control embryos (I and K), whereas in DAPT-treated embryos, expression is seen relatively homogenous covering the entire medio-lateral axis (J and L). Abbreviations: ep: epiphysis, di: diencephalon; f: caudal fin, hb: hindbrain; ll, lower lip; mb: midbrain, MHB: midbrain-hindbrain boundary, n: notochord; oe, olfactory epithelium; pa1-7, pharyngeal arch 1-7; sc: spinal cord, t: telencephalon; ul, upper lip; v: vitellum.

When comparing the expression of all five cloned genes, namely, *LpHes-B*, *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B*, both in DAPT-treated and DMSO-control embryos, an interesting feature can be observed. Expression of all five genes in DAPT-treated embryos is virtually unchanged in the head with respect to DMSO-control embryos, except for a slight upregulation of *LpNgn*, *LpCoe-A* and *LpCoe-B*, and a slight downregulation of *LpHes-B* and *LpAsh* (Figure 8-8C, F; Figure 8-9, C, F; Figure 8-10C, F; Figure 8-11C, D; Figure 8-12C, D). At the hindbrain-spinal cord transition, however, it is clear that expression of *LpHes-B* and *LpNgn* changes radically in DAPT-treated embryos when compared to DMSO-control embryos. *LpHes-B* is completely downregulated from a certain point all along the spinal cord (Figure 8-13A, B), whereas *LpNgn* suffers a drastic change from a pattern of two bands, one dorsal and one ventral, into several spaced clusters of different sizes and intensities of staining (Figure 8-13E, F). Also, expression of *LpCoe-B* in DMSO-control embryos at the posterior hindbrain level is seen as a large patch of staining, and from this point towards the posterior it is seen as a faint band of staining in the dorsal part of the spinal cord (Figure 8-13I). However, in DAPT-treated embryos, *LpCoe-B* expression is conspicuously upregulated in the entire spinal cord with no apparent pattern change (i.e. as a dorsal band of staining) (Figure 8-13J). Such drastic changes in the expression pattern at the hindbrain-spinal cord transition level are not observed for *LpAsh* and *LpCoe-A*, even though there is a mild downregulation of *LpAsh* and a clear upregulation of *LpCoe-A* at more posterior levels (Figure 8-13C, D and G, H).

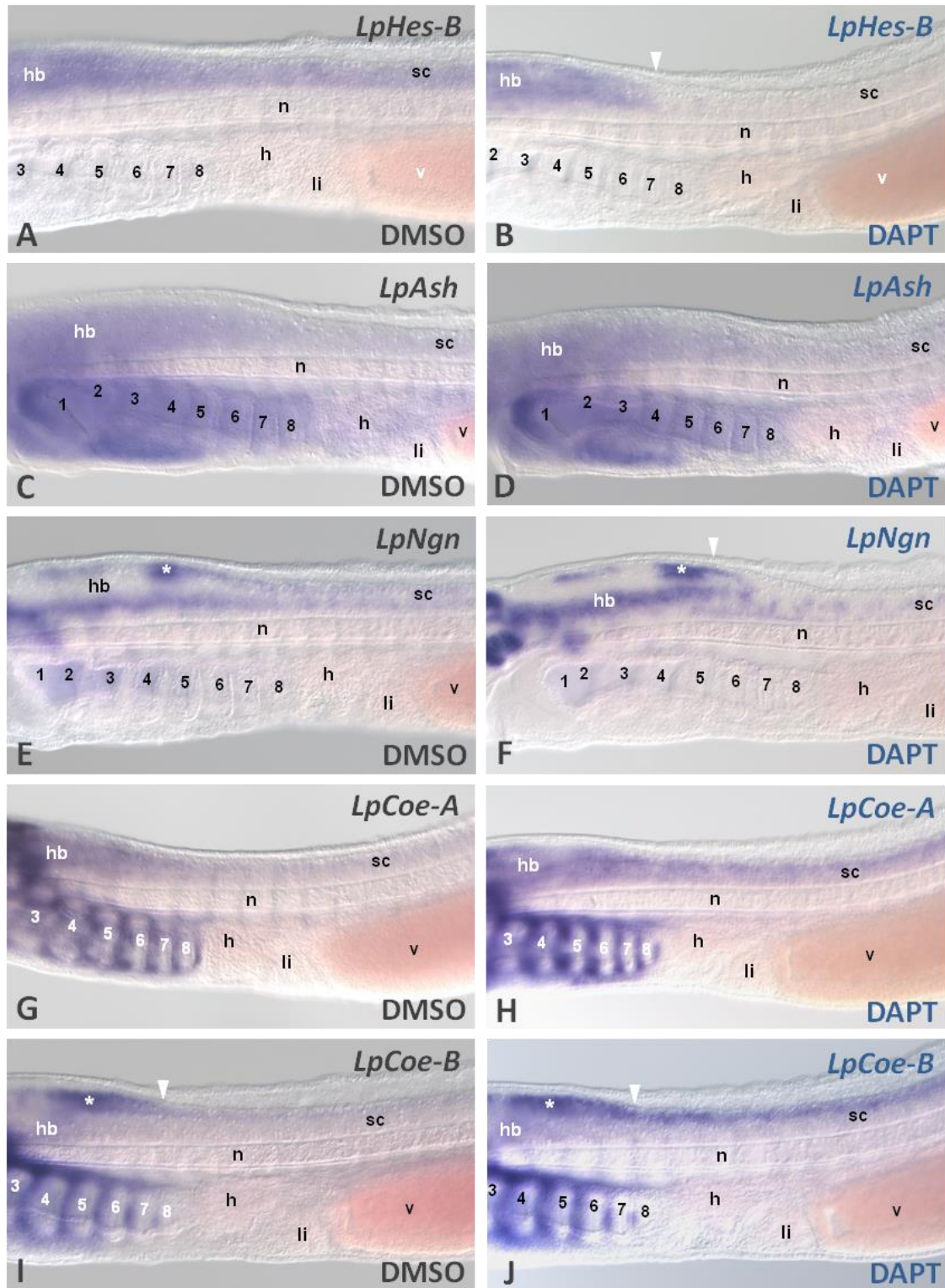


Figure 8-13. Expression of *LpHes-B*, *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B* in treated embryos at the hindbrain-spinal cord transition. (A), (C), (E), (G) and (I) are DMSO-control embryos. (B), (D), (F), (H) and (J) are DAPT-treated embryos. In all images anterior is to the left. Upon DAPT treatment, expression of *LpHes-B*, *LpNgn* and *LpCoe-B* changes radically at the hindbrain-spinal cord transition, from a position just after a big patch of expression of *LpNgn* and *LpCoe-B* (marked by an asterisk in E and F for *LpNgn*, and in I and J for *LpCoe-B*) both in DMSO- and DAPT-treated embryos. This position corresponds approximately to pharyngeal pouch 7 for *LpHes-B* and *LpCoe-B* (marked by an arrowhead in B for *LpHes-B*, and in I and J for *LpCoe-B*), and to pharyngeal pouches 5-6 for *LpNgn* (marked by an arrowhead in F). From this point towards the posterior, *LpHes-B* expression is extinguished, *LpCoe-B* is

notably upregulated preserving the dorsal expression domain, whereas *LpNgn* expression changes from two bands (one dorsal and one ventral) of expression into a series of clusters of staining. Expression of *LpAsh* or *LpCoe-A* does not show this marked expression pattern change. Probes for each gene are as indicated. Numbers indicate pharyngeal pouches 1-8. Abbreviations: h, heart; hb, hindbrain; li, liver; n, notochord; sc, spinal cord; v, vitellum.

Looking in more detail, *LpHes-B* expression changes at the level of pharyngeal pouch 7 upon DAPT treatment (Figure 8-13B, arrowhead). Similarly, *LpCoe-B* expression changes from pharyngeal pouch 7-10 (Figure 8-13I and J, arrowheads). However, expression of *LpNgn* changes more anteriorly at the level of pharyngeal pouch 5-6 (Figure 8-13F, arrowhead). Interestingly, the drastic changes of *LpNgn* and *LpCoe-B* expression occur just after a big patch of staining in the posterior hindbrain (Figure 8-13E, F and I, J, asterisks).

8.6. Discussion

8.6.1. Morphology of DAPT-treated Embryos

The most notable characteristic observed in DAPT-treated embryos is their backwards curvature associated with shorter length. Shorter embryos are also seen in zebrafish embryos treated with DAPT at 50 μ M (Arslanova et al., 2010; Yang et al., 2008, 2010). However, in contrast to these embryos, there was no evidence for disrupted somitogenesis in DAPT-treated lamprey embryos even at 100 μ M (Arslanova et al., 2010). In addition, DAPT does not seem to cause extensive apoptosis (Arslanova et al., 2010; Campbell et al., 2006). Shorter embryos are also observed in zebrafish embryos injected with morpholino oligonucleotides directed against *presenilin1* (Nornes et al., 2007). All together, alteration of the Notch pathway affects

axial growth both in lampreys and zebrafish, but, at least in lampreys, it does not seem to affect the general internal anatomy.

8.6.2. Expression Pattern Changes After DAPT Treatment

The effects of DAPT on the expression of the genes cloned in the present study follow some of the predictions based on data from gnathostomes. An advantage of this study is that the genes analysed are involved in different stages of neuronal differentiation. *Hes* genes are known to inhibit the neuronal fate by inhibiting the activity of the so-called proneural genes of the *ASC* and *Neurogenin* subfamilies (Ishibashi et al., 1995; Ohtsuka et al., 1999). They also help to maintain a neural stem cell state (Ohtsuka et al., 2001). On the other hand, *ASC* and *neurogenin* genes are able to specify ectoderm to a neuronal fate, while *COE* genes are involved in neuronal differentiation (Garcia-Dominguez et al., 2003). Therefore, expression analysis of these genes allows discerning at what point in neuronal differentiation the Notch pathway acts to regulate neural HLH gene expression.

***LpHes-B*:** As expected from expression of mouse *Hes1* and *Hes5* in *Notch1* mutants and *Notch* misexpression (Yang et al., 2006), expression of *LpHes-B* completely disappears in the spinal cord, further indicating that DAPT treatments affect the Notch pathway considering that *Hes* genes are the canonical and direct effectors of this pathway. This effect is also consistent with *LpHes-A* and *LpHes-B* functioning as inhibitors of neurogenesis in the lamprey spinal cord as suggested by sequence information and expression data in normal embryos (Chapter 4). Contrary to these –

inhibitory– *Hes* genes, mouse *Hes6* expression is downregulated by the Notch pathway, in agreement with its role in promoting neuronal differentiation instead of inhibiting neuronal differentiation and/or promoting a proliferative state of neuroblasts (Koyano-Nakagawa et al., 2000).

***LpAsh*:** In contrast to what has been observed with *Mash1* in mouse (de la Pompa et al., 1997; Yang et al., 2006), expression of *LpAsh* is slightly downregulated in the spinal cord, particularly at the posterior end (Figure 8-9).

***LpNgn*:** The intriguing expression pattern of *LpNgn* in DAPT-treated embryos could be correlated with mouse *Hes1;Hes5* double-mutants, two genes well-known to be direct Notch effectors. In these mutants, cell arrangement is severely disorganised and cells lose radial morphology and instead become round in shape (Hatakeyama et al., 2004). The only difference, however, is that this change in morphology and organisation in mouse *Hes1;Hes5* double-mutants occurs in the entire neural tube except in the telencephalon (Hatakeyama et al., 2004), whereas the effects observed in DAPT-treated *L. planeri* embryos seem to be specific to the spinal cord. Interestingly, no significant increase in cell death was observed in these mutants.

Zebrafish embryos injected with morpholino nucleotides directed against zebrafish *presenilin1* and *presenilin2* lead to an up-regulation of *neurogenin1* expression in the spinal cord (Campbell et al., 2006; Nornes et al., 2007, 2009). Zebrafish *presenilin2* is expressed in the ventricular zone of the spinal cord (Groth et al., 2002). However, in contrast to the spinal cord of other vertebrates and the head, *LpNgn* expression is generally downregulated in the spinal cord after treatment with DAPT, which surprisingly suggests that *LpNgn* might, directly or indirectly, instead be up-regulated by Notch signalling except in a particular neuronal population as discussed below.

***LpCoe-A* and *LpCoe-B*:** Both *LpCoe-A* and *LpCoe-B* expression considerably increases in the spinal cord. So far, a direct action of Notch signalling over expression of vertebrate *COE* homologues has only been analysed in *Xenopus* embryos (Dubois et al., 1998; Pozzoli et al., 2001). Co-injection of *Xcoe2* (a gene shown to be a promoter of neurogenesis; Dubois et al., 1998; Pozzoli et al., 2001) and Notch-intracellular domain (NICD) RNA in two-cell stage *Xenopus* embryos causes a complete inhibition of neurogenesis at neural plate stages, suggesting that *Xcoe2* is sensitive to Notch signalling (Pozzoli et al., 2001). On the other hand, co-injection of *Xebf3* (a gene also shown to drive neurogenesis) and NICD RNA leads to ectopic neurogenesis, suggesting that *Xebf3* is resistant to Notch signalling (Pozzoli et al., 2001). Conversely, injection of a dominant negative form of the Notch receptor ligand *Delta* causes an increase in *Xcoe2* expression in the neural plate, while *Xebf3* expression is unaffected (Pozzoli et al., 2001). Additionally, co-injection of *Xcoe2* and dominant negative *Delta* RNA causes ectopic neurogenesis (Dubois et al., 1998). These studies, however, are at neural plate stages, and it cannot be ruled out that *Xcoe2* and *Xebf3* response to Delta/Notch signalling might be different in the spinal cord of older embryos. In any case, the present study shows that both lamprey *COE* genes are sensitive to Notch signalling in the spinal cord in the same way as *XCoe2* at neural plate stages. Additionally, *Zcoe2* also seems to be sensitive to lateral inhibition as seen in the zebrafish mutants *white tail* (*wit*) which are believed to be deficient in lateral inhibition, a process led by Notch signalling (Bally-Cuif et al., 1998). Inhibition of COE/Ebf protein function by Notch signalling has also been observed in B-lymphocytes, where Notch signalling reduces the capacity of Ebf to activate target promoters and also impairs the ability of Ebf to bind DNA (Smith et al., 2005).

Particularly interesting is *LpCoe-B* expression. In control embryos, *LpCoe-B* is weakly expressed in the dorsal part of the spinal cord (Figure 8-12E). In DAPT-treated embryos, expression is highly upregulated preserving the same dorsal position (Figure 8-12F). These results suggest that DAPT treatments only affect expression levels but not cellular migration. Furthermore, expansion of *LpCoe-A* and *LpCoe-B* expression in the middle part of the neural tube (corresponding to the ventricular zone) in DAPT-treated embryos suggests premature neurogenesis in this region. This premature neurogenesis is comparable to mouse *Hes* mutants. Mouse *Hes1* or *Hes5* single-mutants, *Hes1;Hes5* double-mutants and *Hes1;Hes3;Hes5* triple-mutants display premature neurogenesis in a progressively more severe gradient, respectively (Ohtsuka et al., 1999; Hatakeyama et al., 2004). Taken together, the present study suggests that inhibition of Notch signalling in the lamprey spinal cord affects timing of neuronal differentiation without any obvious anatomical disruption.

8.6.3. Interneurons in the Lamprey Spinal Cord

An interesting aspect arises from expression of *LpNgn* in the spinal cord of DAPT-treated embryos. *LpNgn* transcripts are observed as big dots with an irregular distribution as seen from lateral view (Figure 8-10). This pattern is in sharp contrast with the prediction of a general up-regulation of *LpNgn* expression in the spinal cord upon inhibition of Notch signalling as has been observed in other vertebrate species (Geling et al., 2002; Nornes et al., 2009; de la Pompa et al., 1997; Yang et al., 2006). However, the resultant pattern has a great resemblance to the distribution of a particular type of interneurons, the dorsal longitudinal ascending (DoLA) interneurons, found in the zebrafish spinal cord (Nornes et al., 2009; Tamme et al., 2002; Wells et al., 2010).

(Figure 8-14). DoLA interneurons have a seemingly irregular distribution along the rostrocaudal axis of the spinal cord and are specifically marked by *tbx16/spadetail* (Nornes et al., 2009; Tamme et al., 2002). DoLA interneurons migrate anteriorly shortly after their birth in the posterior spinal cord, and therefore their distribution is more compact in the posterior-most spinal cord (Nornes et al., 2009; Wells et al., 2010), which is very similar to *LpNgn* expression in the posterior spinal cord in DAPT-treated embryos (Figure 8-10E, H, I, K).

A direct implication of Notch pathway components in DoLA interneuron development has been documented. Inhibition of *presenilin2*, but not *presenilin1*, in zebrafish leads to an increase in the number of DoLAs (Nornes et al., 2009). However, when splicing of *Presenilin-1* is affected, an increase in DoLA cell numbers is observed (Nornes et al., 2007). The expression pattern of *LpNgn* in DAPT-treated embryos seen as big, strongly stained clusters (which are not observed in normal or control embryos) might be indicative of such an increase in the number of this particular neuronal population. Although there is no further evidence, expression of *LpNgn* in DAPT-treated embryos raises the intriguing possibility of the presence of a particular neuronal population in the lamprey spinal cord, potentially homologous to zebrafish DoLAs, which is revealed when all other *LpNgn* expression domains in the spinal cord are cleared upon DAPT treatment. These results also suggest that Notch signalling does not inhibit *LpNgn* expression in the lamprey spinal cord as in gnathostomes, and instead Notch signalling may promote it such that after application of DAPT *LpNgn* expression is reduced. This regulation, however, might not occur in lamprey DoLA-like cells, where a canonical regulation of *LpNgn* expression by Notch signalling is present (that is, Notch signalling inhibits *LpNgn* expression in these cells).

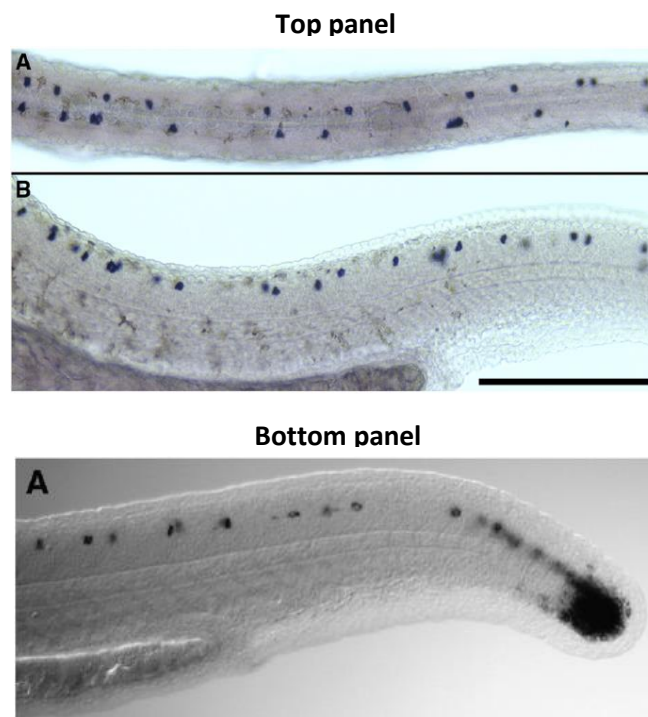


Figure 8-14. Expression of *tbx16* in zebrafish embryos marking dorsal longitudinal ascending (DoLA) interneurons. Top panel, (A, B) Trunk and tail region of a 30 hours post fertilisation (hpf) embryo stained to show the distribution of *tbx16*-expressing cells (DoLA interneurons) from dorsal (A) and lateral (B) views. Rostral is to the left and dorsal is up in B (scale bar: 250 μ m). Bottom panel, (A) Trunk and tail of an embryo at ~24 hpf after in situ hybridisation to detect *tbx16* transcripts in DoLA interneurons in the developing spinal cord. *tbx16* is also transcribed in presomitic mesoderm in the tail tip. Top panel, (A, B) taken from Wells et al. (2010); Bottom panel, (A) taken from Nornes et al. (2009).

8.6.4. The Lamprey Hindbrain-Spinal Cord Transition

An interesting aspect that arises from this study is that DAPT treatments seem to affect mostly the spinal cord, as expression in the head is only mildly either downregulated (*LpHes-B*), or upregulated (*LpNgn*, *LpCoe-A* and *LpCoe-B*), but the expression pattern is maintained overall without any radical change. This suggests an important difference in the sensitivity of neural HLH gene expression to Notch signalling between the brain and spinal cord in the time range of treatment. A possible distinction of brain versus spinal cord sensitivity to Notch signalling is observed in mice double-null mutants for *Presenilin-1* and *Presenilin-2*, in which *Shh* and *Nkx2.2*

expression in the ventral neural tube is absent from the trunk but preserved in the head, in contrast to normal expression in the ventral brain and spinal cord (Donoviel et al., 1999). These authors suggested that the effects of *Presenilin* loss are restricted to the ventral neural tube in the trunk region. In addition, a separate spinal cord developmental programme has been observed with expression of *Cdx* genes in zebrafish. Zebrafish *Cdx1a* and *Cdx4* are able to promote spinal cord development at the expense of the hindbrain by preventing formation of rhombomeres, and they also induce the development of spinal cord cell types and spinal cord-specific genes (Skromne et al., 2007). Therefore, in light of the differential sensitivity of neural HLH gene expression between the brain and spinal cord after interruption of Notch signalling, the present study suggests that lampreys have an independent spinal cord developmental programme and that this is an ancestral feature of vertebrates.

Finally, an interesting correlation of the anterior extent of Notch sensitivity of *LpHes-B*, *LpNgn* and *LpCoe-B* in the spinal cord emerges when compared with the anterior expression boundaries of lamprey *Hox* genes. The expression pattern of *LpHes-B* and *LpCoe-B* in the brain, virtually unaffected after DAPT treatment, seems complementary to expression of lamprey *Hox8p* and *HoxQ8* in the neural tube, taking the pharyngeal pouches as a reference (Takio et al., 2004, 2007). On the other hand, the generally unaffected expression pattern of *LpNgn* in the brain seems complementary to the strongest expression of lamprey *Hox4w* (Takio et al., 2004, 2007). However, any possible functional interaction between *Hox* genes, neural HLH genes and Notch signalling in the spinal cord is an open question.

8.7. Summary of Lamprey Notch Signalling

Developing lamprey embryos were treated with the Notch pathway inhibitor DAPT from stage 24 to stage 26 to study the expression of *LpHes-B*, *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B* in the stage 26 spinal cord. DAPT-treated embryos showed a consistent shorter length and backwards bending of the body. Expression of *LpHes-B* was abolished in the spinal cord, suggesting that these treatments effectively inhibit Notch signalling in the developing lamprey spinal cord. In contrast, both *LpCoe-A* and *LpCoe-B* were upregulated in the spinal cord, and their expression expanded into the ventricular zone, suggesting premature neurogenesis of neuroblasts. Expression of *LpCoe-B* in DAPT-treated embryos suggests that DAPT does not affect cellular migration but the timing of neuronal differentiation. Unexpectedly, expression of *LpNgn* was not upregulated in the spinal cord, and instead it was downregulated except in small patches, possibly corresponding to dorsal longitudinal ascending (DoLA) interneurons. Expression of *LpAsh* was downregulated in a small region of the posterior-most spinal cord but otherwise unchanged. Expression of all genes in the brain was virtually unaffected in DAPT-treated embryos, suggesting that, at least in the time-frame analysed, Notch signalling has a minor role in the regulation of the expression of the analysed genes. This additionally suggests that the spinal cord has an independent developmental programme with respect to the brain. Taken together, these results demonstrate that the Notch signalling pathway is active in the lamprey spinal cord and suggest that it inhibits neuronal differentiation as in gnathostomes.

9. DISCUSSION

The present study improves our understanding of the genetic content of the lamprey HLH gene family. Together with their expression patterns, a better picture of the general anatomy, composition, and divisions of the lamprey nervous system is revealed, along with particular developmental processes that regulate its development. This gets us closer to the structure of the ancestral vertebrate nervous system state.

9.1. The Lamprey HLH Gene Family

Diversity of lamprey HLH genes follows the predictions. The lamprey genome contains a wide variety of HLH gene subfamilies belonging to the traditional groups A, B, C, D, E and F (Table 3-4, Chapter 3). However, analysis of the available information from the lamprey Ensembl genome database, at its current status, does not identify other HLH gene subfamilies such as *Beta3*, *Net*, *AP4*, *Sim* or *Hey* (Table 3-5, Chapter 3), very likely to be present in the lamprey genome as they are found in many chordates and animal phyla (Liu and Zhao, 2010; Satou et al., 2003; Simionato et al., 2007, 2008; Wang et al., 2009). Among others, these subfamilies include *Oligo* and *NSCL*, which are also involved in the development of the nervous system in gnathostomes (Krüger et al., 2006; Li and Richardson, 2008; Li et al., 1999a, 1999b; Ligon et al., 2006; Schmid et al., 2007; Takebayashi et al., 2000).

Therefore, even though a wide variety of HLH gene subfamilies were identified in the lamprey genome in the present study with the use of the Ensembl genome database (Flicek et al., 2014) for *Petromyzon marinus_v7*, and some members of them already experimentally described such as *Ascl1* (Häming et al., 2011), *Neurogenin1* and *NeurogeninA* (Guérin et al., 2009; Sauka-Spengler et al., 2007), *NeuroD2* (Guérin et al., 2009), *Parascleraxis* (Freitas et al., 2006), *Id* and *Id2/3* (Guérin et al., 2009; Meulemans et al., 2003; Sauka-Spengler et al., 2007), *n-Myc* (Sauka-Spengler et al., 2007), *TwistA-C* (Sauka-Spengler et al., 2007), other subfamilies still remain to be identified. Of note, although they were not analysed in much detail, *Neurogenin1* (Guérin et al., 2009) and *NeurogeninA* (Sauka-Spengler et al., 2007) seem to present different expression patterns, suggesting that they are different genes, whereas *Ascl1* (Häming et al., 2011) also seems to be a different gene from *LpAsh* (this study), and *Id2/3* (Guérin et al., 2009) was reported to be a different gene from *Id* (Meulemans et al., 2003). Only one gene prediction was retrieved in the present study for each subfamily (i.e. *Neurogenin*, *ASC* and *Emc*, respectively), suggesting that even within subfamilies that were identified by BLAST searches in the present study, more members could be identified experimentally. Furthermore, *Hand*, a member of the *Hand* subfamily which was not retrieved by BLAST searches in the present study (see Table 3-5, Chapter 3), has been independently identified and experimentally analysed by other groups (Jandzik et al., 2014; Onimaru et al., 2011; Tulenko et al., 2013), lending support to the notion that more lamprey HLH genes will very likely be discovered in the future.

9.2. Lamprey Neural HLH Genes in the CNS

9.2.1. Brain Territories

Expression patterns of the cloned genes, in particular *LpCoe-A*, *LpCoe-B* and *LpNgn*, highlight the regionalised nature of the lamprey brain subdivided into telencephalon, diencephalon, mesencephalon (although not very clear with data from the present study), and hindbrain. Expression of *LpNgn* (Figure 5-7F) demarcates territories in the diencephalon that are consistent with the prosomeric model to some extent. *LpNgn* demarcates regions that could correspond to prosomere 1 (P1) and P2 in the diencephalon. P3 is not easily demarcated when compared to mouse *neurogenin2* (Osório et al., 2010), in particular considering the region delimited by the zona limitans intrathalamica (Figure 5-7F, white asterisk). Expression of *LpNgn* in the lamprey diencephalon thus suggests the absence of P3 at the embryonic stages analysed. The suggested absence of P3 is based on 1) the absence of an *LpNgn* expression domain anterior to the domain delimited by the zona limitans intrathalamica (white asterisk in A), which is observed with mouse *neurogenin2* in the basal part of P3 (Figure 2B in Osório et al., 2010); 2) the position of the epiphysis dorsal to the dorsal thalamus confirming the extent of P2; and 3) a large, unstained ventral domain corresponding to the hypothalamus. However, both a P3 and a prethalamus (that corresponds to the alar part of P3) have been described in the adult lamprey brain with some difficulty, although they are comparatively thinner with respect to P1 and P2 (Pombal and Puelles, 1999; Pombal et al., 2009). In this updated model of embryonic lamprey brain subdivisions, *LpNgn* expression highlights the secondary prosencephalon, which is composed of the hypothalamus and telencephalon (Figure 9-1, dark blue region). These

forebrain subdivisions are more consistent with the lamprey prosomeric model described by Murakami and colleagues (Murakami et al., 2001, 2002; see also Villar-Cheda et al., 2005) than that described by Osório and colleagues (Osorio et al., 2005), particularly considering the position of the epiphysis dorsal to P2 (i.e. dorsal to the dorsal thalamus).

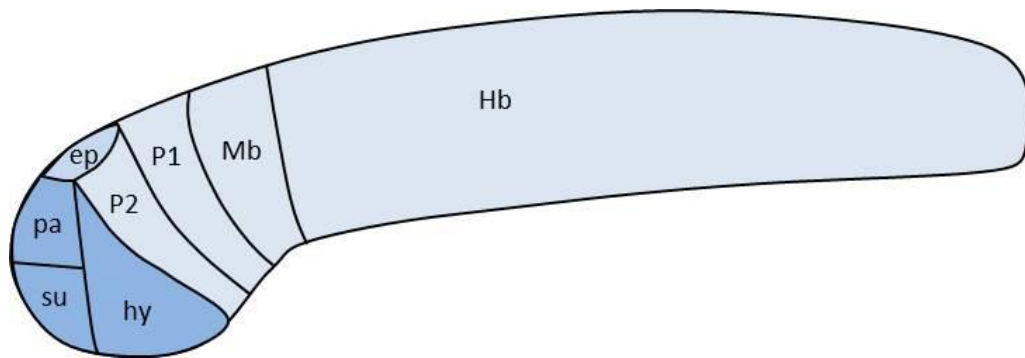


Figure 9-1. Prosomeric model of the embryonic lamprey brain. Schematic representation of the *Lampetra planeri* brain following expression of *LpNgn* at stage 27 and based on mouse *neurogenin2* expression (Osório et al., 2010), highlighting the extent of prosomere (P)1, P2, and the apparent absence of P3 in all stages examined in this study. The alar part of P1 contains the pretectum, while the alar part of P2 contains the dorsal thalamus. In this updated model, the secondary prosencephalon is identified (dark blue) by *LpNgn* expression which contains the entire telencephalon and the hypothalamus. A vertical *LpNgn* expression domain at the telencephalon-diencephalon boundary (Figure 5-7F) may correspond –at least in part– to the eminentia thalami, supraoptic/paraventricular area and anterior hypothalamus as observed with mouse *neurogenin2* in (Osório et al., 2010). *LpNgn* expression does not reveal a clear boundary between the pallium and subpallium in the telencephalon, and this boundary is shown somewhat arbitrarily. Abbreviations: AH, anterior hypothalamus; ep, epiphysis; et, eminentia thalami; Hb, hindbrain; hy, hypothalamus; Mb, midbrain; P1, prosomere 1; P2, Prosomere 2; pa, pallium; SPV, supraoptic/paraventricular area; su, subpallium.

9.2.2. Spinal Cord

Expression of *LpHes-B*, *LpAsh*, *LpNgn*, *LpCoe-A* and *LpCoe-B* reflect active neurogenesis in the developing spinal cord by stage 24 in an anterior to posterior gradient as the embryo elongates. Additionally, expression of *LpHes-B* and *LpNgn* highlights the presence of a ventricular zone all along the neural tube. Previous studies

with antibodies for PCNA have recognised a layer of proliferating cells next to the lumen of the embryonic, larval, and adult lamprey brain (Villar-Cheda et al., 2005, 2006). Subsequent molecular expression studies of a number of proliferation markers such as *PCNA* and *Musashi* and developmental genes have also revealed this layer of proliferating cells (Guérin et al., 2009).

On the other hand, expression of *LpCoe-A* and *LpCoe-B* shows the presence of a complementary marginal zone of post-mitotic neurons, although they do not demarcate it entirely along the dorso-ventral axis. However, with the data obtained in the present study it is difficult to discern whether the lamprey neural tube also possesses a subventricular zone (SVZ), which could be discerned by expression of –for example– *NeuroD* or *NSCL* homologues as in rodents and chicken (Koyano-Nakagawa et al., 2000; Ma et al., 1997; Perez et al., 1999; Theodorakis et al., 2002). Interestingly, chicken *neurogenin1* and *neurogenin2* (Perez et al., 1999), as well as mouse *COE2* (Garel et al., 1997a), are expressed in the SVZ at some point in the development of the spinal cord. However, neither lamprey *COE* genes nor *LpNgn* show a narrow expression pattern reminiscent of a SVZ in the stages and sections analysed. Still, the presence of a SVZ is inferred as this zone is defined by migrating, post-mitotic neuroblasts starting to differentiate, and the identification of a well-defined ventricular zone and a complementary marginal zone suggests that cells in the lamprey spinal cord follow a similar horizontal migratory pattern from the ventricular zone to the marginal zone as they differentiate as occurs in jawed vertebrates. Overall, expression of all cloned genes corroborate the anatomical similarity of the lamprey spinal cord –as seen in cross-sections– with that of gnathostomes.

9.3. Lamprey Neural HLH Genes in the PNS

The present study sheds more light on the distribution and development of lamprey cranial sensory placodes and their respective ganglia. Expression of *LpNgn* at stage 21 in the anterior tip of the forming head suggests that this part of the ectoderm is already specified as the presumptive NHP territory, while the NHP has been morphologically observed only from stage 23 (Kuratani et al., 2001). This makes *LpNgn* the earliest marker of presumptive NHP territory in the anterior lamprey head ectoderm.

Comparison of the expression patterns of *LpHes-B*, *LpNgn*, *LpCoe-A* and *LpCoe-B* highlight both the placodal and sensory ganglia repertoire of the lamprey head. In particular, expression of *LpCoe-B* at stages 24 and 25 marks with relative clarity the trigeminal ophthalmic, trigeminal maxillomandibular, facial and glossopharyngeal ganglia. The fact that *LpCoe-B* expression is seen in the forming cranial sensory ganglia and not at their placodal stage is consistent with the role of *COE* genes in neuronal differentiation and cell cycle exit (Dubois et al., 1998; Garcia-Dominguez et al., 2003; Garel et al., 1999; Green and Vetter, 2011; Pozzoli et al., 2001).

Although *neurogenin* genes are involved in neuronal differentiation in some instances (Ali et al., 2011; Lacomme et al., 2012; Mizuguchi et al., 2001; Novitch et al., 2001), they are also involved in neuronal specification of ectodermal cells such as *Xenopus neurogenin2* (Ma et al., 1996) or zebrafish *neurogenin1* (Blader et al., 1997), and they generally predate expression of differentiation genes such as *NeuroD* (Andermann et al., 2002; Abu-Elmagd et al., 2001; Ma et al., 1998; Sommer et al.,

1996) or *COE* (Burns and Vetter, 2002; Garel et al., 1997a) in cranial placodes/ganglia and in the neural tube. Consistent with these roles, *LpNgn* was observed more strongly at the placodal stage marking epibranchial placodes and weakly in epibranchial ganglia, in agreement with a general role of vertebrate *neurogenins* at early stages of neuronal differentiation.

An important difference that is seen with *LpNgn* is that not all lamprey placodes are marked by *LpNgn*, namely the otic placode, which suggests either the presence of an as yet unidentified lamprey *neurogenin* homologue or that some lamprey placodes use genes of other transcription factor families in their developmental programmes other than *neurogenin* genes as in gnathostomes, in which case would be a novel feature in vertebrate species analysed so far at this level.

One of the main contributions of the present study is the marking of DRG by *LpNgn*. This expression not only shows that, as in gnathostomes, *neurogenin* homologues are expressed in DRG, but also shows the sequence of differentiation of DRG from an anterior to posterior gradient. In addition, *LpNgn* is not expressed in migrating neural crest cells as chicken *neurogenin1* and *neurogenin2*, and is expressed in the condensed DRG.

Significant differences both at the genetic and anatomical level in the nervous system have been observed with the study of lamprey ASC homologues with respect to gnathostomes. As mentioned before in section 1.4.4. in the Introduction, there is no

clear evidence for the presence of sympathetic ganglia in lampreys, expected to be located above the vitellum and below the notochord. *LpAsh* does not show a expression pattern that resembles any sort of ganglia in this region, in agreement with previous reports suggesting that lampreys do not have sympathetic ganglia (Häming et al., 2011). *LpAsh* does not show any expression in enteric neurons either, and enteric ganglia have been described in lampreys (Häming et al., 2011). Conclusions from the present study point to the fact that, from lampreys to gnathostomes, a significant evolutionary change involving ASC genes appeared in the gnathostome autonomic nervous system.

9.4. Regulation of Lamprey Neural HLH Genes by the Notch Signalling Pathway

DAPT treatments revealed overall conservation in the regulation of neural HLH genes analysed in the present study by the Notch signalling pathway. The clearest example of the specificity of DAPT treatments was the complete downregulation of *LpHes-B* expression in the spinal cord. Additionally, the clear upregulation of both *LpCoe-A* and *LpCoe-B* corroborated this specificity. To our knowledge, this is the first time that DAPT has been used in lampreys to study the role of the Notch signalling pathway in the nervous system. However, it cannot be ruled out that DAPT treatments could also interfere with the amyloid precursor protein (APP) pathway, as DAPT acts on the γ -secretase complex which also cleaves APP (Kaether et al., 2006; Tolia and De Strooper, 2009). However, observation of the predicted effects of DAPT treatments on the expression of neural HLH genes strongly favours the specificity of this treatment to inhibit the Notch pathway. Considering this, it is clear that the lamprey spinal cord employs the Notch signalling pathway in a similar manner as in gnathostomes, though it

presents some conspicuous differences concerning the expression of *LpAsh* and *LpNgn*. *LpAsh* does not suffer any major change in its expression apart from a very mild up-regulation at the posterior, and it was expected to be upregulated based on data from gnathostomes (de la Pompa et al., 1997; Yang et al., 2006); on the other hand, expression of *LpNgn* changed from a relatively homogenous pattern into separate clusters of expression instead of being homogenously upregulated as in gnathostomes (Geling et al., 2002; Nornes et al., 2009; de la Pompa et al., 1997; Yang et al., 2006). This suggests a step by step mechanism in the evolution of Notch signalling in the vertebrate spinal cord by gradually modulating the expression of subsets of neural HLH genes towards a gnathostome “canonical” mechanism, although it cannot be ruled out that lampreys display a profile of Notch signalling in the spinal cord not representative of ancestral vertebrates.

9.5. New Aspects of the Lamprey Nervous System Revealed by the Present Study

From all the genes analysed in the present study, *LpNgn* is the most revealing with respect to the anatomy of the lamprey nervous system and showed novel features not previously recognised. As explained in section 9.2.1., *LpNgn* might clarify the prosomeric subdivisions of the lamprey brain with the demarcation of the secondary prosencephalon and the suggestion of absence of prosomere 3. Also, as mentioned in section 9.3., *LpNgn* marks a domain in the anterior-most ectoderm of the head which suggests that the presumptive NHP is already specified at least since stage 21, even though it is morphologically recognised until stage 23.

Another novel feature revealed by *LpNgn* expression is the detection of putative sensory placodes in the middle part of each pharyngeal arch. These putative sensory placodes may represent novel “branchial placodes” which could produce lateral line cells in the larva. However, another possibility is that they may represent neurogenic hypobranchial placodes. Two hypobranchial placodes have been identified in *Xenopus laevis* and in the direct-developing frog *Eleutherodactylus coqui* but not in other vertebrates (Schlosser, 2003; Schlosser and Northcutt, 2000; Schlosser et al., 1999). These placodes are located ventral and slightly caudal to the second and third pharyngeal pouches, and in *Xenopus* it has been shown that they give rise to hypobranchial ganglia (Schlosser, 2003; Schlosser and Northcutt, 2000; Schlosser et al., 1999). However, given that the amphibian hypobranchial placodes are located ventral to the pharyngeal pouches, whereas the putative lamprey “branchial placodes” are located in the middle of each pharyngeal arch, together with the fact that so far no hypobranchial ganglia have been described in lampreys, it seems more likely that the putative lamprey “branchial placodes” are indeed novel sensory placodes.

LpNgn also serves as a novel molecular marker for epibranchial placodes. Lamprey epibranchial placodes, as such, have not been identified by using molecular markers, although they have been recognised morphologically (Damas, 1944). Even though in the present study epibranchial placodes were not analysed morphologically, it is clear that *LpNgn* is a purely neuronal marker, and considering the expression of *neurogenin* genes in the sensory neuron lineage of the PNS in gnathostomes, it is highly unlikely that *LpNgn* will mark a cell population different than a sensory neuron lineage in the lamprey head ectoderm. Thus, its expression dorsal to pharyngeal arches at stages 23-24 (Figure 5-6D, E, G, H) can be confidently taken as staining of epibranchial placodes,

either at the purely placodal stage or also cells just starting to delaminate. Overall, this makes *LpNgn* the first molecular marker of epibranchial placodes in lampreys.

Similarly, *LpNgn* specifically marks DRG. Lamprey DRG have been marked by DiI (Häming et al., 2011), and by acetylated tubulin (Kuratani et al., 1997) and Hu antibodies (Modrell et al., 2014). However, the staining of *LpNgn* can eventually allow making developmental comparisons in terms of neural crest cell migration (that eventually will condense into DRG) or cell type composition (i.e. neurons versus satellite glial cells) of the DRG. As markers for migrating trunk neural crest cells in lampreys (and possibly for DRG glial cells) emerge, *LpNgn* will constitute an excellent reference point for developmental comparisons.

Combined expression of *LpHes-B*, *LpNgn*, *LpCoe-A* and *LpCoe-B* delineate the ventricular and marginal zones. While the embryonic ventricular zone had been recognised by different markers (Guérin et al., 2009), the expression of *LpHes-B*, *LpNgn*, *LpCoe-A* and *LpCoe-B* offer a possibility to obtain a clearer picture of cell populations in the lamprey neural tube and different stages of neuronal differentiation based on their well-established developmental roles in gnathostomes. Finally, DAPT treatments not only revealed for the first time that, similar to gnathostomes, lamprey neural HLH genes are regulated by Notch signalling in the spinal cord, but they also suggest that lampreys have an independent spinal cord developmental programme, a hypothesis that has not been put forward before in lampreys.

9.6. The Ancestral Vertebrate Spinal Cord: A Hypothesis

Cephalochordates and urochordates are the closest living relatives to vertebrates (Delsuc et al., 2006). Their neural tubes, however, are much simpler and contain far fewer cells than those of vertebrates. The ascidian larval central nervous system is composed of the sensory vesicle, the neck, the visceral or motor ganglion, and the tail nerve cord (Lemaire et al., 2002; Meinertzhagen et al., 2004). The *Ciona intestinalis* CNS consists of only 350 cells, of which approximately 80 are neurons (Yamada et al., 2009). Based on morphology and *Hox5* gene expression, the tail nerve cord is considered to be the equivalent of the vertebrate spinal cord (Wada and Satoh, 2001). However, the ascidian tail nerve cord is thought to be composed only of ciliated ependymal cells, distributed in a row of ventral keel cells, left and right lateral rows, and a dorsal row of capstone cells, thus four cells in cross-section, although this might be a derived feature (Lemaire et al., 2002; Meinertzhagen et al., 2004; Wada and Satoh, 2001). On the other hand, the amphioxus larval central nervous system is a tubular nerve cord with any external morphological indications of segmentation. It presents a transient anterior swelling called the cerebral vesicle, while the caudal end of the nerve cord forms a terminal enlargement called the caudal ampulla (Benito-Gutiérrez, 2006; Wicht and Lacalli, 2005). Similar to *C. intestinalis*, the anterior limit of expression of *Hox5* in the amphioxus neural tube suggests a region homologous to the vertebrate spinal cord from the level of the eighth somite to the posterior end (Wada and Satoh, 2001). Aside from the low cell number, a medial zone of neural progenitors and a complementary peripheral zone of differentiated neurons appear to be lacking in either amphioxus or tunicates.

Expression of members of the Notch signalling pathway and neural HLH genes has been analysed in ascidians and cephalochordates. In *Ciona intestinalis*, *Notch* is strongly expressed in the neural tube including the tail nerve cord at mid-tailbud stages, but starts to be downregulated at late-tailbud stages (Imai et al., 2004; Yamada et al., 2009). However, *Delta* does not seem to be expressed in the tail nerve cord at late-tailbud stages (Imai et al., 2004). *Hes-a* might be expressed in the tail nerve cord at mid/late-tailbud stages, but *Hes-b* and *Hes-c* seem not to be expressed in the tail nerve cord (Imai et al., 2004). *Achaete-Scute a-like1* and *Achaete-Scute a-like2* (*Achaete-Scute b* not analysed) are not expressed in the tail nerve cord, whereas *Neurogenin* is strongly observed in the anterior nerve cord at mid/late-tailbud and also in separate clusters posteriorly at early-tailbud (Imai et al., 2004). Finally, *C. intestinalis* *COE* is expressed in a small number of neural cells, mostly in the sensory vesicle and with little or no expression in the tail nerve cord (Imai et al., 2004; Mazet et al., 2005).

In amphioxus, the expression of *Notch*, *Delta*, *Hairy*, *Ngn* and *COE* genes has been analysed and found in the neural tube. *AmphiNotch* is strongly expressed in the neural tube at early neurula stages, but is subsequently downregulated at 28 hours at late neurula stages (Holland et al., 2001). Very mild, if any, expression of *AmphiDelta* is observed in the region equivalent to the vertebrate spinal cord at neurula and larval stages (Rasmussen et al., 2007). *Amphi-hairyA* is strongly expressed in the posterior third of the neural tube at neurula stages (Minguillon et al., 2003). *Amphi-hairyB* expression is strongly observed more anteriorly than *Amphi-hairyA* in the neural tube in neurula-stage embryos, and an additional –separated– domain is observed near the anterior end of the neural tube (Minguillon et al., 2003). *Amphi-hairyC* and *Amphi-hairyD* are observed in separate domains in the neural tube at neurula stages similar to

Amphi-hairyB (Minguillon et al., 2003). Later at larval stages, expression of *Amphi-hairyA* persists in the posterior neural tube although at a lower level (Minguillon et al., 2003). The anterior expression domain of *Amphi-hairyB* persists somewhat strong but the posterior domain diminishes in intensity (Minguillon et al., 2003). However, *Amphi-hairyC* and *Amphi-hairyD* are no longer observed in the neural tube at larval stages (Minguillon et al., 2003). *Amphi-hairyE-H* seem not to be expressed in the embryo or larvae or expressed at undetectable levels (Minguillon et al., 2003). Finally, *AmphiNeurogenin* is expressed in scattered cells of the neural tube at neurula and larval stages (Holland et al., 2000), similar to *AmphiCoe* (Mazet et al., 2004).

Taken together, a coherent expression of Notch pathway components (i.e. *Notch*, *Delta* and *Hairy/E(spl)* genes), together with a subset of neural HLH genes regulated by this pathway in vertebrates (e.g. *Achaete-scute*, *Neurogenin* and *COE* genes), is not observed in protochordate neural tubes as the region equivalent (possibly homologous) to the vertebrate spinal cord develops.

The protochordate region equivalent to the vertebrate spinal cord is thus markedly different both at the anatomical and at the gene expression pattern levels with respect to Notch signalling and neural HLH genes. In contrast, the vertebrate spinal cord is a far more complex structure than the equivalent protochordate neural tube. Its development involves the production of a very large number of different cell types which become well organised into a medial proliferative zone (the ventricular zone) and a peripheral zone (the marginal zone) of differentiated neurons. As explained in Chapter 8, the Notch signalling pathway is active and deployed in the ventricular zone all along the gnathostome spinal cord together with a set of neural HLH genes for a relatively long

time period. Even in lampreys, expression of neural HLH genes in the spinal cord is present very similar to gnathostomes, and they are also regulated by Notch signalling (though with some marked differences, namely, *LpNgn* and *LpAsh* expression), which is also present in the ventricular zone. Thus, the evolution of the vertebrate spinal cord might have required the emergence of a Notch-regulated medial progenitor zone together with a cohesive battery of neural HLH genes expressed in this zone or their derivatives. This zone was maintained for a much longer developmental time with respect to protochordate ancestors. Maintaining neural progenitor cells of the neural tube for longer periods of time might have provided a pool of precursor cells that eventually yielded larger spinal cords after neural tube closure, with more cells and more potential for sensory integration and motor control. This would underline the morphological gulf of the neural tube between living vertebrates and invertebrate chordates. Thus, a medial, proliferative, Notch-regulated ventricular zone might be a character that evolved in the vertebrate lineage before the cyclostome radiation.

9.7. Future Directions

The study of neural HLH genes, both in the lamprey and other vertebrates, is an excellent tool to identify cell populations/territories within the neural tube, patterns of neuronal migration and differentiation along the anterior-posterior, dorsal-ventral and mediolateral axes. It also allows the identification of particular neuronal types such as sensory neurons, interneurons and motoneurons based on studies in gnathostomes. Expression of *LpCoe-A* and *LpCoe-B* already hints at the presence of motoneurons while expression of *LpNgn* suggests sites of sensory neuroblasts. Description and expression analysis of members of other HLH gene subfamilies in the lamprey would allow a better identification of population/cell territories in the neural tube as in

gnathostomes. For example, a combination of expression patterns of *NeuroD* and *NSCL* genes would possibly allow the identification of a subventricular zone in the neural tube if they follow the same expression patterns as in gnathostomes (Koyano-Nakagawa et al., 2000; Ma et al., 1997; Perez et al., 1999; Theodorakis et al., 2002). Analysis of *Atonal* genes, whose expression is very restricted to the dorsal spinal cord, would also allow the identification of dorsal interneurons (Ben-Arie et al., 1996; Gowan et al., 2001).

Particularly interesting would be the identification of *Olig* homologues in lampreys, as they are well-known as important regulators of oligodendrocyte development in gnathostomes (Li and Richardson, 2008; Takebayashi et al., 2000; Zhou et al., 2000), which could provide insights into the evolutionary origin of oligodendrocytes and the evolutionary steps towards vertebrate myelination. Furthermore, Notch signalling is continuously required during development to permit oligodendrocyte development in the zebrafish spinal cord (Kim et al., 2008a). Even though lampreys have *bona fide* neural crest cells (Häming et al., 2011; Hirata et al., 1997; Horigome et al., 1999; McCauley and Bronner-Fraser, 2003, 2006; Meulemans and Bronner-Fraser, 2002; Nikitina et al., 2008; Sauka-Spengler and Bronner-Fraser, 2008; Sauka-Spengler et al., 2007), they do not have myelin either in the CNS or in the PNS (Bullock et al., 1984), although Schwann cell-like cells (which form myelin in the gnathostome PNS) have been reported in the trunk lateral line nerve (Gelman et al., 2009). However, chondrichthyans possess fully-formed myelin both in CNS and PNS (Bakay and Lee, 1966; Rotenstein et al., 2008, 2009; Waehneltd et al., 1984; Yoshida and Colman, 1996), and it is therefore uncertain how myelin appeared in the vertebrate nervous system, although placoderms have been proposed to be the first myelinated vertebrates

(Zalc et al., 2008). Thus, the identification and expression analysis of lamprey *Olig* genes might shed light into the evolutionary origin of myelin.

Identification of lamprey *NSCL* homologues would also possibly help to discern rhombomere boundaries as in chicken and mouse (Theodorakis et al., 2002), which particularly in lampreys, it is difficult to discern between rhombomeres 1 and 2, and rhombomeres 5 and 6 (and probably 7) (Kuratani et al., 1998; Meléndez-Ferro et al., 2003; Osorio et al., 2005).

Fragments of three other neural HLH genes were cloned in the present study with sequence similarity (most likely homology) to the *Atonal* (one fragment) and *NeuroD* (two fragments) gene subfamilies. Analysis of their expression patterns will undoubtedly provide a much finer resolution of cell territories in the neural tube and sites of neuronal differentiation in the nervous system. As mentioned previously, *Atonal* genes mark dorsal interneurons and so the analysis of lamprey *Atonal* could also mark a specific dorsal neuronal population in the lamprey neural tube. On the other hand, at least in sharks, zebrafish and *Xenopus*, *NeuroD* is expressed in all neurogenic cranial placodes/ganglia (Andermann et al., 2002; O'Neill et al., 2007; Schlosser and Northcutt, 2000). Therefore, lamprey *NeuroD* genes could also mark other cranial placodes/ganglia that have not been marked in the present study such as the otic placode or lateral line placodes. A lamprey *NeuroD* homolog, *NeuroD2*, has been cloned and its expression found in the neural tube (only one section at stage 24), as well as in the trigeminal, facial and posterior lateral line ganglia, but this gene has not been analysed

in much detail and only at stages 24 and 26 (Guérin et al., 2009), and thus no definitive conclusions can be obtained from its developmental expression.

Another interesting aspect to follow would be a deeper description of the putative sensory “branchial placodes” identified in the present study. By analysing additional genes known to mark epibranchial and hypobranchial placodes in *Xenopus* such as *Six1*, *Eya1*, *NeuroD* and *Delta-1* (Schlosser, 2003; Schlosser and Northcutt, 2000), or *Phox2a* in chicken (Abu-Elmagd et al., 2001), and following their development by lineage tracing with DiI, it could help us to determine whether, first, they are indeed cranial placodes, and second, whether they are sensory placodes that ultimately will form sensory neurons. So far no specific hypobranchial placode markers have been identified in *Xenopus* if we consider that the putative lamprey “branchial placodes” are instead hypobranchial placodes. Also, the search for putative lamprey hypobranchial ganglia is needed to test this hypothesis, which can be accomplished by examining –for example– immunostaining with neuronal antibodies against Elav-family members HuC/D (Hu) (Modrell et al., 2014), acetylated tubulin (Kuratani et al., 1997), or neurofilament (Hall et al., 2000), in embryos, larvae or adults.

Apart from Notch signalling, the vertebrate spinal cord is regulated by other signalling pathways. Two of those pathways, the fibroblast growth factor (FGF) and the retinoic acid (RA) pathways, have an important role in the patterning of the spinal cord along the anterior-posterior axis: the FGF pathway maintains a proliferative state of neuroblasts in the stem zone while RA signalling, produced from the somites, induces neuronal differentiation (Diez del Corral et al., 2003). Thus, differentiation along the

anterior-posterior axis of the spinal cord is regulated by a balance between anterior RA signaling and posterior FGF signalling. As the embryo grows this interface moves posteriorly and a wave of cell differentiation moves along the spinal cord.

In this context, preliminary treatments have been performed in developing lamprey embryos using the FGF pathway inhibitor SU5402 and the RA pathway inhibitor BMS493, both in the same time period of treatment studied in this project (i.e. from stage 24 to stage 26), and from stage 22 to stage 24. Preliminary analyses of *LpCoe-B* expression in embryos treated with SU5402 from stage 22 to stage 24 revealed an increase of *LpCoe-B* expression in the spinal cord as large clusters. This increase in *LpCoe-B* expression is in agreement with the role of the FGF pathway in the repression of neuronal differentiation in the spinal cord (Diez del Corral et al., 2003). Embryos treated with BMS493 have not been analysed for gene expression. Therefore, more experiments with these embryos are needed, which will considerably expand our knowledge on the regulation of expression of lamprey neural HLH genes along the anterior-posterior axis of the spinal cord by different signalling pathways.

Taken together, there is still a wide area to be exploited by studying the HLH gene family in lampreys, together with pharmacological treatments, to have a much better understanding of the lamprey nervous system and, in comparison with gnathostomes, of the putative ancestral state of the vertebrate nervous system and/or ancient developmental mechanisms. Current techniques such as *in situ* hybridisation already

provide plenty of information, but with the advent of morpholinos and drug treatments that can be used in lampreys, it will certainly provide invaluable information to study the lamprey nervous system from a functional perspective as well, with a particular reference to the neural HLH genes.

10. REFERENCES

- Aaker, J.D., Patineau, A.L., Yang, H., Ewart, D.T., Nakagawa, Y., McLoon, S.C., and Koyano-Nakagawa, N. (2010). Interaction of MTG family proteins with NEUROG2 and ASCL1 in the developing nervous system. *Neurosci. Lett.* **474**, 46–51.
- Aguado-Llera, D., Goormaghtigh, E., de Geest, N., Quan, X.-J., Prieto, A., Hassan, B.A., Gómez, J., and Neira, J.L. (2010). The Basic Helix–Loop–Helix Region of Human Neurogenin 1 Is a Monomeric Natively Unfolded Protein Which Forms a “Fuzzy” Complex upon DNA Binding. *Biochemistry (Mosc.)* **49**, 1577–1589.
- Ahmad, I. (1995). Mash-1 is expressed during ROD photoreceptor differentiation and binds an E-box, Eopsin-1 in the rat opsin gene. *Dev. Brain Res.* **90**, 184–189.
- Akai, J., Halley, P.A., and Storey, K.G. (2005). FGF-dependent Notch signaling maintains the spinal cord stem zone. *Genes Dev.* **19**, 2877–2887.
- Akazawa, C., Sasai, Y., Nakanishi, S., and Kageyama, R. (1992). Molecular characterization of a rat negative regulator with a basic helix-loop-helix structure predominantly expressed in the developing nervous system. *J. Biol. Chem.* **267**, 21879–21885.
- Akazawa, C., Ishibashi, M., Shimizu, C., Nakanishi, S., and Kageyama, R. (1995). A mammalian helix-loop-helix factor structurally related to the product of *Drosophila* proneural gene *atonal* is a positive transcriptional regulator expressed in the developing nervous system. *J. Biol. Chem.* **270**, 8730–8738.
- Alder, M.N., Rogozin, I.B., Iyer, L.M., Glazko, G.V., Cooper, M.D., and Pancer, Z. (2005). Diversity and Function of Adaptive Immune Receptors in a Jawless Vertebrate. *Science* **310**, 1970–1973.
- Alder, M.N., Herrin, B.R., Sadlonova, A., Stockard, C.R., Grizzle, W.E., Gartland, L.A., Gartland, G.L., Boydston, J.A., Turnbough, C.L., and Cooper, M.D. (2008). Antibody responses of variable lymphocyte receptors in the lamprey. *Nat. Immunol.* **9**, 319–327.
- Alders, M., Hodges, M., Hadjantonakis, A.-K., Postmus, J., van Wijk, I., Blik, J., de Meulemeester, M., Westerveld, A., Guillemot, F., Oudejans, C., et al. (1997). The human Achaete-Scute homologue 2 (ASCL2, HASH2) maps to chromosome 11p15. 5, close to IGF2 and is expressed in extravillous trophoblasts. *Hum. Mol. Genet.* **6**, 859–867.
- Aldridge, R.J., Briggs, D.E.G., Smith, M.P., Clarkson, E.N.K., and Clark, N.D.L. (1993). The anatomy of conodonts. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **340**, 405–421.
- Ali, F., Hindley, C., McDowell, G., Deibler, R., Jones, A., Kirschner, M., Guillemot, F., and Philpott, A. (2011). Cell cycle-regulated multi-site phosphorylation of Neurogenin 2 coordinates cell cycling with differentiation during neurogenesis. *Development* **138**, 4267–4277.
- Allende, M.L., and Weinberg, E.S. (1994). The expression pattern of two zebrafish achaete-scute homolog (ash) genes is altered in the embryonic brain of the cyclops mutant. *Dev. Biol.* **166**, 509–530.

- Amemiya, C., Saha, N., and Zapata, A. (2007). Evolution and development of immunological structures in the lamprey. *Curr. Opin. Immunol.* *19*, 535–541.
- Amoyel, M., Cheng, Y.-C., Jiang, Y.-J., and Wilkinson, D.G. (2005). Wnt1 regulates neurogenesis and mediates lateral inhibition of boundary cell specification in the zebrafish hindbrain. *Development* *132*, 775–785.
- Andermann, P., Ungos, J., and Raible, D.W. (2002). Neurogenin1 Defines Zebrafish Cranial Sensory Ganglia Precursors. *Dev. Biol.* *251*, 45–58.
- Anderson, M.K., Pant, R., Miracle, A.L., Sun, X., Luer, C.A., Walsh, C.J., Telfer, J.C., Litman, G.W., and Rothenberg, E.V. (2004). Evolutionary origins of lymphocytes: ensembles of T cell and B cell transcriptional regulators in a cartilaginous fish. *J. Immunol.* *172*, 5851–5860.
- Appel, B., Givan, L.A., and Eisen, J.S. (2001). Delta-Notch signaling and lateral inhibition in zebrafish spinal cord development. *BMC Dev. Biol.* *1*, 13.
- Ardavín, C.F., and Zapata, A. (1987). Ultrastructure and changes during metamorphosis of the lympho-hemopoietic tissue of the larval anadromous sea lamprey *Petromyzon marinus*. *Dev. Comp. Immunol.* *11*, 79–93.
- Ben-Arie, N., McCall, A.E., Berkman, S., Eichele, G., Bellen, H.J., and Zoghbi, H.Y. (1996). Evolutionary conservation of sequence and expression of the bHLH protein Atonal suggests a conserved role in neurogenesis. *Hum. Mol. Genet.* *5*, 1207–1216.
- Arslanova, D., Yang, T., Xu, X., Wong, S.T., Augelli-Szafran, C.E., and Xia, W. (2010). Phenotypic analysis of images of zebrafish treated with Alzheimer's γ -secretase inhibitors. *BMC Biotechnol.* *10*, 24.
- Asamoto, K. (2005). Network of the sympathetic nervous system: Focus on the input and output of the cervical sympathetic ganglion. *Anat. Sci. Int.* *80*, 132–140.
- Atchley, W.R., and Fitch, W.M. (1997). A natural classification of the basic helix–loop–helix class of transcription factors. *Proc. Natl. Acad. Sci.* *94*, 5172–5176.
- Azmi, S., and Taneja, R. (2002). Embryonic expression of mSharp-1/mDEC2, which encodes a basic helix-loop-helix transcription factor. *Mech. Dev.* *114*, 181–185.
- Bae, S., Bessho, Y., Hojo, M., and Kageyama, R. (2000). The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development* *127*, 2933–2943.
- Bae, Y.-K., Shimizu, T., and Hibi, M. (2005). Patterning of proneuronal and inter-proneuronal domains by hairy- and enhancer of split-related genes in zebrafish neuroectoderm. *Development* *132*, 1375–1385.
- Bajoghli, B., Aghaallaei, N., Hess, I., Rode, I., Netuschil, N., Tay, B.-H., Venkatesh, B., Yu, J.-K., Kaltenbach, S.L., Holland, N.D., et al. (2009). Evolution of Genetic Networks Underlying the Emergence of Thymopoiesis in Vertebrates. *Cell* *138*, 186–197.
- Bajoghli, B., Guo, P., Aghaallaei, N., Hirano, M., Strohmeier, C., McCurley, N., Bockman, D.E., Schorpp, M., Cooper, M.D., and Boehm, T. (2011). A thymus candidate in lampreys. *Nature* *470*, 90–94.

- Bakay, L., and Lee, J.C. (1966). Ultrastructural Changes in the Edematous Central Nervous System: III. Edema in Shark Brain*. *Arch. Neurol.* *14*, 644–660.
- Baker, C.V.H., and Bronner-Fraser, M. (2001). Vertebrate Cranial Placodes I. Embryonic Induction. *Dev. Biol.* *232*, 1–61.
- Baker, C.V.H., Modrell, M.S., and Gillis, J.A. (2013). The evolution and development of vertebrate lateral line electroreceptors. *J. Exp. Biol.* *216*, 2515–2522.
- Ball, D.W., Azzoli, C.G., Baylin, S.B., Chi, D., Dou, S., Donis-Keller, H., Cumaraswamy, A., Borges, M., and Nelkin, B.D. (1993). Identification of a human achaete-scute homolog highly expressed in neuroendocrine tumors. *Proc. Natl. Acad. Sci.* *90*, 5648–5652.
- Bally-Cuif, L., Dubois, L., and Vincent, A. (1998). Molecular cloning of Zcoe2, the zebrafish homolog of Xenopus Xcoe2 and mouse EBF-2, and its expression during primary neurogenesis. *Mech. Dev.* *77*, 85–90.
- Bardack, D., and Zangerl, R. (1968). First fossil lamprey: a record from the Pennsylvanian of Illinois. *Science* *162*, 1265–1267.
- Bassham, S., and Postlethwait, J.H. (2005). The evolutionary history of placodes: a molecular genetic investigation of the larvacean urochordate *Oikopleura dioica*. *Development* *132*, 4259–4272.
- Beck, C.W., and Slack, J.M.W. (1998). Analysis of the developing Xenopus tail bud reveals separate phases of gene expression during determination and outgrowth. *Mech. Dev.* *72*, 41–52.
- Beel, A.J., and Sanders, C.R. (2008). Substrate specificity of γ -secretase and other intramembrane proteases. *Cell. Mol. Life Sci.* *65*, 1311–1334.
- Begbie, J., and Graham, A. (2001). The ectodermal placodes: a dysfunctional family. *Philos. Trans. R. Soc. B Biol. Sci.* *356*, 1655–1660.
- Begbie, J., Ballivet, M., and Graham, A. (2002). Early Steps in the Production of Sensory Neurons by the Neurogenic Placodes. *Mol. Cell. Neurosci.* *21*, 502–511.
- Begley, C.G., Lipkowitz, S., Göbel, V., Mahon, K.A., Bertness, V., Green, A.R., Gough, N.M., and Kirsch, I.R. (1992). Molecular characterization of NSCL, a gene encoding a helix-loop-helix protein expressed in the developing nervous system. *Proc. Natl. Acad. Sci.* *89*, 38–42.
- Belanger-Jasmin, S., Llamosas, E., Tang, Y., Joachim, K., Osiceanu, A.-M., Jhas, S., and Stifani, S. (2007). Inhibition of cortical astrocyte differentiation by Hes6 requires amino- and carboxy-terminal motifs important for dimerization and phosphorylation. *J. Neurochem.* *103*, 2022–2034.
- Bellefroid, E.J., Bourguignon, C., Hollemann, T., Ma, Q., Anderson, D.J., Kintner, C., and Pieler, T. (1996). X-MyT1, a Xenopus C2HC-type zinc finger protein with a regulatory function in neuronal differentiation. *Cell* *87*, 1191–1202.
- Benito-Gutiérrez, È. (2006). A gene catalogue of the amphioxus nervous system. *Int J Biol Sci* *2*, 149–160.

- Bertrand, N., Castro, D.S., and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.* **3**, 517–530.
- Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S., and Kageyama, R. (2001). Dynamic expression and essential functions of *Hes7* in somite segmentation. *Genes Dev.* **15**, 2642–2647.
- Blader, P., Fischer, N., Gradwohl, G., Guillemot, F., and Strahle, U. (1997). The activity of *neurogenin1* is controlled by local cues in the zebrafish embryo. *Development* **124**, 4557–4569.
- Blaugrund, E., Pham, T.D., Tennyson, V.M., Lo, L., Sommer, L., Anderson, D.J., and Gershon, M.D. (1996). Distinct subpopulations of enteric neuronal progenitors defined by time of development, sympathoadrenal lineage markers and *Mash-1*-dependence. *Development* **122**, 309–320.
- Boudjelal, M., Taneja, R., Matsubara, S., Bouillet, P., Dolle, P., and Chambon, P. (1997). Overexpression of *Stra13*, a novel retinoic acid-inducible gene of the basic helix-loop-helix family, inhibits mesodermal and promotes neuronal differentiation of P19 cells. *Genes Dev.* **11**, 2052–2065.
- Bray, S.J. (2006). Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell Biol.* **7**, 678–689.
- Bronner, M.E., and LeDouarin, N.M. (2012). Development and evolution of the neural crest: An overview. *Dev. Biol.* **366**, 2–9.
- Brox, A., Puelles, L., Ferreiro, B., and Medina, L. (2004). Expression of the genes *Emx1*, *Tbr1*, and *Eomes* (*Tbr2*) in the telencephalon of *Xenopus laevis* confirms the existence of a ventral pallial division in all tetrapods. *J. Comp. Neurol.* **474**, 562–577.
- Brugmann, S.A., and Moody, S.A. (2005). Induction and specification of the vertebrate ectodermal placodes: precursors of the cranial sensory organs. *Biol. Cell* **97**, 303–319.
- Bullock, T.H., Moore, J.K., and Fields, R.D. (1984). Evolution of myelin sheaths: both lamprey and haghfish lack myelin. *Neurosci. Lett.* **48**, 145–148.
- Burns, A.J., and Thapar, N. (2006). Advances in ontogeny of the enteric nervous system. *Neurogastroenterol. Motil.* **18**, 876–887.
- Burns, C.J., and Vetter, M.L. (2002). *Xath5* regulates neurogenesis in the *Xenopus* olfactory placode. *Dev. Dyn.* **225**, 536–543.
- Burnstock, G. (1969). Evolution of the autonomic innervation of visceral and cardiovascular systems in vertebrates. *Pharmacol. Rev.* **21**, 247–324.
- Butler, A.B., and Hodos, W. (1996). *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation* (New York: Wiley-Liss).
- Campbell, W.A., Yang, H., Zetterberg, H., Baulac, S., Sears, J.A., Liu, T., Wong, S.T.C., Zhong, T.P., and Xia, W. (2006). Zebrafish lacking Alzheimer presenilin enhancer 2 (*Pen-2*) demonstrate excessive p53-dependent apoptosis and neuronal loss. *J. Neurochem.* **96**, 1423–1440.

- Carretero-Paulet, L., Galstyan, A., Roig-Villanova, I., Martinez-Garcia, J.F., Bilbao-Castro, J.R., and Robertson, D.L. (2010). Genome-Wide Classification and Evolutionary Analysis of the bHLH Family of Transcription Factors in Arabidopsis, Poplar, Rice, Moss, and Algae. *PLANT Physiol.* 153, 1398–1412.
- 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., and 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.
- Cau, E., and Wilson, S.W. (2003). *Ash1a* and *Neurogenin1* function downstream of Floating head to regulate epiphysial neurogenesis. *Development* 130, 2455–2466.
- Cau, E., Gradwohl, G., Fode, C., and Guillemot, F. (1997). *Mash1* activates a cascade of bHLH regulators in olfactory neuron progenitors. *Development* 124, 1611–1621.
- Cau, E., Gradwohl, G., Casarosa, S., Kageyama, R., and Guillemot, F. (2000). *Hes* genes regulate sequential stages of neurogenesis in the olfactory epithelium. *Development* 127, 2323–2332.
- Cau, E., Casarosa, S., and Guillemot, F. (2002). *Mash1* and *Ngn1* control distinct steps of determination and differentiation in the olfactory sensory neuron lineage. *Development* 129, 1871–1880.
- Cerny, R., Cattell, M., Sauka-Spengler, T., Bronner-Fraser, M., Yu, F., and Medeiros, D.M. (2010). Evidence for the prepattern/cooption model of vertebrate jaw evolution. *Proc. Natl. Acad. Sci.* 107, 17262–17267.
- Chang, M., Zhang, J., and Miao, D. (2006). A lamprey from the Cretaceous Jehol biota of China. *Nature* 441, 972–974.
- Chavez-Gutierrez, L., Tolia, A., Maes, E., Li, T., Wong, P.C., and de Strooper, B. (2008). Glu332 in the Nicastrin Ectodomain Is Essential for γ -Secretase Complex Maturation but Not for Its Activity. *J. Biol. Chem.* 283, 20096–20105.
- Chen, J.-Y. (2009). The sudden appearance of diverse animal body plans during the Cambrian explosion. *Int. J. Dev. Biol.* 53, 733–751.
- Chen, H., Biel, M.A., Borges, M.W., Thiagalingam, A., Nelkin, B.D., Baylin, S.B., and Ball, D.W. (1997). Tissue-specific expression of human *achaete-scute* homologue-1 in neuroendocrine tumors, transcriptional regulation by dual inhibitory regions. *Cell Growth Differ.* 8, 677–686.
- Chen, H., Kunnimalaiyaan, M., and Van Gompel, J.J. (2005). Medullary thyroid cancer: the functions of *raf-1* and human *achaete-scute* homologue-1. *Thyroid* 15, 511–521.
- Chesnutt, C., Burrus, L.W., Brown, A.M.C., and Niswander, L. (2004). Coordinate regulation of neural tube patterning and proliferation by TGF β and WNT activity. *Dev. Biol.* 274, 334–347.
- Chiara, F., Badaloni, A., Croci, L., Yeh, M.L., Cariboni, A., Hoerder-Suabedissen, A., Consalez, G.G., Eickholt, B., Shimogori, T., Parnavelas, J.G., et al. (2012). Early B-cell factors 2 and 3 (EBF2/3) regulate early migration of Cajal–Retzius cells from the cortical hem. *Dev. Biol.* 365, 277–289.

- Chien, C., Hsiao, C.-D., Jan, L.Y., and Jan, Y.N. (1996). Neuronal type information encoded in the basic-helix-loop-helix domain of proneural genes. *Proc. Natl. Acad. Sci.* *93*, 13239–13244.
- Chung, P.-C., Lin, W.-S., Scotting, P.J., Hsieh, F.-Y., Wu, H.-L., and Cheng, Y.-C. (2011). Zebrafish Her8a Is Activated by Su(H)-Dependent Notch Signaling and Is Essential for the Inhibition of Neurogenesis. *PLoS ONE* *6*, e19394.
- Conway Morris, S., and Caron, J.-B. (2014). A primitive fish from the Cambrian of North America. *Nature*.
- Cooper, M.D., and Alder, M.N. (2006). The Evolution of Adaptive Immune Systems. *Cell* *124*, 815–822.
- Cornell, R.A., and Eisen, J.S. (2002). Delta/Notch signaling promotes formation of zebrafish neural crest by repressing Neurogenin 1 function. *Development* *129*, 2639–2648.
- Corradi, A., Croci, L., Broccoli, V., Zecchini, S., Previtali, S., Wurst, W., Amadio, S., Maggi, R., Quattrini, A., and Consalez, G. (2003). Hypogonadotropic hypogonadism and peripheral neuropathy in Ebf2-null mice. *Development* *130*, 401–410.
- Craven, J. (2007). The autonomic nervous system, sympathetic chain and stellate ganglion. *Anaesth. Intensive Care Med.* *9*, 39–41.
- Crick, F.C., and Koch, C. (2005). What is the function of the claustrum? *Philos. Trans. R. Soc. B Biol. Sci.* *360*, 1271–1279.
- Croci, L. (2006). A key role for the HLH transcription factor EBF2COE2,O/E-3 in Purkinje neuron migration and cerebellar cortical topography. *Development* *133*, 2719–2729.
- Croci, L., Chung, S.-H., Masserdotti, G., Gianola, S., Bizzoca, A., Gennarini, G., Corradi, A., Rossi, F., Hawkes, R., and Consalez, G. (2006). A key role for the HLH transcription factor EBF2COE2,O/E-3 in Purkinje neuron migration and cerebellar cortical topography. *Development* *133*, 2719–2729.
- Daburon, V., Mella, S., Plouhinec, J.-L., Mazan, S., Crozatier, M., and Vincent, A. (2008). The metazoan history of the COE transcription factors. Selection of a variant HLH motif by mandatory inclusion of a duplicated exon in vertebrates. *BMC Evol. Biol.* *8*, 131.
- Damas, H. (1944). Recherches sur le développement de *Lampetra fluviatilis* L. Contribution à l'étude de la céphalogenèse des vertébrés. *Arch Biol* *55*, 1–284.
- Das, S., Hirano, M., Aghaallaei, N., Bajoghli, B., Boehm, T., and Cooper, M.D. (2013). Organization of lamprey variable lymphocyte receptor C locus and repertoire development. *Proc. Natl. Acad. Sci.* *110*, 6043–6048.
- Davis, J.A., and Reed, R.R. (1996). Role of Olf-1 and Pax-6 transcription factors in neurodevelopment. *J. Neurosci.* *16*, 5082–5094.
- Dehal, P., and Boore, J.L. (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLoS Biol.* *3*, e314.
- Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* *439*, 965–968.

- Diez del Corral, R.D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., and Storey, K. (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron* 40, 65–79.
- Donoghue, P.C., Forey, P.L., and Aldridge, R.J. (2000). Conodont affinity and chordate phylogeny. *Biol. Rev. Camb. Philos. Soc.* 75, 191–251.
- Donoghue, P.C.J., Graham, A., and Kelsh, R.N. (2008). The origin and evolution of the neural crest. *BioEssays* 30, 530–541.
- Donoviel, D.B., Hadjantonakis, A.-K., Ikeda, M., Zheng, H., Hyslop, P.S.G., and Bernstein, A. (1999). Mice lacking both presenilin genes exhibit early embryonic patterning defects. *Genes Dev.* 13, 2801–2810.
- Dovey et al., H.F. (2001). Functional gamma-secretase inhibitors reduce beta-amyloid peptide levels in brain. *J. Neurochem.* 76, 173–181.
- Dries, D.R., Shah, S., Han, Y.-H., Yu, C., Yu, S., Shearman, M.S., and Yu, G. (2009). Glu-333 of nicastrin directly participates in gamma-secretase activity. *J. Biol. Chem.* 284, 29714–29724.
- Dubois, L., and Vincent, A. (2001). The COE–Collier/Olf1/EBF–transcription factors: structural conservation and diversity of developmental functions. *Mech. Dev.* 108, 3–12.
- Dubois, L., Bally-Cuif, L., Crozatier, M., Moreau, J., Paquereau, L., and Vincent, A. (1998). XCoe2, a transcription factor of the Col/Olf-1/EBF family involved in the specification of primary neurons in *Xenopus*. *Curr. Biol.* 8, 199–209.
- Eastwood, K., Yin, C., Bandyopadhyay, M., and Bidwai, A. (2011). New insights into the Orange domain of E(spl)-M8, and the roles of the C-terminal domain in autoinhibition and Groucho recruitment. *Mol. Cell. Biochem.* 356, 217–225.
- Eickholt, B.J., Graham, A., Lumsden, A., and Wizenmann, A. (2001). Rhombomere Interactions Control the Segmental Differentiation of Hindbrain Neurons. *Mol. Cell. Neurosci.* 18, 141–148.
- Abu-Elmagd, M., Ishii, Y., Cheung, M., Rex, M., Le Rouëdec, D., and Scotting, P.J. (2001). cSox3 Expression and Neurogenesis in the Epibranchial Placodes. *Dev. Biol.* 237, 258–269.
- Epstein, F.H., Goyal, R.K., and Hirano, I. (1996). The enteric nervous system. *N. Engl. J. Med.* 334, 1106–1115.
- Ernsberger, U., Patzke, H., Tissier-Seta, J.P., Reh, T., Goridis, C., and Rohrer, H. (1995). The expression of tyrosine hydroxylase and the transcription factors cPhox-2 and Cash-1: evidence for distinct inductive steps in the differentiation of chick sympathetic precursor cells. *Mech. Dev.* 52, 125–136.
- Ferreiro, B., Skoglund, P., Bailey, A., Dorsky, R., and Harris, W.A. (1992). XASH1, a *Xenopus* homolog of achaete-scute: a proneural gene in anterior regions of the vertebrate CNS. *Mech. Dev.* 40, 25–36.
- Fields, S., Ternyak, K., Gao, H., Ostraat, R., Akerlund, J., and Hagman, J. (2008). The “zinc knuckle” motif of Early B cell Factor is required for transcriptional activation of B cell-specific genes. *Mol. Immunol.* 45, 3786–3796.

- Figdor, M.C., and Stern, C.D. (1993). Segmental organization of the embryonic diencephalon. *Nature* 363, 630–634.
- Filippi, A., Tiso, N., Deflorian, G., Zecchin, E., Bortolussi, M., and Argenton, F. (2005). The basic helix–loop–helix olig3 establishes the neural plate boundary of the trunk and is necessary for development of the dorsal spinal cord. *Proc. Natl. Acad. Sci. U. S. A.* 102, 4377–4382.
- Fior, R., and Henrique, D. (2005). A novel hes5/hes6 circuitry of negative regulation controls Notch activity during neurogenesis. *Dev. Biol.* 281, 318–333.
- Fischer, A., and Gessler, M. (2007). Delta Notch and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors. *Nucleic Acids Res.* 35, 4583–4596.
- Fisher, A.L., Ohsako, S., and Caudy, M. (1996). The WRPW motif of the hairy-related basic helix–loop–helix repressor proteins acts as a 4-amino-acid transcription repression and protein–protein interaction domain. *Mol. Cell. Biol.* 16, 2670–2677.
- Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. (2014). Ensembl 2014. *Nucleic Acids Res.* 42, D749–D755.
- Florio, M., Leto, K., Muzio, L., Tinterri, A., Badaloni, A., Croci, L., Zordan, P., Barili, V., Albieri, I., Guillemot, F., et al. (2012). Neurogenin 2 regulates progenitor cell-cycle progression and Purkinje cell dendritogenesis in cerebellar development. *Development* 139, 2308–2320.
- Fode, C., Ma, Q., Casarosa, S., Ang, S.-L., Anderson, D.J., and Guillemot, F. (2000). A role for neural determination genes in specifying the dorsoventral identity of telencephalic neurons. *Genes Dev.* 14, 67–80.
- Forey, P., and Janvier, P. (1993). Agnathans and the origin of jawed vertebrates. *Nature* 361, 129–134.
- Fortini, M.E. (2002). Gamma-secretase-mediated proteolysis in cell-surface-receptor signalling. *Nat. Rev. Mol. Cell Biol.* 3, 673–684.
- Fraering, P.C., Ye, W., LaVoie, M.J., Ostaszewski, B.L., Selkoe, D.J., and Wolfe, M.S. (2005). Gamma-Secretase Substrate Selectivity Can Be Modulated Directly via Interaction with a Nucleotide-binding Site. *J. Biol. Chem.* 280, 41987–41996.
- Fraser, S., Keynes, R., and Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* 344, 431–435.
- Freitas, R., Zhang, G., and Cohn, M.J. (2006). Evidence that mechanisms of fin development evolved in the midline of early vertebrates. *Nature* 442, 1033–1037.
- Friedman, R.A., Makmura, L., Biesiada, E., Wang, X., and Keithley, E.M. (2005). Eya1 acts upstream of Tbx1, Neurogenin 1, NeuroD and the neurotrophins BDNF and NT-3 during inner ear development. *Mech. Dev.* 122, 625–634.
- Fritzsch, B., and Northcutt, R.G. (1993). Cranial and spinal nerve organization in amphioxus and lampreys: evidence for an ancestral craniate pattern. *Acta Anat. Basel* 148, 96–109.

- Frontini, A., Zaidi, A.U., Hua, H., Wolak, T.P., Greer, C.A., Kafitz, K.W., Li, W., and Zielinski, B.S. (2003). Glomerular territories in the olfactory bulb from the larval stage of the sea lamprey *Petromyzon marinus*. *J. Comp. Neurol.* **465**, 27–37.
- Frowein, J. v., Campbell, K., and Götz, M. (2002). Expression of *Ngn1*, *Ngn2*, *Cash1*, *Gsh2* and *Sfrp1* in the developing chick telencephalon. *Mech. Dev.* **110**, 249–252.
- Fujii, T. (1982). Electron microscopy of the leucocytes of the typhlosole in ammocoetes, with special attention to the antibody-producing cells. *J. Morphol.* **173**, 87–100.
- Fujimoto, K., Shen, M., Noshiro, M., Matsubara, K., Shingu, S., Honda, K., Yoshida, E., Suardita, K., Matsuda, Y., and Kato, Y. (2001). Molecular Cloning and Characterization of DEC2, a New Member of Basic Helix-Loop-Helix Proteins. *Biochem. Biophys. Res. Commun.* **280**, 164–171.
- Funakoshi, K., and Nakano, M. (2007). The Sympathetic Nervous System of Anamniotes. *Brain. Behav. Evol.* **69**, 105–113.
- Furlong, R.F., and Graham, A. (2005). Vertebrate neurogenin evolution: long-term maintenance of redundant duplicates. *Dev. Genes Evol.* **215**, 639–644.
- Furlong, R.F., and Holland, P.W. (2002a). Were vertebrates octoploid? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **357**, 531–544.
- Furlong, R.F., and Holland, P.W.H. (2002b). Bayesian phylogenetic analysis supports monophyly of ambulacraria and of cyclostomes. *Zoolog. Sci.* **19**, 593–599.
- Futai, E., Yagishita, S., and Ishiura, S. (2009). Nicastrin Is Dispensable for γ -Secretase Protease Activity in the Presence of Specific Presenilin Mutations. *J. Biol. Chem.* **284**, 13013–13022.
- Gans, C., and Northcutt, R.G. (1983). Neural crest and the origin of vertebrates: a new head. *Science* **220**, 268–273.
- Garcia-Dominguez, M., Poquet, C., Garel, S., and Charnay, P. (2003). *Ebf* gene function is required for coupling neuronal differentiation and cell cycle exit. *Development* **130**, 6013–6025.
- Garel, S., Marín, F., Mattéi, M.-G., Vesque, C., Vincent, A., and Charnay, P. (1997a). Family of *Ebf/Olf-1*-related genes potentially involved in neuronal differentiation and regional specification in the central nervous system. *Dev. Dyn.* **210**, 191–205.
- Garel, S., Marín, F., Mattéi, G., Vesque, C., Vincent, A., and Charnay, P. (1997b). Family of *Ebf/Olf-1*-related genes potentially involved in neuronal differentiation and regional specification in the central nervous system. *Dev. Dyn.* **210**, 191–205.
- Garel, S., Marín, F., Grosschedl, R., and Charnay, P. (1999). *Ebf1* controls early cell differentiation in the embryonic striatum. *Development* **126**, 5285–5294.
- Garel, S., Garcia-Dominguez, M., and Charnay, P. (2000). Control of the migratory pathway of facial branchiomotor neurones. *Development* **127**, 5297–5307.
- Garel, S., Yun, K., Grosschedl, R., and Rubenstein, J.L.R. (2002). The early topography of thalamocortical projections is shifted in *Ebf1* and *Dlx1/2* mutant mice. *Development* **129**, 5621–5634.

- Garriga-Canut, M., Roopra, A., and Buckley, N.J. (2001). The Basic Helix-Loop-Helix Protein, SHARP-1, Represses Transcription by a Histone Deacetylase-dependent and Histone Deacetylase-independent Mechanism. *J. Biol. Chem.* **276**, 14821–14828.
- Geling, A., Steiner, H., Willem, M., Bally-Cuif, L., and Haass, C. (2002). A gamma-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO Rep.* **3**, 688–694.
- Geling, A., Plessy, C., Rastegar, S., Strähle, U., and Bally-Cuif, L. (2004). Her5 acts as a prepattern factor that blocks neurogenin1 and coe2 expression upstream of Notch to inhibit neurogenesis at the midbrain-hindbrain boundary. *Development* **131**, 1993–2006.
- Gelman, S., Ayali, A., Tytell, E.D., and Cohen, A.H. (2007). Larval lampreys possess a functional lateral line system. *J. Comp. Physiol. A* **193**, 271–277.
- Gelman, S., Ayali, A., Kiemel, T., Sanovich, E., and Cohen, A.H. (2008). Metamorphosis-related changes in the lateral line system of lampreys, *Petromyzon marinus*. *J. Comp. Physiol. A* **194**, 945–956.
- Gelman, S., Cohen, A.H., and Sanovich, E. (2009). Developmental changes in the ultrastructure of the lamprey lateral line nerve during metamorphosis. *J. Morphol.* **270**, 815–824.
- Gess, R.W., Coates, M.I., and Rubidge, B.S. (2006). A lamprey from the Devonian period of South Africa. *Nature* **443**, 981–984.
- Gilbert, S.F. (2003). *Developmental Biology* (United States: Sinauer).
- Göbel, V., Lipkowitz, S., Kozak, C.A., and Kirsch, I.R. (1992). NSCL-2: a basic domain helix-loop-helix gene expressed in early neurogenesis. *Cell Growth Differ.* **3**, 143–148.
- Gordon, M.K., Mumm, J.S., Davis, R.A., Holcomb, J.D., and Calof, A.L. (1995). Dynamics of MASH1 expression in vitro and in vivo suggest a non-stem cell site of MASH1 action in the olfactory receptor neuron lineage. *Mol. Cell. Neurosci.* **6**, 363–379.
- Gowan, K., Helms, A.W., Hunsaker, T.L., Collisson, T., Ebert, P.J., Odom, R., and Johnson, J.E. (2001). Crossinhibitory activities of Ngn1 and Math1 allow specification of distinct dorsal interneurons. *Neuron* **31**, 219–232.
- Gradwohl, G., Fode, C., and Guillemot, F. (1996). Restricted expression of a novel murine atonal-related bHLH protein in undifferentiated neural precursors. *Dev. Biol.* **180**, 227–241.
- Gratton, M.-O., Torban, E., Jasmin, S.B., Theriault, F.M., German, M.S., and Stifani, S. (2003). Hes6 Promotes Cortical Neurogenesis and Inhibits Hes1 Transcription Repression Activity by Multiple Mechanisms. *Mol. Cell. Biol.* **23**, 6922–6935.
- Grbavec, D., and Stifani, S. (1996). Molecular interaction between TLE1 and the carboxyl-terminal domain of HES-1 containing the WRPW motif. *Biochem. Biophys. Res. Commun.* **223**, 701–705.
- Green, Y.S., and Vetter, M.L. (2011). EBF factors drive expression of multiple classes of target genes governing neuronal development. *Neural Dev* **6**, 8104–8106.

- Groth, C., Nornes, S., McCarty, R., Tamme, R., and Lardelli, M. (2002). Identification of a second presenilin gene in zebrafish with similarity to the human Alzheimer's disease gene presenilin2. *Dev. Genes Evol.* *212*, 486–490.
- Groves, A.K., George, K.M., Tissier-Seta, J.-P., Engel, J.D., Brunet, J.-F., and Anderson, D.J. (1995). Differential regulation of transcription factor gene expression and phenotypic markers in developing sympathetic neurons. *Development* *121*, 887–901.
- Guérin, A., d' Aubenton-Carafa, Y., Marrakchi, E., Da Silva, C., Wincker, P., Mazan, S., and Rétaux, S. (2009). Neurodevelopment Genes in Lampreys Reveal Trends for Forebrain Evolution in Craniates. *PLoS ONE* *4*, e5374.
- Guillemot, F., and Joyner, A.L. (1993). Dynamic expression of the murine Achaete-Scute homologue Mash-1 in the developing nervous system. *Mech. Dev.* *42*, 171–185.
- Guillemot, F., Lo, L.-C., Johnson, J.E., Auerbach, A., Anderson, D.J., and Joyner, A.L. (1993). Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* *75*, 463–476.
- Guillemot, F., Nagy, A., Auerbach, A., Rossant, J., and Joyner, A.L. (1994). Essential role of Mash-2 in extraembryonic development. *Nature* *371*, 333–336.
- Guo, P., Hirano, M., Herrin, B.R., Li, J., Yu, C., Sadlonova, A., and Cooper, M.D. (2009). Dual nature of the adaptive immune system in lampreys. *Nature* *459*, 796–801.
- Hagman, J., and Lukin, K. (2005). Early B-cell factor “pioneers” the way for B-cell development. *Trends Immunol.* *26*, 455–461.
- Hagman, J., Belanger, C., Travis, A., Turck, C.W., and Grosschedl, R. (1993). Cloning and functional characterization of early B-cell factor, a regulator of lymphocyte-specific gene expression. *Genes Dev.* *7*, 760–773.
- Hagman, J., Gutch, M.J., Lin, H., and Grosschedl, R. (1995). EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains. *EMBO J.* *14*, 2907.
- Hall, G.F., Chu, B., Lee, S., Liu, Y., and Yao, J. (2000). The single neurofilament subunit of the lamprey forms filaments and regulates axonal caliber and neuronal size in vivo. *Cell Motil. Cytoskeleton* *46*, 166–182.
- Häming, D., Simoes-Costa, M., Uy, B., Valencia, J., Sauka-Spengler, T., and Bronner-Fraser, M. (2011). Expression of Sympathetic Nervous System Genes in Lamprey Suggests Their Recruitment for Specification of a New Vertebrate Feature. *PLoS ONE* *6*, e26543.
- Hammond, K.L., and Whitfield, T.T. (2006). The developing lamprey ear closely resembles the zebrafish otic vesicle: *otx1* expression can account for all major patterning differences. *Development* *133*, 1347–1357.
- Hans, S., Scheer, N., Riedl, I., v. Weizsäcker, E., Blader, P., and Campos-Ortega, J.A. (2004). *her3*, a zebrafish member of the hairy-E(spl) family, is repressed by Notch signalling. *Development* *131*, 2957–2969.
- Hardisty, M.W. (1961). The growth of larval lampreys. *J. Anim. Ecol.* *30*, 357–371.

- Hardisty, M.W. (2006). Lampreys: Life without jaws (Forrest Text).
- Hatakeyama, J., Bessho, Y., Katoh, K., Ookawara, S., Fujioka, M., Guillemot, F., and Kageyama, R. (2004). Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. *Development* *131*, 5539–5550.
- Heimberg, A.M., Cowper-Sallari, R., Semon, M., Donoghue, P.C.J., and Peterson, K.J. (2010). From the Cover: microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. *Proc. Natl. Acad. Sci.* *107*, 19379–19383.
- Hirano, M., Guo, P., McCurley, N., Schorpp, M., Das, S., Boehm, T., and Cooper, M.D. (2013). Evolutionary implications of a third lymphocyte lineage in lampreys. *Nature* *501*, 435–438.
- Hirata, M., Ito, K., and Tsuneki, K. (1997). Migration and Colonization Patterns of HNK-1-Immunoreactive Neural Crest Cells in Lamprey and Swordtail Embryos. *Zoolog. Sci.* *14*, 305–312.
- Hirsch, M.-R., Tiveron, M.-C., Guillemot, F., Brunet, J.-F., and Goridis, C. (1998). Control of noradrenergic differentiation and Phox2a expression by MASH1 in the central and peripheral nervous system. *Development* *125*, 599–608.
- Holland, L.Z., and Holland, N.D. (1999). Chordate origins of the vertebrate central nervous system. *Curr. Opin. Neurobiol.* *9*, 596–602.
- Holland, L.Z., and Holland, N.D. (2001). Evolution of neural crest and placodes: amphioxus as a model for the ancestral vertebrate? *J. Anat.* *199*, 85–98.
- Holland, L.Z., Schubert, M., Holland, N., and Neuman, T. (2000). Evolutionary Conservation of the Presumptive Neural Plate Markers *AmphiSox1/2/3* and *AmphiNeurogenin* in the Invertebrate Chordate *Amphioxus*. *Dev. Biol.* *226*, 18–33.
- Holland, L.Z., Rached, L.A., Tamme, R., Holland, N.D., Inoko, H., Shiina, T., Burgtorf, C., and Lardelli, M. (2001). 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, N.D., Holland, L.Z., Honma, Y., and Fujii, T. (1993). Engrailed Expression during Development of a Lamprey, *Lampetra japonica*: A Possible Clue to Homologies between Agnathan and Gnathostome Muscles of the Mandibular Arch. *Dev. Growth Differ.* *35*, 153–160.
- Holland, P.W., Garcia-Fernàndez, J., Williams, N.A., and Sidow, A. (1994). Gene duplications and the origins of vertebrate development. *Development* *1994*, 125–133.
- Horigome, N., Myojin, M., Ueki, T., Hirano, S., Aizawa, S., and Kuratani, S. (1999). Development of Cephalic Neural Crest Cells in Embryos of *Lampetra japonica*, with Special Reference to the Evolution of the Jaw. *Dev. Biol.* *207*, 287–308.
- Howard, M.J., Stanke, M., Schneider, C., Wu, X., and Rohrer, H. (2000). The transcription factor dHAND is a downstream effector of BMPs in sympathetic neuron specification. *Development* *127*, 4073–4081.

- Huang, G., Xie, X., Han, Y., Fan, L., Chen, J., Mou, C., Guo, L., Liu, H., Zhang, Q., Chen, S., et al. (2007). The Identification of Lymphocyte-Like Cells and Lymphoid-Related Genes in *Amphioxus* Indicates the Twilight for the Emergency of Adaptive Immune System. *PLoS ONE* 2, e206.
- Hufton, A.L., Groth, D., Vingron, M., Lehrach, H., Poustka, A.J., and Panopoulou, G. (2008). Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement. *Genome Res.* 18, 1582–1591.
- Imai, K.S., Hino, K., Yagi, K., Satoh, N., and Satou, Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* 131, 4047–4058.
- Ishibashi, M., Sasai, Y., Nakanishi, S., and Kageyama, R. (1993). Molecular characterization of HES-2, a mammalian helix-loop-helix factor structurally related to *Drosophila* hairy and Enhancer of split. *Eur. J. Biochem.* 215, 645–652.
- Ishibashi, M., Moriyoshi, K., Sasai, Y., Shiota, K., Nakanishi, S., and Kageyama, R. (1994). Persistent expression of helix-loop-helix factor HES-1 prevents mammalian neural differentiation in the central nervous system. *EMBO J.* 13, 1799.
- Ishibashi, M., Ang, S.L., Shiota, K., Nakanishi, S., Kageyama, R., and Guillemot, F. (1995). Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes Dev.* 9, 3136–3148.
- Jackson, D.J., Meyer, N.P., Seaver, E., Pang, K., McDougall, C., Moy, V.N., Gordon, K., Degnan, B.M., Martindale, M.Q., Burke, R.D., et al. (2010). Developmental expression of COE across the Metazoa supports a conserved role in neuronal cell-type specification and mesodermal development. *Dev. Genes Evol.* 220, 221–234.
- Jandzik, D., Hawkins, M.B., Cattell, M.V., Cerny, R., Square, T.A., and 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.
- Janvier, P. (1996). *Early Vertebrates*. Oxford Monographs on Geology and Geophysics (Oxford: Oxford University Press).
- Janvier, P. (1999). Catching the first fish. *Nature* 402, 21–22.
- Janvier, P. (2006). Modern look for ancient lamprey. *Nature* 443, 921–924.
- Janvier, P. (2008). Early Jawless Vertebrates and Cyclostome Origins. *Zoolog. Sci.* 25, 1045–1056.
- Janvier, P., and Arsenault, M. (2002). Calcification of early vertebrate cartilage. *Nature* 417, 609–610.
- Janvier, P., and Lund, R. (1983). *Hardistiella montanensis* N. gen. et sp. (Petromyzontida) from the Lower Carboniferous of Montana, with Remarks on the Affinities of the Lampreys. *J. Vertebr. Paleontol.* 2, 407–413.

- Jasoni, C.L., and Reh, T.A. (1996). Temporal and spatial pattern of MASH-1 expression in the developing rat retina demonstrates progenitor cell heterogeneity. *J. Comp. Neurol.* **369**, 319–327.
- Jasoni, C.L., Walker, M.B., Morris, M.D., and Reh, T.A. (1994). A chicken achaete-scute homolog (CASH-1) is expressed in a temporally and spatially discrete manner in the developing nervous system. *Development* **120**, 769–783.
- Jeffery, W.R. (2007). Chordate ancestry of the neural crest: New insights from ascidians. *Semin. Cell Dev. Biol.* **18**, 481–491.
- Jhas, S., Ciura, S., Belanger-Jasmin, S., Dong, Z., Llamosas, E., Theriault, F.M., Joachim, K., Tang, Y., Liu, L., Liu, J., et al. (2006). Hes6 Inhibits Astrocyte Differentiation and Promotes Neurogenesis through Different Mechanisms. *J. Neurosci.* **26**, 11061–11071.
- Jiménez-Delgado, S., Crespo, M., Permanyer, J., Garcia-Fernández, J., and Manzanares, M. (2006). Evolutionary genomics of the recently duplicated amphioxus Hairy genes. *Int. J. Biol. Sci.* **2**, 66.
- Jimenez-Guri, E., and Pujades, C. (2011). An ancient mechanism of hindbrain patterning has been conserved in vertebrate evolution. *Evol. Dev.* **13**, 38–46.
- Jimenez-Guri, E., Uchina, F., Colas, J.-F., Sharpe, J., Padrón-Barthe, L., Torres, M., and Pujades, C. (2010). Clonal analysis in mice underlines the importance of rhombomeric boundaries in cell movement restriction during hindbrain segmentation. *PLoS ONE* **5**, e10112.
- Johnels, A.G. (1956). On the peripheral autonomic nervous system of the trunk region of *Lampetra planeri*. *Acta Zool. Stockh.* **37**, 251–286.
- Johnson, J.E., Birren, S.J., and Anderson, D.J. (1990). Two rat homologues of *Drosophila* achaete-scute specifically expressed in neuronal precursors. *Nature* **346**, 858–861.
- Johnson, J.E., Birren, S.J., Saito, T., and Anderson, D.J. (1992). DNA binding and transcriptional regulatory activity of mammalian achaete-scute homologous (MASH) proteins revealed by interaction with a muscle-specific enhancer. *Proc. Natl. Acad. Sci.* **89**, 3596–3600.
- Kaether, C., Haass, C., and Steiner, H. (2006). Assembly, Trafficking and Function of γ -Secretase. *Neurodegener. Dis.* **3**, 275–283.
- Kageyama, R., Ohtsuka, T., and Kobayashi, T. (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development* **134**, 1243–1251.
- Kageyama, R., Ohtsuka, T., and Kobayashi, T. (2008). Roles of Hes genes in neural development. *Dev. Growth Differ.* **50**, S97–S103.
- Kasamatsu, J., Sutoh, Y., Fugo, K., Otsuka, N., Iwabuchi, K., and Kasahara, M. (2010). Identification of a third variable lymphocyte receptor in the lamprey. *Proc. Natl. Acad. Sci.* **107**, 14304–14308.
- Kawano-Yamashita, E., Terakita, A., Koyanagi, M., Shichida, Y., Oishi, T., and Tamotsu, S. (2007). Immunohistochemical characterization of a parainopsin-containing photoreceptor cell involved in the ultraviolet/green discrimination in the pineal organ of the river lamprey *Lethenteron japonicum*. *J. Exp. Biol.* **210**, 3821–3829.

- Kiecker, C., and Lumsden, A. (2005). Compartments and their boundaries in vertebrate brain development. *Nat. Rev. Neurosci.* 6, 553–564.
- Kilarski, W., and Plytycz, B. (1981). The presence of plasma cells in the lamprey (Agnatha). *Dev. Comp. Immunol.* 5, 361–366.
- Kim, H., Shin, J., Kim, S., Poling, J., Park, H.-C., and Appel, B. (2008a). Notch-regulated oligodendrocyte specification from radial glia in the spinal cord of zebrafish embryos. *Dev. Dyn.* 237, 2081–2089.
- Kim, H., Kim, S., Chung, A.-Y., Bae, Y.-K., Hibi, M., Lim, C.S., and Park, H.-C. (2008b). Notch-regulated perineurium development from zebrafish spinal cord. *Neurosci. Lett.* 448, 240–244.
- Kim, H.M., Oh, S.C., Lim, K.J., Kasamatsu, J., Heo, J.Y., Park, B.S., Lee, H., Yoo, O.J., Kasahara, M., and Lee, J.-O. (2006). Structural Diversity of the Hagfish Variable Lymphocyte Receptors. *J. Biol. Chem.* 282, 6726–6732.
- Kimberly, W.T., Xia, W., Rahmati, T., Wolfe, M.S., and Selkoe, D.J. (2000). The transmembrane aspartates in presenilin 1 and 2 are obligatory for γ -secretase activity and amyloid β -protein generation. *J. Biol. Chem.* 275, 3173–3178.
- Klein, T., and Bischoff, R. (2011). Active Metalloproteases of the A Disintegrin And Metalloprotease (ADAM) Family: Biological Function and Structure. *J. Proteome Res.* 10, 17–33.
- Kluge, B., Renault, N., and Rohr, K.B. (2005). Anatomical and molecular reinvestigation of lamprey endostyle development provides new insight into thyroid gland evolution. *Dev. Genes Evol.* 215, 32–40.
- Kopan, R., and Ilagan, M.X.G. (2009). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell* 137, 216–233.
- Korzh, V., Sleptsova, I., Liao, J., He, J., and Gong, Z. (1998). Expression of zebrafish bHLH genes *ngn1* and *nrd* defines distinct stages of neural differentiation. *Dev. Dyn.* 213, 92–104.
- Koyama, H., Kishida, R., Goris, R.C., and Kusunoki, T. (1990). Organization of the primary projections of the lateral line nerves in the lamprey *Lampetra japonica*. *J. Comp. Neurol.* 295, 277–289.
- Koyano-Nakagawa, N., Kim, J., Anderson, D., and Kintner, C. (2000). Hes6 acts in a positive feedback loop with the neurogenins to promote neuronal differentiation. *Development* 127, 4203–4216.
- Kozmik, Z., Holland, N.D., Kalousova, A., Paces, J., Schubert, M., and 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.
- Krüger, M., Schmid, T., Krüger, S., Bober, E., and Braun, T. (2006). Functional redundancy of NSCL-1 and NeuroD during development of the petrosal and vestibulocochlear ganglia. *Eur. J. Neurosci.* 24, 1581–1590.

- Von Kupffer, C. (1895). Studien zur Vergleichenden Entwicklungsgeschichte des Kopfes der Kranioten. III. Die Entwicklung der Kopfnerven von *Ammocoetes planeri* (Munich: J. F. Lehmann).
- Kuraku, S., and Kuratani, S. (2006). Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. *Zoolog. Sci.* 23, 1053–1064.
- Kuraku, S., Hoshiyama, D., Katoh, K., Suga, H., and Miyata, T. (1999). Monophyly of Lampreys and Hagfishes Supported by Nuclear DNA-Coded Genes. *J. Mol. Evol.* 49, 729–735.
- Kuraku, S., Meyer, A., and Kuratani, S. (2008). 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., Aizawa, S., and Hirano, S. (1997). Peripheral development of cranial nerves in a cyclostome, *Lampetra japonica*: morphological distribution of nerve branches and the vertebrate body plan. *J. Comp. Neurol.* 384, 483–500.
- Kuratani, S., Horigome, N., Ueki, T., Aizawa, S., and Hirano, S. (1998). Stereotyped axonal bundle formation and neuromeric patterns in embryos of a cyclostome, *Lampetra japonica*. *J. Comp. Neurol.* 391, 99–114.
- Kuratani, S., Nobusada, Y., Horigome, N., and Shigetani, Y. (2001). Embryology of the lamprey and evolution of the vertebrate jaw: insights from molecular and developmental perspectives. *Philos. Trans. R. Soc. B Biol. Sci.* 356, 1615–1632.
- Kuroda, N., Uinuk-ool, T.S., Sato, A., Samonte, I.E., Figueroa, F., Mayer, W.E., and Klein, J. (2003). Identification of chemokines and a chemokine receptor in cichlid fish, shark, and lamprey. *Immunogenetics* 54, 884–895.
- Lacalli, T.C. (2001). New perspectives on the evolution of protochordate sensory and locomotory systems, and the origin of brains and heads. *Philos. Trans. R. Soc. B Biol. Sci.* 356, 1565–1572.
- Lacomme, M., Liaubet, L., Pituello, F., and Bel-Vialar, S. (2012). *NEUROG2* Drives Cell Cycle Exit of Neuronal Precursors by Specifically Repressing a Subset of Cyclins Acting at the G1 and S Phases of the Cell Cycle. *Mol. Cell. Biol.* 32, 2596–2607.
- Latimer, A.J., Shin, J., and Appel, B. (2005). *her9* promotes floor plate development in zebrafish. *Dev. Dyn.* 232, 1098–1104.
- Ledent, V., and Vervoort, M. (2001). The Basic Helix-Loop-Helix Protein Family: Comparative Genomics and Phylogenetic Analysis. *Genome Res.* 11, 754–770.
- Ledent, V., Paquet, O., and Vervoort, M. (2002). Phylogenetic analysis of the human basic helix-loop-helix proteins. *Genome Biol* 3, 1–18.
- Lemaire, P., Bertrand, V., and Hudson, C. (2002). Early Steps in the Formation of Neural Tissue in Ascidian Embryos. *Dev. Biol.* 252, 151–169.
- Li, H., and Richardson, W.D. (2008). The evolution of Olig genes and their roles in myelination. *Neuron Glia Biol.* 4, 129–135.

- Li, C.-M., Yan, R.-T., and Wang, S.-Z. (1999a). Misexpression of a bHLH gene, cNSCL1, results in abnormal brain development. *Dev. Dyn.* 215, 238–247.
- Li, C.-M., Yan, R.-T., and Wang, S.-Z. (1999b). Misexpression of cNSCL1 disrupts retinal development. *Mol. Cell. Neurosci.* 14, 17–27.
- Li, J., Das, S., Herrin, B.R., Hirano, M., and Cooper, M.D. (2013). Definition of a third VLR gene in hagfish. *Proc. Natl. Acad. Sci.* 110, 15013–15018.
- Li, S., Yin, M., Liu, S., Chen, Y., Yin, Y., Liu, T., and Zhou, J. (2010). Expression of ventral diencephalon-enriched genes in zebrafish. *Dev. Dyn.* 239, 3368–3379.
- Liberg, D., Sigvardsson, M., and Akerblad, P. (2002). The EBF/Olf/Collier Family of Transcription Factors: Regulators of Differentiation in Cells Originating from All Three Embryonal Germ Layers. *Mol. Cell. Biol.* 22, 8389–8397.
- Ligon, K.L., Fancy, S.P.J., Franklin, R.J.M., and Rowitch, D.H. (2006). Olig gene function in CNS development and disease. *Glia* 54, 1–10.
- Lin, H., and Grosschedl, R. (1995). Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* 376, 263–267.
- Lin, Y.C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C.A., Dutkowski, J., et al. (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* 11, 635–643.
- Liu, K.J., and Harland, R.M. (2003). Cloning and characterization of *Xenopus* Id4 reveals differing roles for Id genes. *Dev. Biol.* 264, 339–351.
- Liu, W., and Zhao, C. (2010). Genome-Wide Identification and Analysis of the Chicken Basic Helix-Loop-Helix Factors. *Comp. Funct. Genomics* 2010, 1–12.
- Lo, L., Tiveron, M.-C., and Anderson, D.J. (1998). MASH1 activates expression of the paired homeodomain transcription factor Phox2a, and couples pan-neuronal and subtype-specific components of autonomic neuronal identity. *Development* 125, 609–620.
- Lo, L.C., Johnson, J.E., Wuenschell, C.W., Saito, T., and Anderson, D.J. (1991). Mammalian achaete-scute homolog 1 is transiently expressed by spatially restricted subsets of early neuroepithelial and neural crest cells. *Genes Dev.* 5, 1524–1537.
- Lobe, C.G. (1997). Expression of the helix-loop-helix factor, Hes3, during embryo development suggests a role in early midbrain-hindbrain patterning. *Mech. Dev.* 62, 227–237.
- Longo, A., Guanga, G.P., and Rose, R.B. (2008). Crystal Structure of E47–NeuroD1/Beta2 bHLH Domain–DNA Complex: Heterodimer Selectivity and DNA Recognition. *Biochemistry (Mosc.)* 47, 218–229.
- Lucas, M.E., Müller, F., Rüdiger, R., Henion, P.D., and Rohrer, H. (2006). The bHLH transcription factor hand2 is essential for noradrenergic differentiation of sympathetic neurons. *Development* 133, 4015–4024.

- Lukaszewicz, A.I., and Anderson, D.J. (2011). Cyclin D1 promotes neurogenesis in the developing spinal cord in a cell cycle-independent manner. *Proc. Natl. Acad. Sci.* *108*, 11632–11637.
- Lukin, K., Fields, S., Hartley, J., and Hagman, J. (2008). Early B cell factor: Regulator of B lineage specification and commitment. *Semin. Immunol.* *20*, 221–227.
- Lütolf, S., Radtke, F., Aguet, M., Suter, U., and Taylor, V. (2002). Notch1 is required for neuronal and glial differentiation in the cerebellum. *Development* *129*, 373–385.
- Ma, Q., Kintner, C., and Anderson, D.J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* *87*, 43–52.
- Ma, Q., Sommer, L., Cserjesi, P., and Anderson, D.J. (1997). Mash1 and neurogenin1 expression patterns define complementary domains of neuroepithelium in the developing CNS and are correlated with regions expressing notch ligands. *J. Neurosci.* *17*, 3644–3652.
- Ma, Q., Chen, Z., Barrantes, I. del B., Luis de la Pompa, J., and Anderson, D.J. (1998). neurogenin1 Is Essential for the Determination of Neuronal Precursors for Proximal Cranial Sensory Ganglia. *Neuron* *20*, 469–482.
- Ma, Q., Fode, C., Guillemot, F., and Anderson, D.J. (1999). NEUROGENIN1 and NEUROGENIN2 control two distinct waves of neurogenesis in developing dorsal root ganglia. *Genes Dev.* *13*, 1717–1728.
- Ma, Q., Anderson, D.J., and Fritzsche, B. (2000). Neurogenin 1 Null Mutant Ears Develop Fewer, Morphologically Normal Hair Cells in Smaller Sensory Epithelia Devoid of Innervation. *J. Assoc. Res. Otolaryngol.* *1*, 129–143.
- Maier, E., and Gunhaga, L. (2009). Dynamic expression of neurogenic markers in the developing chick olfactory epithelium. *Dev. Dyn.* *238*, 1617–1625.
- Malgaretti, N., Pozzoli, O., Bosetti, A., Corradi, A., Ciarmatori, S., Panigada, M., Bianchi, M.E., Martinez, S., and Consalez, G.G. (1997). Mmot1, a new helix-loop-helix transcription factor gene displaying a sharp expression boundary in the embryonic mouse brain. *J. Biol. Chem.* *272*, 17632–17639.
- Mallatt, J., and Sullivan, J. (1998). 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. *Mol. Biol. Evol.* *15*, 1706–1718.
- Mallatt, J., and Winchell, C.J. (2007). Ribosomal RNA genes and deuterostome phylogeny revisited: More cyclostomes, elasmobranchs, reptiles, and a brittle star. *Mol. Phylogenet. Evol.* *43*, 1005–1022.
- Malmqvist, B., and Brönmark, C. (1982). Filter Feeding in Larval *Lampetra planeri*: Effects of Size, Temperature and Particle Concentration. *Oikos* *38*, 40–46.
- Matei, V., Pauley, S., Kaing, S., Rowitch, D., Beisel, K.W., Morris, K., Feng, F., Jones, K., Lee, J., and Fritzsche, B. (2005). Smaller inner ear sensory epithelia in Neurog1 null mice are related to earlier hair cell cycle exit. *Dev. Dyn.* *234*, 633–650.
- Matter-Sadzinski, L., Matter, J.-M., Ong, M.-T., Hernandez, J., and Ballivet, M. (2001). Specification of neurotransmitter receptor identity in developing retina: the chick ATH5

- promoter integrates the positive and negative effects of several bHLH proteins. *Development* **128**, 217–231.
- Mayer, W.E., Uinuk-ool, T., Tichy, H., Gartland, L.A., Klein, J., and Cooper, M.D. (2002a). Isolation and characterization of lymphocyte-like cells from a lamprey. *Proc. Natl. Acad. Sci.* **99**, 14350–14355.
- Mayer, W.E., O’huigin, C., Tichy, H., Terzic, J., and Saraga-Babic, M. (2002b). Identification of Two Ikaros-like Transcription Factors in Lamprey. *Scand. J. Immunol.* **55**, 162–170.
- Mazet, F., and Shimeld, S.M. (2005). Molecular evidence from ascidians for the evolutionary origin of vertebrate cranial sensory placodes. *J. Exp. Zool. B Mol. Dev. Evol.* **304**, 340–346.
- Mazet, F., Masood, S., Luke, G.N., Holland, N.D., and Shimeld, S.M. (2004). Expression of *AmphiCoe*, an amphioxus COE/EBF gene, in the developing central nervous system and epidermal sensory neurons. *Genesis* **38**, 58–65.
- Mazet, F., Hutt, J.A., Milloz, J., Millard, J., Graham, A., and Shimeld, S.M. (2005). Molecular evidence from *Ciona intestinalis* for the evolutionary origin of vertebrate sensory placodes. *Dev. Biol.* **282**, 494–508.
- McCauley, D.W., and Bronner-Fraser, M. (2002). Conservation of Pax gene expression in ectodermal placodes of the lamprey. *Gene* **287**, 129–139.
- McCauley, D.W., and Bronner-Fraser, M. (2003). Neural crest contributions to the lamprey head. *Development* **130**, 2317–2327.
- McCauley, D.W., and Bronner-Fraser, M. (2006). Importance of SoxE in neural crest development and the evolution of the pharynx. *Nature* **441**, 750–752.
- McGraw, H.F., Nechiporuk, A., and Raible, D.W. (2008). Zebrafish Dorsal Root Ganglia Neural Precursor Cells Adopt a Glial Fate in the Absence of Neurogenin1. *J. Neurosci.* **28**, 12558–12569.
- McPherson, C.E., Varley, J.E., and Maxwell, G.D. (2000). Expression and Regulation of Type I BMP Receptors during Early Avian Sympathetic Ganglion Development. *Dev. Biol.* **221**, 220–232.
- Medina, L., and Abellán, A. (2009). Development and evolution of the pallium. *Semin. Cell Dev. Biol.* **20**, 698–711.
- Medina, L., Legaz, I., González, G., De Castro, F., Rubenstein, J.L.R., and Puelles, L. (2004). Expression of *Dbx1*, *Neurogenin 2*, *Semaphorin 5A*, *Cadherin 8*, and *Emx1* distinguish ventral and lateral pallial histogenetic divisions in the developing mouse claustroramygdaloid complex. *J. Comp. Neurol.* **474**, 504–523.
- Medina, L., Bupesh, M., and Abellán, A. (2011). Contribution of Genoarchitecture to Understanding Forebrain Evolution and Development, with Particular Emphasis on the Amygdala. *Brain. Behav. Evol.* **78**, 216–236.
- Meierhans, D., el-Ariss, C., Neuenschwander, M., Sieber, M., Stackhouse, J.F., and Allemann, R.K. (1995). DNA binding specificity of the basic-helix-loop-helix protein MASH-1. *Biochemistry (Mosc.)* **34**, 11026–11036.

- Meinertzhagen, I.A., Lemaire, P., and Okamura, Y. (2004). The Neurobiology of the Ascidian Tadpole Larva: Recent Developments in an Ancient Chordate. *Annu. Rev. Neurosci.* 27, 453–485.
- Meléndez-Ferro, M., Villar-Cheda, B., Manoel Abalo, X., Pérez-Costas, E., Rodríguez-Muñoz, R., Degrip, W.J., Yáñez, J., Rodicio, M.C., and Anadón, R. (2002). Early development of the retina and pineal complex in the sea lamprey: Comparative immunocytochemical study. *J. Comp. Neurol.* 442, 250–265.
- Meléndez-Ferro, M., Pérez-Costas, E., Villar-Cheda, B., Rodríguez-Muñoz, R., Anadón, R., and Rodicio, M.C. (2003). Ontogeny of gamma-aminobutyric acid-immunoreactive neurons in the rhombencephalon and spinal cord of the sea lamprey. *J. Comp. Neurol.* 464, 17–35.
- Mella, S., Soula, C., Morello, D., Crozatier, M., and Vincent, A. (2004). Expression patterns of the *coe/ebf* transcription factor genes during chicken and mouse limb development. *Gene Expr. Patterns* 4, 537–542.
- Merchán, P., Bardet, S.M., Puellas, L., and Ferran, J.L. (2011). Comparison of pretectal genoarchitectonic pattern between quail and chicken embryos. *Front. Neuroanat.* 5.
- Meulemans, D., and Bronner-Fraser, M. (2002). Amphioxus and lamprey AP-2 genes: implications for neural crest evolution and migration patterns. *Development* 129, 4953–4962.
- Meulemans, D., McCauley, D., and Bronner-Fraser, M. (2003). Id expression in amphioxus and lamprey highlights the role of gene cooption during neural crest evolution. *Dev. Biol.* 264, 430–442.
- Milán, F.J., and Puellas, L. (2000). Patterns of calretinin, calbindin, and tyrosine-hydroxylase expression are consistent with the prosomeric map of the frog diencephalon. *J. Comp. Neurol.* 419, 96–121.
- Minguillon, C., Jiménez-Delgado, S., Panopoulou, G., and Garcia-Fernández, J. (2003). The amphioxus *Hairy* family: differential fate after duplication. *Development* 130, 5903–5914.
- Mizuguchi, R., Sugimori, M., Takebayashi, H., Kosako, H., Nagao, M., Yoshida, S., Nabeshima, Y., Shimamura, K., and Nakafuku, M. (2001). Combinatorial roles of *olig2* and *neurogenin2* in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron* 31, 757–771.
- Modrell, M.S., Hockman, D., Uy, B., Buckley, D., Sauka-Spengler, T., Bronner, M.E., and Baker, C.V.H. (2014). A fate-map for cranial sensory ganglia in the sea lamprey. *Dev. Biol.* 385, 405–416.
- Morohashi, Y., Kan, T., Tominari, Y., Fuwa, H., Okamura, Y., Watanabe, N., Sato, C., Natsugari, H., Fukuyama, T., Iwatsubo, T., et al. (2006). C-terminal Fragment of Presenilin Is the Molecular Target of a Dipeptidic γ -Secretase-specific Inhibitor DAPT (N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-Butyl Ester). *J. Biol. Chem.* 281, 14670–14676.
- Morrison, S.J. (2001). Neuronal differentiation: Proneural genes inhibit gliogenesis. *Curr. Biol.* 11, R349–R351.
- Mueller, T., and Wullmann, M.F. (2003). Anatomy of neurogenesis in the early zebrafish brain. *Dev. Brain Res.* 140, 137–155.

- Murai, K., Philpott, A., and Jones, P.H. (2011). Hes6 is required for the neurogenic activity of neurogenin and NeuroD. *PLoS One* 6, e27880.
- Murakami, Y., Ogasawara, M., Sugahara, F., Hirano, S., Satoh, N., and Kuratani, S. (2001). Identification and expression of the lamprey Pax6 gene: evolutionary origin of the segmented brain of vertebrates. *Development* 128, 3521–3531.
- Murakami, Y., Ogasawara, M., Satoh, N., Sugahara, F., Myojin, M., Hirano, S., and Kuratani, S. (2002). Compartments in the lamprey embryonic brain as revealed by regulatory gene expression and the distribution of reticulospinal neurons. *Brain Res. Bull.* 57, 271–275.
- Murakami, Y., Pasqualetti, M., Takio, Y., Hirano, S., Rijli, F.M., and Kuratani, S. (2004). Segmental development of reticulospinal and branchiomotor neurons in lamprey: insights into the evolution of the vertebrate hindbrain. *Development* 131, 983–995.
- Murakami, Y., Uchida, K., Rijli, F.M., and Kuratani, S. (2005). Evolution of the brain developmental plan: Insights from agnathans. *Dev. Biol.* 280, 249–259.
- Nagawa, F., Kishishita, N., Shimizu, K., Hirose, S., Miyoshi, M., Nezu, J., Nishimura, T., Nishizumi, H., Takahashi, Y., Hashimoto, S., et al. (2007). Antigen-receptor genes of the agnathan lamprey are assembled by a process involving copy choice. *Nat. Immunol.* 8, 206–213.
- Nakao, T., and Ishizawa, A. (1987). Development of the spinal nerves in the lamprey: III. Spinal ganglia and dorsal roots in 26-day (13 mm) larvae. *J. Comp. Neurol.* 256, 369–385.
- Newman, M.J., and Trewin, N.H. (2001). A new jawless vertebrate from the Middle Devonian of Scotland. *Palaeontology* 44, 43–51.
- Nicol, J.A.C. (1952). Autonomic nervous systems in lower chordates. *Biol. Rev.* 27, 1–48.
- Nieber, F., Pieler, T., and Henningfeld, K.A. (2009). Comparative expression analysis of the neurogenins in *Xenopus tropicalis* and *Xenopus laevis*. *Dev. Dyn.* 238, 451–458.
- Nieminen, P., Liippo, J., and Lassila, O. (2000). Pax-5 and EBF are Expressed in Committed B-Cell Progenitors Prior to the Colonization of the Embryonic Bursa of Fabricius. *Scand. J. Immunol.* 52, 465–469.
- Nieto, M., Schuurmans, C., Britz, O., and Guillemot, F. (2001). Neural bHLH genes control the neuronal versus glial fate decision in cortical progenitors. *Neuron* 29, 401–413.
- Nikitina, N., Sauka-Spengler, T., and Bronner-Fraser, M. (2008). Dissecting early regulatory relationships in the lamprey neural crest gene network. *Proc. Natl. Acad. Sci.* 105, 20083–20088.
- Nikitina, N., Tong, L., and Bronner, M.E. (2011). Ancestral network module regulating *prdm1* expression in the lamprey neural plate border. *Dev. Dyn.* 240, 2265–2271.
- Nilsson, S. (2011). Comparative anatomy of the autonomic nervous system. *Auton. Neurosci.* 165, 3–9.
- Nornes, S., Newman, M., Verdile, G., Wells, S., Stoick-Cooper, C.L., Tucker, B., Frederich-Sleptsova, I., Martins, R., and Lardelli, M. (2007). Interference with splicing of Presenilin

- transcripts has potent dominant negative effects on Presenilin activity. *Hum. Mol. Genet.* **17**, 402–412.
- Nornes, S., Newman, M., Wells, S., Verdile, G., Martins, R.N., and Lardelli, M. (2009). Independent and cooperative action of Psen2 with Psen1 in zebrafish embryos. *Exp. Cell Res.* **315**, 2791–2801.
- Northcutt, R.G. (1997). Evolution of gnathostome lateral line ontogenies. *Brain. Behav. Evol.* **50**, 25–37.
- Novitsch, B.G., Chen, A.I., and Jessell, T.M. (2001). Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron* **31**, 773–789.
- Ogasawara, M., Shigetani, Y., Suzuki, S., Kuratani, S., and Satoh, N. (2001). Expression of Thyroid transcription factor-1 (TTF-1) gene in the ventral forebrain and endostyle of the agnathan vertebrate, *Lampetra japonica*. *Genesis* **30**, 51–58.
- Ohno, S. (1970). *Evolution by gene duplication* (Berlin: Springer).
- Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., and Kageyama, R. (1999). Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. *EMBO J.* **18**, 2196–2207.
- Ohtsuka, T., Sakamoto, M., Guillemot, F., and Kageyama, R. (2001). Roles of the Basic Helix-Loop-Helix Genes Hes1 and Hes5 in Expansion of Neural Stem Cells of the Developing Brain. *J. Biol. Chem.* **276**, 30467–30474.
- O'Neill, P., McCole, R.B., and Baker, C.V.H. (2007). A molecular analysis of neurogenic placode and cranial sensory ganglion development in the shark, *Scyliorhinus canicula*. *Dev. Biol.* **304**, 156–181.
- O'Neill, P., Mak, S.-S., Fritzsche, B., Ladher, R.K., and Baker, C.V.H. (2012). The amniote paratympanic organ develops from a previously undiscovered sensory placode. *Nat. Commun.* **3**, 1041.
- Onimaru, K., Shoguchi, E., Kuratani, S., and Tanaka, M. (2011). Development and evolution of the lateral plate mesoderm: Comparative analysis of amphioxus and lamprey with implications for the acquisition of paired fins. *Dev. Biol.* **359**, 124–136.
- Osorio, J., Mazan, S., and Rétaux, S. (2005). Organisation of the lamprey (*Lampetra fluviatilis*) embryonic brain: Insights from LIM-homeodomain, Pax and hedgehog genes. *Dev. Biol.* **288**, 100–112.
- Osório, J., Mueller, T., Rétaux, S., Vernier, P., and Wullimann, M.F. (2010). Phylotypic expression of the bHLH genes Neurogenin2, NeuroD, and Mash1 in the mouse embryonic forebrain. *J. Comp. Neurol.* **518**, 851–871.
- Ota, K.G., Fujimoto, S., Oisi, Y., and Kuratani, S. (2011). Identification of vertebra-like elements and their possible differentiation from sclerotomes in the hagfish. *Nat. Commun.* **2**, 373.
- Pancer, Z., and Cooper, M.D. (2006). The evolution of adaptive immunity. *Annu. Rev. Immunol.* **24**, 497–518.

- Pancer, Z., Mayer, W.E., Klein, J., and Cooper, M.D. (2004a). Prototypic T cell receptor and CD4-like coreceptor are expressed by lymphocytes in the agnathan sea lamprey. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 13273–13278.
- Pancer, Z., Amemiya, C.T., Ehrhardt, G.R., Ceitlin, J., Gartland, G.L., and Cooper, M.D. (2004b). Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* *430*, 174–180.
- Pancer, Z., Saha, N.R., Kasamatsu, J., Suzuki, T., Amemiya, C.T., Kasahara, M., and Cooper, M.D. (2005). Variable lymphocyte receptors in hagfish. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 9224–9229.
- Pang, K., Matus, D.Q., and Martindale, M.Q. (2004). The ancestral role of COE genes may have been in chemoreception: evidence from the development of the sea anemone, *Nematostella vectensis* (Phylum Cnidaria; Class Anthozoa). *Dev. Genes Evol.* *214*, 134–138.
- Panopoulou, G., and Poustka, A. (2005). Timing and mechanism of ancient vertebrate genome duplications – the adventure of a hypothesis. *Trends Genet.* *21*, 559–567.
- Pasini, A., Henrique, D., and Wilkinson, D.G. (2001). The zebrafish Hairy/Enhancer-of-split-related gene *her6* is segmentally expressed during the early development of hindbrain and somites. *Mech. Dev.* *100*, 317–321.
- Perez, S.E., Rebelo, S., and Anderson, D.J. (1999). Early specification of sensory neuron fate revealed by expression and function of neurogenins in the chick embryo. *Development* *126*, 1715–1728.
- Perron, M., Opdecamp, K., Butler, K., Harris, W.A., and Bellefroid, E.J. (1999). *X-ngnr-1* and *Xath3* promote ectopic expression of sensory neuron markers in the neurula ectoderm and have distinct inducing properties in the retina. *Proc. Natl. Acad. Sci.* *96*, 14996–15001.
- Pierani, A., and Wassef, M. (2009). Cerebral cortex development: From progenitors patterning to neocortical size during evolution: Neocortical development and evolution. *Dev. Growth Differ.* *51*, 325–342.
- Pissarra, L., Henrique, D., and Duarte, A. (2000). Expression of *hes6*, a new member of the Hairy/Enhancer-of-split family, in mouse development. *Mech. Dev.* *95*, 275–278.
- Podlisny, M.B., Citron, M., Amarante, P., Sherrington, R., Xia, W., Zhang, J., Diehl, T., Levesque, G., Fraser, P., Haass, C., et al. (1997). Presenilin proteins undergo heterogeneous endoproteolysis between Thr291 and Ala299 and occur as stable N- and C-terminal fragments in normal and Alzheimer brain tissue. *Neurobiol. Dis.* *3*, 325–337.
- Pogoda, H.-M., von der Hardt, S., Herzog, W., Kramer, C., Schwarz, H., and Hammerschmidt, M. (2006). The proneural gene *ascl1a* is required for endocrine differentiation and cell survival in the zebrafish adenohypophysis. *Development* *133*, 1079–1089.
- Poitras, L., Ghanem, N., Hatch, G., and Ekker, M. (2007). The proneural determinant MASH1 regulates forebrain *Dlx1/2* expression through the *I12b* intergenic enhancer. *Development* *134*, 1755–1765.
- Pombal, M.A., and Puelles, L. (1999). Prosomeric map of the lamprey forebrain based on calretinin immunocytochemistry, Nissl stain, and ancillary markers. *J. Comp. Neurol.* *414*, 391–422.

- Pombal, M.A., Megías, M., Bardet, S.M., and Puellas, L. (2009). New and Old Thoughts on the Segmental Organization of the Forebrain in Lampreys. *Brain. Behav. Evol.* **74**, 7–19.
- De la Pompa, J.L., Wakeham, A., Correia, K.M., Samper, E., Brown, S., Aguilera, R.J., Nakano, T., Honjo, T., Mak, T.W., Rossant, J., et al. (1997). Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* **124**, 1139–1148.
- Porteus, M.H., Bulfone, A., Liu, J.K., Puellas, L., Lo, L.C., and Rubenstein, J.L. (1994). DLX-2, MASH-1, and MAP-2 expression and bromodeoxyuridine incorporation define molecularly distinct cell populations in the embryonic mouse forebrain. *J. Neurosci.* **14**, 6370–6383.
- Pozzoli, O., Bosetti, A., Croci, L., Consalez, G.G., and Vetter, M.L. (2001). Xebf3 Is a Regulator of Neuronal Differentiation during Primary Neurogenesis in *Xenopus*. *Dev. Biol.* **233**, 495–512.
- Preston, A.R., and Eichenbaum, H. (2013). Interplay of Hippocampus and Prefrontal Cortex in Memory. *Curr. Biol.* **23**, R764–R773.
- Puelles, L. (2001). Thoughts on the development, structure and evolution of the mammalian and avian telencephalic pallium. *Philos. Trans. R. Soc. B Biol. Sci.* **356**, 1583–1598.
- Puelles, L., and Rubenstein, J.L. (1993). Expression patterns of homeobox and other putative regulatory genes in the embryonic mouse forebrain suggest a neuromeric organization. *Trends Neurosci.* **16**, 472–479.
- Puelles, L., and Rubenstein, J.L.R. (2003). Forebrain gene expression domains and the evolving prosomeric model. *Trends Neurosci.* **26**, 469–476.
- Puelles, L., Kuwana, E., Puellas, E., Bulfone, A., Shimamura, K., Keleher, J., Smiga, S., and Rubenstein, J.L. (2000). Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes *Dlx-2*, *Emx-1*, *Nkx-2.1*, *Pax-6*, and *Tbr-1*. *J. Comp. Neurol.* **424**, 409–438.
- Purnell, M.A., and von Bitter, P.H. (1992). Blade-shaped conodont elements functioned as cutting teeth. *Nature* **359**, 629–631.
- Putnam, N.H., Butts, T., Ferrier, D.E.K., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.-K., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**, 1064–1071.
- Raft, S., Koundakjian, E.J., Quinones, H., Jayasena, C.S., Goodrich, L.V., Johnson, J.E., Segil, N., and Groves, A.K. (2007). Cross-regulation of *Ngn1* and *Math1* coordinates the production of neurons and sensory hair cells during inner ear development. *Development* **134**, 4405–4415.
- Raible, D.W., and Ungos, J.M. (2006). Specification of sensory neuron cell fate from the neural crest. *Adv. Exp. Med. Biol.* **589**, 170–180.
- Rasmussen, S.L.K., Holland, L.Z., Schubert, M., Beaster-Jones, L., and Holland, N.D. (2007). *Amphioxus* *AmphiDelta*: evolution of delta protein structure, segmentation, and neurogenesis. *Genesis* **45**, 113–122.
- Richardson, M.K., Admiraal, J., and Wright, G.M. (2010). Developmental anatomy of lampreys. *Biol. Rev.* **85**, 1–33.

- Rogozin, I.B., Iyer, L.M., Liang, L., Glazko, G.V., Liston, V.G., Pavlov, Y.I., Aravind, L., and Pancer, Z. (2007). Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat. Immunol.* 8, 647–656.
- Ronan, M., and Northcutt, R.G. (1987). Primary projections of the lateral line nerves in adult lampreys. *Brain. Behav. Evol.* 30, 62–81.
- Rossner, M.J., Dorr, J., Gass, P., Schwab, M.H., and Nave, K.-A. (1997). SHARPs: Mammalian Enhancer-of-Split- and Hairy-Related Proteins Coupled to Neuronal Stimulation. *Mol. Cell. Neurosci.* 10, 209.
- Rotenstein, L., Herath, K., Gould, R.M., and de Bellard, M.E. (2008). Characterization of the Shark Myelin Po Protein. *Brain. Behav. Evol.* 72, 48–58.
- Rotenstein, L., Milanes, A., Juarez, M., Reyes, M., and de Bellard, M.E. (2009). Embryonic development of glial cells and myelin in the shark, *Chiloscyllium punctatum*. *Gene Expr. Patterns* 9, 572–585.
- Sailsbery, J.K., Atchley, W.R., and Dean, R.A. (2012). Phylogenetic analysis and classification of the fungal bHLH domain. *Mol. Biol. Evol.* 29, 1301–1318.
- Saka, Y., and Smith, J.C. (2001). Spatial and Temporal Patterns of Cell Division during Early *Xenopus* Embryogenesis. *Dev. Biol.* 229, 307–318.
- Sanders, T.A., Lumsden, A., and Ragsdale, C.W. (2002). Arcuate plan of chick midbrain development. *J. Neurosci.* 22, 10742–10750.
- Sanes, D.H., Reh, T.A., and Harris, W.A. (2006). *Development of the Nervous System* (London: Academic Press).
- Sansom, R.S. (2008). The origin and early evolution of the Osteostraci (Vertebrata): A phylogeny for the Thyestiida. *J. Syst. Palaeontol.* 6, 317–332.
- Sansom, R.S. (2009). Phylogeny, classification and character polarity of the Osteostraci (Vertebrata). *J. Syst. Palaeontol.* 7, 95–115.
- Sansom, R.S., Gabbott, S.E., and Purnell, M.A. (2011). Decay of vertebrate characters in hagfish and lamprey (Cyclostomata) and the implications for the vertebrate fossil record. *Proc. R. Soc. B Biol. Sci.* 278, 1150–1157.
- Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila* hairy and Enhancer of split. *Genes Dev.* 6, 2620–2634.
- Sato, A., Uinuk-ool, T.S., Kuroda, N., Mayer, W.E., Takezaki, N., Dongak, R., Figueroa, F., Cooper, M.D., and Klein, J. (2003). Macrophage migration inhibitory factor (MIF) of jawed and jawless fishes: implications for its evolutionary origin. *Dev. Comp. Immunol.* 27, 401–412.
- Sato, T., Diehl, T.S., Narayanan, S., Funamoto, S., Ihara, Y., De Strooper, B., Steiner, H., Haass, C., and Wolfe, M.S. (2007). Active gamma-secretase complexes contain only one of each component. *J. Biol. Chem.* 282, 33985–33993.

- Satou, Y., Imai, K.S., Levine, M., Kohara, Y., Rokhsar, D., and Satoh, N. (2003). A genomewide survey of developmentally relevant genes in *Ciona intestinalis*: I. Genes for bHLH transcription factors. *Dev. Genes Evol.* **213**, 213–221.
- Sauka-Spengler, T., and Bronner-Fraser, M. (2008). Insights from a sea lamprey into the evolution of neural crest gene regulatory network. *Biol. Bull.* **214**, 303–314.
- Sauka-Spengler, T., Meulemans, D., Jones, M., and Bronner-Fraser, M. (2007). Ancient Evolutionary Origin of the Neural Crest Gene Regulatory Network. *Dev. Cell* **13**, 405–420.
- Schlosser, G. (2002). Development and evolution of lateral line placodes in amphibians I. Development. *Zoology* **105**, 119–146.
- Schlosser, G. (2003). Hypobranchial placodes in *Xenopus laevis* give rise to hypobranchial ganglia, a novel type of cranial ganglia. *Cell Tissue Res.* **312**, 21–29.
- Schlosser, G. (2005). Evolutionary origins of vertebrate placodes: insights from developmental studies and from comparisons with other deuterostomes. *J. Exp. Zool. B Mol. Dev. Evol.* **304B**, 347–399.
- Schlosser, G. (2006). Induction and specification of cranial placodes. *Dev. Biol.* **294**, 303–351.
- Schlosser, G., and Ahrens, K. (2004). Molecular anatomy of placode development in *Xenopus laevis*. *Dev. Biol.* **271**, 439–466.
- Schlosser, G., and Northcutt, R.G. (2000). Development of neurogenic placodes in *Xenopus laevis*. *J. Comp. Neurol.* **418**, 121–146.
- Schlosser, G., Kintner, C., and Northcutt, R.G. (1999). Loss of ectodermal competence for lateral line placode formation in the direct developing frog *Eleutherodactylus coqui*. *Dev. Biol.* **213**, 354–369.
- Schmid, T., Krüger, M., and Braun, T. (2007). NSCL-1 and -2 control the formation of precerebellar nuclei by orchestrating the migration of neuronal precursor cells. *J. Neurochem.* **102**, 2061–2072.
- Semon, M., and Wolfe, K.H. (2006). Rearrangement Rate following the Whole-Genome Duplication in Teleosts. *Mol. Biol. Evol.* **24**, 860–867.
- Sendrowski, K., and Sobaniec, W. (2013). Hippocampus, hippocampal sclerosis and epilepsy. *Pharmacol. Rep.* **65**, 555–565.
- Shah, S., Lee, S.-F., Tabuchi, K., Hao, Y.-H., Yu, C., LaPlant, Q., Ball, H., Dann, C.E., Südhof, T., and Yu, G. (2005). Nicastrin functions as a γ -secretase-substrate receptor. *Cell* **122**, 435–447.
- Shen, M., Kawamoto, T., Yan, W., Nakamasu, K., Tamagami, M., Koyano, Y., Noshiro, M., and Kato, Y. (1997). Molecular characterization of the novel basic helix-loop-helix protein DEC1 expressed in differentiated human embryo chondrocytes. *Biochem. Biophys. Res. Commun.* **236**, 294–298.
- Shintani, S., Terzic, J., Sato, A., Saraga-Babic, M., O’hUigin, C., Tichy, H., and Klein, J. (2000). Do lampreys have lymphocytes? The Spi evidence. *Proc. Natl. Acad. Sci.* **97**, 7417–7422.

- Shu, D.G., Luo, H.L., Morris, S.C., Zhang, X.L., Hu, S.X., Chen, L., Han, J., Zhu, M., Li, Y., and Chen, L.Z. (1999). Lower Cambrian vertebrates from south China. *Nature* **402**, 42–46.
- Shu, D.-G., Conway Morris, S., Han, J., Zhang, Z.-F., Yasui, K., Janvier, P., Chen, L., Zhang, X.-L., Liu, J.-N., Li, Y., et al. (2003). Head and backbone of the Early Cambrian vertebrate *Haikouichthys*. *Nature* **421**, 526–529.
- Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., Degnan, B.M., and Vervoort, M. (2007). Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* **7**, 33.
- Simionato, E., Kerner, P., Dray, N., Le Gouar, M., Ledent, V., Arendt, D., and Vervoort, M. (2008). *atonal*- and *achaete-scute*-related genes in the annelid *Platynereis dumerilii*: insights into the evolution of neural basic-Helix-Loop-Helix genes. *BMC Evol. Biol.* **8**, 170.
- Siponen, M.I., Wisniewska, M., Lehtio, L., Johansson, I., Svensson, L., Raszewski, G., Nilsson, L., Sigvardsson, M., and Berglund, H. (2010). Structural determination of functional domains in early B-cell factor (EBF) family of transcription factors reveals similarities to Rel DNA-binding proteins and a novel dimerization motif. *J. Biol. Chem.* **285**, 25875–25879.
- Sippel, R.S., Carpenter, J.E., Kunnimalaiyaan, M., and Chen, H. (2003). The role of human *achaete-scute* homolog-1 in medullary thyroid cancer cells. *Surgery* **134**, 866–871.
- Skromne, I., Thorsen, D., Hale, M., Prince, V.E., and Ho, R.K. (2007). Repression of the hindbrain developmental program by Cdx factors is required for the specification of the vertebrate spinal cord. *Development* **134**, 2147–2158.
- Smith, J.B., and Alloway, K.D. (2010). Functional Specificity of Claustrum Connections in the Rat: Interhemispheric Communication between Specific Parts of Motor Cortex. *J. Neurosci.* **30**, 16832–16844.
- Smith, E.M.K., Åkerblad, P., Kadesch, T., Axelson, H., and Sigvardsson, M. (2005). Inhibition of EBF function by active Notch signaling reveals a novel regulatory pathway in early B-cell development. *Blood* **106**, 1995–2001.
- Smith, J.B., Radhakrishnan, H., and Alloway, K.D. (2012). Rat Claustrum Coordinates But Does Not Integrate Somatosensory and Motor Cortical Information. *J. Neurosci.* **32**, 8583–8588.
- 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., et al. (2013). Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat. Genet.* **45**, 415–421.
- So, J.-H., Chun, H.-S., Bae, Y.-K., Kim, H.-S., Park, Y.-M., Huh, T.-L., Chitnis, A.B., Kim, C.-H., and Yeo, S.-Y. (2009). *Her4* is necessary for establishing peripheral projections of the trigeminal ganglia in zebrafish. *Biochem. Biophys. Res. Commun.* **379**, 22–26.
- Sölter, M., Locker, M., Boy, S., Taelman, V., Bellefroid, E.J., Perron, M., and Pieler, T. (2006). Characterization and function of the bHLH-O protein XHes2: insight into the mechanisms controlling retinal cell fate decision. *Development* **133**, 4097–4108.
- Sommer, L., Shah, N., Rao, M., and Anderson, D.J. (1995). The cellular function of MASH1 in autonomic neurogenesis. *Neuron* **15**, 1245–1258.

- Sommer, L., Ma, Q., and Anderson, D.J. (1996). neurogenins, a Novel Family of atonal-Related bHLH Transcription Factors, Are Putative Mammalian Neuronal Determination Genes That Reveal Progenitor Cell Heterogeneity in the Developing CNS and PNS. *Mol. Cell. Neurosci.* 8, 221–241.
- Squire, L., Berg, D., Bloom, F., du Lac, S., Ghosh, A., and Spitzer, N. (2008). *Fundamental Neuroscience* (London: Academic Press).
- Stock, D.W., and Whitt, G.S. (1992). Evidence from 18S ribosomal RNA sequences that lampreys and hagfishes form a natural group. *Science* 257, 787–789.
- Stolfi, A., Gainous, T.B., Young, J.J., Mori, A., Levine, M., and Christiaen, L. (2010). Early Chordate Origins of the Vertebrate Second Heart Field. *Science* 329, 565–568.
- Streit, A. (2007). The preplacodal region: an ectodermal domain with multipotential progenitors that contribute to sense organs and cranial sensory ganglia. *Int. J. Dev. Biol.* 51, 447–461.
- De Strooper, B., and Annaert, W. (2001). Presenilins and the intramembrane proteolysis of proteins: facts and fiction. *Nat. Cell Biol.* 3, E221–E225.
- Sugahara, F., Aota, S. -i., Kuraku, S., Murakami, Y., Takio-Ogawa, Y., Hirano, S., and 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, 1217–1226.
- Sun, H., and Taneja, R. (2000). Stra13 expression is associated with growth arrest and represses transcription through histone deacetylase (HDAC)-dependent and HDAC-independent mechanisms. *Proc. Natl. Acad. Sci.* 97, 4058–4063.
- Sun, Y., Nadal-Vicens, M., Misono, S., Lin, M.Z., Zubiaga, A., Hua, X., Fan, G., and Greenberg, M.E. (2001). Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms. *Cell* 104, 365–376.
- Suzuki, Y., Mizoguchi, I., Nishiyama, H., Takeda, M., and Obara, N. (2003). Expression of Hes6 and NeuroD in the olfactory epithelium, vomeronasal organ and non-sensory patches. *Chem. Senses* 28, 197–205.
- Swanson, L.W. (2000). What is the brain? *Trends Neurosci.* 23, 519–527.
- Taelman, V., Van Wayenbergh, R., Solter, M., Pichon, B., Pieler, T., Christophe, D., and Bellefroid, E.J. (2004). Sequences downstream of the bHLH domain of the *Xenopus* hairy-related transcription factor-1 act as an extended dimerization domain that contributes to the selection of the partners. *Dev. Biol.* 276, 47–63.
- Tahara, Y. (1988). Normal stages of development in the lamprey, *Lampetra reissneri* (Dybowski). *Zoolog. Sci.* 5, 109–118.
- Takebayashi, H., Yoshida, S., Sugimori, M., Kosako, H., Kominami, R., Nakafuku, M., and Nabeshima, Y. (2000). Dynamic expression of basic helix-loop-helix Olig family members: implication of Olig2 in neuron and oligodendrocyte differentiation and identification of a new member, Olig3. *Mech. Dev.* 99, 143–148.

- Takebayashi, K., Sasai, Y., Sakai, Y., Watanabe, T., Nakanishi, S., and Kageyama, R. (1994). Structure, chromosomal locus, and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-1. Negative autoregulation through the multiple N box elements. *J. Biol. Chem.* **269**, 5150–5156.
- Takebayashi, K., Akazawa, C., Nakanishi, S., and Kageyama, R. (1995). Structure and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-5. Identification of the neural precursor cell-specific promoter element. *J. Biol. Chem.* **270**, 1342–1349.
- Takezaki, N. (2003). Molecular Phylogeny of Early Vertebrates: Monophyly of the Agnathans as Revealed by Sequences of 35 Genes. *Mol. Biol. Evol.* **20**, 287–292.
- Takio, Y., Pasqualetti, M., Kuraku, S., Hirano, S., Rijli, F.M., and Kuratani, S. (2004). Evolutionary biology: lamprey Hox genes and the evolution of jaws. *Nature* **429**.
- Takio, Y., Kuraku, S., Murakami, Y., Pasqualetti, M., Rijli, F.M., Narita, Y., Kuratani, S., and Kusakabe, R. (2007). Hox gene expression patterns in *Lethenteron japonicum* embryos—Insights into the evolution of the vertebrate Hox code. *Dev. Biol.* **308**, 606–620.
- Takke, C., Dornseifer, P., v Weizsacker, E., and Campos-Ortega, J.A. (1999). *her4*, a zebrafish homologue of the *Drosophila* neurogenic gene *E (spl)*, is a target of NOTCH signalling. *Development* **126**, 1811–1821.
- Tamme, R., Wells, S., Conran, J.G., and Lardelli, M. (2002). The identity and distribution of neural cells expressing the mesodermal determinant *spadetail*. *BMC Dev. Biol.* **2**, 9.
- Tanaka, M., Gertsenstein, M., Rossant, J., and Nagy, A. (1997). *Mash2* Acts Cell Autonomously in Mouse Spongiotrophoblast Development. *Dev. Biol.* **190**, 55–65.
- Tanaka, Y., Saito, Y., and Gotoh, H. (1981). Vascular architecture and intestinal hematopoietic nests of two cyclostomes, *Eptatretus burgeri* and *ammocoetes* of *Entosphenus reissneri*: A comparative morphological study. *J. Morphol.* **170**, 71–93.
- Theodorakis, K., Kyriakopoulou, K., Wassef, M., and Karagogeos, D. (2002). Novel sites of expression of the bHLH gene *NSCL1* in the developing nervous system. *Mech. Dev.* **119**, S103–S106.
- Tolia, A., and De Strooper, B. (2009). Structure and function of γ -secretase. *Semin. Cell Dev. Biol.* **20**, 211–218.
- Tomita, K., Ishibashi, M., Nakahara, K., Ang, S.-L., Nakanishi, S., Guillemot, F., and Kageyama, R. (1996). Mammalian hairy and Enhancer of Split Homolog 1 Regulates Differentiation of Retinal Neurons and Is Essential for Eye Morphogenesis. *Neuron* **16**, 723–734.
- Travis, A., Hagman, J., Hwang, L., and Grosschedl, R. (1993). Purification of early-B-cell factor and characterization of its DNA-binding specificity. *Mol. Cell. Biol.* **13**, 3392–3400.
- Treiber, N., Treiber, T., Zocher, G., and Grosschedl, R. (2010a). Structure of an Ebf1:DNA complex reveals unusual DNA recognition and structural homology with Rel proteins. *Genes Dev.* **24**, 2270–2275.

- Treiber, T., Mandel, E.M., Pott, S., Györy, I., Firner, S., Liu, E.T., and Grosschedl, R. (2010b). Early B Cell Factor 1 Regulates B Cell Gene Networks by Activation, Repression, and Transcription- Independent Poising of Chromatin. *Immunity* 32, 714–725.
- Tulenko, F.J., McCauley, D.W., MacKenzie, E.L., Mazan, S., Kuratani, S., Sugahara, F., Kusakabe, R., and Burke, A.C. (2013). Body wall development in lamprey and a new perspective on the origin of vertebrate paired fins. *Proc. Natl. Acad. Sci.* 110, 11899–11904.
- Turbeville, J.M., Schulz, J.R., and Raff, R.A. (1994). Deuterostome phylogeny and the sister group of the chordates: evidence from molecules and morphology. *Mol. Biol. Evol.* 11, 648–655.
- Turner, D.L., and Weintraub, H. (1994). Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev.* 8, 1434–1447.
- Uchida, K., Murakami, Y., Kuraku, S., Hirano, S., and Kuratani, S. (2003). Development of the adenohypophysis in the lamprey: evolution of epigenetic patterning programs in organogenesis. *J. Exp. Zool. B Mol. Dev. Evol.* 300, 32–47.
- Uinuk-ool, T., Mayer, W.E., Sato, A., Dongak, R., Cooper, M.D., and Klein, J. (2002). Lamprey lymphocyte-like cells express homologs of genes involved in immunologically relevant activities of mammalian lymphocytes. *Proc. Natl. Acad. Sci.* 99, 14356–14361.
- Vasiliauskas, D., and Stern, C.D. (2000). Expression of mouse HES-6, a new member of the Hairy/Enhancer of split family of bHLH transcription factors. *Mech. Dev.* 98, 133–137.
- Vilas-Boas, F., and Henrique, D. (2010). HES6-1 and HES6-2 Function through Different Mechanisms during Neuronal Differentiation. *PLoS ONE* 5, e15459.
- Villar-Cheda, B., Abalo, X.M., Anadón, R., and Rodicio, M.C. (2005). The tegmental proliferation region in the sea lamprey. *Brain Res. Bull.* 66, 431–435.
- Villar-Cheda, B., Pérez-Costas, E., Meléndez-Ferro, M., Abalo, X.M., Rodríguez-Muñoz, R., Anadón, R., and Rodicio, M.C. (2006). Cell proliferation in the forebrain and midbrain of the sea lamprey. *J. Comp. Neurol.* 494, 986–1006.
- Viloria, A., Rodríguez-Alonso, M., Costas, V., Pérez-Fernández, J., Pombal, M.A., and Megías, M. (2008). Developmental changes of calretinin immunoreactivity in the lamprey spinal cord. *Brain Res. Bull.* 75, 428–432.
- Wada, H., and Satoh, N. (2001). Patterning the protochordate neural tube. *Curr. Opin. Neurobiol.* 11, 16–21.
- Waehneltdt, T.V., Kiene, M.-L., Malotka, J., Kiecke, C., and Neuhoff, V. (1984). Nervous system myelin in the electric ray, *Torpedo marmorata*: Morphological characterization of the membrane and biochemical analysis of its protein components. *Neurochem. Int.* 6, 223–235.
- Wang, S.S. (2004). Genetic disruptions of O/E2 and O/E3 genes reveal involvement in olfactory receptor neuron projection. *Development* 131, 1377–1388.
- Wang, M.M., and Reed, R.R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature* 364, 121–126.

- Wang, S.S., Tsai, R.Y., and Reed, R.R. (1997). The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development. *J. Neurosci.* *17*, 4149–4158.
- Wang, S.S., Betz, A.G., and Reed, R.R. (2002). Cloning of a Novel Olf-1/EBF-like Gene, O/E-4, by Degenerate Oligo-based Direct Selection. *Mol. Cell. Neurosci.* *20*, 404–414.
- Wang, X., Chu, L.T., He, J., Emelyanov, A., Korzh, V., and Gong, Z. (2001). A novel zebrafish bHLH gene, neurogenin3, is expressed in the hypothalamus. *Gene* *275*, 47–55.
- Wang, Y., Chen, K., Yao, Q., Zheng, X., and Yang, Z. (2009). Phylogenetic Analysis of Zebrafish Basic Helix-Loop-Helix Transcription Factors. *J. Mol. Evol.* *68*, 629–640.
- Van Wayenbergh, R., Taelman, V., Pichon, B., Fischer, A., Kricha, S., Gessler, M., Christophe, D., and Bellefroid, E.J. (2003). Identification of BOIP, a novel cDNA highly expressed during spermatogenesis that encodes a protein interacting with the orange domain of the hairy-related transcription factor HRT1/Hey1 in *Xenopus* and mouse. *Dev. Dyn.* *228*, 716–725.
- Webb, K., Coolen, M., Gloeckner, C., Stigloher, C., Bahn, B., Topp, S., Ueffing, M., and Bally-Cuif, L. (2011). The Enhancer of split transcription factor Her8a is a novel dimerisation partner for Her3 that controls anterior hindbrain neurogenesis in zebrafish. *BMC Dev. Biol.* *11*, 27.
- Wells, S., Conran, J.G., Tamme, R., Gaudin, A., Webb, J., and Lardelli, M. (2010). Cryptic organisation within an apparently irregular rostrocaudal distribution of interneurons in the embryonic zebrafish spinal cord. *Exp. Cell Res.* *316*, 3292–3303.
- Whitlock, K.E. (2004). A New Model for Olfactory Placode Development. *Brain. Behav. Evol.* *64*, 126–140.
- Wicht, H., and Lacalli, T.C. (2005). The nervous system of amphioxus: structure, development, and evolutionary significance. *Can. J. Zool.* *83*, 122–150.
- Winchell, C.J., Sullivan, J., Cameron, C.B., Swalla, B.J., and Mallatt, J. (2002). Evaluating hypotheses of deuterostome phylogeny and chordate evolution with new LSU and SSU ribosomal DNA data. *Mol. Biol. Evol.* *19*, 762–776.
- Wolfe, M.S. (2003). Gamma-secretase--intramembrane protease with a complex. *Sci. Aging Knowl. Environ.* *2003*, PE7.
- Wolfe, M.S., Xia, W., Ostaszewski, B.L., Diehl, T.S., Kimberly, W.T., and Selkoe, D.J. (1999). Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* *398*, 513–517.
- Wullimann, M.F., and Mueller, T. (2002). Expression of Zash-1a in the postembryonic zebrafish brain allows comparison to mouse Mash1 domains. *Gene Expr. Patterns* *1*, 187–192.
- Wullimann, M.F., Rink, E., Vernier, P., and Schlosser, G. (2005). Secondary neurogenesis in the brain of the African clawed frog, *Xenopus laevis*, as revealed by PCNA, Delta-1, Neurogenin-related-1, and NeuroD expression. *J. Comp. Neurol.* *489*, 387–402.
- Xu, H., Dude, C.M., and Baker, C.V.H. (2008). Fine-grained fate maps for the ophthalmic and maxillomandibular trigeminal placodes in the chick embryo. *Dev. Biol.* *317*, 174–186.

- Yamada, S., Hotta, K., Yamamoto, T.S., Ueno, N., Satoh, N., and Takahashi, H. (2009). Interaction of notochord-derived fibrinogen-like protein with Notch regulates the patterning of the central nervous system of *Ciona intestinalis* embryos. *Dev. Biol.* 328, 1–12.
- Yan, R.-T., He, L., and Wang, S.-Z. (2009). Pro-photoreceptor activity of chick neurogenin1. *Invest. Ophthalmol. Vis. Sci.* 50, 5567–5576.
- Yang, T., Arslanova, D., Gu, Y., Augelli-Szafran, C., and Xia, W. (2008). Quantification of gamma-secretase modulation differentiates inhibitor compound selectivity between two substrates Notch and amyloid precursor protein. *Mol. Brain* 1, 15.
- Yang, T., Arslanova, D., Xu, X., Li, Y.-M., and Xia, W. (2010). In vivo manifestation of Notch related phenotypes in zebrafish treated with Alzheimer's amyloid reducing γ -secretase inhibitors. *J. Neurochem.* 113, 1200–1209.
- Yang, X., Tomita, T., Wines-Samuelson, M., Beglopoulos, V., Tansey, M.G., Kopan, R., and Shen, J. (2006). Notch1 Signaling Influences V2 Interneuron and Motor Neuron Development in the Spinal Cord. *Dev. Neurosci.* 28, 102–117.
- Yao, J., Wang, L., Chen, L., Zhang, S., Zhao, Q., Jia, W., and Xue, J. (2006). Cloning and developmental expression of the DEC1 ortholog gene in zebrafish. *Gene Expr. Patterns* 6, 919–927.
- Yoshida, M., and Colman, D.R. (1996). Parallel evolution and coexpression of the proteolipid proteins and protein zero in vertebrate myelin. *Neuron* 16, 1115–1126.
- Youson, J.H. (1997). Is lamprey metamorphosis regulated by thyroid hormones? *Am. Zool.* 37, 441–460.
- Yu, C., Ehrhardt, G.R.A., Alder, M.N., Cooper, M.D., and Xu, A. (2009). Inhibitory signaling potential of a TCR-like molecule in lamprey. *Eur. J. Immunol.* 39, 571–579.
- Zaidi, A.U., Kafitz, K.W., Greer, C.A., and Zielinski, B.S. (1998). The expression of tenascin-C along the lamprey olfactory pathway during embryonic development and following axotomy-induced replacement of the olfactory receptor neurons. *Dev. Brain Res.* 109, 157–168.
- Zalc, B., Goujet, D., and Colman, D. (2008). The origin of the myelination program in vertebrates. *Curr. Biol.* 18, R511–R512.
- Zapata, A., Ardavin, C.F., Gomariz, R.P., and Leceta, J. (1981). Plasma cells in the ammocoete of *Petromyzon marinus*. *Cell Tissue Res.* 221, 203–208.
- Zhang, G., and Cohn, M.J. (2008). Genome duplication and the origin of the vertebrate skeleton. *Curr. Opin. Genet. Dev.* 18, 387–393.
- Zhang, Z., Cotta, C.V., Stephan, R.P., and Klug, C.A. (2003). Enforced expression of EBF in hematopoietic stem cells restricts lymphopoiesis to the B cell lineage. *EMBO J.* 22, 4759–4769.
- Zhao, G., Liu, Z., Ilagan, M.X.G., and Kopan, R. (2010). Gamma-Secretase Composed of PS1/Pen2/Aph1a Can Cleave Notch and Amyloid Precursor Protein in the Absence of Nicastrin. *J. Neurosci.* 30, 1648–1656.

- Zhou, Q., Wang, S., and Anderson, D.J. (2000). Identification of a novel family of oligodendrocyte lineage-specific basic helix–loop–helix transcription factors. *Neuron* 25, 331–343.
- Zhu, M., Yu, X., and Janvier, P. (1999). A primitive fossil fish sheds light on the origin of bony fishes. *Nature* 397, 607–610.
- Zimmerman, K., Shih, J., Bars, J., Collazo, A., and Anderson, D.J. (1993). XASH-3, a novel *Xenopus* achaete-scute homolog, provides an early marker of planar neural induction and position along the mediolateral axis of the neural plate. *Development* 119, 221–232.
- Zordan, P., Croci, L., Hawkes, R., and Consalez, G.G. (2008). Comparative analysis of proneural gene expression in the embryonic cerebellum. *Dev. Dyn.* 237, 1726–1735.

APPENDIX A. Ricardo's *in situ* Protocol for Lampreys

Some notes before starting the *in situ*:

- I do the entire *in situ* in Axygen 2 ml "flat" tubes because it is easier to change the solutions without disturbing the embryos much.
- I use 1 ml of any solution in each tube.
- When changing solutions, either pour the solution in a container –of course being careful not to drop the embryos- or remove it with a pipette (a p1000 is good to remove a large part of the solution first, and then use a p200 to remove the rest). It is better to use a pipette with PBS-based solutions because PBS has superficial tension, and there is a greater risk to drop the embryos if you just pour it. By contrast, you can easily pour solutions that contain a detergent or no PBS at all, as the embryos will remain at the bottom of the tube without much superficial tension. Regardless of the way you pour the solution, at the end I remove any traces of the solution with the pipette almost until the last drop. It is OK if the embryos remain without the solution for a few seconds and are in direct contact with the air while changing solutions; they will not break, but do not leave them like that for too long. They should not be completely dried though. A very important step here is that, if the tubes were incubated rocking, there will always be traces of solution in the cap. To get very clean embryos at the end, I ALWAYS take any remaining liquid in the cap with a pipette (p200). This ensures that the next solution gets less contaminated by the previous solution.
- Before using any solution, I shake it very well first so that you take the right concentration of it.
- I do not do too many tubes in any one experiment, first because it is too tiring considering the large number of washes, and second because I care more about changing solutions carefully instead of doing many tubes at a time quickly. That depends on the speed of the person of course. For me doing 8-10 tubes per *in situ* is optimal. You can also mix stages (I have mixed 3 stages, 5 embryos each in 1 ml) that you can recognise morphologically so that you do not have to do too many experiments if working with a few tubes per *in situ*.
- Wear new gloves for the entire *in situ*. Change them whenever they get very dirty (which happens mostly with the final washes).

In situ

All washes are done rocking at room temperature (RT) with the tubes in a horizontal position unless specified otherwise.

Fixation of embryos

After collecting the embryos either remove the chorion off with fine forceps or just fix the already-hatched embryos depending on the stage. I use 15 ml Falcon tubes.

- Fix in 4% paraformaldehyde (PFA) in PBS-DEPC pH ~7.5 at 4 °C overnight or more (I have left them even a month in the fridge and they still work). Static.
- Wash twice with PBS-DEPC for 30 minutes
- Wash with PBS-DEPC : Methanol 50:50 for 30 minutes
- Wash with PBS-DEPC : Methanol 25:75 for 30 minutes
- Wash 3 times with 100% methanol
- Store at -20 °C in 100% methanol (fresh from the previous 3 washes, i.e. a fourth methanol)

Permeabilisation and hybridisation

- Treat the embryos with a mixture of Hydrogen Peroxide : Methanol 1:1.5 overnight at 4 °C (in the fridge).

* **Note:** *take the stored embryos (in 100% Methanol) from the Falcon tube with a plastic sterile pipette (if the embryos are too big simply cut the tip with a clean razor blade in a Petri dish) and place them in the Axygen tube already containing H₂O₂ and Methanol mixed. In my experience, sometimes is hard to let the embryos drop into the tube from the pipette - they can get stuck in the walls of the pipette and is very common to let many drops of methanol fall into the tube. As many drops will eventually dilute the H₂O₂ too much, I do a few things to avoid this: first, if you only use 5 embryos per tube it is OK letting 1-2 drops of methanol fall per embryo, but in general always try to let as few drops as possible fall into the tube; second, I always try to let 2 or 3 embryos fall together in a single drop of methanol. If you can let all 5 embryos fall in a single drop (which is possible), perfect!; third, if mixing stages and using up to 15 embryos per tube (per 1 ml), too many drops will fall and the tube will inevitably contain too much methanol for the H₂O₂-Methanol treatment, so in this case place all embryos in the tube and later remove the methanol with a pipette and immediately add the H₂O₂-Methanol solution (which can be prepared in separate).*

- Wash 3 times with PBS-DEPC 5 minutes each.

- Treat with ~0.2N HCl in PBS-**DEPC** for 10 minutes.
- * **Note:** *is OK if you spend more time while changing solutions, I just care about treating the embryos for 10 minutes rocking at RT.*
- Incubate the embryos in 5µg/ml of Proteinase K in PBS-**DEPC** at 37° C. Static.
- * **Note:** *is very important not to rock tubes at this stage because you can severely damage the embryos. The length of time depends on the stage and activity of the enzyme. For stages 22-23 I do 12 minutes; for stages 24 or older I do 25 minutes. However, with time the enzyme starts to loose activity and you have to increase the time of treatment. If you are working with early stages increase by ~2 min; if it is stage 24 or older increase by ~5 min. However, if using a new enzyme, start calibrating with 18-20 minutes for stages 24 or older, and increase or decrease time by 1 minute if necessary as a new enzyme is more active. Ideally the enzyme should be stored in 10 µl aliquots so that each tube is used only once and then discarded. From that you can calculate the time for earlier stages.*
- * **Note:** *One way to monitor the effectiveness of this treatment is to look at the notochord when dealing with genes that are not expressed in this tissue. The probe is frequently trapped in the notochord and you can eliminate that background by increasing the time of proteinase K treatment. If you do not see any staining in the notochord that indicates that you have avoided the background (again, only for genes that are not expressed in the notochord). Also, because the enzyme works at 37°C, after I add proteinase K in a tube I place it straight away on ice to hold the activity of the enzyme until you have finished changing solutions in all tubes. Subsequently, after incubation, place all tubes back on ice right away to stop the enzyme from working while you change the next solution. Is better to change solutions of young embryos first because they are more fragile than older ones. At the end you have to calibrate the amount of time of proteinase K treatment with the amount of time it takes you to get the tubes, place them on ice, and change solutions.*
- Refix for 20 minutes in 4% PFA, 0.2% glutaraldehyde in PBS-**DEPC** static on ice. Is OK if you leave them longer in PFA while changing solutions.
- Wash 3 times in PBS-**DEPC** 10 minutes each.
- Wash 2 times in hybridisation buffer 10 minutes each.
- Prehybridise in for >2 hrs. at 65 °C rocking.
- * **Note:** *you can wait the 2 hrs. but I have found that is better to leave them overnight to get nicer in situs, plus, is less tiring. I have even left them 2 days prehybridising and is still OK.*

- Incubate in hybridisation buffer with 600-700 ng of probe per 1 ml for ~2-2.5 days at 65 °C rocking.
- * **Note:** *here is a balance between probe concentration and time of incubation; the more concentrated the probe, the less time you require of incubation, although this also depends on the gene and probe length. What I do is that after synthesising my probes, I spec them and dilute them so that at the end I have 600-700 ng of probe per 2 µl. This is convenient for me because I keep the probes at -70 °C in 10 µl aliquots, so that I can finish one tube in a few experiments without freezing and thawing the probes too many times, which eventually can damage the probes and affect the in situ.*

Washing and coloring

- Wash twice in wash A for 30 minutes at 65 °C rocking.
- Substitute the solution gradually with TST 50% and 75% in wash A, and 100% (twice) for 10 minutes each at 65 °C rocking.
- * **Note:** *you can carry on with the next washes or stop here and put the tubes in the fridge in TST for next day. I have found that is better to leave them in TST overnight until next day (in situs get cleaner), plus is less tiring, otherwise you will spend day and night finishing all the washes because once you continue with the next RNAase treatment is advised not to stop again and finish all the washes. After the second wash I store the embryos in the fridge with a new wash and in the next day I apply the RNAase.*
- Add RNAase A at 0.05 mg/ml for 30 minutes.
- * **Note:** *As with proteinase K, while changing the RNAase solution I place the tubes on ice until I have finished all tubes, and when the treatment has finished I place them back on ice as well while changing the next wash.*
- Wash twice in wash B for 30 minutes each at 65 °C rocking.
- Wash twice in wash C for 30 minutes each at 65 °C rocking.
- Wash twice in wash D for 30 minutes each at 65 °C rocking.
- Block with TST containing heat-treated sheep serum 1:4 for >2 hrs. *However, I have found that is better to leave them overnight in the fridge.*
- Incubate with antibody solution at 4 °C overnight. *(Do not forget to pre-absorb first!).*
- Wash 5 times with TST 60 minutes each (first wash can be 30 minutes).

- Incubate with 6 μ l/ml of NBT/BCIP in NTMT overnight in the dark (covered with foil). You can start with this concentration and then increase it according to the speed of staining.

- * **Note:** *The time of the staining depends on the gene and probe length. For probes longer than 1 kb it takes about two days; for probes ~600-700 it takes about three days. However, I recommend first to do a staining until you get background so that you identify all the expression sites of your gene, and then in another in situ monitor carefully and stop the staining until just before you expect to get the background. A good way that I have found to monitor the overall background is the colour of the vitellum and the embryo itself. A good-stained embryo is the one that still has the vitellum pink, so that the blue-purple colour of the digoxigenin contrasts really nicely with it. If after staining the vitellum is not pink and gets darker, brownish, that means that the embryo is overstained and is better to do another in situ taking care of the pink colour. Monitoring this colour carefully will give you a nicely stained embryo. Also, change the staining solution everyday. On the next day of staining, the staining solution will be all purple, so if you want to keep staining for another day change the solution, you will notice the contrast with the fresh solution's yellow colour. Although you can vary the concentration of the NBT/BCIP, I have found that 6 μ l/ml is perfect, as it allows you to make only one change a day and gives you a perfect time frame to avoid overstaining, given that you can "measure" the progress of the staining in days and not in hours, which becomes quite handy after a while. I also use 4-well Nunc dishes with 500 μ l of the staining solution in each well and only a few embryos per well. If you have lots I advice to split them into different wells so that you have 5-8 embryos per well, although perhaps more embryos per well should not be a problem either. To do this, take the embryos with a plastic sterile pipette as in the beginning. As opposed to methanol, with TST is easy to take all five embryos in a single drop. If you still cannot do this, after you drop all embryos in the dish, you can remove excess TST with a normal pipette (it does not matter if the embryos stay in direct contact with the air for a few seconds) and immediately add NBT/BCIP.*

- After specific staining is developed, wash the embryos with PBT 3 times 10 minutes each.

- Fix with 4% PFA >30 minutes static on ice, although I always fix them overnight at 4 °C.

- Wash with PBT 3 times 10 minutes each.

- Clear embryos by gradually adding 25%, 50%, 75% and 99% glycerol in PBS. Static.

- * **Note:** *When placing the embryos in glycerol, wait until the embryo sinks before changing to the next solution to avoid tissue damage by drastic changes of different concentrations of glycerol. With 25% and 50% glycerol, the embryos usually sink very quickly, but I still wait at least 15 minutes before changing the next solution.*

- Store embryos in 99% glycerol covered with foil at 4 °C indefinitely.

Taking pictures

- Mount embryos in 99% glycerol first. I place the embryos in a glass microscope slide specially made with pieces of broken slides glued in specific positions, so that they form an empty square where you can put the embryos immersed in glycerol.
- All pictures were taken with an **Axioskop 2 plus** (Carl Zeiss) and **AxioCam HRc**. The best settings that I have found that give you nice pink and blue-purple colours are:
 - Reflector turret set to **2**.
 - Differential Interference Contrast (DIC) set to **II**.
- Images were taken with a 4X or 10X objective.
- Images were processed using the AxioVision 4 software.

SOLUTIONS

Treat solutions with **diethylpyrocarbonate (DEPC)** which destroys RNases added to 0.1% and left overnight. Autoclave next day to break down DEPC to H₂O and CO₂. TRIS **cannot** be treated with DEPC, and you **cannot** treat solutions that you **cannot** autoclave. For TRIS, use DEPC-water and weigh with an RNase-free spatula (which can be done by putting it in chloroform and then let it dry). Solutions that need to be treated with DEPC are marked.

- **Hydrogen peroxide 1:1.5 Methanol:** 1 ml of 30% H₂O₂ + 450 µl MeOH
- **~0.2 N HCl:** 36 µl concentrated HCl + 49.964 ml of PBS-DEPC = 50 ml

Hybridisation buffer

Usual stock solution	Final concentration	Volume for 50 ml
100% Formamide	50% Formamide	25 ml
20X SSC-DEPC pH 7.0	5X SSC pH7.0	12.5 ml
10% Tween 20	0.1% Tween 20	500 µl
10% CHAPS	0.1% CHAPS	500 µl
30 mg/ml yeast RNA	1mg/ml yeast RNA	~1.67 ml
10mg/ml Heparin	100 µg/ml Heparin	500 µl
DEPC-H ₂ O		9.334 ml

Washing solutions: all in DEPC-Water

- **Wash A:** 50% Formamide, 5X SSC pH 4.5, 1% SDS
 - **Wash B:** 50% Formamide, 2X SSC pH 4.5
 - **Wash C:** 2X SSC pH 4.5, 0.3% CHAPS
 - **Wash D:** 0.2X SSC pH 4.5, 0.3% CHAPS
- **TST:** 10 mM Tris-HCl pH 7.5, 0.5 M NaCl, 0.1% Tween 20.
 - **Heat-treated Sheep serum:** incubate sheep serum for 30 min. at 65 °C. Is important not to incubate it higher than 65 °C otherwise it will solidify.
 - **Block solution:** TST 1:4 in heat-treated sheep serum.
 - **Antibody solution:** Dilute 1:4000 of Ab in Block solution, preabsorb on ice >2 hours before using.
 - **PBT:** PBS with 0.1% Tween 20.
 - **NTMT:** 100 mM NaCl, 100 mM Tris-HCl pH 9.5, 50 mM MgCl₂, 1% Tween 20.