NEURAL CREST CELL DEVELOPMENT AND EVOLUTION

by

TYLER ALEXANDER SQUARE
B.A., University of Colorado, 2011

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
Department of Ecology and Evolutionary Biology
2017

This thesis entitled:
Vertebrate neural crest cell development and evolution
written by Tyler Alexander Square
has been approved for the Department of Ecology and Evolutionary Biology

_____

Daniel M. Medeiros

_____ _____

_____ _____

David Stock, Kristin Artinger, Andrew Martin, and Michael Klymkowsky

Date April 10th, 2017

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IACUC protocol # 2392

ABSTRACT

Square, Tyler Alexander (Ph.D. Ecology and Evolutionary Biology)
Neural Crest Cell Development and Evolution
Thesis directed by Associate Professor Daniel Meulemans Medeiros

Neural crest cells play a major role in vertebrate development and evolution. Both the appearance and diversification of these cells have been evoked as a major driver of vertebrate success. This cell type is unique because it is capable of contributing to extremely diverse tissues: structural tissues like cartilage and bone, signal transducing cells of the peripheral nervous system, and multiple types of pigment cells found throughout the body -- in every vertebrate hair, feather, and scale. Elaborations of the biological systems that rely on these cell types were critical in the specialization of vertebrates, as evidenced in fossil and extant animals and biological systems. The present work aims to identify the cellular mechanisms responsible for both the appearance of multiple neural crest cell types, as well as lineage-specific modifications that have been made within cell types in order to achieve the vast array of vertebrate diversity.

ACKNOWLEDGEMENTS

I would like to thank the following individuals for helping with different aspects of this work – be that everything from generating data, to advice on protocols and life – all of which was critical for the completion of this dissertation work: Daniel Medeiros, David Stock, Michael Klymkowsky, Kristin Artinger, Andrew Martin, David Jandzik, Aaron Garnett, Maria Cattell, Marek Romášek, Haley Stein, Jr Kai Yu, Cheng-Yi Chen, Jeff Peng, Andrew Hansen, Amrita Purkayastha, James Massey, Zachary Root, Bilge Birsoy, Jianli Shi, Nolan Kane, Silas Tittes, Kyle Keepers, Scott Davis, Stacey Smith, William Bowman, Deane Bowers, Robert Cerny, Jeremiah Smith, Stephanie Bryant, Geoff Legault, Daniel Brunelle, John Basey, Alexander Coe, Om Neelay, Harnek Gulati, Tobias Albrigtsen, David Clouthier, Ingo Braasch, Craig Miller, David McCauley, Stephen Green, Josiah Workman, Jill Skarstad, Linda Bowden, Melanie Green, Cora Fagan-Edminster, Ned Friedman, Claire Figel, Diana Oliveras, Carol Kearns, Randy Didomenico, Marcelo Sousa, Mathew Sharples, and Claire Altier.

Additionally, I would like to thank the following agencies, companies, and other non-humans for their help with funding, data analysis, and general contributions to this work: the National Science Foundation, the National Institutes of Health, The CU-Boulder graduate school, the Evo-Devo-Eco Network, the Undergraduate Research Opportunities Program, Quintara Biosciences, The BioFrontiers sequencing facility, Xenbase, Genghis Khan (R.I.P.), Nugget, Q-Tip, Tyrone (R.I.P.), and the Catacombs (R.I.P.).

CONTENTS

CHAPTERS

Tables

Figures

CHAPTER VI

CHAPTER I: Introduction

Neural crest cells (NCCs) have played a prominent role in the history of vertebrate evolution and radiation (Gans and Northcutt, 1983). This vertebrate-specific cell type contributes to diverse structures such as cartilage and bone of the head skeleton, pigment cells, components of the heart, and neurons and glia of the peripheral nervous system (Simoes-Costa and Bronner, 2015). The head skeleton in particular was instrumental in the success of the first vertebrates for its general functions, such as pharyngeal pumping and protection of the sense organs. Thereafter in the vertebrate lineage, the radiation of neural crest derivatives contributed massively to their spread through niches, morphospace, and the globe. Using a multi-species comparative approach, we are able to discern the core features of the genetic networks involved in both neural crest cell fate, and the patterning of NCC-derived tissues. This informs us as to the genetic and developmental changes that occurred some 500-600 million years ago, during early vertebrate evolution and radiation. Besides informing us about how neural crest evolution did occur, compiling such broad developmental systems allows us to see which aspects of development are most deeply shared, and thus which genetic circuits have been most extensively engrained in the development of all living vertebrates. Furthermore, it is from such a 'lowest common denominator' view of development can we best discern when and how lineage specific modifications have occurred, and how they might influence the developmental process in a broader context.

To fully understand the origin of a tissue or structure, it is critical to dissect its developmental origin within its phylogenetic context. While ontogeny does not reflect phylogeny in a literal sense (as Haeckel famously posited), considering phylogeny when addressing ontogeny is extremely valuable. Sampling and testing developmental systems in a range of organisms with a phylogeny-forward approach allows for the most careful and robust conclusions to be drawn with respect to the history and function of a given developmental system (of a specific organism, or a shared process).

In the case of neural crest cells, their phylogenetic context is quite clear: all living vertebrate embryos utilize neural crest cells to make the bulk of their facial skeleton, chambered heart, peripheral nervous system, and pigmentation. Conversely, no living invertebrate within the phylum Chordata has any one cell population with such a history; only multipotent stem cells of early embryos display such a diverse repertoire of fates. It is this developmental potency that many use to distinguish NCCs from other cell types (Hall, 2000; Le Douarin and Dupin, 2003). Thus *bona fide* NCCs made their evolutionary debut in stem vertebrates, in other words, at some point between the divergence of tunicates (Urochordata) and vertebrates, and the deepest divergence within crown vertebrates, being that between the living jawless vertebrates and the living jawed vertebrates. Therefore, comparing non-vertebrate chordates to vertebrates

can give us the best insights as to the developmental origins of NCCs, and what occurs in this unique embryonic region to make it capable of such a developmental feat.

The most closely related invertebrate chordates, lancelets (Cehpalochordata) and tunicates, both have many functionally and molecularly indistinguishable cell types to those derived from NCCs, including melanocytes, migratory neurons, and cellular cartilage (Jandzik et al., 2015; Yu et al., 2010). However, none of these cell types are known to be derived from the same region of the embryo (as we see in the vertebrate neural crest), nor are they derived from a multipotent cell population. Therefore, it seems that vertebrates' earliest neural crest cells gained the capacity to differentiate into a number of diverse cell types that were previously unique to other germ layers. Thus, both the appearance of a multipotent, migratory cell type, and their cooption of the genetic programs that drive these disparate fates occurred in stem vertebrates. In order to better understand the evolution of neural crest cells, first a comparison between vertebrates and invertebrate chordates is essential.

Developmental origins of neural crest cells: a comparison to invertebrate chordates

During gastrula stages of chordate development, newly-specified mesoderm involutes, and induces the overlying ectoderm to form the neural plate by secreting TGFβ inhibitors and other ligand gradients (De Robertis and Kuroda, 2004). These signals delineate the anteroposterior (A/P) and dosroventral (D/V) axes within the future ectoderm, and thus dictate the early patterning of the neural plate and non-neural ectoderm by modulating expression via secreted ligands such as BMP, FGF, and WNT (De Robertis and Kuroda, 2004; Garnett et al., 2012; Marchant et al., 1998; McCauley and Bronner-Fraser, 2004; Meulemans and Bronner-Fraser, 2004; Patthey et al., 2008; Yu et al., 2008b). From this juncture, the future morphology of all major ectodermally-derived tissues is set in motion, including the brain and spinal cord. Across chordates, this process of neural specification can appear outwardly quite different: the shape, size, and number of cells present during this process vary widely between the major chordate groups. Despite these differences in the 'mechanical' acts of gastrulation and neurulation, the genetic underpinnings of these events are strikingly conserved (Basch et al., 2006; Fujimi et al., 2006; Grinblat and Sive, 2001; Holland and Holland, 1998; Nandadasa et al., 2009; Nishida and Stach, 2014; Theveneau and Mayor, 2012; Van Otterloo et al., 2012; Yu et al., 2008b). Transcription factor families activated during gastrulation such as Zic, Pax, Msx, Dlx, and Tfap2 delineate three main regions of the early ectoderm: the non-neural ectoderm, the neural border, and the neural plate (Yu et al., 2008; reviewed by (Milet and Monsoro-Burq, 2012). This high level of conservation indicates that the specification of all three of these ectodermal domains as transcriptionally-distinct cell populations is a

synapomorphy of chordates, and likely evolved prior to the appearance of multipotent neural crest cells (Yu et al., 2008b).

As the neural tube approaches closure in vertebrates, NCCs begin delaminating from the neural plate border and migrate throughout the body along different stereotyped routes. Just prior to delamination of these cells, another wave of expression begins throughout all future vertebrate NCCs: SoxE proteins, Prdm1, Snail, Foxd3, Id, cMyc, and Twist transcription appear nearly in concert, and drive neural crest specification (reviewed by (Milet and Monsoro-Burq, 2012)). These genes help to keep NCCs undifferentiated, and also make them competent to receive later signals secreted by the migrating cells' environments. These later signals will contribute greatly to the eventual position and commitment of a specific migrating cell (Dupin and Le Douarin, 2014; Le Douarin and Dupin, 2003), though there is some evidence that NCCs point of origin might contribute to their morphology in a position-independent manner (Noden, 1978). With surprisingly few exceptions, all vertebrates specify migratory NCCs using a conserved set of genetic interactions (Green et al., 2015; Meulemans and Bronner-Fraser, 2004). Conversely, no cell population with such a coexpression of genes is known in any living invertebrate chordate. At the neural border, the Florida amphioxus *(Branciostoma floridae*) and tunicates both express only one of these genes, *snail*, and the tunicate *Ciona intestinalis* does deploy *foxD* in a subset of their neural border (Imai et al., 2009; Yu et al., 2008b). With this slight similarity in expression in mind, it makes sense that the qualitatively closest cell type in an invertebrate can be found arising from the tunicate neural border, where a small number of cells with limited developmental potential delaminate and display migratory behavior (Imai et al., 2009; Stolfi et al., 2015). Despite their lack of expression in or around the neural plate border (save *snail*), copies of all neural crest specifier gene families have been found in invertebrate chordate genomes. However the vast majority of these genes are never expressed in neural tissue. Instead, many but not all of these are expressed in early mesoderm (which is also the case in vertebrates), intriguingly the germ layer where some NCC cell types like smooth muscle and cartilage probably first evolved.

Conversely, invertebrate chordates do not possess recognizable gene copies of many of the signaling ligands that NCCs use to modulate their migration and differentiation, including Endothelin signaling ligands. In fact, there are at least nine cell fate specification ligands which are "new" to vertebrates (Martinez-Morales et al., 2007), indicating that neural crest cells as a whole arose first by the cooption of pre-existing transcription factors to the neural border (Meulemans and Bronner-Fraser, 2005). Thereafter this nascent cell type relied heavily on 'new' or highly divergent and specialized vertebrate-specific signaling cascades to migrate, become patterned, and differentiate into the plethora of derivatives we see today.

The specification and patterning of different neural crest derivatives is extensive in jawed vertebrates.

Aside from the production of multipotent, migratory cells form the neural border, vertebrates also managed another developmental feat: sending these different cells types to unique and complex sets of embryonic destinations in unison. By molecularly segregating the fates of different NCC lineages both between and within cells types, this allows for the modularity of these tissues. For example, it becomes useful in evolution to be able to differentially control the placement of pigment, cartilage, and neurons. This modularity is further bootstrapped by patterning schemes within each tissue type, which thereafter allow for even more modularity within a structure (e.g. dorsal vs ventral head skeleton derivatives). These fate-specification and patterning mechanisms apparently utilized both the above-mentioned 'new' signals that we only find in vertebrate genomes, as well as ancient families of transcription factors that predate the chordate common ancestor.

Especially in living jawed vertebrates (gnathostomes) we have distilled the function, location, and developmental timing of many pathways responsible for NCC migration, patterning, and differentiation. In jawed vertebrates, these pathways have been tested and characterized extensively from all corners of the clade: animals such as sharks and rays (Compagnucci et al., 2013; Gillis et al., 2013), the ray-finned fish *Polypterus* (Kralovic et al., 2010), zebrafish (Dutton et al., 2001; Ellies et al., 1997; Lee et al., 2004), cichlids (Renz et al., 2011), Axolotl (Cerny et al., 2004), *Xenopus* (Aybar et al., 2003; Marchant et al., 1998; Square et al., 2015a), chicken (Gordon et al., 2010; Kos et al., 2001), and mouse (Abu-Issa et al., 2002; Depew et al., 2002; Zhao et al., 1997) have all revealed many shared and divergent aspects of neural crest cell (NCC) developmental evolution, though conservation of these mechanisms appears to be a rule instead of an exception. This extensive sampling of jawed vertebrates allows for a rather robust picture of early jawed vertebrate development; namely the mechanisms deployed in the last common ancestor to all living gnathostomes.

Research direction and motivation

From these previous studies, we know that these migration and patterning mechanisms evolved sometime between the appearance of crown chordates and the appearance of crown gnathostomes, events that happened approximately 100 million years apart (Brazeau and Friedman, 2015; Mallatt and Chen, 2003; Morris and Caron, 2014; Shu et al., 1999; Shu et al., 2003). Many extinct fossil vertebrates originated during this time (between the mid Cambrian and early Ordovician periods), and diversified during the Devonian (a period informally known as "the Age of the Fishes"). For instance, this is where

we see the evolution of bone, odontodes, paired appendages, predatory lifestyles, and the jaw to name a few key traits that arose during this time (Janvier, 1996); all of these traits somehow involved NCCs. Today, the only extant jawless vertebrates comprise lampreys and hagfishes, which form a natural group, the cyclostomes (Delarbre et al., 2002; Heimberg et al., 2010; Oisi et al., 2013; Stock and Whitt, 1992). Cyclostomes diverged during this period, and correspondingly they possess most, but not all of the same neural crest cell derivatives as modern jawed vertebrates. Namely, they are without bone, the sympathetic nervous system, NCC-derived cranial ganglia, myelin-secreting Schwann cells, and dentine-producing odontoblasts. Cyclostomes most likely diverged from stem gnathostomes before these cell types arose, but given the general prevalence of evolutionary loss in developmental systems, this cannot be absolutely assumed. Despite these specific differences, cyclostome NCCs are multipotent (McCauley and Bronner-Fraser, 2003; Nikitina et al., 2008) and are still capable of becoming the four major types of tissues (being cartilage-related skeletal tissues, smooth muscle, neurons of the peripheral nervous system, and pigment cells).

The divergence between cyclostomes and gnathostomes is the most ancient of any divergence within living vertebrates, making modern cyclostomes are phylogenetically positioned in a very informative place for understanding basal vertebrate features. By addressing the extent of these patterning networks in NCCs of the sea lamprey (*Petromyzon marinus*), a tractable jawless vertebrate, we can ascertain which of these mechanisms arose in unison with the different neural crest derivatives, and which mechanisms are specifically jawed vertebrate innovations. So, while the genetic changes underlying neural crest cell specification as a whole have been determined to some degree (Green et al., 2015), the processes that arose in the first vertebrates for modulating the major neural crest cell types, dictating their migration, and patterning them into three-dimensional structures and tissues remain largely unaddressed. Lampreys are particularly informative to these ends: they offer us a 'developmental sample' of vertebrate growth that has evolved separately from all gnathostomes, having diverged just before the major fish radiations occurred in the Devonian. This allows us to discern some of the ancestral gene expression patterns and functions of early NCC pathways. This motivates the present work: by working with amphioxus (*Branchiostome floridae*), the sea lamprey (*Petromyzon marinus*), and the African clawed frog (*Xenopus laevis*), we have put together a more complete picture of the evolutionary story of stem vertebrates and their early neural crest cells. With special focus on the head skeleton, and using a phylogeny forward approach, this body of works aims to better understand the patterning schemes and signaling processes that evolved during the evolution of the first vertebrates approximately 550 million years ago.

CHAPTER II: The origin and diversification of the developmental mechanisms that pattern the vertebrate head skeleton

Introduction

Vertebrates emerged during the Cambrian explosion more than 500 million years ago. They have since attained an incredible degree of specialization and morphological diversity, which likely played a role in their becoming the most species-rich and geographically dispersed deuterostomes on the planet. This success is thought to have been made possible, in large part, by the origin and elaboration of the vertebrate head skeleton. In the earliest vertebrates, the facilitation of pharyngeal pumping by the head skeleton seems most probably to be a major source of early vertebrate success, by simply increasing the rates of respiration and filter feeding (Gans and Northcutt, 1983; Northcutt and Gans, 1983). While many modern vertebrates (namely tetrapods) have discarded this pharyngeal pumping strategy in their adult forms, the head skeleton still performs many basic functions in all vertebrates by supporting and protecting the brain and anterior sense organs. Aside from its ancestral and shared functions, the head skeleton is also extraordinarily evolvable, having proven itself capable of taking on a wide array of adaptive shapes and compositions for respiration, feeding, communication, and sensing the environment. The fossil record suggests this flexibility arose very early in the vertebrate lineage, with an impressive diversity of both jawless and jawed forms arising within 100 million years after the first vertebrates appeared (Fig. 1).

Though the fossil record shows when the adult head skeleton arose and how it has diversified, it tells us nothing about the developmental and genetic bases of its extreme evolvability or its origin. The modern comparative or "evolutionary developmental" biology approach allows us to deduce conserved and divergent features of development, and functionally link these to similarities and differences in adult morphology. In this review we examine what is known about early head skeleton development in a broad sampling of living vertebrates, including the only living jawless vertebrates, the cyclostomes. The cyclostomes consist of two groups: hagfishes and lampreys. Historically, there has been some debate about whether cyclostomes are monophyletic or paraphyletic, however various sequence-based analyses support the former scenario (Delarbre et al., 2002; Heimberg et al., 2010; Stock and Whitt, 1992). The fossil record reveals that cyclostomes represent only a fraction of agnathan diversity (see Fig. 1) and suggests that modern adult lampreys and hagfishes possess highly derived skeletal morphologies and lifestyles, a phenomenon that is also easily observed in many gnathostome groups (e.g. the evolution of

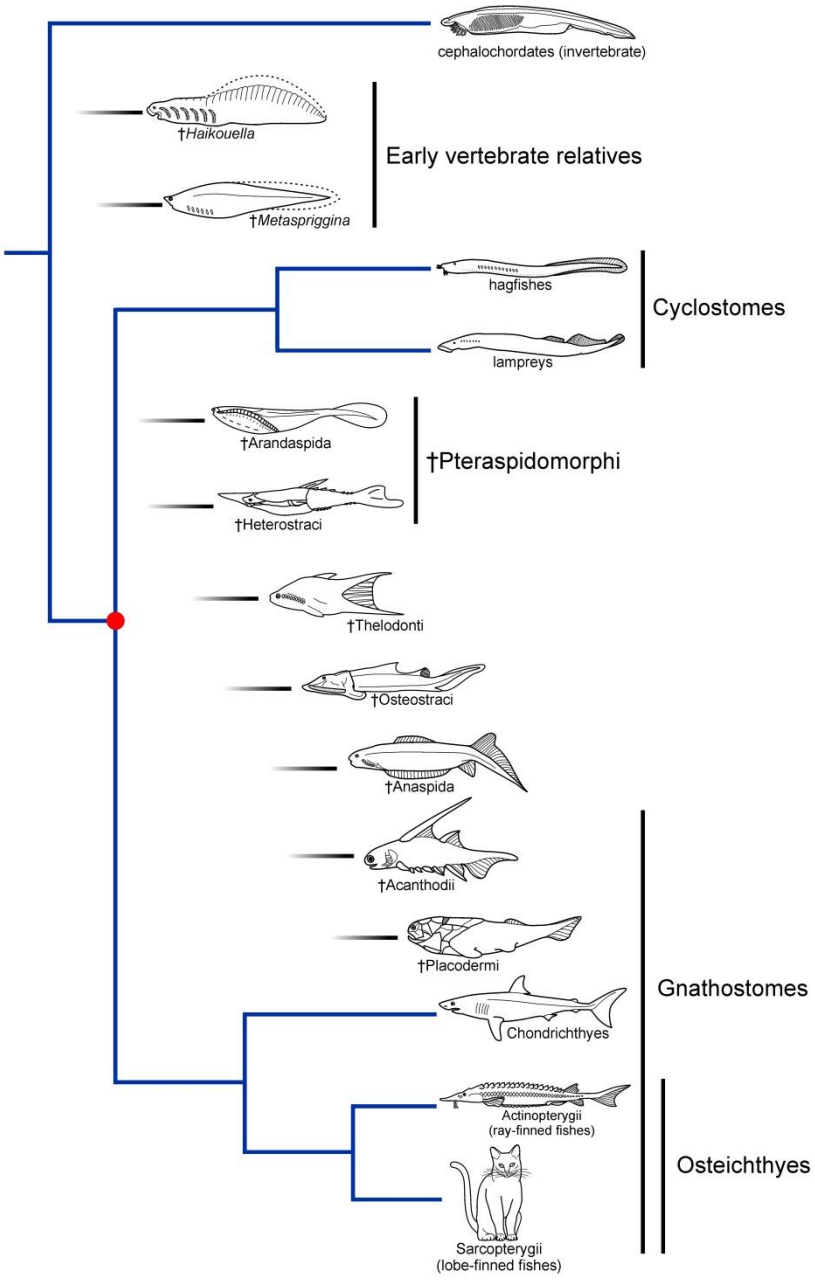

Fig. 1. A cladogram of major vertebrate groups. Extinct groups are indicated with a (†). Relatedness of only the extant groups is depicted with a blue cladogram. The node representing the position of the last common ancestor to all living vertebrates is colored red. The position along the X-axis for each fossil group's line shows their relative approximate time of presence in the known fossil record. The length of every line is the same for these extinct groups, and is not representative of the known duration of each group's history on earth (which is extremely variable considering these groups). Some of these groups may be paraphyletic. Some groups are very diverse; only a single living member or well-characterized fossil from each group is pictured. Some major groups not shown include Conodonta and Galeapsida. Some illustrations are after those within (Hildebrand and Goslow, 2001) and (Janvier, 1996).

the middle ear bones from pharyngeal arch structures in mammals). Nevertheless, modern cyclostomes offer our only window into agnathan developmental genetics. Furthermore, larval lampreys possess a simple head skeleton composed mainly of simple, cartilaginous PAs, and appear grossly similar in morphology to the earliest fossil agnathans. Thus larval lampreys likely retain many ancestral features lost or masked in gnathostomes.

In this phylogenetic context, here we use comparisons of cyclostome and gnathostome head skeleton development in combination with the fossil record to support several key conclusions about the head skeleton of the most recent common ancestor of modern vertebrates, found at the node bearing the red circle in Fig. 1: 1) it formed mainly from a set of three distinct streams of cranial neural crest cells (CNCCs), 2) after migration into the pharynx, these cells activated a 'cartilage gene regulatory network (GRN)' of at least three genes, as well as a combinatorial code of at least 10 transcription factors in particular oropharyngeal subdomains, and 3) these subpopulations of skeletal precursors differentiated into an endoskeleton consisting mainly of pharyngeal arches (PAs), throughout which histologically distinct cartilage types were deployed. Taken together, the conserved aspects of head skeleton formation and patterning suggest that major differences in head skeleton morphology are likely due to changes in differentiation and morphogenetic programs downstream of a conserved developmental prepattern. However, some striking differences in this prepattern are evident, and may be tied to some specific large-scale differences in morphology between modern jawless vertebrates (agnathans) and jawed vertebrates (gnathostomes), and also within gnathostomes.

Modern vertebrate head skeletons are derived from CNCCs that migrate as three streams

In both modern cyclostomes and gnathostomes, vital dye labeling and gene expression suggest all cartilaginous PA (pharyngeal arch) and pre-oral skeletal elements are derived mainly from CNCCs (Kuratani et al., 2016; McCauley and Bronner-Fraser, 2003; Noden, 1978; Santagati and Rijli, 2003). After being specified by highly conserved gene regulatory interactions (Sauka-Spengler et al., 2007; Simoes-Costa and Bronner, 2015), CNCCs migrate ventrally from the neural border during and after neurulation, populating the pharynx and oral region. In both gnathostomes and lamprey, CNCCs are specified at all positions along the neural plate border from the midbrain through the hindbrain, but congregate into three main streaming populations as they proceed towards the pharynx: pre-oral and PA1 cells in the 1st stream, PA2 cells in the 2nd stream, and branchial arch (PAs 3+) cells in the 3rd stream (Fig. 2; reviewed by (Theveneau and Mayor, 2012)). The CNCC-negative regions (between the streams) are consistently situated beneath rhombomeres 3 and 5 of the brain in both lampreys and gnathostomes

(Minoux and Rijli, 2010). These spatial similarities mark high conservation of this general migratory architecture.

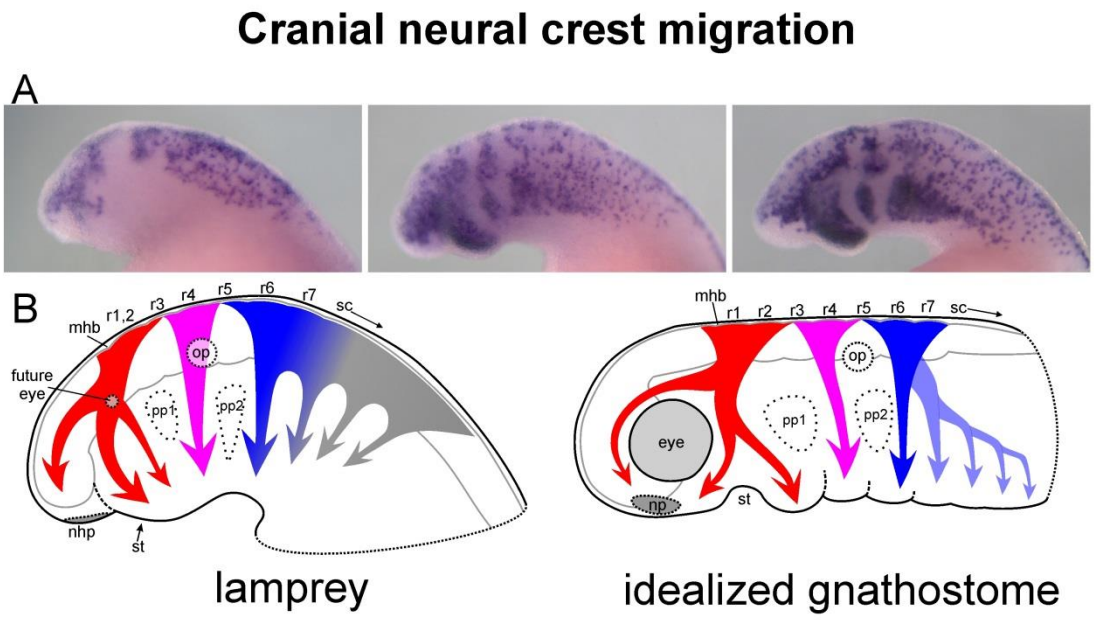

Fig. 2. Vertebrate cranial neural crest migration occurs as three topographically conserved streams. (A) Left lateral views of early pharyngula stage lamprey embryos (*Petromyzon marinus;* Tahara stage 21-23(Tahara, 1988)) stained via *in situ* hybridization for *ednrb* transcripts (see Square *et al.*, 2016 for a broader staging series of *ednrb in situ* hybridizations). This gene marks migratory skeletogenic cranial neural crest, among other neural crest derivatives. (B) A cartoon of early cranial neural crest migration in lamprey and an idealized gnathostome. The three homologous populations of CNCCs are indicated by different colors (red, pink, and blue); the sub-arrows within each of these streams are meant to depict general directions of cell migration, but do not explicitly indicate stream subdivisions. Each uniquely colored stream is molecularly defined by *hox* expression (see text). The 3rd stream in lamprey appears to have some contribution from the trunk; this is partially colored gray. The posteriormost arches of the gnathostome are slightly transparent to indicate that these splitting events occur after the population of cells is already moving ventrally (a process coincident with pharyngeal pouch formation [see text]). An outline of the brain is shown in gray. mhb, midbrain/hindbrain boundary; nhp, nasohypophyseal plate; np, nasal placode; op, otic placode; pp1 and pp2, pharyngeal pouches 1 and 2; r1-7, rhombomeres 1-7; st, stomodeum.

While CNCC migration patterns are generally conserved across vertebrates, there are some notable differences between the lamprey and gnathostome 3rd CNCC streams. In gnathostomes the 3rd stream originates from rhombomeres 5-7, and becomes progressively subdivided as it is still migrating and after migration (Minoux and Rijli, 2010). In contrast, the 3rd 'stream' of CNCCs in lamprey behaves more like a sheet, and emerges from a broad domain that includes 5-7th rhombomeres (Kuratani et al., 1998a) and part of the presumptive spinal cord, a region that would give rise to trunk neural crest in gnathostomes (Fig. 2). The reason for this is unclear, though it may simply reflect the fact that lamprey has a relatively long pharynx consisting of seven branchial arches (PA3-PA9). Nevertheless, in both lineages, the posterior arch neural crest cells become segregated in concert with pharyngeal pouch morphogenesis, and end up as separated 'tubes' of ectomesenchyme surrounding a mesodermal core (Cerny et al., 2004), flanking each gill slit.

Notably, the posterior PAs (the branchial arches) tend to be more homogenous both within and among species with regard to their gene expression patterns and eventual morphology. This contrasts with PA1 and PA2, which each come from unique streams of CNCCs; these are far more specialized in their gene expression and eventual morphology both within a given vertebrate's PAs, and among vertebrates. How and when separate 1st, 2nd, and 3rd streams, evolved, and began to give rise to morphologically distinct derivatives is unclear. However, it is possible that the first CNCC to migrate from the neural tube was already divided into molecularly distinct AP populations by hox gene expression carried over from the CNS. Physical segregation of 3rd stream CNCCs from each other may evolved later in concert with the otic capsule, which is a conserved landmark for the boundary between the 2nd versus 3rd streams in all vertebrates.

When the three streams began giving rise to morphologically distinct derivatives is even more speculative as all extant vertebrates display highly specialized PA1 and PA2 and posterior PA morphologies. Looking to the fossil record, some extinct vertebrate relatives also possess specialized cartilages around the mouth, as does the invertebrate amphioxus (Jandzik et al., 2015), but the developmental origin of these fossil structures (specifically whether they are PA-derived or not) is extremely difficult to assess. Irrespective of the timing of PA1 and PA2 specialization, gene expression (see below) and their resemblance to the PAs of fossil chordates indicates that PAs derived from the 3rd stream are likely the least developmentally derived, and might offer the best possible reconstruction of ancestral PA patterning at their first appearance in the vertebrate stem (discussed below).

Transcription factor expression in the vertebrate head skeleton

A conserved cellular environment, a conserved cartilage GRN, and a conserved tissue type

In both cyclostomes and gnathostomes, CNCCs entering the pharynx and oral region are exposed to a range of intercellular signals that regulate proliferation, differentiation, and morphogenesis. These include bone morphogenetic proteins (BMPs), Fibroblast growth factors (FGFs), Endothelins (Edns), retinoic acid (RA), and Hedgehog (Hh) (reviewed by (Santagati and Rijli, 2003)). In lamprey, the expression of key pathway members has been documented in and/or surrounding the pre-skeletal CNCCs for each of these signal types (Campo-Paysaa et al., 2015; Jandzik et al., 2014b; Kuraku et al., 2010; McCauley and Bronner-Fraser, 2004; Medeiros and Crump, 2012; Square et al., 2016a; Sugahara et al., 2011). Furthermore, the function of FGF and RA signaling has been addressed in lamprey, and found to be generally conserved with gnathostomes (Jandzik et al., 2014b; Kuratani et al., 1998b). In addition to entering a similar intercellular signaling environment, CNCCs of both lampreys and gnathostomes differentiate into collagen-containing cellular cartilage and related skeletal tissues, such as joint tissue (gnathostomes) and soft mucocartilage (lamprey) (Cattell et al., 2011; Crump et al., 2004a; Medeiros and Crump, 2012; Zhang et al., 2006). Consistent with this, gnathostome and lamprey CNCCs activate the same core set of chondrogenic regulators, including SoxE, Twist, and Ets (Meulemans and Bronner-Fraser, 2004). These genes are activated in the developing amphioxus oral cirri skeleton(Jandzik et al., 2015), a non-vertebrate cartilage, and *soxE* genes are expressed in the cartilage of horseshoe crabs and cuttlefish (Tarazona et al., 2016). This indicates that a core cellular cartilage differentiation program was already present before vertebrates arose, with a rudimentary version of this gene regulatory cascade likely predating the protostome/deuterostome divergence.

Combinatorial expression of *alx*, *hand*, *msx*, and *prrx*, define spatially conserved precursor populations in the PAs of all vertebrates

In addition to a core set of transcription factors that appear to drive skeletal differentiation in all CNCC, there are several others expressed only in subsets of head skeleton precursors in both cyclostomes and gnathostomes (Fig. 3). For simplicity, we use the combined expression of all known paralogs of a given gene group when defining which regions are positive for a gene type (e.g. an *msx* positive domain means there is at least one *msx* gene expressed there, but there could be multiple). Together, these genes appear to act combinatorially to confer 'module-specific' identity upon CNCC subpopulations. In gnathostomes, most of these genes are known to affect different regions of head skeleton development in unique ways, though more work is needed to understand precisely how these genes confer regional shape and morphology. The conserved expression of these factors in all modern vertebrates suggests they mark

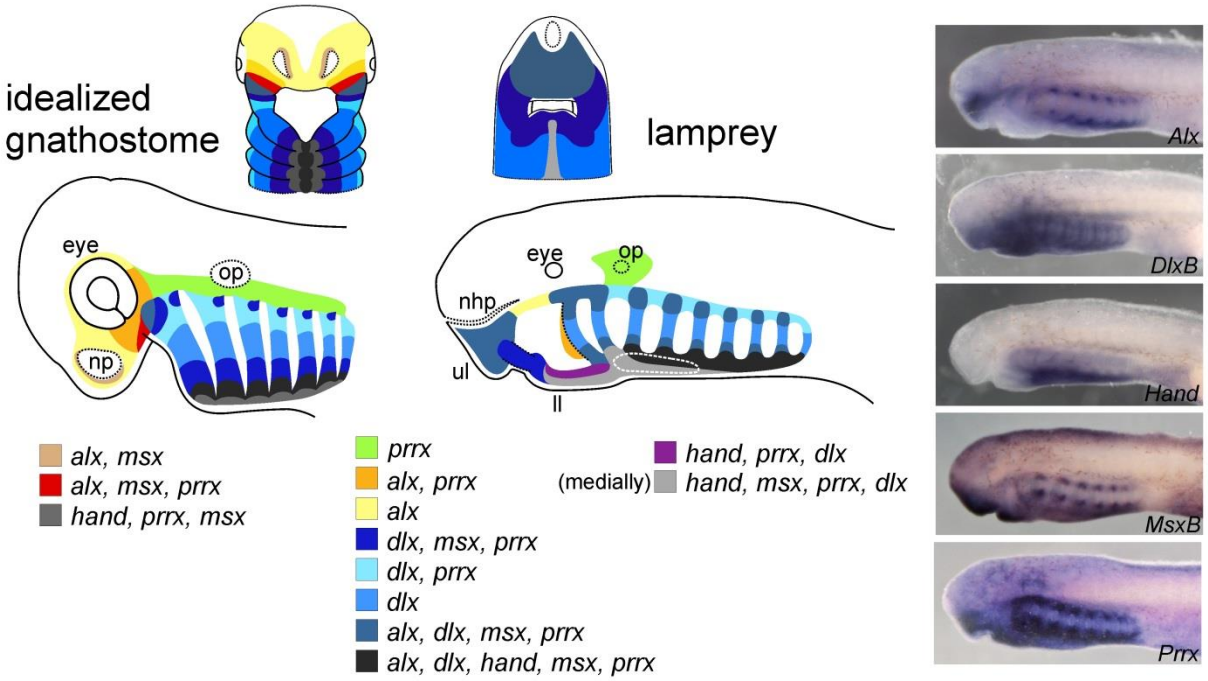

Fig. 3. A comparison of the gene expression schemes of lamprey and gnathostome nascent head skeletons. Top are oral views, below which are left lateral views. The key below is arranged such that the combinatorial domains in the left and right columns are found only in gnathostomes or lamprey, respectively, while the combinatorial domains in the middle column are found in both lineages. For gene types with multiple paralogs, these domains represent the expression of all paralogs (e.g. an *msx* positive domain could express any one of the *msx* genes in that lineage). On the right are example *in situ* hybridizations of each gene depicted in the lamprey expression map. The gnathostome map represents a simplification and slight modification to the map found in (Square et al., 2015a). The ventrally-positioned white dotted oval in the lamprey represents the position of the endostyle, which is derived from endoderm and expresses none of the genes addressed here. ll, lower lip; nhp, nasohypophyseal plate; np, nasal placode; op, otic placode; ul, upper lip.

evolutionarily conserved subpopulations of skeletal precursors present in their most recent common ancestor. Although these developmental modules of CNCCs are likely deeply homologous, if not homologous *sensu stricto* (or 'historically' homologous), it is less clear if the adult skeletal structures derived from them can be considered homologous in either a deep or strict/historical sense (for a discussion of deep homology, see (Shubin et al., 2009)).

The dorsal and ventral poles of all vertebrate PAs examined to date express a similar combination of transcription factors: *alx* (Beverdam and Meijlink, 2001; Cattell et al., 2011; Compagnucci et al., 2013; Dee et al., 2013; McGonnell et al., 2011; Square et al., 2015a), *hand* (Cerny et al., 2010; Charite et al., 2001; Compagnucci et al., 2013; Firulli, 2003; Square et al., 2015a), *msx* (Antonopoulou et al., 2004; Cerny et al., 2010; Compagnucci et al., 2013; Square et al., 2015a; Swartz et al., 2011), and *prrx* (Compagnucci et al., 2013; Hernandez-Vega and Minguillon, 2011; Square et al., 2015a; ten Berge et al., 1998) transcripts can be found in both lamprey and gnathostomes in broadly similar patterns (Fig. 3; Fig. S1). In the ventral PAs all four of these genes are transcribed, though in lamprey ventral *alx* expression is absent from PA1 and PA2 (Cattell et al., 2011). Skeletogenic mesenchyme in the dorsal PAs of sharks, *Xenopus*, and mouse also express *msx* and *prrx*, though this domain appears to be only *prrx* positive in at least zebrafish. The intermediate domain between the poles (the light blue '*dlx* only' module in Fig. 3) in all vertebrates is marked by the exclusion of these four polar PA gene transcripts, but this domain also consistently overlaps the region wherein the highest number of *dlx* genes are expressed (discussed below; Fig. 4). Combined, *alx*, *hand*, *msx*, and *prrx* represent an ancient PA polarity scheme that molecularly designates CNCCs to dorsalmost, intermediate, and ventralmost identities. The function of all four of these genes are shown in various gnathostome models (Antonopoulou et al., 2004; Beverdam et al., 2001; Cerny et al., 2010; ten Berge et al., 1998; Yanagisawa et al., 2003), however functional data from any cyclostome is lacking.

Nested *dlx* expression marks dorsal, ventral and intermediate skeletal precursor populations in the PAs of all living vertebrates

In gnathostomes, the *dlx* genes comprise six orthology groups (1-6), and are typically found as tandem duplicates in the genome (*dlx1* and *-2*, *dlx3* and *-4*, and *dlx5* and *-6*) (Stock et al., 1996). This paired genomic architecture arose from *cis*-duplication at the first *dlx* locus in the pre-vertebrate chordate lineage; this apparently occurred in stem olfactores (tunicates + vertebrates) after the divergence of cephalochordates (Wada and Makabe, 2006) (Fig. 4). This paired architecture has been retained at most gnathostome *dlx* loci, but this arrangement has not been confirmed in cyclostomes for any of their six *dlx* loci. In early vertebrates, the first tandem pair underwent whole genome and/or regional *trans*-duplications (and potentially some losses) giving rise to the six genes we find in most modern vertebrate

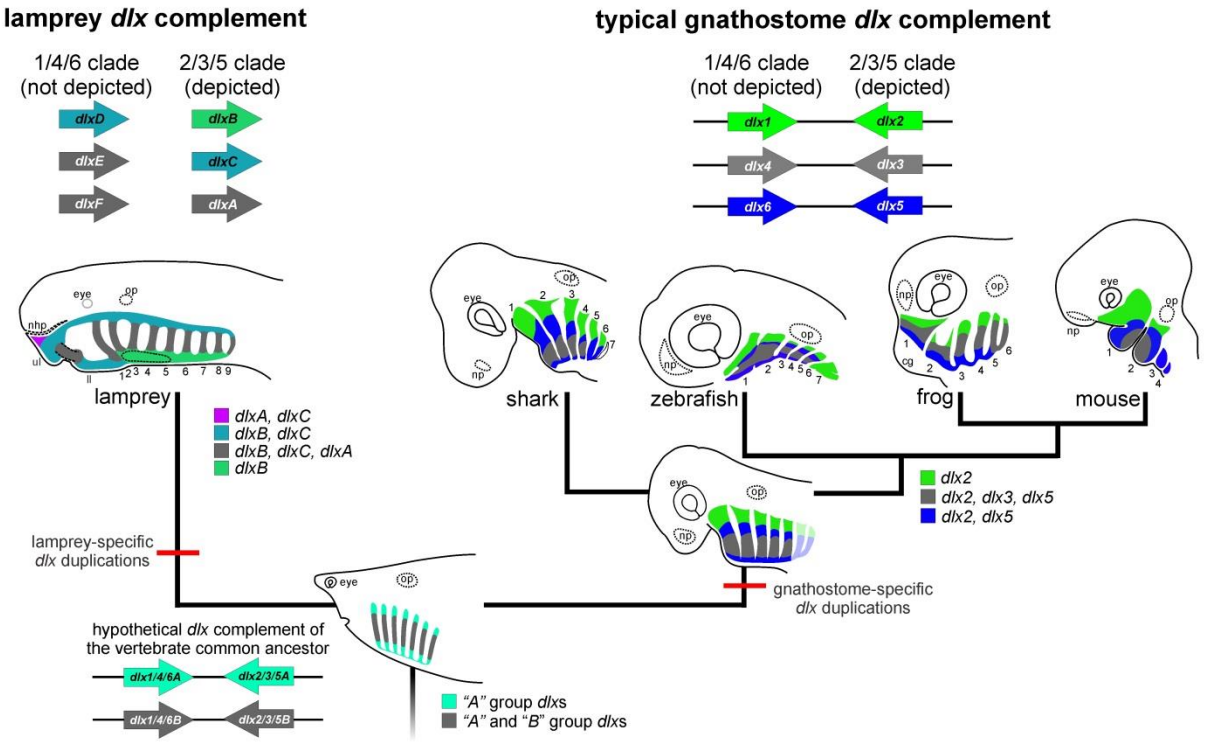

Fig. 4. *dlx* evolution in vertebrates. The colors of the genes are meant to reflect nested expression categories, rather than the relatedness of these genes, with the gray color being shared by all groups to indicate that all *dlx2/3/5* genes are expressed there. From bottom to top: The common ancestor to all living vertebrates likely had two pairs of *dlx* genes, which arose by first a *cis*-duplication, and later a *trans*-duplication (see *dlx* section in main text). Hypothesized *dlx* gene expression is depicted on a representation of this common ancestor's head skeleton, with all four genes expressed in the intermediate PAs (gray domain), and only one pair of *dlx* genes expressed throughout the PAs (turquoise domain). Thereafter (above), the cyclostome and gnathostome lineages diverged, each of which duplicated and retained at least two more *dlx* genes independently (thus lamprey and gnathostome *dlx* gene domains are colored differently, other than the gray domain where all six are expressed). Modern gnathostome *dlx* expression is depicted for a shark (*Scyliorhinus canicula*) (Compagnucci et al., 2013), zebrafish (*Danio rerio*) (Talbot et al., 2010), frog (*Xenopus laevis*)(Square et al., 2015a), and mouse (*Mus musculus*) (Jeong et al., 2008). The 'average' of these four vertebrates is depicted below as an idealized gnathostome common ancestor. The ventrally-positioned black dotted oval in the lamprey represents the position of the endostyle, which is derived from endoderm and expresses no *dlx* genes. The PAs are numbered in each extant vertebrate at their base. ll, lower lip; nhp, nasohypophyseal plate; np, nasal placode; op, otic placode; ul, upper lip.

lineages (Stock, 2005; Takechi et al., 2013). It is important to note that despite cyclostomes and most gnathostomes having six main types of *dlx* genes, there is strong evidence that hagfish, lamprey, and gnathostome *dlx* genes are not all directly orthologous to each other (Cerny et al., 2010; Fujimoto et al., 2013; Kuraku et al., 2010; Stock, 2005; Takechi et al., 2013), although the exact history of duplication and retention is difficult to address in the face of differential paralog loss (Kuraku, 2013). The current *dlx* assemblages have thus arisen from some number of lineage-specific *trans*-duplications in each of these three groups.

In gnathostomes, tandem pairs of *dlx* genes typically share some aspects of their expression domains, although notably the *dlx1/4/6* group is consistently expressed in a more restricted manner than its tandem duplicate in the *dlx2/3/5* clade (Compagnucci et al., 2013; Depew et al., 2002; Gillis et al., 2013; Square et al., 2015a; Talbot et al., 2010). This likely reflects a regulatory condition that arose at the original *dlx1/4/6* and *dlx2/3/5* locus after the initial tandem duplication. In gnathostomes, some *dlx* functions are relegated to specific sets of *dlx* genes, such as *dlx2* in migratory CNCCs (Compagnucci et al., 2013; Crump et al., 2004a; Square et al., 2015a), *dlx2, -3, -4,* and *-5* in developing tooth germs (Borday-Birraux et al., 2006; Renz et al., 2011; Zhao et al., 2000), and *dlx1, -2, -5,* and *-6* in the forebrain (Renz et al., 2011; Zerucha et al., 2000). Subsets of lamprey *dlx* genes are also found in migrating CNCCs (*dlxA, -B, -C,* and *-D*), as well as the forebrain (*dlxA, -C, -D,* and *-E*) (Kuraku et al., 2010). Despite the ambiguity surrounding the exact level of gene orthology here, the presence of these specialized expression domains combined with information from phylogenetic analyses indicate that the full sets of lamprey and hagfish *dlx* genes do not stem from only within-cyclostome duplications (Fujimoto et al., 2013). Furthermore, recent evidence suggests that cyclostomes diverged after at least one whole genome duplication in vertebrates (Kuraku, 2013; Kuraku et al., 2009; Smith and Keinath, 2015), which is the process assumed to have given rise to the *trans*-duplication of *dlx* genes(Stock, 2005; Stock et al., 1996; Takechi et al., 2013). Thus, the vertebrate common ancestor likely had at least two pairs of *dlx* genes (Stock, 2005; Takechi et al., 2013) (Fig. 4). It also seems most probable that these two pairs of *dlx* genes were differentially expressed in the vertebrate common ancestor given the high degree of specialization we see across all living vertebrate *dlx* complements.

At mid-pharyngula stages the *dlx* genes appear in nested PA expression domains in both sea lamprey(Cerny et al., 2010) and gnathostomes (Brown et al., 2005; Compagnucci et al., 2013; Depew et al., 2002; Gillis et al., 2013; Renz et al., 2011; Square et al., 2015a; Talbot et al., 2010) (Fig. 4). In both lineages, this nested pattern is deployed and refined soon after the CNCCs cease their migration, with the highest number of *dlx* paralogs consistently expressed in CNCCs occupying the intermediate and/or ventral-intermediate domain of the PAs. In gnathostomes, this nesting is referred to as the "*dlx* code" for its role in specifying identity along the DV axis in PA1 (a.k.a. the proximal/distal axis)(Depew et al.,

2002). There is also strong evidence that hagfish *dlx* genes are at least expressed differentially in pre-skeletogenic mesenchyme (Fujimoto et al., 2013), though the presence or absence of nesting in the PAs remains to be addressed. Importantly, the CNCC population expressing the most *dlx* paralogs is generally situated more ventrally in gnathostomes as compared to the centrally-nested lamprey scheme, although the exact focal point of nesting seems to vary slightly between groups (Fig. 4). For example, most gnathostomes show distinct dorsal boundaries of *dlx3* and *dlx5* expression, though the amphibian *Xenopus laevis* shows a shared dorsal boundary of these two genes, making their dlx code more centered along the dorsoventral axis (Fig. S2).

There are many other lineage-specific modifications to the *dlx* code in gnathostome PAs, especially with regard to the *dlx1/4/6* group. To list a few examples: *Xenopus laevis dlx2* is absent from the dorsal PA2 (Square et al., 2015a), mouse *dlx5* is expressed further ventrally (distally) than *dlx2* (Jeong et al., 2008), sharks exhibit an early restriction of *dlx1* to the dorsal PA1 (Compagnucci et al., 2013; Gillis et al., 2013), and reptilian *dlx4* seems to have been pseudogenized (Brown et al., 2005; Takechi et al., 2013). Furthermore, the posteriormost one to three PAs frequently show a delay or change in the dlx code in the pharyngula stages addressed here. This is likely due to either temporal differences in the specification of PA CNCCs, or simply degeneration of *dlx* expression in these less prolific PAs. Aside from these nuances, cyclostomes and gnathostomes still each possess three main *dlx*-positive domains in the PAs marked by different combinations of *dlx2/3/5* gene expression (Fig. 4).

Thus the common ancestor of all vertebrates likely also specified at least three different PA subdomains of CNCCs using a rudimentary *dlx* nesting scheme: the dorsal and ventral domains with transcripts from only one pair of *dlx*s, and an intermediate region with transcripts from two pairs of *dlx*s (Fig. 4). Given the important role of *dlx*s in specifying dorsoventral PA identity in zebrafish (Talbot et al., 2010) and mouse (Depew et al., 2002), it seems plausible that nested *dlx* expression has an ancient role in conferring position-specific skeletal morphology or histology that evolved before the appearance of PA subdivisions in gnathostomes, including the primary jaw joint in PA1. Comparative studies on *dlx* transcriptional regulation across vertebrates, including cyclostomes, should shed more light on ancestral vertebrate dlx expression.

Like the domains delineated by *alx, hand, msx,* and *prrx,* these main *dlx* domains do not seem to have a 1:1 correspondence with any skeletal elements, but instead mark groups of entire future elements, and occasionally parts of future elements, comprising larger-scale morphological modules (e.g. the elements of the future lower jaw skeleton are marked by *dlx5* and *-6*). This is shown by vital dye labeling (Gillis et al., 2013), reporter expression (Ruest et al., 2003), and the functional transformation and truncation of multiple skeletal elements when *dlx* genes are perturbed (Depew et al., 2002; Talbot et al., 2010). Therefore, these *dlx* domains instead seem to grant more a general identity to these different

domains in the PAs along the dorsoventral axis that contain the CNCCs belonging to multiple nascent skeletal elements.

Combinatorial gene expression defines conserved skeletal precursor populations outside of the PAs

Within gnathostomes are found four main groups of *alx* genes (*1-4*), though some of these have been lost in evolution by different groups (McGonnell et al., 2011). In lamprey only a single ortholog is known(Cattell et al., 2011). In addition to the PAs (discussed above), *alx* gene expression is found throughout the presumptive chondrocranium of gnathostomes, where these genes play critical roles in the differentiation of these non-PA skeletal structures (Antonopoulou et al., 2004; Beverdam et al., 2001; Beverdam and Meijlink, 2001; Compagnucci et al., 2013; Dee et al., 2013; McGonnell et al., 2011). Notably, these skeletal elements are composed of CNCCs, mesoderm, or a mix of both cell types, and the location of the 'boundary' between these two cell types seems to be somewhat plastic in evolution (Kague et al., 2012; Piekarski et al., 2014). Cyclostomes do possess cartilages dorsal to and surrounding the mouth, some of which comprise their neurocranium (Johnels, 1944; (Oisi et al., 2013)), though these individual skeletal elements lack obvious homology to the well-conserved set of gnathostome neurocranial elements. Even comparing lampreys and hagfishes does not reveal obvious affinities between these structures, though their genetic and developmental affinities can be addressed in this context (Kuratani et al., 2016). In lamprey, these 'trabecular' skeletal elements express *alx* during their development (Cattell et al., 2011). Lamprey and most gnathostomes thus have a domain near the eye that is *alx* positive, but *dlx*, *hand*, *msx*, and *prrx*-negative. Based on its conserved position and gene expression, we posit that this "*alx*-only" module (yellow in Fig. 3) is likely homologous between cyclostomes and gnathostomes, having been greatly expanded in the gnathostome lineage to form other components of the ventral braincase. Interestingly, this is one of the few head skeletal regions in lamprey that is thought to originate from mesoderm (Kuratani et al., 2016). This region of the lamprey head skeleton therefore may offer potential insights into the evolution of the braincase.

prrx genes are related to *alx* genes, but this duplication predates the divergence of echinoderms and chordates (Howard-Ashby et al., 2006), and thus these genes have diverged substantially in both their amino acid sequences and their expression patterns. In gnathostomes are two *prrx* genes, though both seem to have very similar expression patterns. We identified a single *prrx* gene transcript in lamprey, which according to phylogenetic analysis is the outgroup to all gnathostome Prrxs (Fig. S3). As in gnathostomes, this *P. marinus prrx* gene is expressed in a mesenchymal domain near the otic placode (green in Fig. 3) which likely gives rise to the otic capsule (ten Berge et al., 1998). Around the mouth, all vertebrates addressed also have a single region where *alx* and *prrx* are coexpressed in the absence of any

dlx, hand, or *msx* gene. In lamprey, this *alx + prrx* module becomes the medial velar skeleton (the opposable "flap" of the velum; orange in Fig. 4), while in gnathostomes this domain is found within the future chondrocranium, laterally to the future palate and below/partially surrounding the eye. This region is not part of the gnathostome PAs, but still receives a CNCC contribution. Thus in both lineages, this *alx + prrx* module is adjacent to and strongly associated with PA1, but has a distinct expression profile compared to PA tissue, namely in that it is *dlx*-negative (Kuraku et al., 2010). In lamprey, the position and function of the medial velar skeleton indicates developmental affinity for PA1, however upon closer inspection this structure also shows a clear histological association to the more dorsal 'chondrocranial region' near the notochord (Fig. S3). This indicates that it might instead be derived from non-PA CNCCs or mesoderm of the rudimentary lamprey "chondrocranium".

Nested *hox* expression defines PA identities in all living vertebrates

Aside from within-PA patterning, nested *hox* gene expression in the pharynx along the anteroposterior axis occurs in both lampreys and gnathostomes; these expression patterns were recently shown to be deployed via conserved regulatory machinery (Parker et al., 2014). As in other tissues, these expression patterns are collinear, and confer each PA with a unique *hox* identity. In all modern vertebrates, pre-oral and PA1 mesenchyme is *hox*-negative, PAs 2+ are *hox2* positive, and PAs 3+ are *hox3* positive (Hunt et al., 1991a; Hunt et al., 1991b; Hunt et al., 1991c; Lyon et al., 2013; Minoux et al., 2009; Takio et al., 2007). Interestingly, these highly conserved *hox* domains correspond to the three migrating populations of CNCCs (discussed above), and generally reflect the expression profile of the brain region they delaminate from (Parker et al., 2016). In gnathostomes, this expression confers PA-specific morphology: exogenous or depleted Hox function can lead to homeotic transformations of the PAs (reviewed by (Minoux et al., 2009)), however no work on any cyclostome Hox function has been published to date. Both lamprey and gnathostomes also have at least one other hox gene expressed in a subset of more posterior arches, though the similarity is not as strong: the arctic lamprey expresses *Hox4x* in a gradient from posterior to anterior (Takio et al., 2007), whereas gnathostomes show clearer nesting of *hox4* in PAs 4+ (Lyon et al., 2013; Minoux et al., 2009), and *hox5* in PAs 5+ (if present) (Lyon et al., 2013). It is useful to note that, like the *dlx* genes, *hox* genes have undergone multiple *trans*-duplications in vertebrate lineages due to whole genome and other large-scale duplications (Smith and Keinath, 2015), and each modern *hox* cluster seems to have nuanced expression between gnathostome groups (Hunt et al., 1991a; Hunt et al., 1991b; Hunt et al., 1991c; Lyon et al., 2013; Minoux et al., 2009). Despite this, we can still safely conclude that the vertebrate common ancestor deployed *hox2, hox3,* and *hox4* genes in the PAs to confer specific transcriptional identities to CNCCs of the nascent viscerocranium, with PA1 CNCCs specifically being *hox*-negative.

In the *hox*-negative pre-oral and PA1 mesenchyme of lamprey and gnathostomes (derived from the 1st stream of CNCCs), *otx* is uniquely transcribed prior to CNCC migration (Acampora et al., 1995; Kuratani et al., 1998b; Tomsa and Langeland, 1999; Zhang et al., 2014). Thus an '*otx* and *hox*' CNCC patterning scheme has been suggested to encompass all PAs, with *otx* in first stream CNCCs being implicated as a possible avenue for jaw evolution (Kuratani, 2004). Importantly, the role of *otx* in this tissue seems to strongly differ from *hox* in that this expression is not maintained throughout most 1st stream migratory or post-migratory skeletogenic CNCCs. At the pharyngula stages when we find persisting *hox* expression in PAs 2+, the majority of PA1 and pre-oral mesenchyme is *otx*-negative. In all vertebrates addressed to date, *otx* is found in a small subset of migratory and post-migratory CNCCs (Kudoh et al., 2001; Matsuo et al., 1995; Tomsa and Langeland, 1999; Zhang et al., 2014), which usually appears to correspond to CNCC-derived ganglia. To our knowledge, the function of *otx* in craniofacial skeletogenesis has been addressed only in the mouse (Matsuo et al., 1995), however, the knockout method used does not separate this gene's potential role in general CNCC specification from a possible later role in the identity of anterior ectomesenchyme, unlike some temporally-inducible *hox* experiments performed in *Xenopus* and mouse (Pasqualetti et al., 2000; Santagati et al., 2005). Thus PA1 truncations and deformities seen in *otx*-null or heterozygous mice might simply arise from the improper specification of the neural plate and neural plate border (including premigratory CNCCs) at the position of the anterior midbrain rather than a more specific perturbance in the skeletal identity of PA1 and pre-oral CNCCs. Furthermore, no homeotic transformation or suggestion of a different identity was seen in these mutants; only a truncation or absence of certain skeletal elements. Thus, while *otx* might have a deeply conserved, early role in 1st stream CNCC differentiation via specification of anterior neural tissues, this function appears to be more general and occurs distinctly earlier than *hox* gene function in post-migratory skeletogenic CNCCs.

Differences in transcription factor expression between cyclostomes and gnathostomes in the nascent head skeleton

While there are many similarities in gene expression in the head skeletal precursors of lamprey and gnathostomes, there are also some clear differences. Lacking an appropriate outgroup for rooting these comparisons, these differences could reflect gain, modification, or loss in either the gnathostome or cyclostome lineages. Thus, despite gross similarity between lamprey larvae and fossils of early vertebrate relatives, it should not be automatically assumed that a trait is ancestral if it is present in lamprey but not gnathostomes.

The *dlx* genes, though nested in lampreys and gnathostomes, show some definite differences in expression between these two major groups. The presence of a *dlx*-negative ventralmost domain in

gnathostomes (Compagnucci et al., 2013; Jeong et al., 2008) is not mirrored in lampreys (Cerny et al., 2010), where *dlxB* is expressed throughout the ventralmost extent of ectomesenchyme in the head (below the endostyle). This might somehow relate to the differences in morphology of these ventral PAs between both lineages, where lampreys have no distinct ventral cartilage elements but gnathostomes have medial basibranchials and a basihyal that are composed of CNCCs derived from both the left and right sides of the head. Outside the PAs, an interesting and previously noted difference between gnathostomes and cyclostomes is the presence of *dlx*-positive pre-oral mesenchyme in the latter (Kuratani et al., 2013; Shigetani et al., 2002). Mesenchyme in this region of the lamprey head has an expression profile more similar to the oro-maxillary region in gnathostomes (dorsal/proximal PA1). This difference between lamprey and gnathostomes has been hypothesized to stem from a heterotopic shift in the signals received by the CNCCs in the oral region (Kuratani et al., 2013; Shigetani et al., 2002), which could have changed the locations at which CNCCs were induced to have a *dlx*-positive expression profile. However with no appropriate outgroup, whether this pre-oral dlx expression is ancestral and lost in gnathostomes or a derived expression gain in lampreys will remain difficult to address.

goosecoid (*gsc*) genes exhibit unique expression profiles across vertebrates (Cerny et al., 2010; Gaunt et al., 1993; Schultemerker et al., 1994; Square et al., 2015a), but are expressed within the same broad populations of ectomesenchyme. In the early pharynx, this gene's expression is relegated to mainly to PA1, PA2, and a subset of non-PA mesenchyme just ventral to the brain in all lineages. However even within gnathostomes, *gsc* expression is plastic, exhibiting clear differences between different major groups. For example, zebrafish *gsc* transcription is observed in dorsal PA1 (Schultemerker et al., 1994), but this domain is not reflected in any other gnathostome addressed to date. Generally speaking, in gnathostomes this gene is expressed in the ventral PA1 and PA2, and also in mesenchyme of the future ventral chondrocranium (beneath the eye, near the palate). Lamprey also expresses *gsc* intricately in PA1, PA2, and the upper lip, although this gene is essentially absent from the ventralmost domain, save a small spot of expression in the distalmost lower lip (Cerny et al., 2010). So, overall, it seems that *gsc* has an ancient role in specializing the regional expression profiles of at least PA1, PA2, and some subset of dorsal or pre-oral non-PA mesenchyme in vertebrates, though the precise expression pattern of *gsc* in the vertebrate common ancestor is unclear.

Lamprey expresses alx in the dorsal and ventral portions of most PAs (Cattell et al., 2011), similar to prrx expression (Fig. 3, Fig. S1). This set of alx expression domains is not seen in any extant gnathostome addressed to date, which instead express alx genes only in the ventral PAs and chondrocranium adjacent to the dorsal PAs (Beverdam and Meijlink, 2001; Cattell et al., 2011; Compagnucci et al., 2013; Dee et al., 2013; McGonnell et al., 2011; Square et al., 2015a). Interestingly, *alx* and *prrx* are related genes; thus if old gene regulatory machinery was coopted for these pharyngeal

expression domains, the alx-negative dorsal PAs in gnathostomes might reflect a loss of expression in the ancestral gnathostome alx gene, with prrx having retained this expression. Given that some expression is conserved between alx and prrx, this notion could be tested by work on alx and prrx gene regulation in both gnathostomes and cyclostomes.

Other genes that are expressed dissimilarly in lamprey and gnathostomes include *nkx3.2* (*bapx*) (Cerny et al., 2010; Nichols et al., 2013), *barx* (Cerny et al., 2010; Nichols et al., 2013), *mef2* (Jandzik et al., 2014b), and *emx* (Fig. S1). These genes offer a range of differences: some are found in a mesenchyme with different skeletogenic potentials (*barx*), CNCCs in dissimilar locations (*nkx3.2*), or absent from CNCCs in lamprey (*emx* and *mef2*). In the case of lamprey *emxA*, it would appear that the dorsoventral cue that positions the expression in the pharynx might be conserved with *emx2* (an "intermediate" PA patterning gene in gnathostome CNCCs), but the lamprey ortholog is being deployed in epithelial ectoderm and endoderm instead of the future skeleton (Fig. S1). Similarly, lamprey *barx* is found in the innermost, CNCC-derived mesenchyme of the PAs (nearest the cavity of the pharynx), but this tissue does not give rise to larval skeletal elements. With some functional assays, these curious differences in expression might offer clues as to how these genes are activated, and thus allow us to understand how cyclostomes and gnathostomes become so different both morphologically and histologically during development.

The early vertebrate head: general characteristics and evolutionary trends

Fossilized early vertebrate relatives such as *Haikouella* (Mallatt and Chen, 2003), *Haikouichthys* (Shu et al., 1999; Shu et al., 2003), *Myllokunmingia* (Shu et al., 1999), and *Metaspriggina* (Morris and Caron, 2014) arose during the early to mid-Cambrian, and are generally assumed to predate the most recent common ancestor of all living vertebrates. While phylogenies based on available character states usually place these fossil animals at the base of all vertebrates, there is some uncertainty: occasionally these fossils show a greater affinity for gnathostomes, lampreys, or hagfish, depending on tree reconstruction parameters and the traits chosen for analysis (Mallatt and Chen, 2003; Morris and Caron, 2014; Sansom et al., 2010; Shu et al., 1999). Regardless of disagreements over their exact phylogenetic position(Donoghue and Purnell, 2009; Gess et al., 2006; Morris and Caron, 2014; Sansom et al., 2010; Shu et al., 1999; Shu et al., 2003), these fossils inform us as to the timing of evolutionary events, confirming that a given character state had arisen by a given epoch.

Like modern lamprey larvae, all of the aforementioned fossil vertebrate relatives appear to have possessed a cartilaginous endoskeleton composed mainly of simple pharyngeal cartilage rods (within the PAs), with some oral and/or cranial chondroid tissue (Mallatt and Chen, 2003; Morris and Caron, 2014; Shu et al., 1999; Shu et al., 2003). Dissimilar to this condition, hagfishes display extremely derived

reduction of the pharyngeal skeleton, having apparently lost all but two of their branchial arches despite still possessing a high number of gill slits (between 6 and 12 pairs). Despite their reduced number, hagfish branchial arch cartilages are also fused to the rest of their head skeleton like in lampreys, in a basket-like condition (albeit a small basket) (Oisi et al., 2013). Notably, the fossilized cartilage bars found in specimens such as *Metaspriggina* (Morris and Caron, 2014) and *Haikouichthys* (Shu et al., 2003) do not appear to be connected by dorsal and ventral horizontal bars, nor are they juxtaposed at their apices, suggesting the fused branchial basket might be a derived character of cyclostomes. Like most modern gnathostomes, the recently reported Cambrian fossil *Metaspriggina* (Morris and Caron, 2014) also appears to have distinct dorsal and ventral PA skeletal elements. If *Metaspriggina* diverged before the gnathostome and cyclostome lineages split, this would raise the possibility that cyclostomes acquired single vertical rods in each PA by secondarily fusing ancestrally separate dorsal and ventral skeletal elements. Interestingly, modern anurans possess both of these cyclostome-like characters in their branchial arches: these lack any dorsoventral segregation, and are also fused at their apices in a basket-like condition(Rose, 2014). Together, this could explain why lamprey has gnathostome-like patterning of its skeletal primordia, while also forming contiguous rods in each PA: despite reverting to a fused morphology, the patterning scheme is mostly retained (Cerny et al., 2010) (as is also seen in *Xenopus* branchial arches (Square et al., 2015a)). An alternative hypothesis is that cyclostomes and gnathostomes diverged earlier than presumed, and *Metaspriggina* split from the lineage leading to gnathostomes after the evolution of distinct dorsal and ventral skeletal elements in that lineage. It is also possible that *Metaspriggina* and gnathostomes independently acquired separate dorsal and ventral elements. Regardless of the precise scenario, the presence of gene expression reminiscent of the gnathostome condition in the developing lamprey head skeleton and the bipartite pharyngeal skeleton of *Metaspriggina* strongly support the idea that the developmental mechanisms needed to distinguish the dorsal from ventral pharyngeal skeleton precursors predate jaws by 100 million years or more.

Looking above the PAs, one general trend in vertebrate head skeletal evolution was the expansion of the chondrocranium (forming the brain case, or neurocranium), which does not seem to be prominent in the earliest vertebrate fossils, if present at all (Mallatt and Chen, 2003; Morris and Caron, 2014; Shu et al., 1999; Shu et al., 2003). In modern gnathostomes, these more dorsal structures are partially or sometimes largely derived from mesoderm, unlike the facial skeleton and PAs (the viscerocranium) which are mainly derived from CNCCs (Kague et al., 2012; Piekarski et al., 2014). While cyclostomes have small cartilages near the brain, including “trabeculae”, these do not have any clear structural homology to any specific gnathostome cartilages, though it is suggested by their general position (Kuratani et al., 2016). Nevertheless, a prominent and three-dimensionally complex ethmoid/trabecular skeleton and the presence of an extensive braincase are clearly derived characters of the gnathostome group. This means

that the head skeleton of the vertebrate common ancestor was composed of mainly PAs and some oral-associated mesenchyme, the majority of which were probably derived from CNCCs, similar to the condition in modern lampreys.

Cyclostomes and the earliest fossil vertebrate relatives lack cranial bones (or any bones for that matter), which first appear in the fossil record in jawless vertebrates of the subclass Pteraspidomorphi (Janvier, 1996) (Fig. 1). This means that the embryonic cartilaginous head endoskeleton in bony vertebrates (osteichthyans) is more similar compositionally to the earliest vertebrate head skeletons, rather than the adult bony fish head skeleton, which comprises mainly dermal bones with some ossified elements of the endoskeleton. However, it is interesting to note that in modern osteichthyans, many bones of the head develop from the same pool of CNCCs that give rise to the larval cartilaginous head skeleton (Hirasawa and Kuratani, 2015). Importantly, some key aspects of genetic patterning that occur within the common progenitor pool of CNCCs affect the initial larval head skeleton as well as the later ossified skeleton, such as the *dlx*(Depew et al., 2002) and *hox* (Minoux et al., 2009) codes. So despite the striking histological and morphological differences between a given osteichthyan's larval and adult head skeletons, these two complex structures are strongly tied to each other developmentally. Thus, the patterning schemes described here presumably have effects on all modern head skeletons, both larval and adult.

Differences in vertebrate head skeletons

What is the output of this extensive patterning via transcription factor expression? Given that these modules do not correspond to individual skeletal elements in gnathostomes, and head skeleton elements are histologically contiguous in cyclostomes, the relationship between this patterning scheme and eventual skeletal morphology is quite abstracted, i.e. there is poor correspondence in shape and size between individual gene regulatory modules (like those shown in Fig. 3) and the delineation of cartilage and/or bony elements that will arise from a given region later in development. All modern vertebrate head skeletons are very complex, displaying intricate three dimensional structure within skeletal elements, and also with respect to the juxtaposition of these elements. In order to generate such any sort of three-dimensional variety in a given skeletal region, this requires that cells within a given primordium divide at a different rate or on a different axis than cells in other primordia, or in other parts of the same primordium (Kimmel et al., 1998). While some studies are beginning to understand how cell division is controlled to these ends in the head skeleton (Le Pabic et al., 2014), much work is still needed to place these events within the context of vertebrate transcription factor patterning in the head. In gnathostomes, mutations in the transcription factors that pattern head skeleton precursors cause the loss of joints as well as changes in the shape of skeletal elements, showing that these patterning genes drive some part of both

the overall shape and composition of the head skeleton. With modern developmental genetic tools, the processes linking early patterning to later cell division and histological composition will become more evident.

One general difference between cyclostomes and gnathostomes relates to the high level of dorsoventral symmetry in transcription factor expression within posterior lamprey PAs as compared to gnathostomes. Interestingly, in lamprey, this symmetry starts at the level of gene expression (Fig. 3), and is later reflected in their highly symmetrical pharyngeal basket. Gnathostomes instead have more polarized PAs with respect to both gene expression and eventual morphology. This is at least in part dictated by the ventral secretion of Edn1 protein in gnathostomes, which works to specify the ventral and intermediate domains. Conversely, no sea lamprey *edn* ligand is restricted to all or most ventral PAs; instead, *ednA* and *ednE* transcripts are found in a more centralized pattern in the PAs (Square *et al.*, 2016), removing the possibility that any *edn* ligands in sea lamprey work as ventral specifiers. We speculate that the level of symmetry in these gene expression domains might somehow contribute to the symmetry of the lamprey pharyngeal basket by driving a similar identity, and thus morphology of these domains in lamprey (namely the dark blue *alx, dlx, msx, prrx* domains in Fig. 3). More functional studies in cyclostomes are needed to link head skeleton gene expression and morphology, but in gnathostomes the *dlx* genes alone have been shown to drive the general identity of entire regions of PAs, which supports this notion given the expression patterns known across vertebrates (Fig. 4).

Another obvious aspect of derived morphology is the specialization of the oral skeleton (pre-oral mesenchyme, PA1, and PA2) in all modern vertebrate lineages. In lamprey, oral skeletogenic mesenchyme gives rise to the oral hood, velum, and elongated lower lip, whereas in gnathostomes it forms the mandible, maxilla, palate and nasal region. How these differentiated structures are related to each other has been a topic of debate for more than a century. Despite striking differences in final morphology, developmental gene expression has the potential to show how the precursor cell populations that generate disparate skeletal structures might be related, at least on the level of transcriptional identity. The lamprey upper lip is a skeletal element positioned just rostral to the mouth, similar to the maxillary/nasal elements of the gnathostome head skeleton. Despite this spatial similarity, nearly the entire lamprey upper lip has a transcriptional identity similar to the dorsalmost anterior PA1 of gnathostomes, while most of the lower lip has the transcriptional identity of the gnathostome ventral PA1. Interestingly, the lower lip demonstrates an anterior expansion perpendicular to the dorsoventral axis of PA1, rather than an extension ventrally/distally with a bent axis.

The head skeletons of both gnathostomes and cyclostomes are composed of multiple morphologically distinct skeletal elements (though these elements are connected in cyclostomes), and several histologically distinct skeletal tissue types (Cattell et al., 2011). These include joint tissue in

gnathostomes, various types of mucocartilage in lampreys, and soft and rigid collagen-containing, proteoglycan-rich cellular cartilage in both groups. Future work in both gnathostomes and cyclostomes will help elucidate if the conserved targets of the genes that confer these local transcriptional identities include structural genes such as collagens, lecticans, and glypicans.

Based on work in gnathostomes, it is likely that patterning via transcription factor expression in stem vertebrates provided a framework for position-specific morphogenesis and tissue differentiation, though precisely how this occurred is unclear. Given the conserved aspects of early head skeleton development, it is likely changes developmentally downstream of genes such as *hox, alx, dlx, hand, msx,* and *prrx* were major players in the generation of the different major vertebrate forms. These genes' expression patterns seem to have evolved prior to the most recent common ancestor of all living vertebrates – at least in large part. However, the nuanced differences we do see in these genes might contribute to large differences amongst vertebrates in the shape or composition of some regions of the head skeleton, such as less symmetrical *alx* in the gnathostome PAs, or *dlx* expression in the lamprey upper lip.

Conclusion

Based on detailed comparisons of living jawed and jawless vertebrates, it is likely that the head skeleton of the most recent common ancestor of all living vertebrates formed mainly from CNCCs. These cells migrated as three molecularly distinct populations, and expressed a core set of chondrogenic transcription factors. After migration into the oropharyngeal region, those cells activated a set of transcription factors which conferred regional identities upon different skeletal precursor subpopulations along multiple axes, including a set of at least 10 transcription factors (*hox1*, *hox2*, *hox3*, *hox4*, *dlx2/3/5A*, *dlx2/3/5B*, *alx*, *hand*, *msx*, and *prrx*). Similarities in this head skeleton patterning system across jawed and jawless vertebrates likely reflect homology of many of these precursor subpopulations, but it is important to recognize that these regulatory cassettes could be deployed in new regions of the head, in which case it would be more appropriate to consider them deeply homologous (Shubin et al., 2009). It is unclear, however, if the specific skeletal structures derived from these developmental units should be considered homologous *sensu stricto*, even if they are similarly juxtaposed with other elements and tissues in the head. By further dissecting the gene expression and function of different genes in different lineages, we will continue to build our understanding of what all heads share, how they are developmentally constrained, and also how genetic changes underlie changes in morphology.

CHAPTER III: a gene expression map of the larval *Xenopus laevis* head reveals developmental changes underlying the evolution of new skeletal elements

Introduction

The vertebrate head skeleton protects and supports the anterior sense organs, brain, and the feeding and breathing structures of the mouth and pharynx. Despite performing these basic functions in all vertebrates, the morphology of the vertebrate head skeleton is highly plastic, with the number, size, and shape of its components varying between groups. This remarkable capacity to evolve in response to new selective pressures was likely key to vertebrate success. Understanding the constrained and evolutionarily labile aspects of vertebrate head skeleton development is thus essential to understanding the mechanistic bases of vertebrate diversification.

In all modern vertebrates the bulk of the embryonic head skeleton is derived from cranial neural crest cells (CNC), with a smaller contribution from mesoderm-derived mesenchyme (Kague et al., 2012; Piekarski et al., 2014). CNC are specified at the neural border and migrate ventrally into the head and pharynx. After CNC migration, subpopulations of CNC and mesodermal mesenchyme acquire distinct molecular identities as they activate particular combinations of transcription factors. These factors include members of the *hox* (Creuzet et al., 2002), *dlx* (Depew et al., 2002; Talbot et al., 2010), *msx* (Antonopoulou et al., 2004), *hand* (Yanagisawa et al., 2003), *nkx3.2* (*bapx*) (Miller et al., 2003), *emx* (Compagnucci et al., 2013), *alx* (Beverdam et al., 2001), *prrx* (ten Berge et al., 1998), *tbx2/3* (Compagnucci et al., 2013), *satb2* (Fish et al., 2011; Sheehan-Rooney et al., 2010), *gsc* (Gaunt et al., 1993; Schultemerker et al., 1994), *mef2* (Verzi et al., 2007), and *pou3* (Hauptmann and Gerster, 2000; Jeong et al., 2008) families. Interfering with the function of most of these genes in mouse and/or zebrafish has specific effects on skeletal element morphology, including partial homeotic transformations. This suggests these factors are developmentally upstream of gene programs controlling skeletal element morphogenesis.

While functional perturbations and gene expression data have provided mechanistic depth to our understanding of head skeleton development in model vertebrates, how these processes are modified during evolution to generate new morphologies is poorly understood. Recent work suggests that differences in the intercellular signals emanating from cranial epithelia can explain differences in skeletal element size and shape between closely related species (Abzhanov and Tabin, 2004; Mallarino et al.,

2012). Whether similar changes also underlie the differences in head skeleton morphology seen at larger evolutionary distances is unknown. Provocatively, inter-specific CNC transplantations between ducks and quails have shown that CNC will generate donor-specific skeletal morphology regardless of the host's intercellular signaling environment (Fish et al., 2014; Schneider and Helms, 2003). This indicates that, rather than being determined solely by extrinsic signals, much of vertebrate head skeleton morphology is hardwired in the morphogenetic programs operating in CNC. Alterations to these programs, and/or the transcription factors that control them, are thus strong candidates for the types of changes driving head skeleton diversification.

A first step in understanding the mechanistic bases of head skeleton evolution is identifying changes in head development corresponding to lineage-specific skeletal novelties. Such correlations can then be validated by gain- and loss-of-function genetic manipulations in species displaying the ancestral and derived conditions. In addition to being a tractable developmental model system, *Xenopus laevis* is a representative of an ancient and highly specialized tetrapod group, the anurans. Like other anurans, the *X. laevis* head skeleton possesses a unique combination of ancestral gnathostome features and anuran-specific alterations in skeletal element number, size, and shape. As in basal gnathostomes, *X. laevis* forms its primary jaw joint from the first pharyngeal arch (PA1). However, unlike teleosts or mammals, the *X. laevis* second arch (PA2) skeleton resembles that of early tetrapods, with a dorsal element modified to form the columella (stapes), and a large ventral element, the ceratohyal. *X. laevis* larvae also possess two anuran-specific cartilages around the mouth, the suprarostral and infrarostral. The infrarostral cartilage articulates with the lower jaw at a novel paired joint called the intramandibular joint (reviewed by (Svensson and Haas, 2005). The suprarostral is fused to the trabecular cartilage in *X. laevis* and its close relatives, but articulates with the trabecular cartilage at a movable joint in other frogs (Pugener et al., 2003; Trueb and Hanken, 1992; Zhang et al., 2013). Finally, the tadpoles of *X. laevis* and most other anurans display a posterior pharyngeal arch skeleton fused to form a branchial basket (Pugener et al., 2003).

Here we present a detailed gene expression map of the head mesenchyme in *X. laevis* larvae, focusing on 13 families of transcription factors with known functions in vertebrate head skeleton development. We then compare this map to homologous gene expression in zebrafish, mouse, and shark embryos to deduce the conserved and evolutionarily plastic aspects of vertebrate head skeleton development. While we observed broad conservation between *X. laevis* and other gnathostomes, we also identified several divergent aspects that correlate to lineage-specific novelties. Soon after migration of CNC into the pharynx, we noted a conspicuous change in the transcription factor code of PA2, with *dlx1/2* expression becoming reduced and *emx2* expression expanding dorsally. As *dlx1/2* is necessary for chondrogenic fate, we propose that this shift may underlie the development of the reduced dorsal PA2

cartilage typical of modern anamniote tetrapods. Around the mouth and in the posterior pharynx, we noted highly conserved transcription factor expression between *X. laevis* and other vertebrates at early larval stages. However, shortly before skeletal differentiation, we observed a shift in the expression of the joint inhibitor *barx1*, and new expression of the joint marker *gdf5* in the forming mandible. This suggests that the anuran-specific infrarostral cartilages evolved by the partitioning of Meckel's cartilage with a new paired joint. Taken together, our comparisons illustrate that the early patterning of head skeleton precursors is largely conserved across vertebrates, with subtle changes to this pre-pattern corresponding to changes in skeletal element size. We further propose that one way new skeletal elements evolve is by changes in the expression of genes regulating later developmental processes, like tissue differentiation.

Methods

Gene nomenclature and riboprobe synthesis

Per XenBase conventions (Bowes et al., 2008) *X. laevis*-specific duplicates are distinguished by the suffixes *-a* and *-b*. However, for simplicity, we sometimes refer to two *X. laevis* duplicates by their generic gene name (e.g. *dlx3-a* and *dlx3-b* are sometimes collectively called "*dlx3*").

Xenopus laevis genes were identified using GenBank annotated sequences or BLAST searches of *X. laevis* ESTs and/or genome assembly. Fragments of these genes were amplified using the primers listed in Tab. S1 (Square et al., 2015a) and subcloned. Digoxygenin-labelled antisense *in situ* hybridization riboprobes were synthesized using linearized plasmid or PCR product as template and SP6, T7, or T3 RNA Polymerase (Promega).

In situ hybridization

Frog husbandry, staging, and *in situ* hybridizations (ISH) were carried out as described previously (Cerny et al., 2010), with few modifications. Proteinase K treatment was adjusted to 100µg/ml for 1-5 min for st. 27-37 larvae, and up to 20 min for st. 41-45 larvae. Hybridization was performed at 60°C. For double ISH, two digoxygenin-labelled riboprobes were added to a single hybridization solution, and both were developed simultaneously. For signal development, 0.05% NBT + 0.35% BCIP (Roche) was used as a substrate. Prehybridization and hybridization steps were performed in SSC pH 7.0; 5 mM EDTA, pH 8.0; 500 µg/mL tRNA; 0.1% Tween-20; 0.1% CHAPS; and 50 µg/mL heparin.

Sectioning

After photography, whole-mount hybridized *X. laevis* embryos were washed for 30 minutes in 15% sucrose at room temperature, transferred to 7.5% gelatin (90-110 Bloom, Sigma) in 15% sucrose, kept shaking at 37°C for 4-6 hours, and finally transferred to 20% gelatin in 15% sucrose and shaken

overnight at 37°C. Before sectioning the larvae were embedded in fresh 20% gelatin in 15% sucrose in silicone molds and frozen for at least 1h at -70°C. The blocks were sectioned to 18µm using a Leica CM3000 cryostat. The sections were mounted on FisherBrand Superfrost Plus microscope slides and dried overnight at 37°C. Gelatin was removed by washing in 3% gelatin/38% ethanol for 45 seconds at 37°C followed by two rinses in distilled water. Sections were counterstained with Nuclear fast red (Vector Laboratories) for 45 seconds, rinsed twice in distilled water, dried overnight and cover-slipped using DePeX (VWR Int.) mounting medium.

Imaging

Whole-mount *in situ* hybridized *X. laevis* embryos and larvae were photographed using a Carl Zeiss Axiocam MRc5, Carl Zeiss Discovery V8 dissecting microscope, and Axiovision 4.6 software. Sections were photographed using a Carl Zeiss Imager A2 compound microscope.

Results

Expression of *dlx* paralogs and *mef2c*

We isolated fragments of all known *X. laevis dlx* paralogs, and compared their expression in post-migratory CNC by *in situ* hybridization from st. 33-38 (Fig. 1; S1A-BB). As in other jawed vertebrates, *X. laevis dlx2* is the most broadly expressed *dlx* paralog, marking migratory and post-migratory CNC along the full DV extent of the pharynx. This expression is apparent in every PA throughout early larval development, with the exception of PA2, where *dlx2* transcripts are largely excluded from the dorsal aspect (Fig. 1C). The closely linked gene *dlx1* is transcribed in a pattern similar to *dlx2*, though its expression is weaker in the dorsal-most and ventral-most aspects of the pharynx (Fig. 1C, D). *dlx4* displays the most restricted pattern of all *dlx* paralogs and is limited to a sharply defined domain in the middle portion of all PAs (Fig. 1G, H). *dlx3-a* also strongly marks this intermediate domain, though its expression domain is broader (Fig. 1E, F), with weak expression extending slightly ventrally and dorsally. The *X. laevis* specific duplicate *dlx3-b* is expressed in a pattern identical to *dlx3-a* and is not shown. *dlx5* and *dlx6-a* are transcribed in the intermediate and ventral PAs (Fig. 1I, J; K, L), though *dlx5* expression extends slightly more ventrally. We were unable to detect expression of the *X. laevis dlx6* duplicate *dlx6-b*.

In mouse, proper expression of *dlx3, dlx4, dlx5,* and *dlx6* is dependent on the activity of the transcription factor *mef2c* (Miller et al., 2007; Verzi et al., 2007). We isolated a fragment of *X. laevis mef2c* and analyzed its expression before and during *dlx* gene expression. Consistent with its role in mouse, we observed expression of *X. laevis mef2c* throughout migrating CNC and in post-migratory CNC

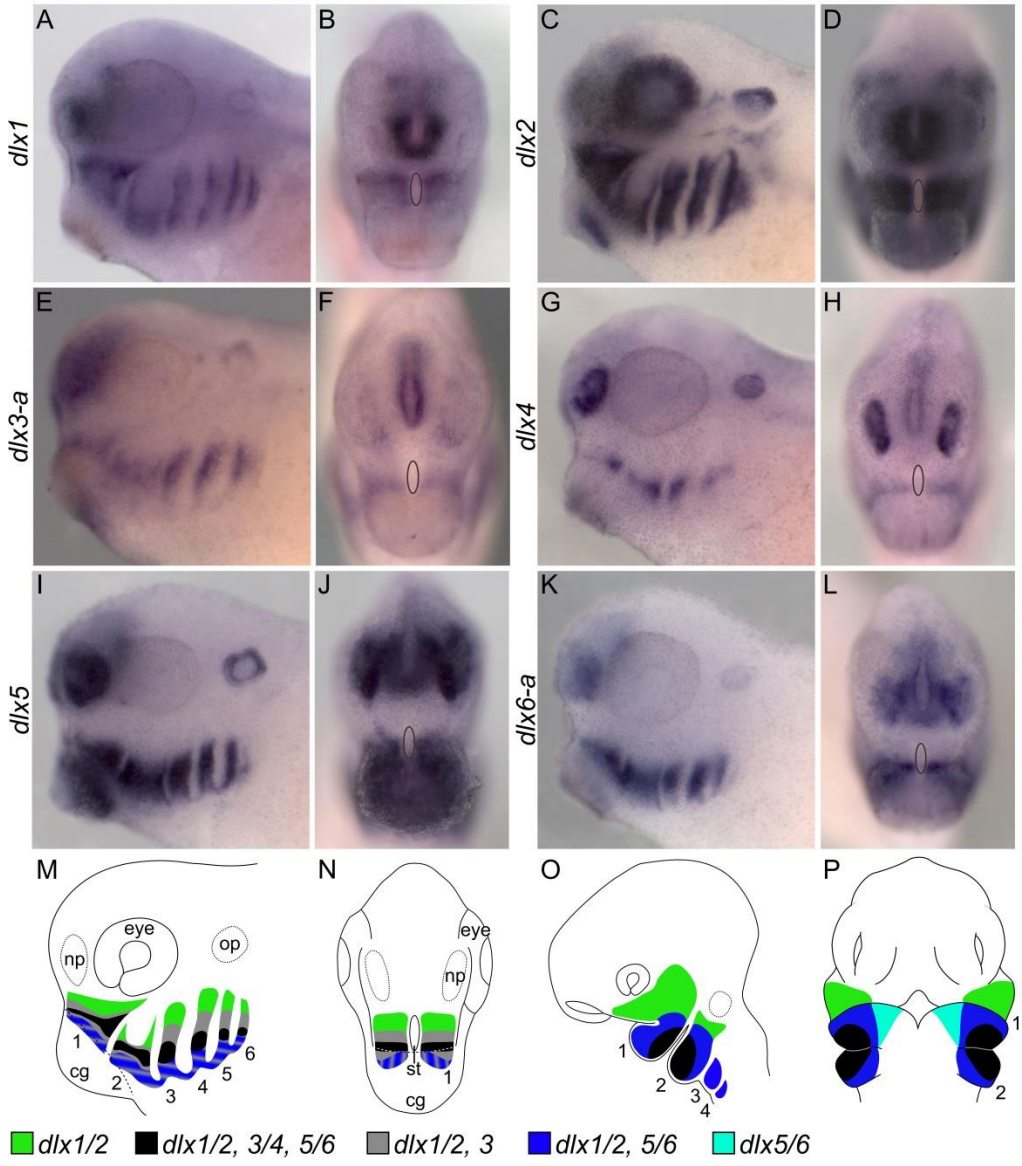

Fig. 1. *dlx* expression in *X. laevis* at st. 33/34. Lateral views with anterior to left (A, C, E, G, I, K) and anterior views (B, D, F, H, J, L) of larvae stained by ISH for (A and B) *dlx1*, (C and D) *dlx2*, (E and F) *dlx3-a*, (G and H) *dlx4*, (I and J) *dlx5*, and (K and L) *dlx6-a*. The position of the stomodeum is indicated by an oval in all anterior views. (M-P) Schematic of *dlx* expression in the *X. laevis* (M and N) and mouse (O and P) cranial neural crest cells. See figure for key. Lateral views in (M) and (O), anterior views in (N) and (P). For *X. laevis*, the ventral blue domain is shown with grey stripes to indicate the expression of *dlx3* in this region weakly. For the mouse, data were used from Jeong et al., 2008. cg, cement gland; np, nasal placode; op, otic placode; st, stomodeum. Each PA is labeled by number.

in a pattern reminiscent of *dlx3-a* (Fig. S1CC-HH; compare Fig. S1K to DD). *mef2c* expression also marks the intersomitic regions, and portions of the head mesoderm.

Expression of *emx2* and *nkx3* genes in the intermediate pharynx

In *X. laevis*, as in zebrafish, CNC occupying approximately the middle third of the pharyngeal arches expresses members of all 6 *dlx* paralogy groups (Talbot et al., 2010). The transcription factor *emx2* has also been shown to mark this "intermediate domain" in zebrafish, mouse, and shark (Compagnucci et al., 2013; Thisse et al., 2004). We observed intermediate domain expression of *X. laevis emx2* in all PAs from st. 33-38 (Fig. 2A, B; S2A-E). This expression was reminiscent of *dlx4*, except in PA2, where *emx2* expression extends into the dorsal domain (arrow in Fig. 2A).

In all gnathostomes examined to date, the intermediate domain of PA1 expresses *nkx3.2* (*bapx*) homologs (Compagnucci et al., 2013; Miller et al., 2003; Tucker et al., 2004; Wilson and Tucker, 2004). We isolated fragments of two *X. laevis nkx3.2* duplicates, *nkx3-2-a* and *nkx3-2-b*, as well as the closely related genes *nkx3-3* (*zax*) (Newman and Krieg, 1999) and *nkx3-1-a*, and visualized their expression from st. 33-38. *nkx3-2-a* (data not shown), *nkx3-2-b* (Fig. 2C, D; S2F-I, S5I, J), and *nkx3-3* (Fig. 2E, F; S2J-N, S5H) were all transcribed in CNC within the intermediate domain of PA1, though in slightly different patterns. Specifically, *nkx3-3* mRNA marks a cluster of CNC in the posterior intermediate aspect of PA1 (arrow in Fig. 2E), while *nkx3-2-a* (data not shown) and *nkx3-2-b* are expressed more broadly in the ventro-medial CNC of PA1 (arrow in Fig. 2C), overlapping with *nkx3-3* posteriorly. At st. 41 *nkx3-2-b* transcripts were also detected in the ventral midlines of PA1 and PA2, as has been reported for zebrafish *nkx3.2* (white arrowhead and arrow in Fig. 6E, respectively). (Miller et al., 2003; Schwend and Ahlgren, 2009). In addition to CNC, both *nkx3-2* genes and *nkx3-3* are expressed in pharyngeal endoderm, with *nkx3-3* transcripts also marking stripes of ectoderm in the posterior pharynx, and flanking the stomodeum (white arrowheads in Fig. 2F). *nkx3-1-a* was never detected in CNC from st. 24-45, but strongly marked the cement gland and paraxial mesoderm (data not shown). The expression of the *X. laevis* duplicate *nkx3-1-b* was not examined.

Expression of *hand, msx, satb2, tbx2/3* and *gsc* in the ventral pharynx

hand, *msx*, *satb2*, and *tbx2/3* transcripts mark CNC throughout the ventral pharynx in mouse and zebrafish (Charite et al., 2001; Firulli, 2003; Fish et al., 2011; Hohimer et al., 1993; Mesbah et al., 2008; Miller et al., 2003; Ribeiro et al., 2007; Thomas et al., 1998). *gsc* also marks ventral CNC, though this expression is largely restricted to the anterior pharyngeal arches (Gaunt et al., 1993; Miller et al., 2003; Schultemerker et al., 1994). We isolated fragments of *X. laevis hand1, hand2-a, msx1-b, msx2, satb2, gsc-b, tbx2-a*, and *tbx3-a*. As in other jawed vertebrates, *hand1*, *hand2-b*, *msx1-b*, *msx2*, and *satb2* transcripts

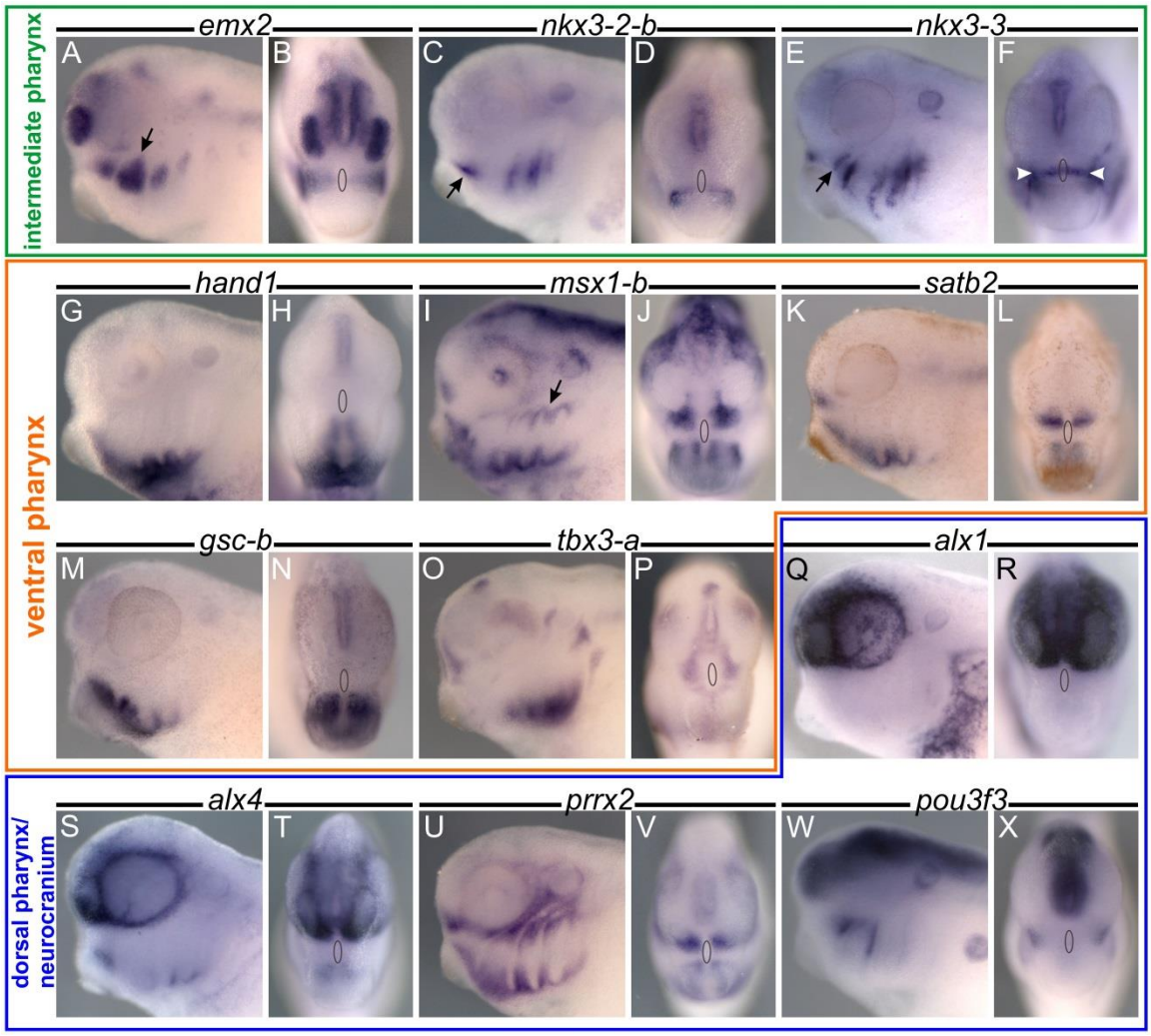

Fig. 2. Transcription factor expression in the nascent *X. laevis* head skeleton at st. 33/34. Lateral views with anterior to left (A, C, E, G, I, K, M, O, Q, S, U, W), and anterior views (B, D, F, H, J, L, N, P, R, T, V, X) of st. 33/34 larvae stained by ISH. The position of the stomodeum is indicated by an oval in all anterior views. (A and B) *emx2* is expressed centrally in PAs. (C and D) *nkx3-2-b* transcripts are found both in ventro-medial PA1 mesenchyme (arrow), and endoderm in more posterior PAs. (E and F) *nkx3-3* is expressed in the region of the nascent jaw joint (arrow), as well as ectoderm around the stomodeum and more posterior arches, and endorderm in the pharyngeal pouches. (G and H) *hand1* transcripts are found in ventral CNC-derived mesenchyme. (I and J) *msx1-b* expression in the neurocranium, PAs, and FNP. Arrow in (I) indicates dorsal PA CNC expression. (K and L) *satb2* is expressed contiguously throughout ventral PAs, and at the junction of the antero-dorsal PA1 and FNP. (M and N) *gsc-b* is detected ventrally in the anteriormost 4 PAs. (O and P) *tbx3-a* transcripts are found in anterior PA1/frontonasal mesenchyme, as well as both CNC and endoderm of PAs 1, 3, 4, 5, and 6 (see Fig. S3Q). (Q and R) *alx1* is expressed throughout the future neurocranium and the antero-dorsal PA1. (S and T) *alx4* is expressed throughout the future neurocranium and antero-dorsal PA1, as well as the ventral PAs. (U and V) *prrx1* is detected in the neurocranium and FNP, as well as the dorsal and ventral PAs. (W and X) *pou3f3* is expressed in a domain branching PA1 and PA2, and strongly in the developing brain.

were detected in the ventral portion of all pharyngeal arches (Fig. 2G-L; S2O-II; S3A-F). *tbx3-a* is also expressed broadly in the ventral pharynx, though it is excluded from PA2, and only limited expression is seen in the anterior aspect of PA1 (Fig. 2O, P; S3M-Q). *gsc-b* expression was observed in ventral PA1 and PA2, as in mouse and zebrafish, with additional ventral expression noted in PA3 and PA4 at st. 33/34. This posterior expression was lost by st. 35/36 (Fig. 2M, N; S3G-L). *tbx2-a* expression was not restricted to the ventral domain, but marked CNC at all dorso-ventral levels (Fig. S3R-U).

Expression of *alx, prrx, msx, gsc, pou3f3, satb2,* and *tbx3a* in the dorsal pharynx, maxillary region, and neurocranium

The *alx* genes, *alx1, alx3,* and *alx4* mark the presumptive neurocranium in mouse, chick, and zebrafish embryos (Beverdam et al., 2001; Beverdam and Meijlink, 2001; Dee et al., 2013; McGonnell et al., 2011). *X. laevis* lacks *alx3*, and a recent report described *X. laevis alx1* and *alx4* expression in patches of mesenchyme around the eye and in the frontonasal process (FNP) at st. 24 and 35 (McGonnell et al., 2011). We observed much broader expression of *alx1* and *alx4* in the FNP, maxillary region, and in mesenchyme around the eye and overlying the brain between st. 33-38 (Fig. 2Q-T; S4A-H). Additional expression of *alx4* was seen in clusters of CNC in the ventral-most aspect of all PAs (Fig. 2S; see also Fig. S4E). Sectioning revealed expression of *alx4* throughout the future suprarostral, ethmoid plate, and trabecular cartilages (Fig. 3C).

prrx genes are *alx*-related transcription factors that mark the presumptive neurocranium in mouse and zebrafish (Hernandez-Vega and Minguillon, 2011; ten Berge et al., 1998). Unlike *alx*, however, *prrx* expression is also seen widely in the pharynx. In *X. laevis* we observed *prrx1/2*-positive mesenchyme in the FNP, and surrounding the eye and otic pit (Fig. 2U, V; S4I-T; 3D). *prrx1/2* transcripts were also detected in the maxillary domain, contiguous with expression in the future suprarostal cartilage (the FNP)(Fig. 3D). In the pharynx, *prrx1/2* expression marks the dorsal and ventral domains of each pharyngeal arch.

In addition to the ventral expression mentioned above, zebrafish and mouse *msx* homologs mark mesenchyme in the maxillary region, the FNP, and overlying the brain (Antonopoulou et al., 2004; Swartz et al., 2011). All of these domains express *msx1-b* and *msx2* in *X. laevis* (Fig. 2I, J; S2Y-II; 3F). However, we also detected *msx1-b* and *msx2* transcripts in the dorsal aspect of each PA (arrow in Fig. 2I; for *msx2* see also Fig. S2II), an expression domain not seen in zebrafish or mouse, but present in shark and lamprey (Cerny et al., 2010; Compagnucci et al., 2013).

In zebrafish, *gsc* marks the trabecular cartilage and the dorsal aspect of PA2, while *pou3f3* is expressed in dorsal PA1 and PA2 in both zebrafish and mouse (Gaunt et al., 1993; Hauptmann and Gerster, 2000; Jeong et al., 2008; Schultemerker et al., 1994; Tucker et al., 2004). As in zebrafish, we

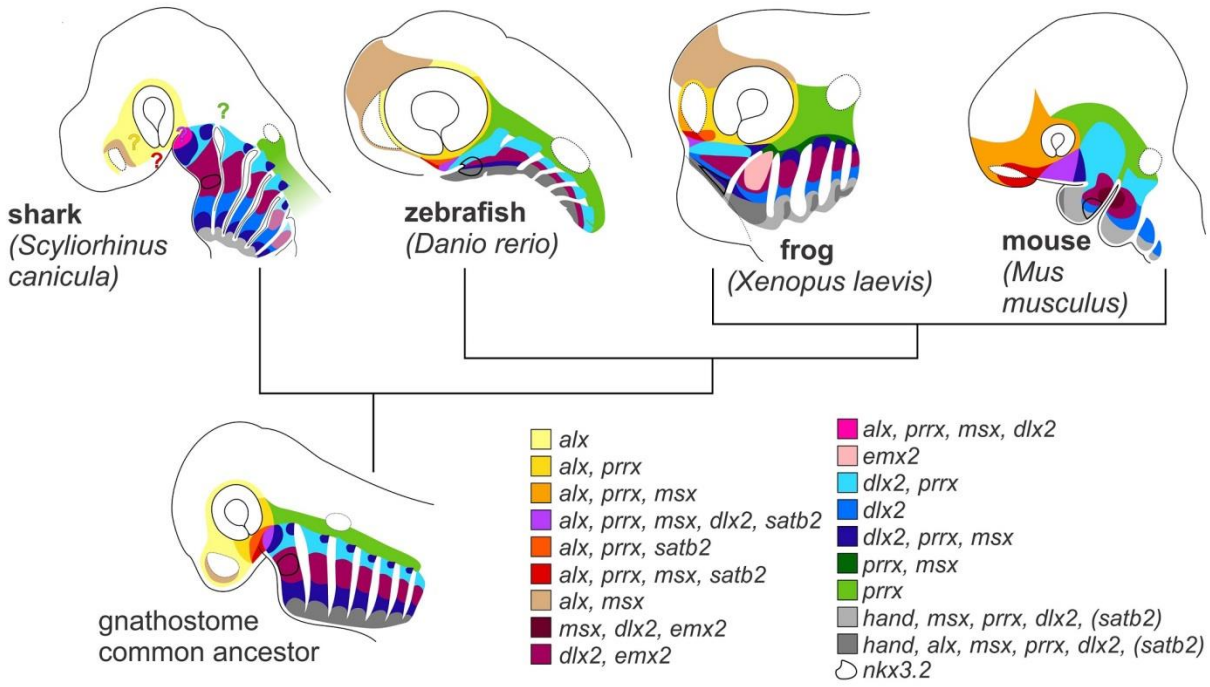

Fig. 3. Expression maps of larval gnathostome head skeletons. Combinatorial transcription factor expressions mapped to the head of a representative shark (*Scyliorhinus canicula*), ray-finned fish (zebrafish; *Danio rerio*), frog (*Xenopus laevis*), and mouse (*Mus musculus*). All schemes are depicted with anterior to left. Data for the shark, fish, and mouse were taken from current literature (see main text for citations); these maps represent our interpretation of published expression data. Extrapolating from these four organisms, we have reconstructed a hypothetical expression map for the gnathostome common ancestor. *nkx3.2* expression in PA1 is uniquely indicated by a black outline. Colored question marks on the shark indicate our prediction of correspondingly colored expression domains which may be present in the shark head. *satb2* in parentheses indicates a lack of this data in shark only (in both light and dark grey domains). *pou3f3*, *gsc*, and *tbx2/3* were excluded from this map for simplicity.

detected *X. laevis gsc-b* expression in the trabecular cartilages (Fig. 3E; see also Fig. S3J, K). However, expression in dorsal PA2 was not observed. *X. laevis pou3f3* transcripts marked clusters of cells in the dorsal and medial portions of PA1 and PA2 (Fig. 2W, X; S4U-F).

In addition to marking the ventral pharynx, *satb2* and *tbx2/3* transcripts mark the maxillary/premaxillary region of mouse and zebrafish (Fish et al., 2011; Leoyklang et al., 2007; Ribeiro et al., 2007; Sheehan-Rooney et al., 2010; Zirzow et al., 2009). *satb2* and *tbx3-a* are both expressed in the maxillary region of *X. laevis* with additional expression in FNP mesenchyme (Fig. 2K, L; O, P; S3A-E; M-P; 3A, B). *tbx3-a* expression in this domain extends dorsally over the length of the nasal capsule, mirroring *msx1/2* expression (compare Figs. 2J and S2Z to 2P).

Markers of cartilage and joint differentiation: *barx1* and *gdf5*

Combinatorial expression of transcription factors mark subpopulations of skeletal precursors in the *X. laevis* head, which are largely conserved in other vertebrates (Fig. 4). To better correlate this transcription factor map with skeletal differentiation, we examined the expression of *barx1* and *gdf5* from st. 33 until st. 45, when the head skeleton has largely chondrified, and compared this expression to alcian blue reactivity (Fig. S6). In zebrafish and amniotes, *barx1* is expressed broadly in CNC where it is essential for chondrogenesis (Barlow et al., 1999; Sperber and Dawid, 2008) and recent work in zebrafish has also shown that it must be downregulated for joint tissue to form (Nichols et al., 2013). Consistent with a general role for *barx1* in chondrogenesis, we observed *barx1* transcripts throughout pharyngeal CNC but excluded from the intermediate domain of PA1 (Fig. S4Y-DD), which gives rise to the primary jaw joint. *barx1* expression was also reduced in the dorsal aspect of PA2 (arrow in Fig. 5F). We also noted that *barx1* expression largely recapitulates *Sox9* expression (Fig. S4EE-HH) and anticipates alcian blue reactivity (Fig. S6), further supporting a role in cartilage differentiation.

gdf5 is a TGFß signaling peptide essential for joint formation in the limbs and head (Settle et al., 2003; Storm and Kingsley, 1999). In *X. laevis*, *gdf5* expression is apparent at st. 41 in the nascent primary jaw joint (arrow in Fig. 6C; see also Fig. S5), which arises from the intermediate domain of PA1. Aside from the primary jaw joint, additional expression is seen flanking the ventral midline of PA2 (Fig. S5G), and associated with the presumptive suprarostral cartilages (arrow in Fig. S5B). Spots of expression are observed weakly in the lower jaw at st. 41 (arrowheads Fig. 6C, G), and more intensely at st. 45 (arrowhead in Fig. 6D), presaging formation of the intramandibular joints.

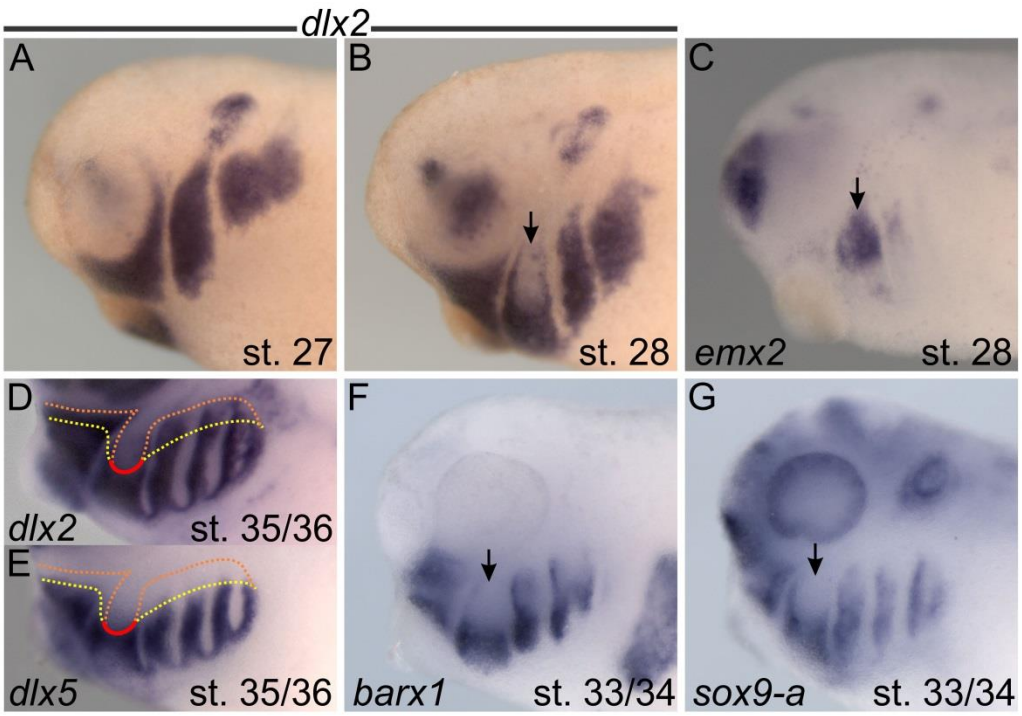

Fig. 4. Unique gene expression in the *X. laevis* PA2. All larvae are shown in lateral view with anterior to left. (A, B) *dlx2* is expressed throughout migratory PA CNC at (A) st. 27, but at (B) st. 28 it becomes sharply downregulated in the dorsal PA2 (the future columella; arrow in B). (C) Coinciding with *dlx2* down-regulation, *emx2* is expressed at st. 28 in the dorsal PA2 (arrow). At st. 35/36, (D) *dlx2* and (E) *dlx5* uniquely share part of their dorsal boundary in PA2. Orange and yellow dotted lines show the boundary of *dlx2* and *dlx5*, respectively, while the red solid line indicates their shared border. (F) *barx1*, and (G) *sox9-a* also display conspicuous down-regulation in the dorsal PA2 at st. 33/34 (arrows in F and G).

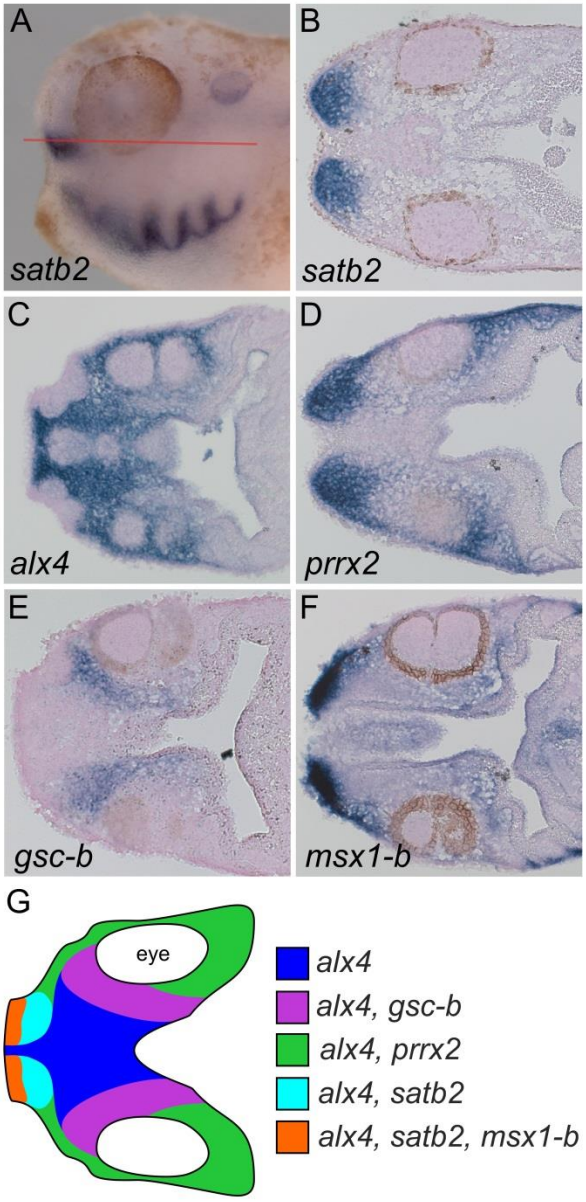

Fig. 5. Expression domains in the forming suprarostral plate, ethmoid plate, and trabecular cartilages at st. 37/38. Anterior to left in all panels. (A) Wholemount lateral view of *satb2* at st. 37/38; red line depicts the approximate plane of section used to assess gene expression in all subsequent panels. (B) *satb2* is expressed dorsal to the PAs in the suprarostral plate at st. 37/38. (C) *alx4* marks all aspects of the future neurocramium ventral to the eyes and brain. (D) *prrx2* is found only in a lateral subset of the suprarostal plate, ethmoid plate, and trabeculae. (E) *gsc-b* is detected in the trabeculae and a small portion of the lateral ethmoid plate. (F) *msx1-b* is expressed strongly in the future suprarotal plate. (G) A summary scheme of the expressions shown in B-F.

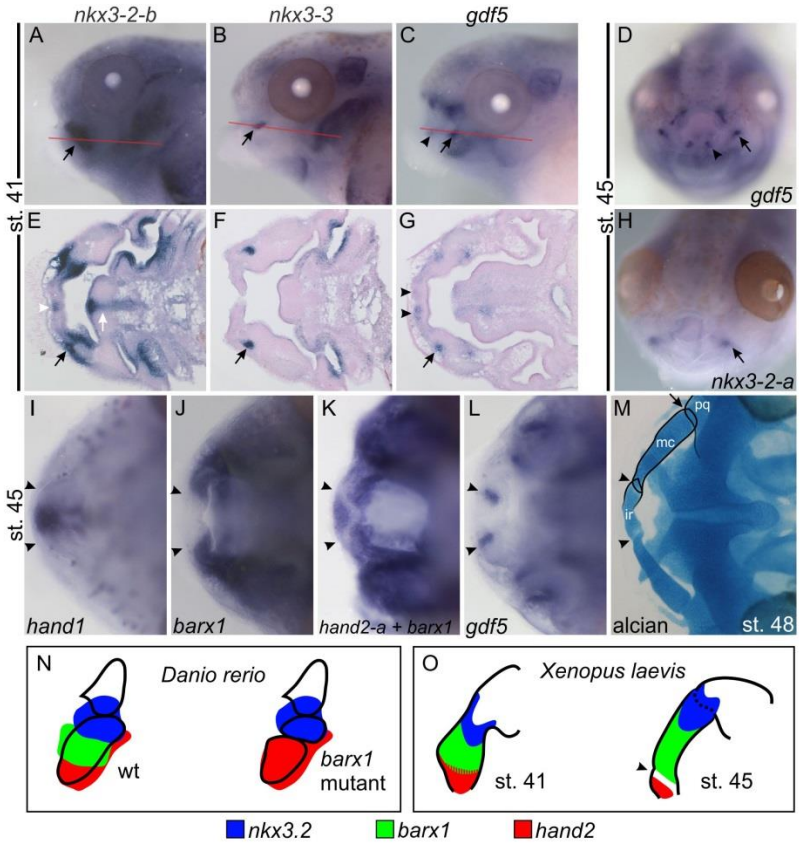

Fig. 6. Gene expression in the intramandibular and primary jaw joint. Lateral views (A-C), dorsal views (I-M), and sections (E-G) are all shown with anterior to the left. Anterior views in (D, H). All black arrows mark the primary jaw joint, and all black arrowheads mark the intramandibular joint. (A-C) lateral views of wholemount ISH for (A) *nkx3-2-b*, (B) *nkx3-3*, and (C) *gdf5* at st. 41. Red lines depict the plane of section in the corresponding panel below each wholemount photo, black arrows mark the future primary jaw joint, and the black arrowhead marks the future intramandibular joint. (D) Anterior view of wholemount ISH for *gdf5* at st. 45. (E-G) Sections of st. 41 larvae showing expressions in the future intramandibular and primary jaw joints. (E) *nkx3-2-b* is expressed in a broad domain surrounding the nascent primary jaw joint (black arrow), and also medially in the future infrarostral (white arrowhead) and future basihyal (white arrow). (F) *nkx3-3* expression is found in the primary jaw joint (black arrow). (G) *gdf5* expression is found in the nascent intramandibular joints (black arrowheads), and primary jaw joint (black arrow). (H) st. 45 wholemout ISH for *nkx3-2-a*, shows expression in the primary jaw joint (black arrow). (I-L) dorsal views of st. 45 wholemount ISH showing mandibular expression of (I) *hand1*, (J) *barx1*, (K) *hand2-a* + *barx1* (double ISH; see methods), and (L) *gdf5*. Black arrowheads mark the future intramandibular joints. (M) Dorsal view with a ventral focal plane of a st. 48 *X. laevis* larva stained with alcian blue. Black arrowheads mark the intramandibular joints, black arrow marks the right primary jaw joint. (N-O) schematics of expressions showing *hand2* (red), *barx1* (green), and *nkx3.2* (blue) in the mandibular arch. (N) wt and *barx1* mutant zebrafish mandibular expression (adapted from Fig. 9B of Nichols et al., 2013). (O) st. 41 and st. 45 *X. laevis* mandible expression. The *barx1* mutant zebrafish takes on a larval morphology strikingly similar to that of an anuran, with an extra joint forming within Meckel's cartilage. Black arrowhead indicates the position of the future intramandibular joint. ir, infrarostral; mc, Meckel's cartilage; pq, palatoquadrate.

Discussion

dlx expression in X. laevis highlights conserved and divergent aspects of the pharyngeal "*dlx* code"

Gnathostome *dlx* genes are typically distributed in the genome as linked, co-regulated gene pairs representing 6 different paralogy groups (MacDonald et al., 2010; Sumiyama et al., 2002). In post-migratory CNC, *dlx* genes are expressed in a nested pattern along the dorso-ventral axis of each pharyngeal arch, with the expression boundaries of the more restricted paralogs encompassed by the boundaries of the more broadly-expressed paralogs (Cerny et al., 2010; Coffin Talbot et al., 2010; Ellies et al., 1997). As in other gnathostomes, *X. laevis* displays nested expression of *dlx* genes in post-migratory CNC, with the *dlx3/4* pair having the most restricted expression pattern, *dlx5/6* expression being less restricted, and *dlx1/2* being the most broadly expressed (Fig. 1M, N). However, this nested pattern differs from the dorso-ventrally tiered *dlx* expression scheme described in mouse (Depew et al., 2002b). Specifically, the *dlx3/4* gene pair is most strongly expressed in the intermediate domain of the *X. laevis* pharynx, with reduced expression of *dlx3* and no detectable expression of *dlx4* in the ventral-most aspect of the PAs. This pattern is reminiscent of zebrafish and cichlids where *dlx3/4* expression initially extends into the ventral domain, and then becomes restricted to the intermediate domain (Coffin Talbot et al., 2010; Renz et al., 2011). Recent work in elasmobranchs has shown that *dlx3/4* expression is nested within the *dlx5/6* expression domain, however it is unclear if the ventral boundary shifts dorsally at later stages (Compagnucci et al., 2013; Debiais-Thibaud et al., 2013; Gillis et al., 2013; Takechi et al., 2013). Coexpression of all *dlx* paralogs is also seen in the intermediate domain of the sea lamprey, with one and two paralogs expressed in the ventral-most and dorsal-most domains, respectively. Taken together, *dlx* expression in *X. laevis*, teleosts, and lamprey suggest that an intermediate pharyngeal domain expressing all or most *dlx* paralogs flanked dorsally by CNC expressing *dlx1/2*, and ventrally by CNC expressing *dlx1/2* and *dlx5/6*, represents the ancestral gnathostome *dlx* expression pattern. It is less clear if this exact pattern was present in the vertebrate ancestor, as the orthology of lamprey and gnathostome *dlx* paralogs is still unresolved.

While all vertebrates display nested *dlx* expression in the pharynx, surprisingly little is known about the precise role of *dlx* genes in CNC. The most frequently cited model of *dlx* function in postmigratory CNC is the "*dlx* code" in which *dlx* genes perform homeotic selector functions in CNC cell subpopulations along the proximal-distal axis, analogous to the *hox* code along the antero-posterior axis (Beverdam et al., 2002; Depew et al., 2002; Ellies et al., 1997). While this model is supported by the morphology of PA1 and PA2 skeletal elements in *dlx* knock-out mice, there is limited evidence that a *hox*-type *dlx*-code operates in other vertebrates. In zebrafish, *dlx1/2* has a general role in skeletogenic

CNC survival and chondrogenesis (Sperber and Dawid, 2008), while simultaneous *dlx3/4* and *dlx5* depletion causes the loss of joints and hypomorphic skeletal elements (Coffin Talbot et al., 2010). In addition, there is no evidence from mouse or zebrafish that *dlx* paralogs encode functionally distinct proteins with different transcriptional targets, a requisite feature of a *hox*-like selector code. Rather, *dlx5/6* and *dlx1/2* have been shown to act redundantly in mouse, while *dlx3/4* and *dlx5/6* are at least partially redundant in zebrafish (Coffin Talbot et al., 2010; Jeong et al., 2008). Finally, in zebrafish, the boundaries of *dlx* expression do not correspond to individual cartilage anlagen, as would be expected of a *hox*-type selector code (Coffin Talbot et al., 2010; Medeiros and Crump, 2012). Based on the observed functional redundancy of *dlx1/2* and *dlx5/6* in mouse, an alternative to the *dlx*-code model was proposed. This "quantitative model" posits that rather than acting as homeotic selectors of skeletal element morphology, nested *dlx* expression creates zones of high, medium, and low Dlx dosage (Jeong et al., 2008). These different levels of collective Dlx activity, together with unknown modulators of *Dlx* function, then work to pattern CNC along the proximal-distal axis.

In *X. laevis*, *dlx* expression boundaries in the first and second arches appear to correspond to presumptive skeletal elements and joints, consistent with both the "*dlx*-code" and the "quantitative" models of skeletal element specification. However, in the posterior arches, nested *dlx* paralog expression does not presage the formation of individual skeletal elements. Instead, CNC subpopulations expressing different combinations of *dlx* genes, differentiate into a fused pharyngeal basket. A similar situation is seen in lamprey where CNC displaying nested *dlx* expression give rise to a fused, unjointed skeleton rather than distinct skeletal elements (Cerny et al., 2010). In these instances, *dlx* genes likely do not function as a "*dlx* code" to determine skeletal element morphology or to position joints. In lamprey, three separate mesenchymal condensations form in each posterior PA before overt cartilage differentiation begins. The middle condensation is the first to differentiate into cartilage, with chondrogenesis gradually spreading into the dorsal and ventral condensations (Morrison et al., 2000). In *X. laevis*, chondrogenesis also appears to begin in the intermediate domain of each arch and spread dorsally and ventrally as shown by *aggrecan* transcription (Suzuki et al., 2012) and alcian blue staining (Fig. S6). However, it is unclear if this is preceded by the formation of separate mesenchymal condensations as in lamprey. Thus, it is possible that in lamprey and *X. laevis*, *dlx* genes regulate of the positioning of early mesenchymal condensations and/or the timing of skeletal differentiation along the dorso-ventral axis. In either case, our data suggest that *dlx* genes do not always act as homeotic selectors of skeletal element morphology. Instead, they appear to function more generally to confer positional identities upon CNC subpopulations in the pharynx. Whether this patterning regulates skeletal element morphology, mesenchyme condensation, or the timing of differentiation, can vary between species, or even between the anterior and posterior PAs in the same species.

Reconstructing the ancestral features of gnathostome skeletogenic mesenchyme patterning

While *dlx* genes are essential for proper head skeleton development, they are only one part of the complex system that patterns the nascent vertebrate head skeleton. Members of several other transcription factor families mark subpopulations of head mesenchyme and result in skeletal defects when perturbed in mouse and/or zebrafish. We compared the expression of these factors in *X. laevis* with that of their orthologs in mouse, zebrafish, shark, and lamprey to identify conserved features of vertebrate and gnathostome embryonic head skeleton patterning (Fig. 4).

In *X. laevis*, zebrafish, lamprey, and shark CNC in the intermediate pharyngeal domain coexpress all *dlx* paralogs (Cerny et al., 2010; Coffin Talbot et al., 2010; Compagnucci et al., 2013; Ellies et al., 1997; Takechi et al., 2013). In *X. laevis*, mouse, zebrafish, and shark, this region is also marked by *emx2* transcripts (Compagnucci et al., 2013; Kawahara and Dawid, 2002; Thisse et al., 2004) and, in PA1, *nkx3.2* (Compagnucci et al., 2013; Miller et al., 2003; Tucker et al., 2004). Dorsal to this domain, presumptive CNC in all gnathostomes examined express *prrx* and *dlx1/2*, with *msx* expression marking the dorsal aspect of PA1 (Antonopoulou et al., 2004; Compagnucci et al., 2013; Hernandez-Vega and Minguillon, 2011; Medeiros and Crump, 2012; Swartz et al., 2011; ten Berge et al., 1998; Thisse et al., 2004; Thomas et al., 1998). In shark, *X. laevis*, and lamprey, dorsal *msx* expression is also seen in the posterior arches (Cerny et al., 2010; Compagnucci et al., 2013). These data are consistent with coexpression of *dlx1/2*, *prrx1/2*, and *msx* in the dorsal pharynx in the gnathostome common ancestor, with a homologous *dlx/msx* positive domain present in the vertebrate common ancestor. Dorsal *msx* expression would have then been lost from the posterior arches of amniotes and zebrafish.

In the ventralmost aspect of most PAs, zebrafish, mouse, and frog co-express *dlx1/2*, *dlx5/6*, *msx*, *prrx*, *hand*, *satb2*, and *tbx2/3* (reviewed by Medeiros and Crump, 2012; Mesbah et al., 2008; Ribeiro et al., 2007). This ventral domain appears conserved in shark and lamprey (Cerny et al., 2010; Compagnucci et al., 2013), though *prrx* and *tbx2/3* expression has not been described in lamprey, and *satb2* has not been described in either lamprey or shark. *alx* transcripts mark a subset of ventral CNC in the posterior arches of *X. laevis* and lamprey, and ventral PA1 in amniotes (Cattell et al., 2011). In zebrafish, ventral *alx* expression is restricted to PA1 at 24h (Dee et al., 2013) though *alx4b* appears enriched in the ventral posterior arches at 48h (Thisse et al., 2004). Taken together these data suggest *alx* was expressed in the ventralmost aspect of all PAs in the vertebrate common ancestor, and early gnathostomes, then lost from the posterior arches in amniotes. To date, only a single *alx* paralog has been isolated from shark and ventral expression is only observed in PA1 (Compagnucci et al., 2013). Additional *alx* data from chondrichthyians should help resolve the ancestral gnathostome *alx* expression pattern.

Between the ventralmost and intermediate domains, all gnathostomes examined appear to have a "ventral-intermediate" domain bordered dorsally by *emx2*-expressing CNC and ventrally by *hand*

expression (Compagnucci et al., 2013; Morita et al., 1995; Thisse et al., 2004). In *X. laevis* and mouse, this domain expresses *dlx* paralogs but not *msx*. However, in zebrafish, shark, and lamprey it includes *msx/dlx* coexpressing cells, suggesting an *msx/dlx*-positive "ventral-intermediate" domain was present in the vertebrate common ancestor and lost in tetrapods (Cerny et al., 2010; Coffin Talbot et al., 2010; Compagnucci et al., 2013).

In sum, our data are consistent with presence of 4 dorso-ventral domains in the pharyngeal arches of the vertebrate common ancestor and early gnathostomes (Fig. 4), as previously proposed (Cerny et al., 2010). In PA1, this basic pattern shows some elaboration, with the dorsal aspect of PA1 expressing *satb2* and *alx*, the ventral-most aspect expressing *gsc*, and the intermediate domain expressing *nkx3.2* (Fig. 4). Interestingly, dorsal expression of *alx* is seen in all pharyngeal arches in lamprey (Cattell et al., 2011) suggesting dorsal *alx* expression may have been lost from the posterior arches in gnathostomes. Alternately, expansion of *alx* expression into the dorsal pharynx may be a derived feature of lamprey head skeleton patterning.

Outside of the pharyngeal arches, a combination of CNC and mesodermal mesenchyme give rise to the skeletal elements of the ethmoid plate, trabeculae, neurocranium, and the suprarostal cartilages in anurans. In all vertebrates examined, this head mesenchyme lacks *dlx* expression and expresses various combinations of *alx*, *prrx*, *msx*, *gsc*, *tbx2/3*, and *satb2* paralogs (Antonopoulou et al., 2004; Beverdam et al., 2001; Dee et al., 2013; Hernandez-Vega and Minguillon, 2011; Swartz et al., 2011; ten Berge et al., 1998; Thomas et al., 1998)(Compagnucci et al., 2013; Gaunt et al., 1993; Schultemerker et al., 1994; Sheehan-Rooney et al., 2010; Thisse et al., 2004). Immediately dorsal to the posterior arches, all gnathostomes have a domain of *prrx*-positive mesenchyme surrounding the otic capsule. This domain extends to just behind the eye in *X. laevis*, zebrafish and mouse, though the single *prrx* homolog described in shark is limited to the otic region at the stages assayed (Compagnucci et al., 2013). In all gnathostomes and lamprey, *alx* expression is seen in mesenchyme surrounding the eye and nasal pit/capsule (Beverdam and Meijlink, 2001; Cattell, 2011; Dee et al., 2013; McGonnell et al., 2011). In *X. laevis*, zebrafish, and mouse, this expression overlaps with *prrx* in the posterior peri-orbital mesenchyme, and with *msx* above the eye, in the region of the presumptive cranial vault (Hernandez-Vega and Minguillon, 2011; ten Berge et al., 1998; Thisse et al., 2004). Anteriorly, all gnathostomes also have a region of *msx/alx* positive mesenchyme associated with the nasal pit. Finally, in zebrafish and *X. laevis*, coexpression of *prrx*, *alx*, *gsc*, *msx*, and *satb2* is seen in portions of the forming trabecular cartilages/ethmoid plate (Dee et al., 2013; Hernandez-Vega and Minguillon, 2011; Schultemerker et al., 1994; Swartz et al., 2011; Sheehan-Rooney et al., 2010; Thisse et al., 2004). Taken together, our comparisons support a pan-gnathostome head skeleton patterning matrix consisting of 4 molecularly distinct PA domains (corresponding to the light blue, burgundy, dark blue, and grey domains in Fig. 4),

unique dorsal and intermediate PA1 domains, and 6 domains of presumptive neurocranial precursors (Fig. 6).

Gene expression in the anuran second arch reveals developmental novelties associated with the evolution of the middle ear

While most features of head mesenchyme patterning are conserved across all gnathostomes, our comparisons revealed some differences in gene expression between groups. We asked if anuran-specific alterations to the deduced gnathostome ground state could be linked to specific skeletal novelties. One difference in CNC patterning gene expression between *X. laevis* and the basal gnathostome condition is observed in PA2. In most other gnathostomes, PA2 has the expression profile of the other arches, with *dlx1/2* marking its dorsal aspect (Akimenko et al., 1994; Compagnucci et al., 2013) (Fig. 4). This pattern is initially conserved in *X. laevis* (Fig 5A). However, shortly after CNC migration, *dlx1/2* is downregulated in dorsal PA2 (compare Fig. 5A, B; arrow in B). As a result, the dorsal boundaries of *dlx1/2* and *dlx3/5/6* are approximately the same in PA2 of *X. laevis* (Fig. 5D, E). This reduction in dorsal PA2 *dlx* expression is accompanied by an upregulation of *emx2* expression in this domain, showing that reduced *dlx1/2* expression is not due to an absence of CNC (arrow in Fig. 5C), but a change in PA2 CNC gene expression.

In zebrafish and mouse, knockdown of *dlx1/2* causes a general inhibition of chondrogenesis, with dorsal PA1 and PA2 elements especially affected (Coffin Talbot et al., 2010; Sperber et al., 2008). This difference in size between dorsal and ventral elements is reminiscent of the PA2 skeleton of modern anamniote tetrapods, which have a highly reduced dorsal element, the columella (stapes), and a fully formed ventral PA2 element, the ceratohyal. It is tempting to speculate that reduction in *dlx1/2* expression may have driven evolution of the stapes by decreasing the skeletogenic capacity of CNC in PA2. Consistent with this, *sox9-a* expression is downregulated in the *X. laevis* dorsal PA2, and *barx1* transcripts are largely absent from this domain (Fig. 5F, G). It is also worth noting that the amphibian stapes differentiates very late in development, during metamorphosis (Witschi, 1949). Thus, reduced expression of *dlx2* in dorsal PA2 during early *X. laevis* development may also reflect the delayed differentiation of CNC in this domain. In either case, the consequences of attenuating *dlx2* expression during evolution could be tested by phenocopying the ancestral gnathostome *dlx2* expression pattern in *X. laevis* using synthetic *dlx2* mRNA. Provocatively, over-expression of *dlx2* in chick, an amniote in which both the dorsal and ventral PA2 skeletal elements are highly reduced, leads to the formation of ectopic cartilage nodules near PA1 and PA2 -derived cartilages (Gordon et al., 2010). It is possible that modulating the expression of *dlx* genes in CNC subpopulations is one mechanism by which size of skeletal elements is altered during evolution.

The anuran suprarostral and infrarostral cartilages likely evolved by partitioning of pre-existing mesenchymal subpopulations

Much of the variation in vertebrate head skeleton morphology involves changes in the size and/or shape of evolutionarily conserved skeletal elements. However, novel cartilages and bones with no obvious homologs in other groups have also arisen in many lineages. The mechanisms by which such elements evolve are unclear. The suprarostral and infrarostral cartilages are anuran novelties associated with the evolution of larval herbivory, and have been linked to the success of this group (reviewed by (Svensson and Haas, 2005)). Svensson and Haas (2005) proposed three models for how these elements could have evolved: 1) duplication of existing elements, e.g. duplication of Meckel's cartilage and primary jaw joint in the case of the infrarostral; 2) partitioning of pre-existing cartilages by the formation of new joints; and 3) *de novo* initiation of cartilage condensations non-homologous to preexisting elements. We asked whether the gene expression profiles of the nascent suprarostral and infrarostral cartilages were consistent with any of these scenarios. In the presumptive suprarostral cartilage, we noted *msx, satb2, alx* and *prrx* co-expression, mimicking gene expression in the anterior ethmoid plate/FNP of other gnathostomes (Dee et al., 2013; Hernandez-Vega and Minguillon, 2011; Schultemerker et al., 1994; Swartz et al., 2011; Sheehan-Rooney et al., 2010; Thisse et al., 2004). Similarly, we found that the nascent infrarostral cartilage of *X. laevis* has the same transcription factor expression profile as the ventralmost aspect of PA1, which gives rise to the mandible in other gnathostomes. Importantly, we did not see novel transcription factor expression consistent with *de novo* evolution of infrarostral or suprarostral cartilage anlage. We also did not see reiteration of ventral or intermediate PA1 gene expression patterns (i.e. tiered *hand, dlx3/4, nkx3* expression) in the presumptive infrarostral cartilage, as would be predicted by the duplication of PA1 domains. These observations suggest that the suprarostral and infrarostral cartilages are derived from the same precursor populations as the anterior ethmoid plate and Meckel's cartilage, respectively. However little is known how combinatorial expression of transcription factors in head mesenchyme is ultimately translated into skeletal morphology. Thus, we cannot exclude the possibility that some form of duplication or *de novo* evolution of skeletal anlagen occurred via mechanisms developmentally downstream of early mesenchymal patterning.

In *X. laevis* and other pipoid frogs, the suprarostral cartilage is fused to the ethmoid plate, while the infrarostral articulates with Meckel's cartilage at the intramandibular joint (Pugener et al., 2003; Trueb and Hanken, 1992; Zhang et al., 2013). It has been postulated that evolution of this joint was driven by the duplication and redeployment of *nkx3* (*bapx/zax*) genes, which are necessary for jaw joint formation in other vertebrates (Svensson and Haas, 2005). We examined the expression of both *X. laevis nkx3-2* paralogs, and the closely related gene *nkx3-3* in the region of the nascent intramandibular joint and did not observe any expression consistent with this scenario (Fig. 6). This suggests that some factor

downstream of *nkx3.2* induces intramandibular joint formation. We thus examined the expression of *gdf5*, a TGFß signaling ligand regulated by *nkx3.2* in PA1 that is essential for primary jaw joint development (Miller et al., 2003; Settle et al., 2003; Storm and Kingsley, 1999). In contrast to *nkx3-2,* we noted bilateral expression of *gdf5* in ventral PA1 precisely presaging formation of the intramandibular joints. No equivalent expression has been reported in any other gnathostome, suggesting novel *nkx3-2*-independent expression of *gdf5* was associated with the evolution the anuran intramandibular joint/infrarostral cartilage.

Anuran-specific expression of *gdf5* in the nascent intramandibular joint could have arisen by *cis*-regulatory changes at the *gdf5* locus, or by changes in the expression of upstream regulators. In addition to *nkx3.2, barx1* has been shown to regulate joint differentiation in PA1. However, unlike *nkx3.2, barx1* acts as a repressor of joint formation and *barx1* mutants have ectopic paired joints in Meckel's cartilage (Nichols et al., 2013). Due to the similarity of this ectopic joint and the frog intramandibular joint, Nichols et al. (2013) postulated that a change in *barx1* function in PA1 could have occurred in anurans. Presumably, this would have involved a reduction or loss of *barx1* in the ventral PA1, creating an additional *barx1*-free domain capable of supporting joint differentiation. At st. 33 and 41, we observed zebrafish-type patterning of *X. laevis* PA1, with *barx1* expression abutting *hand* expression ventrally/anteriorly, and *nkx3.2* dorsally/posteriorly (Fig. 6N, O; S5C, D). However, between st. 41 and st. 45 a clear gap between the *hand* and *barx1* expression domains becomes apparent (Fig. 6I-K). This gap corresponds precisely to the paired spots of *gdf5* expression that mark the forming intramandibular joint (Fig. 6L, M). Conceivably, this new domain could represent a dorsal shift in *barx1* expression, a ventral shift in *hand* expression, or both. However, the fact that a loss of *hand2* function alone is insufficient to create ectopic joints in zebrafish (Miller et al., 2003), supports a central role for *barx1* restriction in anuran intramandibular joint evolution.

Assuming zebrafish PA1 approximates PA1 of early tetrapods, our results suggest reduced *barx1* expression and recruitment of *gdf5* created a novel joint-forming domain in the anuran mandible. Changes in *barx* and *gdf5* expression have also been implicated in the evolution of the primary jaw joint, as lamprey lacks any *gdf5* homolog expression in PA1, and expresses *barx* contiguously through the ventral and intermediate PA1 (Cerny et al., 2010). *gdf5* homologs also mark joints in tetrapod limbs, which can vary dramatically in number between lineages (Settle et al., 2003; Storm and Kingsley, 1999). These observations raise the possibility that altering *barx1* and *gdf5* expression may be one way to partition skeletal elements and position new joints during evolution.

Conclusions

We generated a gene expression map of CNC and mesoderm-derived mesenchyme in *X. laevis* and compared it to equivalent data from amniotes, zebrafish, and shark. These comparisons show that the head skeletons of all gnathostomes are built upon an ancient and well-conserved pre-pattern of molecularly distinct precursor subpopulations, despite differences in skeletal morphology. We also present evidence that skeletal element size may be altered during evolution by changes in the skeletogenic capacity of precursor subpopulations early in head skeleton development. Finally, we propose that new joints and skeletal elements in anurans likely evolved from preexisting skeletal precursor domains via late changes in genes that regulate skeletal differentiation. Together these results support a model in which early patterning mechanisms divide the head mesenchyme into a highly conserved set of skeletal precursor subpopulations. While subtle changes in this early patterning system can affect skeletal element size, changes to this matrix do not appear to underlie the evolution of new skeletal elements. In contrast, later expression of the genes that regulate skeletal element differentiation can be linked to specific anuran skeletal innovations. We posit that changes in the expression of these downstream regulators, including *barx1* and *gdf5*, are one mechanism by which skeletal element number is altered during evolution. Further work in a broader range of vertebrates with divergent head skeleton morphologies will establish if this model of vertebrate head skeleton evolution is broadly applicable.

CHAPTER IV: Embryonic expression of *endothelins* and their receptors in lamprey and frog reveals stem vertebrate origins of complex Endothelin signaling

Introduction

The NCCs of all living vertebrates are divided into distinct subpopulations with unique migration routes and developmental fates (McCauley and Bronner-Fraser, 2003; Meulemans and Bronner-Fraser, 2004). Modern NCCs presumably evolved step-wise from a more homogenous population of migratory neural tube cells with limited developmental potential (Abitua et al., 2012; Stolfi et al., 2015). When and how NCCs acquired their multipotency, patterning, and stereotyped migration routes, is unclear.

In modern jawed vertebrates, Endothelins are key regulators of NCC differentiation, migration, and patterning (Baynash et al., 1994; Bonano et al., 2008; Clouthier et al., 2010; Kawasaki-Nishihara et al., 2011; Kempf et al., 1998; Krauss et al., 2014; Miller et al., 2000; Pla and Larue, 2003; Sanchez-Mejias et al., 2010; Shin et al., 1999; Yanagisawa et al., 1998). Gnathostomes possess multiple (2-6) endothelin receptors (*ednr*s) (Braasch and Schartl, 2014) as well as multiple (3-6) endothelin ligands (*edn*s) (Braasch et al., 2009). Phylogenetic analysis places the ligands into four paralogy groups (Braasch et al., 2009): *edn1*, *edn2*, *edn3*, and *edn4*, while the receptors form three groups (Braasch and Schartl, 2014): *ednra*, *ednrb1*, and *ednrb2*. Importantly, most modern vertebrates possess subsets of these (discussed below).

Endothelin ligands are translated as ~200 aa (amino acid) polypeptides (Preproendothelins) that are then processed into short, and invariably 21 aa signaling ligands. This occurs as Preproendothelins obtain disulphide linkages between cysteine residues 1 and 15, and 3 and 11 (present on all known Edn ligands), and lose their amino and carboxyl termini via cleavage by a furin endopeptidase, making an intermediate product known as "Big Endothelin". Big Endothelins are subsequently trimmed on the carboxyl end again by an Endothelin converting enzyme (ECE) (Johnson et al., 2002; Khimji and Rockey, 2010; Kido et al., 1998), leaving the 21 aa polypeptide. In jawed vertebrates there are two main groups of ECEs, ECE-1 and -2, which have unique biochemical properties (Emoto and Yanagisawa, 1995) as well as unique developmental roles in mouse (Yanagisawa et al., 2000). Despite these differences, both ECE-1 and ECE-2 have a similar hierarchy of Big Endothelin cleavage rates *in vitro* (Edn1 >> Edn2 > Edn3)(Emoto and Yanagisawa, 1995) (Edn4 has not been assayed in this context). Once fully processed and secreted by the cell, mature Edn ligands interact with the various Ednrs to drive cell fate decisions and provide positional information to NCCs (Baynash et al., 1994; Bonano et al., 2008; Clouthier et al.,

2010; Kawasaki-Nishihara et al., 2011; Kempf et al., 1998; Krauss et al., 2014; Miller et al., 2000; Pla and Larue, 2003; Shin et al., 1999; Yanagisawa et al., 1998). In cell culture assays, each Ednr appears to have a unique set of binding affinities for the different mature Edn ligands (Karne et al., 1993; Sakamoto et al., 1993; Yanagisawa, 1994), though the early developmental relevance of these differences remains unknown.

The first role described for Endothelin signaling was in vasoconstriction and circulatory development (Hirata et al., 1988; Itoh et al., 1988; Tomobe et al., 1988; Yanagisawa et al., 1988). Subsequently, Endothelin pathways were found to perform a variety of functions in NCC including cell fate determination, migratory pathfinding, and patterning of cranial NCC. Specifically, the Edn1/Ednra pathway works to pattern skeletogenic cranial NCC into distinct skeletal progenitor populations along the dorso-ventral axis of the pharynx, thereby positioning the jaw joint, and also has a function in cardiac NCC development (Clouthier et al., 2010; Clouthier et al., 1998; Miller et al., 2000; Sato et al., 2008a; Sato et al., 2008b). In non-skeletogenic NCCs, the Edn3/Ednrb pathway is necessary for pigment cell specification and migration (Baynash et al., 1994; Kawasaki-Nishihara et al., 2011; Krauss et al., 2014; RM et al., 2006), as well as enteric neuron development (Baynash et al., 1994; Sanchez-Mejias et al., 2010). *edn2*, though present in all major vertebrate lineages, does not have known expression in early development to our knowledge. *edn4* is retained only in some lineages of ray-finned fishes (Braasch et al., 2009), and its expression and function in early development is unknown. Interestingly, *ednrb2* has been lost in both zebrafish and therian mammals (Braasch and Schartl, 2014), despite this receptor's importance in NCC-derived pigment cell development in other vertebrates (Kawasaki-Nishihara et al., 2011; Krauss et al., 2014; RM et al., 2006).

When the various roles for endothelin signaling in vertebrate NCC development evolved is unknown. *ednr*s appear to be unique to chordates, and are not found in the genomes of protostomes or non-chordate deuterostomes (Braasch and Schartl, 2014): no tunicate genome appears to contain any *edn*- or *ednr*-like genes (Braasch and Schartl, 2014; Braasch et al., 2009). Amphioxus (*Branchoistoma floridae*) also lacks any Edn ligands (Braasch and Schartl, 2014; Braasch et al., 2009), though it possesses a single *ednr-like* gene that is not expressed during early embryogenesis (unpublished results, (Yu et al., 2008a)). The Japanese lamprey, *Lethenteron japonicum*, a jawless vertebrate, possesses at least one *ednr* and six putative *edn* ligands (Kuraku et al., 2010), suggesting that at least a simple form of Endothelin signaling (i.e. a single receptor with its particular ligand affinities) arose in the common ancestor of jawed and jawless vertebrates.

After exhausting available genomic assemblies and multiple transcript assemblies from adult and larval tissues, here we identify and characterize the expression of what is likely the complete set of *ednr* and *edn* ligands in the sea lamprey, *Petromyzon marinus*. We find dynamic embryonic and larval

expression of two *ednr*s which are likely directly orthologous to gnathostome *ednra* and *ednrb*. We also identify six *P. marinus edn* ligand genes, four of which are expressed in temporo-spatially restricted patterns in embryos and early larvae, with a fifth transcribed diffusely throughout the animal. We then use comparisons with *Xenopus laevis ednr* and *edn* expression to assign tentative functional overlap between some lamprey and gnathostome Endothelin receptors and ligands. Our results suggest that at least one *ednr* duplication occurred prior to the divergence of cyclostomes and gnathostomes, and that sophisticated NCC pattering by duplicated and sub- or neofunctionalized *ednr*s and *edn*s likely evolved in stem vertebrates. We posit that the deployment of Ednr signaling in NCC, followed by duplication and divergence of Edn signaling components, paved the way for fine-tuned control over NCC fate determination, migration, and patterning in modern vertebrates.

Methods

Gene cloning and sequence acquisition

Polymerase chain reactions (PCRs) were performed according to standard protocols using the primers listed in Tab. S1 (Square et al., 2016a). Fragments corresponding to *P. marinus ednA*, *ednC*, *ednE*, *ednra*, and *ednrb* were amplified from a 5' RACE library made from combined cDNA. These genes were named based on their deduced orthology to *L. japonicum* sequences (see below) except *ednrb*, which has not been identified in *L. japonicum* to our knowledge. Notably, part of the *ednrb* sequence fragment we cloned via RACE had been previously identified (Braasch et al., 2009) in the 2007 *P. marinus* genomic assembly on contig56749, though only ~350 nt of this sequence is represented there. Unfortunately, this contig is only ~4.9 kilobases long, and does not contain any other predicted coding sequences that could be used to determine synteny. This sequence is absent from the 2010 *P. marinus* genomic assembly. Using the nucleotide sequences published for the Japanese lamprey (Kuraku et al., 2010) as a reference, fragments corresponding to the putative *P. marinus ednB, ednD*, and *ednF* transcripts were amplified from genomic DNA or cDNA by designing primers against sequenced transcripts (Bryant et al., 2016). In the case of *ednD*, our *P. marinus* gene fragment appears to be completely situated within the 3' UTR. Similarly, seven *P. marinus* metalloprotease transcript fragments (with sequence similarity to gnathostome ECEs) were identified in transcriptome assemblies derived from adult (Bryant et al., 2016) and combined adult and larval tissues (unpublished data). All sequences have been deposited to GenBank.

As is typical within the tetraploid *X. laevis* genome, we found two duplicates of each *edn* and *ednr* gene, except in the case of *ednrb2*, for which there appears to be three copies (-a: NM_001086238.1; -b: NM_001085878.1; -c: KU680750). *edn1-b, edn3-a, ednra-a, ednra-b, ednrb2-a, and ednrb2-b* were

previously described and/or sequenced (Bonano et al., 2008; Kawasaki-Nishihara et al., 2011). The riboprobe template for *X. laevis ednra-b* was a gift from the Mayor lab; the rest of these were subcloned from gDNA or cDNA using available sequence information. Using *X. laevis* genomic assemblies 7.1-8.0 (www.xenbase.org), we identified and cloned fragments of *edn1-a*, *edn2-a*, *edn2-b*, *edn3-b*, *ednrb1-a*, *ednrb1-b*, and *ednrb2-c* out of genomic DNA or cDNA. These novel *edn* and *ednr* sequence fragments have been deposited to GenBank. ClustalW alignments with translated sequence fragments, and subsequent ML trees confirmed the identity of these *Xenopus* sequences.

In situ hybridizations, cryosectioning, and imaging

Using riboprobes generated from the amplified gene fragments described above, *in situ* hybridization (ISH) was performed as previously described (Cerny et al., 2010; Square et al., 2015a). Developmental staging for lamprey followed Tahara, 1988, while *X. laevis* staging followed Nieuwkoop and Faber, 1994. For simplicity, we use "st. 33," "st. 35," and "st. 37" to describe the *X. laevis* developmental windows usually referred to as "st. 33/34," "st. 35/36," and "st. 37/38," respectively. Interestingly, riboprobes designed against the different *X. laevis*-specific gene duplicates of *ednra* (*ednra-a* and *ednra-b*) and *edn2* (*edn2-a* and *edn2-b*) displayed differential staining despite high sequence similarity and regional overlap of the probe binding sites, as previously noted for other *X. laevis*-specific gene duplicates (Square et al., 2015a) (see below). Cryosectioning and imaging was performed as previously described (Jandzik et al., 2014b; Square et al., 2015a).

Phylogenetic reconstruction

To assign the orthology of the *P. marinus edn* ligand fragments described above to those previously published for *L. japonicum edn* genes, a ClustalW nucleotide alignment was constructed using only endothelin ligand DNA sequences derived from both lamprey species. The Maximum Likelihood (ML) method of phylogeny reconstruction was employed in MEGA6 thereafter (Fig. S1). The *P. marinus* genes were named corresponding to their orthology with their closest *L. japonicum* relative, for which all ortholog pairs show strong support in a bootstrap analysis.

In an attempt to address the identity of the lamprey Endothelin signaling components amongst all vertebrate Endothelin ligands (save EdnB and EdnD), Endothelin receptors, and Endothelin converting enzymes, ClustalW protein alignments were built using inferred aa sequences derived from transcript or genomic data, and thereafter used to build ML trees (Figs. S2, S3, and S4). For the Edn ligands, an ML tree was first built by using aa sequences derived from a variety of gnathostomes (specifically excluding those that were missing the conserved region containing and surrounding the functional ligand), testing many parameters for the aa alignment and ML tree preliminarily. Given the heterogeneity in sequence

conservation across Edn aa sequences, these trees were sometimes structured very differently from each other when different parameters were employed, as reflected by the range of bootstrap values and overall tree topology in Fig. S2. We thus selected parameters that would maximize our tree to reflect the accepted relationships of 1) vertebrate taxa, and 2) the synteny analysis on *edn* genes performed previously (Braasch et al., 2009) which supported *edn1/2* and *edn2/4* clades. We then added *P. marinus* aa sequences to the dataset, realigned, and rebuilt the tree, and bootstrapped it 100 times using these gnathostome-optimized parameters (*P. marinus* EdnB and -D, and *L. japonicum* EdnB, -D, -E, and -F were discounted from this analysis due to their incomplete nature). An endothelin receptor ML tree was similarly constructed, using amphioxus Ednr-like as an outgroup (Fig. S3). To address the identity of putative *P. marinus ECEs* and other related metallopeptidase (Bianchetti et al., 2002; Bland et al., 2008; Rawlings and Barrett, 1995) genes, a phylogeny was built using the inferred translations from each of seven sea lamprey M13 family metallopeptidase transcripts, along with aa sequences for these genes from an assortment of deuterostomes, using three LTA4H sequences as an outgroup (an M1 family peptidase(Rawlings and Barrett, 1995); Fig. S4).

Results

Cloning and sequence analysis of endothelin pathway genes

In order to understand the relative timing of the appearance and diversification of Endothelin pathway gene groups, we gathered available sequences from a wide range of deuterostomes (mainly gnathostomes). No gene with moderate sequence similarity to an Edn has been identified in any invertebrate by our own, or previous analyses (Braasch and Schartl, 2014; Braasch et al., 2009), however *ECE-like1* metalloproteases do appear to exist in amphioxus and urchin (Fig. S4), and an amphioxus *ednr-like* gene has been previously identified (Braasch and Schartl, 2014). It is important to note that ECE-like1 (also known as X-converting enzyme, XCE) is a distinct group of proteins found throughout deuterostomes, and is closely related to, but separate from the clade of proteins containing gnathostome ECE-1 and ECE-2 (Fig. S4). Conversely, the single *Ednr-like* gene in amphioxus appears to be a true ortholog of vertebrate *Ednr*s(Braasch and Schartl, 2014); there is no known clade of "*ednr-like*" genes in vertebrates.

The ML trees generated here fail to strongly support strict orthology of lamprey and gnathostome Edn ligands and ECE peptidases (Figs. S2 and S4). Interestingly, Sea lamprey EdnA clusters with the tetrapod Edn1 group with moderate bootstrap support (71%), however in this analysis, the ray-finned fish Edn1 sequences were excluded from this group. Conversely, *P. marinus* Ednra and Ednrb cluster separately with their putative gnathostome orthologs, with moderate bootstrap support (59%, and 58%

respectively; see Fig. S3). Thus while a parsimonious view would suggest that these two lamprey receptors do belong to these two main groups of Ednrs, it remains possible that one or both of these lamprey Ednrs are actually duplicates that were lost in gnathostomes, which would make them a unique Ednr subtype. Sea lamprey appear to possess five *ECE* genes, which we have named *ECE-A*, *-B*, *-C*, *-D*, and *-E*. Notably, ECE-A and -B cluster with the gnathostome ECE-1/2 clade with high bootstrap support (see Fig. S4), whereas ECE-C, -D, and -E have low support, and thus these latter three independent duplicates might not actually code for enzymes capable of processing Edns. We have named them as such simply because they show the strongest sequence similarity to ECEs. All accession numbers corresponding to these sequenced transcripts can be found in Tab. S2 within (Square et al., 2016a).

Expression of *P. marinus* endothelin signaling components

We assayed the expression of *edn* and *ednr* transcripts at stages 15, 17, and 21-28 via *in situ* hybridizations. This developmental series extends from mid-neurulation until the initial differentiation of the head skeleton. The expression of *L. japonicum ednA*, *ednC*, *ednE*, and *ednra* was previously described at mid-pharyngula stages (Kuraku et al., 2010). While the expression patterns of these genes in these two lamprey species are generally similar, our analysis revealed additional expression domains not apparent in *L. japonicum*. These are detailed below.

Expression of ednA-F

ednA expression was first detected in ectoderm surrounding the forming mouth (stomodeum) weakly at st. 22, and more strongly at st. 22.5 (Fig. 1A). This expression domain was also found in *L. japonicum* (Kuraku et al., 2010). This ectoderm is fated to eventually cover the nasohypophyseal plate, contributing to the epithelium of the nostril (Kuratani et al., 2001). At st. 23.5, expression around the mouth appeared in a more distinct "comma" shape, and expression in the anteriormost dorsomedial pharyngeal arches (PAs) was first detected (Fig. 1B). Through st. 25.5, *ednA* expression in the pharynx expanded to the posteriormost PAs, while transcription became reduced around the mouth and in the anteriormost PAs (Fig. 1C). Sectioning at this stage revealed expression in both the center of the posteriormost seven PAs (in the mesodermal 'core', Fig 1C') as well as the ectoderm overlying these PAs. At st. 26.5, expression was reduced to a small patch of mesoderm within each of these PAs (Fig. 1D). This expression persisted until st. 27.5, also reappearing in the region of the dorsal PA1 (not shown). *ednB* expression was detected weakly in the upper lip, ear, heart, and somites at st. 25.5, and later in the brain of st. 28 larvae (Fig. S5A and B). *ednC* expression was first strongly detected at st. 21.5 along the flank in the ectoderm, and weakly in a medial spot within the anterior brain and in the ectoderm overlying lateral portions of the head (Fig. 1E). By st. 23, the ectodermal flank expression largely faded, while the

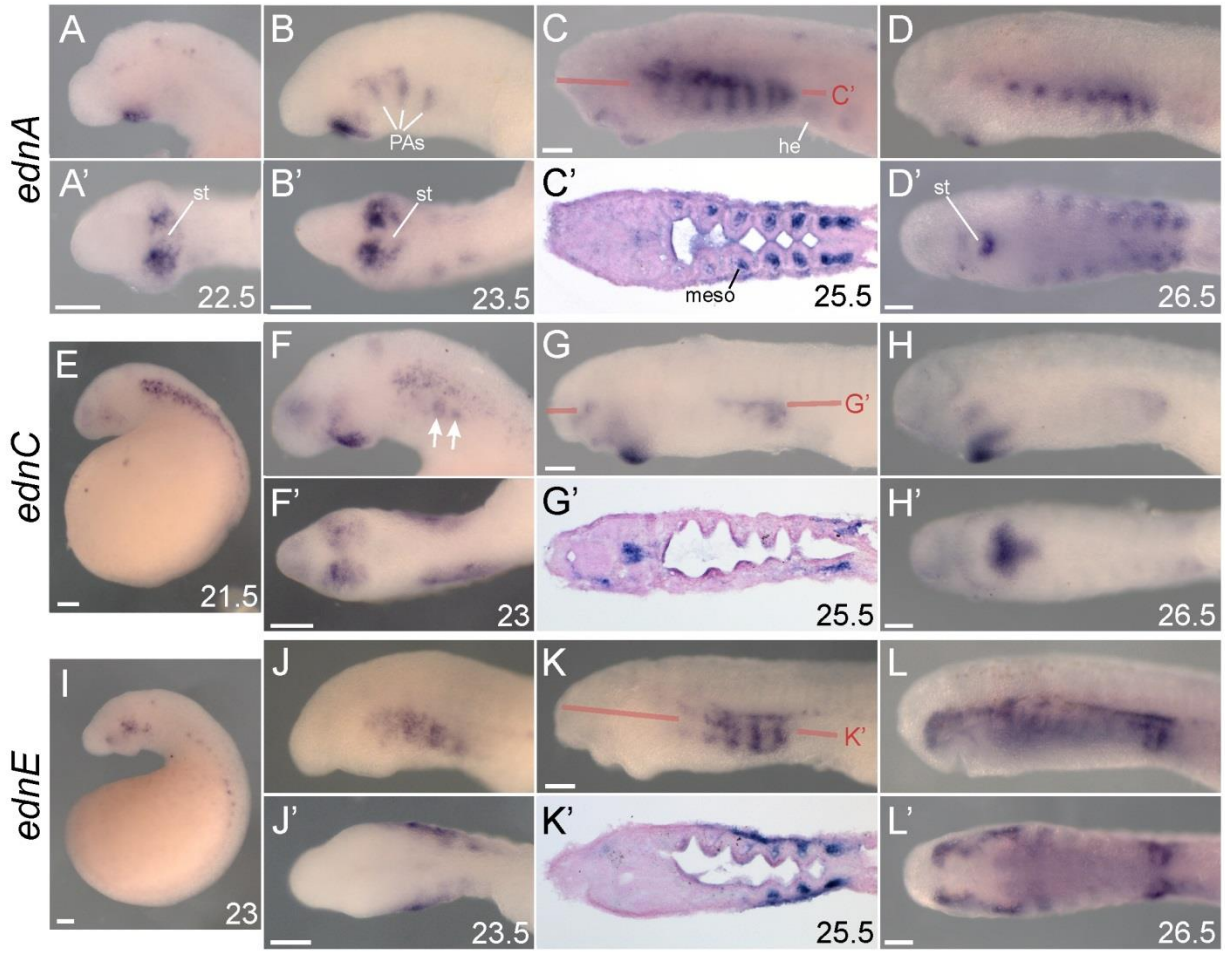

Fig. 1. Expression summary of *edn*s in *P. marinus*. Left lateral views in all non-prime panels. All prime lettered panels show ventral/oral views of the correspondingly lettered panel, save C', G', and K' which show horizontal sections. Developmental stage (Tahara, 1988) is indicated for each specimen in the bottom right of a panel. (A-D) *ednA* expression in *P. marinus* surrounds the stomodeum and marks the pharyngeal arches. (E-H) *ednC* expression in *P. marinus* is found around the stomodeum and on the lateral sides of the head. Arrows in F point to mesenchymal *ednC* expression below the superficial expression in the ectoderm. (I-L) *ednE* expression in *P. marinus* is found along the somites, and throughout the facial and pharyngeal skeleton. All scale bars represent 100 μm. he, heart; meso, mesoderm; PAs, pharyngeal arches; st, stomodeum.

brain expression expanded, and the expression flanking the head shifted more ventrally around the stomodeum to occupy a similar region as *ednA* transcripts (compare Fig. 1B' to 1F'). Deeper mesenchymal expression in the posterior future PAs was also first detected at st. 23 underneath the faded flank expression (arrows in Fig. 1F). At st. 25.5, expression around the mouth shifted to mesenchyme of the lower lip, but still remained in tissues on the lateral sides of the mouth (Fig. 1G); only this lower lip expression was previously characterized in *L. japonicum* (Kuraku et al., 2010). Sectioning revealed that all expression at st. 25.5 was mesenchymal (Fig.1G'). At stages 26.5 (Fig. 1H) and 27.5 (not shown), the posterior PA expression was progressively diminished, and the more anterior expression around the mouth and in PA1 became progressively refined. *ednD* expression was not detected at any embryonic or larval stage assayed here. *ednE* transcripts were first visualized weakly at st. 22 in head ectoderm (not shown), and more robustly in the st. 23 head ectoderm and along the boundary between the anterior yolk and the somites (Fig. 1I). Up to st. 25.5, this superficial PA expression in the head proceeded in an anterior to posterior fashion, and was only present in the most posterior PA by st. 28 (Fig. 1J-L; st. 28 not shown). At st. 25.5 expression began in deeper mesenchymal tissues within the PAs (Fig. 1K'), and in a horizontal stripe dorsal to the PAs, presaging pigment deposition (Fig. 1K); this latter expression domain was previously described in *L. japonicum* (Kuraku et al., 2010). *ednF* expression was detected at st. 25.5 diffusely throughout the entire head, but by st. 28 was mainly expressed in the brain (Fig. S5C-F).

Expression of P. marinus ednra *and* ednrb

P. marinus ednra was first detected in a small paired patches of anterior mesoderm at st. 21 (Fig. S5G and H), and more robustly at st. 22 (Fig. 2A). At this stage, expression also began at the base of the pharynx where the heart will eventually develop (arrowhead in Fig. 2A). These expression domains were maintained until st. 27. ISH revealed *ednra* expression in late migratory and post-migratory skeletogenic NCCs in the head beginning at st. 24, first staining the ventralmost NCCs (future mucocartilage), as well as the upper and lower lips, and the more anterior PAs (Fig. 2B). PA expression thereafter proceeded to initialize in an anterior to posterior wave (Fig. 2B-D). By st. 26.5, this NCC expression was found throughout all PAs (Fig. 2D and D'). At st. 26.5 *ednra* was also detected in nascent lateral plate mesoderm (arrows in Fig. 2D), similar to expression of lamprey *Lbx-A* (Kusakabe et al., 2011). *ednra* transcription was maintained throughout st. 29, strongly marking the future head skeleton as cartilage began to differentiate (Fig. 2E). Most of these expression domains have also been observed in the Japanese lamprey (Kuraku et al., 2010). *ednrb* expression was first detected at st. 21 in pre- and early migratory NCCs (Fig. 2F). By st. 22, these *ednrb*-positive cells in the head were migrating ventrally towards their destinations in the pharynx (Fig. 2G), while the majority of those in the trunk were still

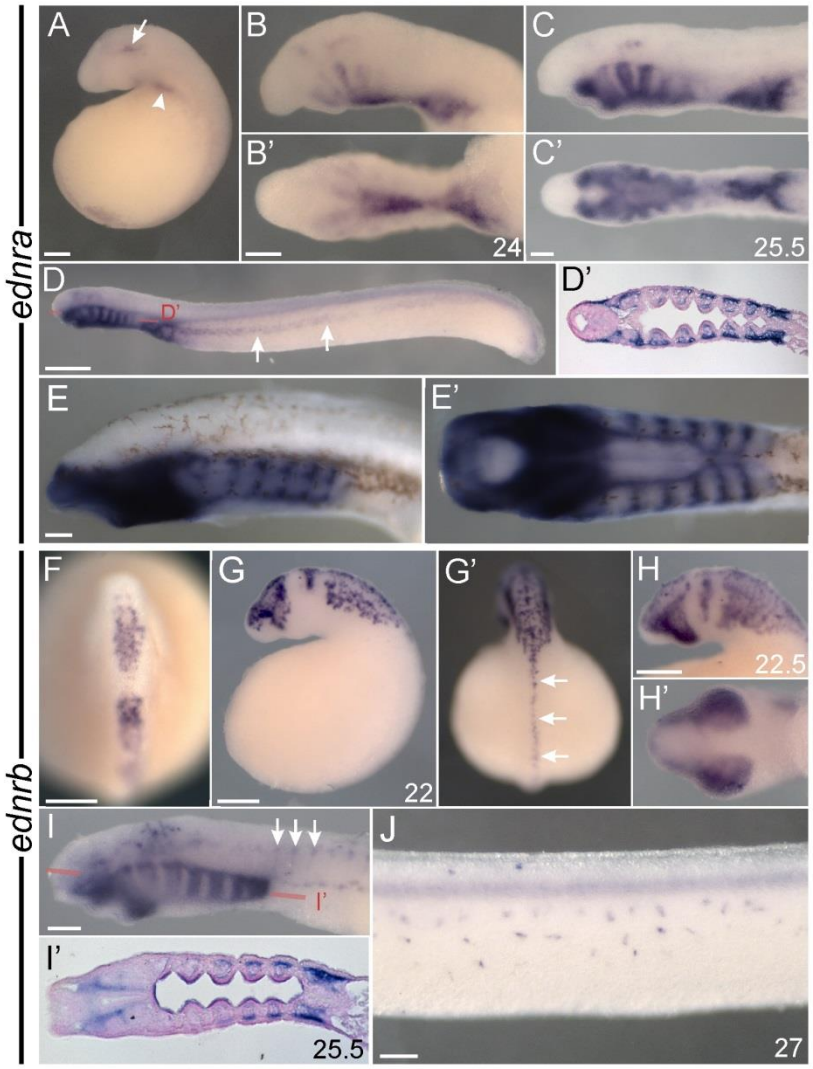

Fig. 2. Expression summary of *ednr*s in *P. marinus*. Left lateral views in all non-prime panels save F, which is a dorsal view with anterior to the top. All prime lettered panels show a ventral view of the correspondingly lettered panel save D' and I', which are horizontal sections, and G', which is a dorsal view with anterior facing away from the page. Developmental stage(Tahara, 1988) is indicated for each specimen in the bottom right of a panel. (A-E) *ednra* expression in *P. marinus* is found in mesoderm, the heart, and the post-migratory future head skeleton. Arrow in A indicates anterior mesoderm expression of *ednra.* Arrowhead in A indicates *ednra* expression in the nascent heart. Arrows in D indicate *ednra* expression consistent with lateral plate mesoderm. Red line labeled D' in D indicates the approximate plane of section in D'. Note expression in the future head skeleton (CNCCs) shown in D'. (F-G) *ednrb* expression in *P. marinus* marks migratory NCCs. Arrows in I indicate expression consistent with dorsal root ganglia. Arrowhead in I indicates expression consistent with the earliest migrating pigment cells. Red line labeled I' in I indicates the approximate plane of section in I'. Note expression in the future head skeleton (CNCCs) shown in I'. Arrows in G' indicate trunk NCCs still poised at the neural crest. All scale bars represent 100 μm, save the scale bar in panel D which is 500 μm. NCCs, neural crest cells; st, stomodeum.

poised at the dorsal neural tube (arrows in Fig. 2G'). This migratory NCC expression continued through st. 24, becoming progressively extended ventrally and more diffuse as more NCCs had begun their migration (st. 23 shown in Fig. 2H). *ednrb* transcripts at st. 25.5 were detected around the mouth and in pre-skeletal NCCs in the PAs (Figs. 2I and I'), as well as future pigment cells (arrowhead in I) and peripheral nervous system components scattered along the lateral sides of the head and anterior yolk and somites, including those that resemble dorsal root ganglia (arrows in Fig. 2I). At st. 26.5 expression in the pharynx was largely lost, though the expression in pigment cells was still apparent (not shown). Bleached larvae revealed this expression persists in melanophores until st. 28, being found in cells on the dorsal ridge and flank of the animals (st. 27 shown in Fig. 2J).

Expression of X. laevis endothelin signaling components

We performed ISH for all genes spanning st. 17 to st. 40, (mid-neurulation to the onset of head skeleton differentiation). Except in the cases of *edn2* and *ednra*, all expressed *X. laevis*-specific gene duplicates analyzed here were discovered in completely overlapping domains at all stages assayed. We failed to clone the *X. laevis* duplicate *ednrb2-b*.

Expression of X. laevis edn1-3

X. laevis edn1 was first detected at st. 19 in two patches of ectoderm just above the future stomodeum, overlying the developing telencephalon (Fig. 3A; see Fig. S6A-C for st. 28 sections). This expression persisted until st. 33, eventually becoming restricted to the ectoderm surrounding each nasal placode. At st. 26 we first observed *edn1* transcription within the ventral PAs, beginning in the region of the ventral hyoid stream. At later stages, expression of *edn1* spread posteriorly, then anteriorly to the rest of the ventral PA endoderm, mesoderm, and ectoderm, but not NCC-derived mesenchyme (Figs. 3B-D and S6D-F). By st. 37 *edn1* was detected ventrally in each PA (Fig. 3D). To assay the expression of *edn2-a*, first a probe corresponding to 387 bp of coding sequence was amplified from cDNA ("cds probe"). This probe produced signal in many different tissues in the head, though this was found to be largely background staining in the brain cavities and in the ear, and possibly also in the notochord (Fig. S6H, I). Sectioning revealed that mesenchymal staining within the eye and surrounding the eye at st. 35 and 37 was within cells. We were unable to clone a fragment of the identified *edn2-b* coding region out of embryonic cDNA. Thereafter, riboprobes were designed mainly against a portion of the 3' UTR of each *edn2-a* (645 nt) and *edn2-b* (684 nt), also including the last ~70 bases of coding sequence ("3' UTR probe"). These were amplified from genomic DNA. Using the UTR probes, *edn2-a*, but not *edn2-b* was detected within the future pronephros from st. 30-37 (Fig. S6G), subsequently in dorsal pharyngeal

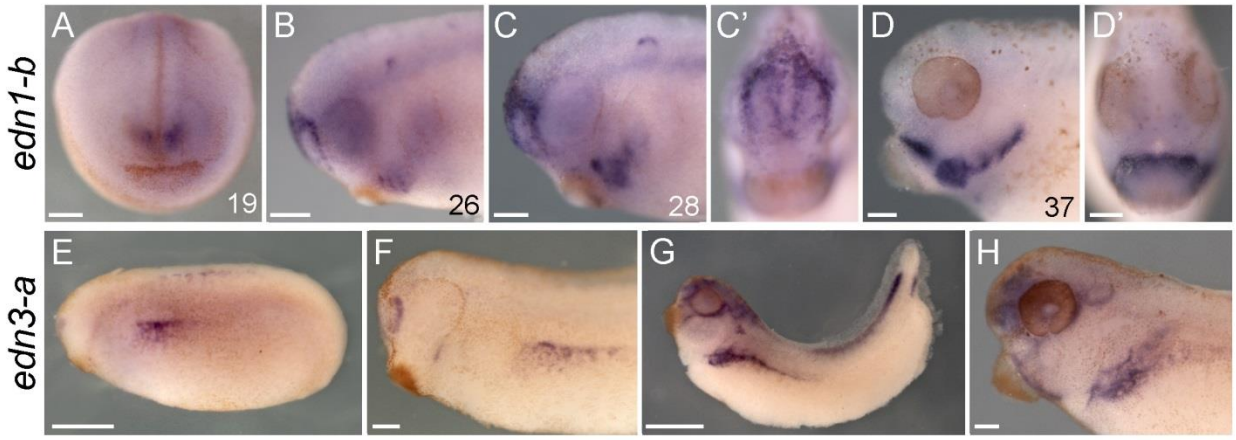

Fig. 3. Expression summary of *edn*s in *X. laevis*. Left lateral views in all non-prime panels save A, which is an anterior view. Prime panels show anterior views of the correspondingly lettered panel. Developmental stage(Nieuwkoop and Faber, 1956) for each specimen is indicated in the bottom right corner of a panel. (A-D) *edn1-b* expression in *X. laevis* marks ectoderm around the nasal placodes, and ventral non-NCC mesenchyme and epithelia in the pharyngeal arches. (E-H) *edn3-a* expression in *X. laevis* marks ectoderm where future *ednrb2*-positive cells will migrate. All scale bars represent 100 μm, save those in E and G which represent 500 μm.

mesenchyme surrounding the eye from st. 35-37, and within the eye. Interestingly, our *edn2-a* 5' UTR and cds probes produced different staining patterns, despite having been designed to detect the same transcript. This may reflect differences in splice variants of the *edn2-a* gene. *edn2-b* transcripts were never detected. *edn3* expression was first seen at st. 23 surrounding in the nasal placodes, along the flank, and in the anterior trunk of the neural tube (Fig. 3E). Expression along the flank expanded and was maintained in lateral plate mesoderm surrounding the pronephros through st. 37, and ceased thereafter (Figs. 3E-H and S6K). ISH staining around the dorsal neural tube proceeded in an anterior to posterior fashion, and was no longer detected by st. 37. In the head, expression in the nasal placode expanded and was joined by other expression in mesenchyme overlying the border between PAs 1 and 2 and around the eye (Fig. 3G and H).

Expression of X. laevis endothelin receptors

As previously described, *ednra-a* but not *ednra-b* expression was detected along the neural plate border during mid-neurulation (Fig. 4A). Thereafter, transcripts from both *ednra-a* (not shown) and *ednra-b* (Fig. 4B) genes were detected in completely overlapping regions. Expression was found in the otic placodes, uniformly throughout migratory cranial neural crest at st. 23, and along the flank in migratory trunk NCCs. These trunk NCCs could later be seen positioned between the boundaries of the somites, and aggregating along the future lateral line (Fig. 4C). By st. 31, expression had begun in the heart region, which continued throughout the entire series assayed here (Fig. 4C-E; st. st. 38-40 not shown). At stages 33-37, transcripts in the PA NCCs were apparently depleted ventrally compared to dorsal mesenchyme (Fig. 4D and E). Transcripts were also detected strongly in the medial fin at stage 40 (Fig. 4G). *ednra* transcripts were also detected in nephrostomes at stages 30-40 (Fig. 4D-F; arrows in Fig. S6R and S). *ednrb1* expression was observed only in the medial PAs from st. 31-40, and in the eye from st. 35-40 (Figs. 4H-J and S6L-N; stages 38-40 not shown). *ednrb2-a* and *ednrb2-c* were both detected in pre-migratory and migratory nascent pigment cells, cranial ganglia, the forebrain, and peripheral nervous system derivatives of the trunk (Figs. 4K-M and S6O-Q).

Discussion

Duplication and specialization of endothelin signaling pathways occurred in stem vertebrates

To better understand the ancestral roles of the endothelin pathway in vertebrate development, we performed an exhaustive search of lamprey genomic and transcriptome sequences for *edn* and *ednr* homologs. We then determined their expression throughout early development. To facilitate comparisons with gnathostomes, we also characterized and analyzed the embryonic and larval expression of all *edn*s

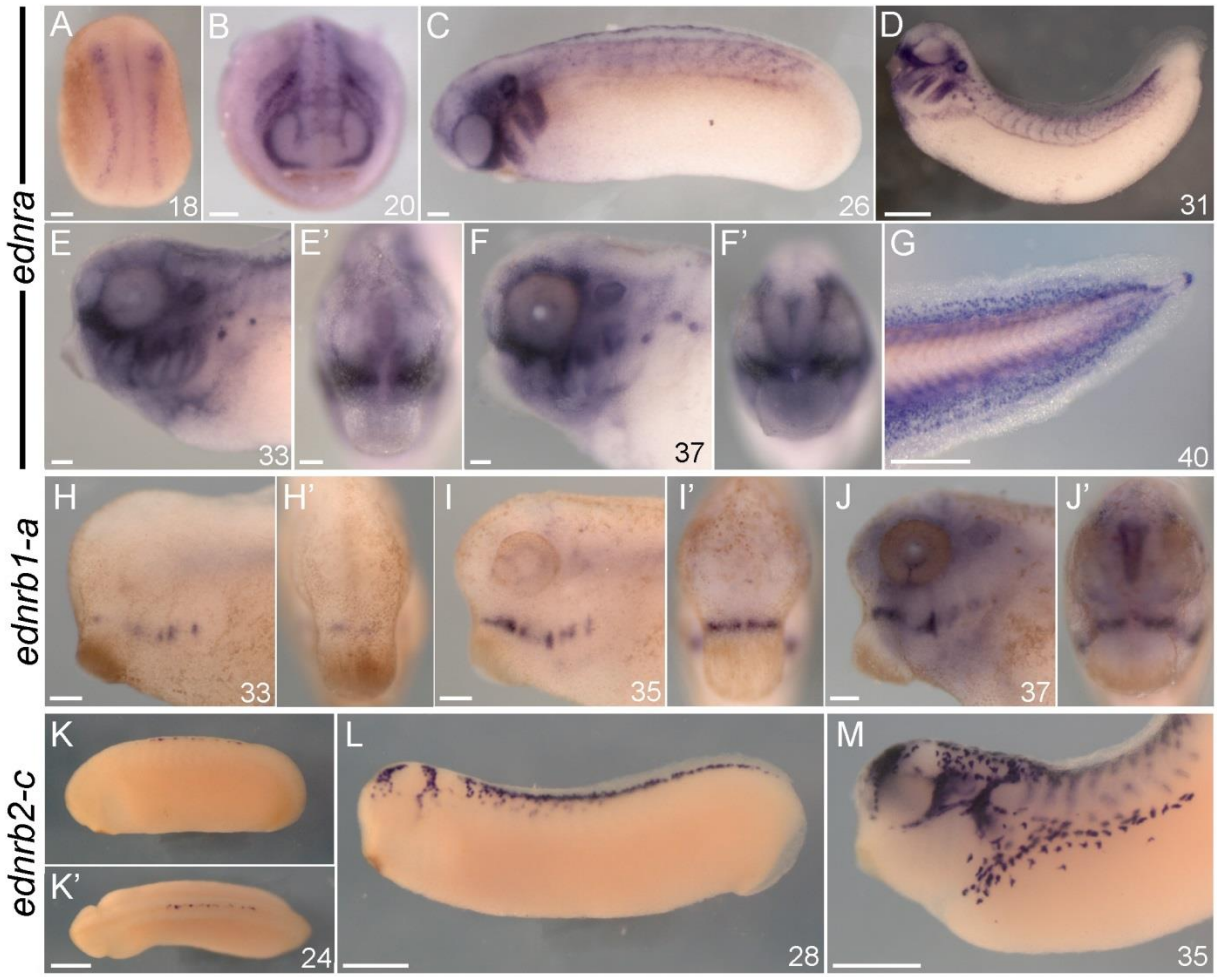

Fig. 4. Expression summary of *ednr*s in *X. laevis.* Left-lateral views in all non-prime panels save A and B, which are dorsal (A) an anterior (B) views. All prime panels show anterior/oral views of the correspondingly lettered panel save K', which is a dorsal view. Developmental stage(Nieuwkoop and Faber, 1956) for each specimen is indicated in the bottom right corner of a panel. (A-G) *ednra* expression in *X. laevis.* (A) *ednra-a* uniquely marks pre-migratory NCCs. (B-G) *ednra-b* expression in migratory and post-migratory NCC derivatives, nephrostomes, the heart, and fin mesenchyme. (H-J) *ednrb1-a* expression in *X. laevis* marks the eye and medial pharyngeal arch NCCs. (K-M) *ednrb2-c* expression in *X. laevis* marks migratory pigment cells, nerves, and the forebrain. All scale bars represent 100 μm, save those in D, G, K', L, and M which represent 500 μm.

and *ednr*s in *X. laevis*. To our knowledge, this is the first report detailing the embryonic and larval expression of all *edn*s and *ednr*s in a single gnathostome.

Phylogenetic analysis lends weak support to a scenario wherein a single ancestral *ednr* gene was duplicated in stem vertebrates, with *ednrb* later duplicated in stem gnathostomes (Fig. S3). Our data showing similar expression of lamprey and gnathostome *ednra* and *ednrb* paralogs supports this, with *ednra*s marking the heart and post-migratory skeletal NCCs, and *ednrb*s marking migrating pigment cells, cranial glia, and other peripheral nervous system precursors. Overall, duplication and functional specialization of *ednr*s is similar to that of other regulators of NCC development, including SoxE (McCauley and Bronner-Fraser, 2006), Tfap2 (Van Otterloo et al., 2012), Dlx (Cerny et al., 2010), and Id (Kee and Bronner-Fraser, 2005; Liu and Harland, 2003), which all have multiple paralogs expressed in NCC. Furthermore, like SoxE (McCauley et al., 2014), Tfap2 (Van Otterloo et al., 2013; Van Otterloo et al., 2012), Id (Kee and Bronner-Fraser, 2005; Liu and Harland, 2003), and FoxD (Ono et al., 2014), there is evidence that duplication of *ednr*s was accompanied by biochemical subfunctionalization and/or neofunctionalization. In mouse, Ednra binds most strongly to Edn1 and Edn2, while mouse Ednrb (a type B1 receptor) binds all Edns with similar affinity (Yanagisawa, 1994); in *X. laevis* Ednrb2 preferentially binds Edn3 (Karne et al., 1993). Furthermore, rescue experiments have shown that mouse Ednrb cannot rescue the function of Ednra in maxillary skeletogenic cranial NCC, due to the inability of Ednrb to signal through $G\alpha_q/G\alpha_{11}$ proteins (Sato et al., 2008a). Whether these differences reflect the loss of Ednrb functionality or a gain of Ednra functionality awaits testing of amphioxus Ednr, the only invertebrate Ednr known, and the only appropriate outgroup.

It is unknown if lamprey Ednra and Ednrb have diverged with regard to ligand binding affinities or signal transduction outputs. However, their divergent expression patterns support the hypothesis that distinct Ednra and Ednrb signaling pathways evolved in stem vertebrates. Because of the unique roles different Ednrs play in NCC migration and fate determination, it is tempting to speculate that duplication and divergence of vertebrate *ednr*s facilitated the evolution of the highly patterned and multipotent NCC of modern jawed and jawless vertebrates.

Given that endothelin cleaving enzymes are responsible for the final processing of endothelin ligands, their appearance was a critical evolutionary step in the advent of the first vertebrate-type Endothelin pathway. In an attempt to understand when the first ECE appeared, we analyzed M13 metallopeptidases from a variety of deuterostomes. Given the available genomic and transcriptomic resources, we were unable to identify coding sequences most closely resembling ECE-1/2 in any invertebrate. However we did find coding sequences resembling the closely related ECE-like1 and Neprilysin in amphioxus (*Branchiostoma floridae*) and sea urchin (*Strongylocentrotus purpuratus*), and PHEX in sea urchin only. It seems probable then that the gene duplications giving rise this suite of genes

(*ECE-like1*, *Nep*, and *PHEX*) predate the divergence of echinoderms and chordates. Given that recombinant in vitro experiments have shown that ECE-like1 is unable to cleave Big Endothelin 1 in an in vitro assay (Valdenaire and Schweizer, 2000), and Neprilysin degrades Big Endothelin 1 in a manner that yields no detectable Edn1 ligand (Abassi et al., 1993), it seems plausible that these invertebrate genes identified here do not have a role in any type of rudimentary Endothelin ligand processing. Thus, in combination with the apparent absence of any *edn*-like gene outside of vertebrates (Braasch et al., 2009), there is no evidence that true Endothelin ligands or Endothelin converting enzymes exist outside of vertebrates based on deposited these deuterostome sequences. Despite the overall sequence similarity of amphioxus Ednr-like to vertebrate Ednrs, this invertebrate Ednr exhibits many divergent features, including the absence of one highly conserved lysine residue within the 2nd transmembrane domain known to be important for ligand binding (Adachi et al., 1994a; Braasch and Schartl, 2014). While it is possible that invertebrate chordate *edn*s or *ECE*s may have been overlooked during genome sequencing or annotation, the available data strongly suggest that early vertebrates greatly expanded on the elaborate biochemistry underlying these pathways. If no modern invertebrate ECE or Edn ligand truly exists, this would suggest that the Endothelin signaling pathway itself is a vertebrate synapomorphy, which thereafter diversified into a complex system of different ligands, processing enzymes, and receptors before the divergence of cyclostomes and gnathostomes.

Differences in *ednr* expression reveals flexibility in the timing and extent of ednr transcription in NCC

While *P. marinus* and gnathostomes express their *ednra* and *ednrb* paralogs similarly, there are some clear differences (see Tab. 1). For one, lamprey *ednra* is restricted to NCCs in the head, while zebrafish (Nair et al., 2007), *X. laevis*, chicken (Nataf et al., 1998), and possibly mouse (Yanagisawa et al., 1998) all express *ednra* in trunk NCCs. Furthermore, while zebrafish (Nair et al., 2007), Xenopus, and mouse (Clouthier et al., 1998; Yanagisawa et al., 1998) express *ednra* in some migrating NCCs, lamprey and chicken (Nataf et al., 1998) do not express *ednra* in NCCs until they have reached their destinations. Whether these differences reflect an expansion of the ancestral *ednra* expression pattern in gnathostomes, or a restriction of *ednra* to post-migratory cranial NCC in lamprey and chicken is unclear.

P. marinus expresses *ednrb* more broadly in NCCs than any gnathostome, with strong expression apparent in the peripheral nervous system, pigment cells, and all skeletogenic NCCs during and briefly after migration. By contrast, Zebrafish (Arduini et al., 2009; Lister et al., 2006) and mouse (RM et al., 2006; Shin et al., 1999) restrict *ednrb*-type expression to pigment cells and peripheral nervous system derivatives, though early zebrafish *ednrbb* expression has not been well-described, leaving open the possibility of more widespread *ednrb* expression in this species. *X. laevis* expresses *ednrb*-type receptors

in pigment cells, peripheral nervous system derivatives, and a minor subset of post-migratory skeletogenic NCCs. As in lamprey, avian *ednrb1* (Nataf et al., 1996; Pla and Larue, 2003) appears to broadly mark early migratory cranial NCCs, although it is unclear if this expression persists after those cells reach the pharynx. Based on these comparisons, we speculate that lamprey-like *ednrb* expression in all, or most, migrating NCC represents the ancestral vertebrate state, with the loss of some expression domains occurring after duplication of *ednrb* in the gnathostome lineage.

Tab. 1 A summary of *ednr* subfunctionalization in vertebrates.

	Migratory, pre-skeletal NCCs	**Post-migratory, Pre-skeletal NCCs**	**Migratory trunk NCCs (non-melanophores)**	**Post-migratory DRGs**	**Heart (post-migratory NCCs?)**	**Migratory and post-migratory melanophores**
Sea lamprey	*ednrb*	*ednrb*, *ednra*	*ednrb*	*ednrb*	*ednra*	*ednrb*
Zebrafish	*ednra1*(Nair et al., 2007)	*ednra1* + *ednra2*(Nair et al., 2007)	*ednra1*(Nair et al., 2007)	*ednrb1*(Arduini et al., 2009; Lister et al., 2006)	*ednra*(Nair et al., 2007)	*ednrb1*(Arduini et al., 2009; Lister et al., 2006)
Xenopus laevis	*ednra*	*ednra* + *ednrb1* (minor subset)	*ednra*	*ednrb2*	*ednra*	*ednrb2*
Chick/quail	*ednrb1*(Nataf et al., 1996)	*ednra*(Nataf et al., 1998)	*ednrb1*(Nataf et al., 1996)	*ednrb1*(Nataf et al., 1996) + *ednrb2*(Lecoin et al., 1998)	*ednra*(Nataf et al., 1998)	*ednrb2*(Lecoin et al., 1998)
Mouse	*ednra*(Clouthier et al., 1998; Yanagisawa et al., 1998)	*ednra*(Clouthier et al., 1998; Yanagisawa et al., 1998)	*ednra*(?)(Yanagisawa et al., 1998)[33] + *ednrb1*(RM et al., 2006; Shin et al., 1999)	*ednrb1*(RM et al., 2006; Shin et al., 1999)	*ednra*(Clouthier et al., 1998; Yanagisawa et al., 1998)	*ednrb1*(RM et al., 2006; Shin et al., 1999)

Tab. 1 note: Genes are separated by commas to indicate a temporal change, while a + sign indicates they are coexpressed. For *X. laevis*, "*ednra"* indicates that both the "-*a*" and "-*b*" copies are coexpressed. All zebrafish duplicates are listed by their specific name. Mouse *ednrb* is referred to here as *ednrb1* since it is an *ednrb1* gene (see fig S3). DRGs, dorsal root ganglia; NCCs, neural crest cells.

ednrb duplicates in X. laevis highlight divergent subfunctionalization in different gnathostome lineages

Recent phylogenetic analysis suggests that *ednrb* was duplicated in stem gnathostomes, giving rise to *ednrb1* and *ednrb2* (Braasch and Schartl, 2014). Most modern gnathostomes retain both duplicates,

with the exception of zebrafish and therian mammals, which have both lost *ednrb2* (Braasch and Schartl, 2014). In *X. laevis*, *ednrb2* is expressed in pigment cells, while *ednrb1* is only expressed in the eye, and a small population of post-migratory NCCs in the intermediate domain of the PAs, reminiscent of *dlx4* (Square et al., 2015a). In quail and chicken embryos, *ednrb2* is also expressed in pigment cells. However, unlike *X. laevis*, both avian *ednrb* receptors are expressed in dorsal root ganglia and Schwann cells, though at different times (Lecoin et al., 1998; Nataf et al., 1996). Interestingly, both mouse and zebrafish appear to have compensated for the loss of *ednrb2* by expressing *ednrb1* in NCC-derived pigment cells (Arduini et al., 2009; Lister et al., 2006; RM et al., 2006; Shin et al., 1999), an expression domain not seen in any other gnathostome examined to date. The differential ligand binding properties of *ednrb1* (Yanagisawa, 1994) and *ednrb2* (Karne et al., 1993) open up the possibility that these divergent expression patterns may reflect lineage specific-differences in the migration patterns and/or fate of *ednrb*-expressing NCC, in particular, pigment cells.

Endothelin receptor expression in pre-migratory neural crest supports an ancestral function in NCC specification

Only *P. marinus* and *X. laevis* transcribe an *ednr* in pre-migratory NCC in a pattern reminiscent of neural crest specifiers such as *Sox9*/*SoxE2* and *FoxD3*/*FoxD-A* (Bonano et al., 2008; Sauka-Spengler et al., 2007). Work in *X. laevis* suggests a role for early Ednra signaling in the maintenance of NCC identity and regulation of neural crest specifiers (Bonano et al., 2008). In *P. marinus*, *ednrb* rather than *ednra* is expressed broadly in premigratory NCC (Fig. 2H), however it is unknown if it performs a similar function.

Similar expression domains support shared function of some lamprey and gnathostome Endothelin ligands

Despite poor sequence similarity, gene expression patterns lend support to shared functions of some lamprey and gnathostome Endothelin ligands. Specifically, gnathostome *edn1* and lamprey *ednA* have similar expression patterns around the future mouth, pharynx, and in the presumptive nasal epithelium (Kuratani et al., 2001) (Fig. 5; arrows in Fig. 5B and F), while gnathostome *edn3* and lamprey *ednE* have broadly similar expression in domains populated by *ednrb*-expressing melanophores (Fig. 6). While these similarities may be due to the direct orthology of these genes, it could also be due to similar regulatory subfunctionalization of non-orthologous Edns produced by duplications before or after the gnathostome/agnathan split. Such convergent subfunctionalization could have been facilitated by the organization of the ancestral Edn cis-regulatory landscape.

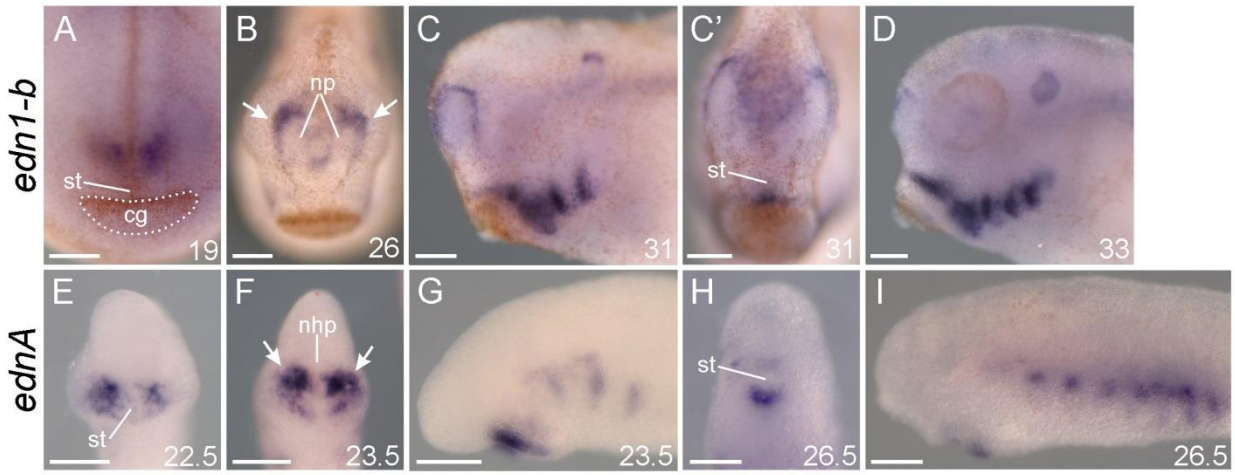

Fig. 5. Lamprey *ednA* expression is reminiscent of gnathostome *edn1* expression, despite low support for strict orthology. Oral/anterior views in A, B, C', E, F, and H. Left lateral views in C, D, G, and I. Developmental stage(Nieuwkoop and Faber, 1956; Tahara, 1988) for each specimen is indicated in the bottom right corner of a panel. (A-D) *X. laevis* expression of *edn1-b* from stage 19 to 33. Dotted outline in A surrounds the cement gland. Arrows in B indicate expression in ectoderm which will eventually contribute to the nostrils (E-I) *P. marinus* expression of *ednA* from stage 22.5 to 26.5. Arrows in F indicate expression in ectoderm which will eventually move rostrally to cover the nasohypophyseal plate, contributing to the future medial nostril. In a similar sequence, both genes are expressed in bilateral patches above the future mouth, surrounding the nasal placodes/nasohypophyseal plate, in pharyngeal arch mesoderm and ectoderm, and just below the stomodeum. All scale bars represent 100 μm. cg, cement gland; nhp, nasohypophyseal plate; np nasal placode; st, stomodeum.

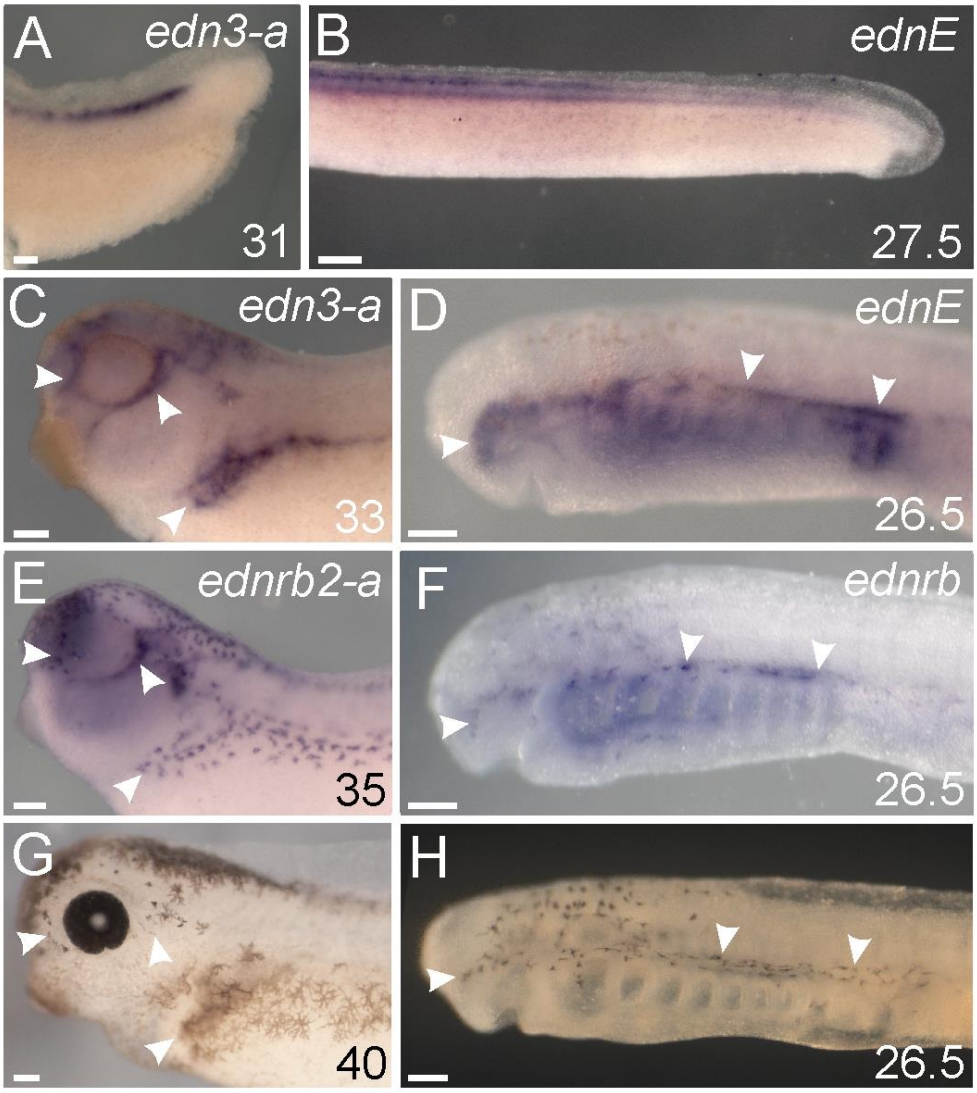

Fig. 6. Lamprey *ednE* and gnathostome *edn3* expression patterns both predict neural crest-derived melanophore migration routes. All panels show left lateral views. Developmental stage (Nieuwkoop and Faber, 1956; Tahara, 1988) for each specimen is indicated in the bottom right corner of each panel. ISH panels (A-F) are labeled with the targeted gene transcript in the top right corner. Arrowheads in C and D mark regions where *edn3-a* (C) or *ednE* (D) are expressed. Arrowheads in E and F indicate cells expressing *ednrb2-a* (E) and *ednrb* (F). Panels G and H show light microscopy photos of larval *X. laevis* (G) and *P. marinus* (H). Arrowheads in G and H indicate areas where pigment cells migrated. All scale bars represent 100 μm.

Phylogenetic analysis suggests a fourth Endothelin ligand gene, *edn4*, was lost in lobe-finned fish; its embryonic expression pattern has not been described. It is possible one of the other lamprey *edn*s is an *edn4* ortholog. Embryonic *edn2* expression is only known in *X. laevis* (this work), and does not strongly resemble any of the lamprey *edn*s alone (save the dorsal pharyngeal expression found in most lamprey *edn*s). Further analyses of *edn2* and *edn4* may resolve these apparent discrepancies.

Differences in edn1/ednA expression correlate with differences in gnathostome and agnathan head skeleton patterning

While *edn1 and ednA* have broadly similar expression patterns, there are some clear differences. In gnathostomes, *edn1* is expressed in the ectoderm, mesoderm, and endoderm of the ventralmost pharynx, where it acts through Ednra to drive expression of the ventral specifier genes *hand1* and *dlx5/6* (Clouthier et al., 2010; Clouthier et al., 1998; Miller et al., 2000; Nair et al., 2007; Sato et al., 2008b). In lamprey, *ednA* transcripts are restricted to the pharyngeal mesoderm and ectoderm in the dorsal-intermediate domain of the pharynx. Unlike gnathostomes, this expression does not abut *hand*-expressing NCC in ventral pharynx (Cerny et al., 2010) (Fig. 1). The significance of this positional difference in *ednA* and *edn1* is unclear, though it is tempting to speculate that it may relate to differences in the dorso-ventral patterning of the pharyngeal skeleton between these groups, namely the pronounced asymmetry in gnathostome PAs as compared to lamprey.

Conclusions

To our knowledge, lamprey has the most *edn* ligands expressed during early development of any vertebrate. This could be the result of cyclostome- or lamprey-specific duplications, and/or the loss of *edn* ligand genes in gnathostomes. Regardless, the presence of multiple *edn*s and *ednr*s in *P. marinus* strongly suggests that sophisticated Endothelin signaling in NCCs predates the split between modern jawed and jawless vertebrates. Given the essential involvement of these pathways in NCC guidance and differentiation, we posit that Endothelin receptor duplication and subfunctionalization played a role in the evolution of the multipotent and highly patterned NCC of modern vertebrates. New methods for high-efficiency mutagenesis (Square et al., 2015b) in lamprey should allow us to better deduce the ancestral roles of Edn signaling in vertebrates and test this hypothesis.

CHAPTER V: CRISPR-mediated mutagenesis in the sea lamprey

Introduction

The first vertebrates had simple cartilaginous head skeletons that lacked proper joints. Today, only two groups of jawless (agnathan) vertebrates survive: lampreys and hagfish. Comparisons between lamprey, hagfish, and jawed vertebrates (gnathostomes) have been used for over a century to identify ancestral vertebrate features and deduce potential gnathostome novelties. However, due to the inaccessibility of hagfish, which live and spawn on the ocean floor (Ota et al., 2007), most studies of agnathan biology have focused on lampreys. Work on lamprey embryos, larvae, and adults have provided important insights into the evolution of key vertebrate features including the adaptive immune system (Pancer et al., 2004), the endocrine system (Sower et al., 2009), regenerative capacity (Smith et al., 2011), the head skeleton (McCauley and Bronner-Fraser, 2006) and large-scale gene duplications (Smith et al., 2013). There is also high interest in understanding lamprey biology to support efforts to conserve native lamprey species, and control invasive lampreys (Close et al., 2010; reviewed by Sower, 2003).

Among lamprey species, the sea lamprey, *Petromyzon marinus*, has been used extensively to study the evolution of vertebrate development (Beamish, 1980; McCauley and Kuratani, 2008; Piavis, 1971). Several features make the sea lamprey particularly well-suited for such work. During their summer spawning season, large numbers of sexually mature adults can be trapped in shallow freshwater streams. Ripe, captured females can be held and manually stripped of tens of thousands of eggs over the course of several days (Nikitina et al., 2009). Once fertilized, lamprey embryos develop slower than most gnathostomes, allowing for fine-grained staging. Lamprey embryos are also similar in size to amphibian embryos, making them amenable to microsurgical grafting, vital dye labeling, and microinjection. Lamprey embryos have been injected with both synthetic mRNA and DNA to achieve gain-of-function phenotypes and analyze *cis*-regulatory sequences (Parker et al., 2014; Sauka-Spengler et al., 2007). Translation-blocking morpholinated anti-sense oligonucleotides (morpholinos) have also been used to test the function of genes involved in early developmental process, like neural crest specification (McCauley and Bronner-Fraser, 2006; Sauka-Spengler et al., 2007). However, morpholino-mediated gene knockdown is not a permanent perturbation in any organism, making analyses of later developmental functions difficult. This difficulty is pronounced in lampreys given their relatively long developmental time compared to other model vertebrates (~4 days until gastrulation, ~6 days until neural crest migration). Furthermore, the effectiveness of different morpholinos can vary dramatically in lamprey,

with a large proportion producing no effect or early embryonic death (Lakiza et al., 2011; Nikitina et al., 2011; Sauka-Spengler et al., 2007)(unpublished results).

The CRISPR/Cas system is a bacterial immune response mechanism that targets and degrades foreign DNA (Barrangou et al., 2007). Bacteria deploying CRISPR/Cas transcribe RNAs from genomic regions called 'clustered regularly interspaced short palindromic repeats' (CRISPRs), some segments of which correspond to sequences present in pathogenic bacterial viruses. Upon infection, these small RNAs direct sequence-specific cleavage of the viral genome by the CRISPR-associated nuclease Cas9. The specificity of the CRISPR/Cas mechanism has made it a powerful tool for targeted mutagenesis of eukaryotic genomes (Mali et al., 2013). To do this, a synthetic RNA called a 'guide RNA' (gRNA) is designed against a target sequence in the host genome immediately upstream of an endogenous protospacer-adjacent motif (PAM; 5' NGG 3'). The gRNA is then introduced into the cell together with Cas9 mRNA or protein, allowing the Cas9 protein and gRNA to form a complex that finds and cuts the target sequence. Once cleaved by the gRNA-guided Cas9 enzyme, the DNA is repaired by endogenous repair mechanisms, leading to deletions and/or insertions via non-homologous end joining. CRISPR/Cas9 mutagenesis has been successfully applied to several traditional developmental model organisms, including *Drosophila melanogaster* (Bassett et al., 2013), *Mus musculus* (Wang et al., 2013), *Danio rerio (Hwang et al., 2013), Xenopus tropicalis* (Guo et al., 2014; Nakayama et al., 2013), nematodes (Lo et al., 2013), sea anemone (Ikmi et al., 2014), and seed plants (Belhaj et al., 2013). Though versatile, CRISPR/Cas9-mediated mutagenesis is not equally efficient in all organisms. Not surprisingly, model vertebrates such as zebrafish and *Xenopus tropicalis* seem to require different concentrations of mRNA and gRNA for efficient F0 gene knockdown (Jao et al., 2013; Guo et al., 2014).

In hopes of improving the tractability of the sea lamprey for developmental genetic studies, we tested the efficacy of the CRISPR/Cas9 system in sea lamprey embryos and larvae. As a proof-of-principle, we targeted two genes with easily scored knock-down phenotypes, but different levels of pleiotropy: *Tyrosinase* (*Tyr*) and *FGF8/17/18*. Tyrosinase is an enzyme needed for melanin synthesis in all vertebrates. Successful mutagenesis of the lamprey *Tyr* gene should reduce pigmentation, but have no other effects on development. In contrast, gene expression and pharmacological inhibition have shown that the *FGF8/17/18* gene family is involved in several conserved developmental processes including mesoderm specification, somitogenesis, CNS patterning, pharyngeal segmentation, and skeletal differentiation (Abzhanov and Tabin, 2004; Jandzik et al., 2014a; Mason et al., 2000; Reifers et al., 1998; Sun et al., 1999). Depending on the timing and penetrance of CRISPR/Cas9-mediated mutagenesis in lamprey embryos, we expected mutation of the single lamprey *FGF8/17/18* homolog to produce a range of phenotypes mimicking the effects of pharmacological FGF signaling inhibition. We found that

CRISPR/Cas9-mediated mutagenesis of both genes was highly effective, validating the method in this evolutionarily informative species.

Materials and Methods

Lamprey spawning and husbandry

Adult spawning phase *Petromyzon marinus* were supplied by Hammond Bay Biological Station (Millersburg, MI) in spring and kept in holding tanks until fully ripened. Eggs and sperm were stripped manually from ripe adults (Nikitina et al., 2009). For *Tyr* and *Fgf8/17/18* gRNA injections, 6 and 7 different females (respectively) were used during 8 and 10 different fertilization events (respectively). Sperm from 2-4 males was typically combined per fertilization event, though some fertilization events had a single sire. Embryos and larvae were kept at 17-18ºC either in a closed, recirculating system, or in single 0.5-1L dishes containing deionized water supplemented with 400-600 ppm artificial sea salt. Water was changed as necessary, and dead embryos were removed from each dish every day to prevent contamination. Staging was performed according to Tahara (Tahara, 1988). Injected embryos (including negative controls) typically exhibited a slight developmental delay; to account for this delay, all staging was performed based on morphological landmarks of unaffected embryos rather than actual age (usually resulting in a 0.5-1 day delay in fixation compared to WT siblings).

Identification and cloning of *P. marinus Tyrosinase*

Using gnathostome Tyrosinase protein sequences as a query, we searched the March 2007 *P. marinus* genome assembly for a *Tyrosinase* homolog (Smith et al., 2013). We identified four putative *Tyr* exons on contig 46507, amplified one of these via PCR, and ligated it into the pJet1.2 vector (Thermo Fisher Scientific, Waltham, Massachusetts). Orthology of the protein encoded by this exon to gnathostome Tyrosinase was confirmed by alignment and phylogenetic analysis using the Maximum Likelihood method in the software package MEGA6 (Tamura et al., 2013) (Fig. S1). The corresponding sequence was deposited to GenBank (accession number KR150760).

RNA design and synthesis

Human codon-optimized *Streptococcus pyogenes* Cas9 mRNA (Cong et al., 2013) was transcribed as previously described (Guo et al., 2014). After digestion with DNAse, unincorporated nucleotides were removed using a Sephadex G-50 column (GE healthcare, Wauwatosa, Wisconsin). The mRNA was further purified by phenol/chloroform extraction and precipitation per standard procedures. Precipitated mRNA was resuspended in RNAse-free water for injection.

CRISPR/Cas9 target sites with the sequence 5' "GG(18N)NGG" 3' were chosen based on three criteria: 1) 50-80% GC content, 2) maximal proximity to the presumptive start codon, and 3) no off-target matches to the known *P. marinus* genome or EST databases showing more than 80% similarity by BLAST, or with less than 3 mismatches in the 10 bases proximal to the PAM sequence. The sequences of all target sites can be found in Tab. S1 within (Square et al., 2015b). Complementary oligonucleotides (Life Technologies, Carlsbad, California) corresponding to the selected genomic target sequences and containing 4 nt overhangs were annealed, phosphorylated *in vitro*, then ligated into the BsaI-digested DR274 plasmid (Hwang et al., 2013). gRNAs were produced as previously described (Blitz et al., 2013) using full or half-sized reactions of the T7 High Yield Kit (New England Biolabs, Ipswich, Massachusetts). gRNAs were purified for injection as described above for Cas9 mRNA. Two target sites within the same exon were chosen per targeted gene.

Preparation of Protein

Cas9 protein was purchased from PNA Bio Inc. (Thousand Oaks, CA). Cas9 protein was resuspended per the manufacturer's instructions, aliquoted, and stored at -70° C. 1 µg of Cas9 protein was incubated with 400 ng of gRNA (a 1:2.5 ratio by mass) for 10 minutes on ice, then brought up to 5 ul with LRD and H_2O for injection (see below).

Microinjection

Microinjections were performed in embryo system water in a room kept at 19ºC-20ºC. Embryos were injected at the 1-2 cell stage with approximately 5nL of injection solution containing 500 or 1000 pg Cas9 mRNA, 400 or 800 pg gRNA, and 10% lysinated rhodamine dextran (LRD). For injections with Cas9 protein we used 400 *Tyr* gRNA 2 plus 1000 pg of Cas9 protein. For some injections, 125 pg GFP mRNA was included to test for degradation of injected RNAs. Embryos were sorted by fluorescence at stage 21-26, and LRD-negative embryos were discarded.

Phenotypic analyses of mutant embryos and larvae

To assess the efficiency of *Tyr* mutagenesis, we fixed *Tyr* gRNA + Cas9 mRNA/Cas9 protein -injected lamprey larvae at stage 30 (approximately 30 days post-fertilization). Larvae were compared to wildtype siblings and scored as having: 1) wildtype pigmentation, 2) <50% pigment reduction, 3) 50 to 90% pigment reduction, or 4) 90% to 100% pigment reduction. *FGF8/17/18* gRNA + Cas9 mRNA-injected larvae were fixed when their injected, unaffected siblings reached st. 23.5 or st. 26.5 and scored for gross morphology. Phenotypic scoring for head size (by total estimated volume) was split into 5 mutant classes, as follows: 1) no or almost no head; only a small bump is visible, 2) head is highly reduced with no discernable morphological landmarks 3) head is 50-75% of its normal volume, with

visible (but reduced) first stream neural crest mesenchyme, 4) Head is slightly reduced in volume, with most head characters discernable, though the upper lip is usually truncated, 5) appears grossly wildtype. Changes in the expression of *Eng*, *Ednra*, *SoxE1*, and *Mef2* were visualized by *in situ* hybridization. It should be noted that there was some variability in the severity and frequency of non-specific defects between injection batches due to differences in how eggs from different females tolerated injection. Females with particularly robust eggs showed survivorship, and frequencies of development delay and deformity identical to uninjected wildtype embryos. In contrast, females with sensitive eggs showed a higher percentage of death and non-specific defects upon injection compared to their wildtype siblings. Such variation is not unexpected, given the genetic and environmental variation inherent in wild populations, and is effectively mitigated by injecting each gRNA into two or more females per season.

Genotyping

To confirm mutagenesis, we isolated genomic DNA from gRNA + Cas9 mRNA-injected embryos and larvae per standard methods. The target loci were amplified via PCR and ligated into the pJet1.2 vector. All genotyping primer sequences can be found in Tab. S1 within (Square et al., 2015b). After transformation and plating per standard methods, single clones were amplified, purified, and sequenced.

In situ hybridizations

Riboprobe templates were amplified via PCR using primers incorporating an SP6 RNA polymerase site. Riboprobes were then synthesized with SP6 RNA polymerase, and in situ hybridizations were carried out as previously described (Cerny et al., 2010). In situ hybridizations were performed in parallel on wildtype embryos and larvae to normalize signal development times. Stage 25+ larvae probed for Tyr expression were bleached in 1X SSC/5% formamide/1% H2O2 to allow for visualization of signal in pigmented cells.

Results

Tyrosinase expression

Lamprey possesses neural crest-derived melanocytes that emerge from the dorsal neural tube around st. 25.5. The lamprey eye begins depositing melanin as the larva approaches stage 27. At stage 26.5 we observed *Tyr* mRNA expression in presumptive neural crest-derived melanocytes, and in the eye, presaging melanin deposition in the retina (Fig. 1A and B).

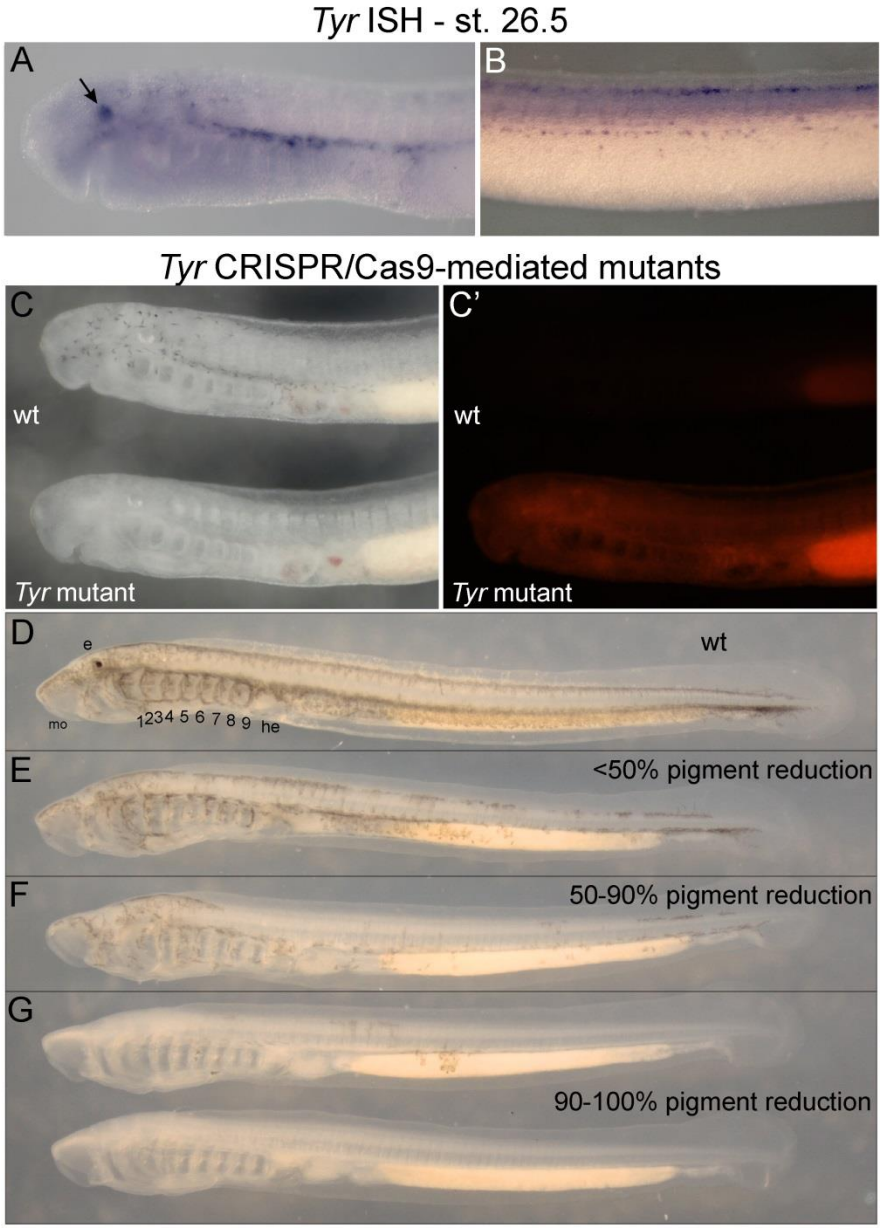

Fig. 1. *Tyrosinase* expression and disruption via CRISPR/Cas9. Lateral views with anterior to left in all panels. (A) A st. 26.5 larva showing expression in neural crest-derived melanocytes near the pharynx, and retinal expression in the presumptive eye (arrow). (B) Lateral flank view of a st. 26.5 larva showing neural crest-derived melanocytes in the dorsal neural tube and on the dorsal surface of the yolk. (C) Brightfield and (C') red channel fluorescence images showing the absence of pigment cells in an embryo that is positive for the LRD lineage tracer. (D-G) Example larvae of each mutant class at st. 30. Pharyngeal arches are labeled by number (1-9); e, eye; he, heart; mo, mouth.

Optimization of *Tyrosinase* mutagenesis

We synthesized two different gRNAs targeting an exon of lamprey *Tyrosinase* homologous to exon 1 of human *Tyrosinase*. We then co-injected 500 pg of Cas9 with 200pg *Tyr* gRNA 1 and 200pg *Tyr* gRNA 2, as described for *Xenopus tropicalis* (Guo et al., 2014). Using these quantities, 37% of larvae displayed an obvious reduction in pigmentation of 10% or more. We then increased the amount of Cas9 mRNA to 1 ng and the amount of gRNA to 400 pg *Tyr* gRNA 1 and 400 pg *Tyr* gRNA 2. While the occurrence of albino or nearly albino larvae was highest with this treatment (59%), we also observed a severe developmental delay or abnormal morphology in 24% of the larvae. We then injected individuals with the same amount of Cas9 mRNA (1 ng), but reduced the gRNA concentration to that of the initial trial (400 pg). This treatment produced a high frequency of albino/nearly albino (21%) and partially pigmented (44%) larvae, and reduced larval deformity to the level typical for healthy uninjected controls (~5-8%) (Tab. 1; See Fig. 1D-G for scoring examples). After determining the highest-tolerated RNA quantities for injection, we sought to test the relative effectiveness of the two *Tyr* gRNAs. To do this we injected zygotes from a single fertilization event with 400 pg of each gRNA singly, along with 1 ng Cas9 mRNA to directly compare the efficiency of each gRNA (g1 n=58, g2 n=66). After scoring, a Wilcoxon Rank-Sum test was performed to test for differences in the ability of each of these gRNAs to produce a mutant phenotype; this resulted in a p-value of 0.7865, indicating that their activity is very similar. Prior to scoring, deformed embryos were removed from this experiment only.

Table 1. Mutagenesis Efficiency Vs. Deformities of *Tyr* gRNA and Cas9 Mixtures

Injection mix	0-10% affected	>10-50% affected	>50-90% affected	>90% affected	Total deformed
500 pg Cas9 mRNA, 400 pg gRNA 1+2 (n=226)	66%	14%	13%	7%	5%
1 ng Cas9 mRNA, 400 pg gRNA 2 (n= 231)	35%	24%	20%	21%	6%
1 ng Cas9 mRNA, 800 pg gRNA 1+2 (n=301)	16%	13%	12%	59%	24%
1 ng Cas9 mRNA, 400 pg negative control gRNA (n=218)	100%	0%	0%	0%	6%
1 ng Cas9 protein, 400 pg gRNA 2 (n=174)	0%	0%	1%	99%	12%

To verify disruption of the *Tyr* locus, two nearly-albino individuals injected with *Tyr* gRNA 1 and two nearly albino embryos injected with *Tyr* gRNA 2 were pooled and genotyped. Of the 7 unique alleles observed, one was wild type, five contained a deletion, and one contained both a deletion and an insertion (Fig. S2A). To look for a correlation between the prevalence of mutant alleles and the phenotype generated, we genotyped an LRD positive individual from each of the 4 mutant categories (Fig. 2). We found that the individual that appeared to have a WT level of pigmentation (but was LRD positive) displayed 0/14 mutated sequences, the individual showing ~25% reduction in pigmentation had 5/13 mutant alleles (2 of which were frameshifts), the individual with ~75% reduction in pigmentation showed 9/10 mutant alleles (5 of which were frameshifts) and the completely albino individual showed 13/13 mutant alleles (9 of which were frameshifts).

Recently, Cas9 protein has become commercially available. To test if an injection solution using Cas9 protein instead of mRNA is capable of generating a higher frequency of mutant lamprey, we injected 400 pg of *Tyr* gRNA 2 in combination with 1 ng of Cas9 protein. This resulted in an astounding 99% completely albino individuals (173/174), but a higher level (12%) of developmentally perturbed lamprey (21/174).

Mutagenesis of *FGF8/17/18*

As with *Tyrosinase*, we found that injection of 1 ng Cas9 mRNA and 800 pg *FGF8/17/18* gRNA caused high mortality. However, at lower RNA concentrations (500 pg Cas9 mRNA, 400 pg total gRNA), survivorship was similar to that of uninjected sibling controls, and surviving embryos displayed phenotypes consistent with *FGF8/17/18* loss-of-function (Jandzik et al., 2014a; Reifers et al., 1998) (Crump et al., 2004b). The earliest and most common mutant phenotype was a reduction in head size (Tab. 2; Fig. 3). This effect was observed at a high frequency in embryos injected with Cas9 mRNA and either *FGF8/17/18* gRNA alone, or in combination (*FGF8/17/18* gRNA 1+ gRNA 2). A Wilcoxon Rank-Sum test revealed that these gRNAs do differ significantly in their ability to produce a mutant phenotype (g1 n=170, g2 n=224; p-value=0.0108). Mutant scores for each gRNA can be seen separately in Tab. 2. To verify that animals displaying these phenotypes were alive and injected rather than exhibiting autofluorescence due to cell death, we documented the relative brightness on both GFP and LRD absorption channels. Dead embryos fluoresced more strongly on the GFP channel than injected embryos, whereas they fluoresced less strongly on the LRD channel (Fig. S3). This supports the idea that these strong, consistent phenotypes observed are due to actual genetic perturbations rather than a

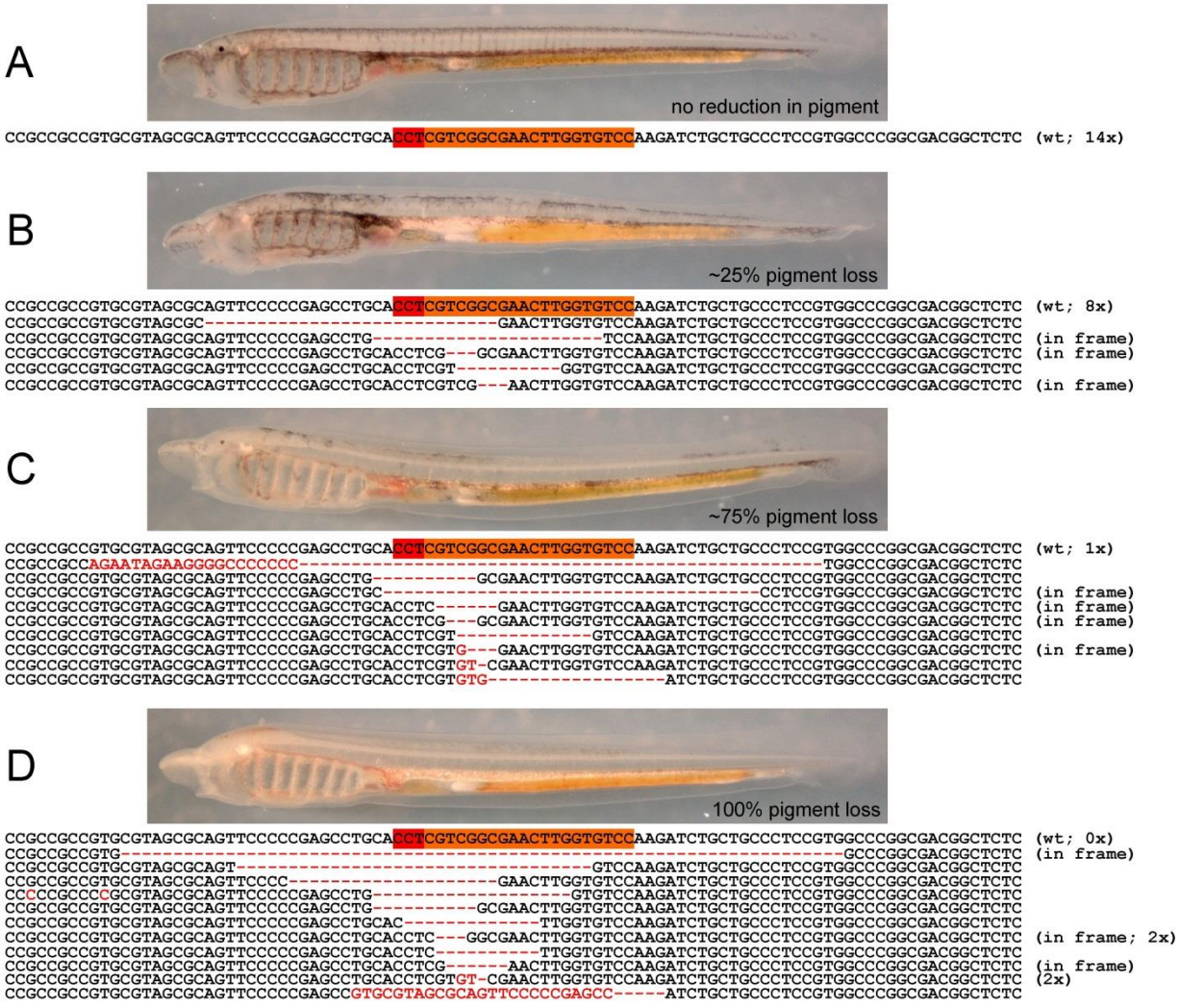

Fig. 2. Genotyped *Tyr* loci of injected individuals. All sequences show the non-template strand in 5' to 3' orientation with respect to the *Tyr* gene. The target sequence and PAM are shown in orange and red, respectively, in the wt sequence. Deletions relative to wt are shown as red dashes, while polymorphisms and insertions are shown as red letters (despite some inserted letters matching the wt sequence). All inferred insertions are stacked on the left side of the deletion region. (A) Sequencing of an injected, LRD positive individual showing no reduction in pigmentation returned 14/14 wt sequences. (B) Sequencing of a *Tyr* mutant displaying ~25% pigmentation reduction returned 5/13 mutated sequences, 2 of which were frameshifts. (C) Sequencing of a *Tyr* mutant displaying ~75% pigmentation reduction returned 9/10 mutated sequences, 5 of which were frameshifts. (D) Sequencing of a completely albino individual returned 13/13 mutated sequences, 10 of which were frameshifts.

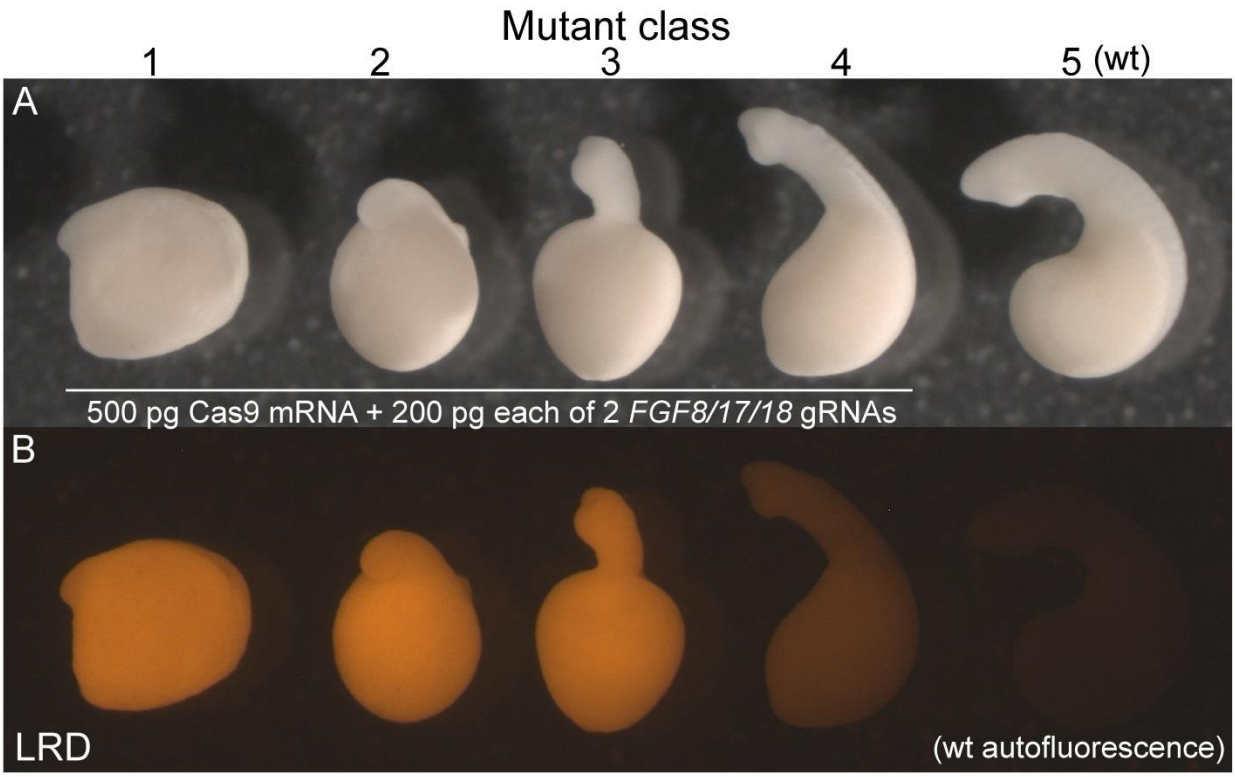

Fig. 3. Mutant classes used for phenotypic analysis of FGF8/17/18 CRISPR/Cas9-mediated mutants. (A) brightfeild image of sibling lamprey embryos. The leftmost four were injected with 500 pg Cas9 mRNA and 200 pg each of 2 FGF8/17/18 gRNAs. The rightmost animal is uninjected. (B) Red channel fluorescence showing LRD deposition in embryos. Mutant class is indicated at the top of the figure; see Tab. 2 for information on the distribution of observed mutant classes in *FGF8/17/18* experiments.

Table 2: Phenotypes of *FGF8/17/18* gRNA and Cas9-injected larvae

Mutant class	1	2	3	4	5 (wt)
gRNA 1 total % per mutant class (n = 170)	9.5%	17%	24.6%	21.1%	27.8%
gRNA 2 total % per mutant class (n = 224)	6.7%	14.3%	18.8%	25.4%	34.8%
Eng ISH (totals per mutant class)	3	7	7	7	3
# with completely absent *Eng* expression	3	7	2	0	0
# with reduced or unilateral *Eng* expression	0	0	4	3	0
# with WT *Eng* expression	0	0	1	4	3

termination of development. We further analyzed *FGF8/17/18* gRNA + Cas9 mRNA-injected embryos and larvae for changes in gene expression by *in situ* hybridization (Figs. 4 and 5). We examined three genes known to be dependent on FGF signaling in lamprey and gnathostomes: *Engrailed* (*Eng*), which is expressed in the midbrain-hindbrain boundary (Reifers et al., 1998), *Ednra*, an Endothelin receptor expressed in post-migratory neural crest cells (Jandzik et al., 2014a), and *SoxE1*, a *sox9* ortholog that marks pre-chondrocytes (Jandzik et al., 2014a; McCauley and Bronner-Fraser, 2006). We also looked at the expression of *Mef2*, a marker of formed somites and pharyngeal mesoderm, to monitor somitogenesis and pharyngeal segmentation (Jandzik et al., 2014a). Consistent with previous reports, expression of *Engrailed* (Fig. 4), *Ednra* (not shown), *and SoxE1* (Fig. 5A-D) was lost or reduced in presumptive *FGF8/17/18* mutants, while *Mef2* expression revealed a loss of somites and disruptions in somite patterning, as seen in the zebrafish FGF8 mutant *acerebellar* (Reifers et al., 1998) (Fig. 5E-K). *Mef2, Ednra, and SoxE1* expression in the pharyngeal arches also revealed a loss of pharyngeal segmentation, a consequence of *FGF8/17/18* loss-of-function in both gnathostomes and lamprey (Jandzik et al., 2014a; Reifers et al., 1998) (Crump et al., 2004b). Occasionally these disruptions were unilateral, presumably resulting from localization of the injection solution to a one of two blastomeres when injections were performed at mid-cleavage (e.g. Fig. 4E). We also assayed *HhA* expression, a lamprey ortholog of *sonic hedgehog*, which is a marker of the *zona limitans intrathalamica* (ZLI) within the forebrain. As expected, *HhA* expression in this signaling center was unaffected in *FGF8/17/18* mutant larvae, indicating that loss of *Eng* expression reflects a specific loss of FGF8/17/18 in the midbrain-hindbrain boundary, rather than a non-specific disruption of neural patterning (Fig. 5L and M).

To verify disruption of the *FGF8/17/18* locus in putative mutants, two embryos with reduced heads (mutant class 3) that were injected with Cas9 mRNA + *FGF8/17/18* gRNA 1 alone, and two

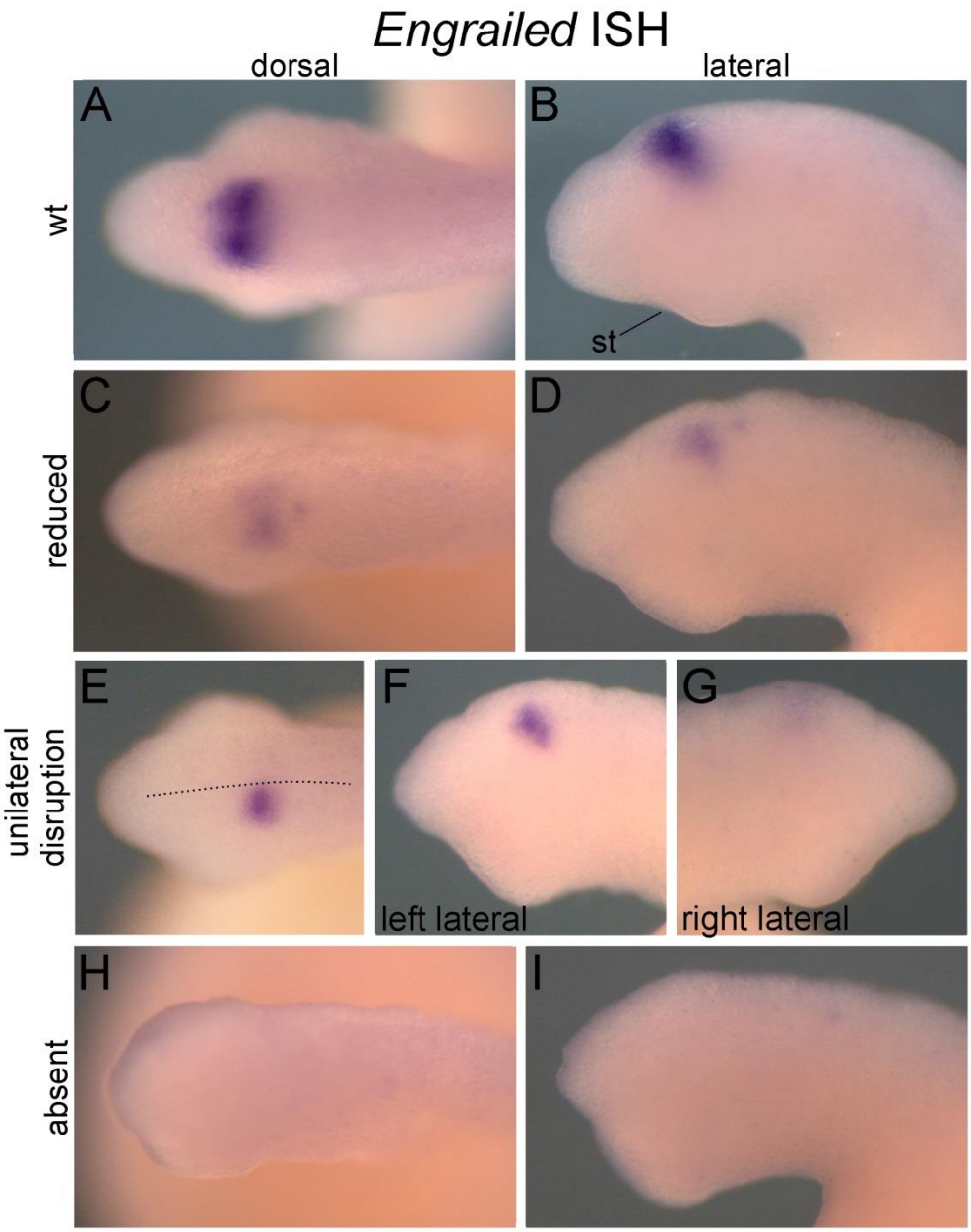

Fig. 4. *Engrailed* ISH on *FGF8/17/18* CRISPR/Cas9-mediated mutants at stage 23. Anterior to left in all panels except G, where anterior is to right. (A and B) wt *Eng* expression in the midbrain-hindbrain boundary of a control embryo. (C-I) *Eng* expression in *FGF8/17/18* CRISPR/Cas9-mediated mutants. Reduction (C and D), unilateral disruptions (E-G), and total absence (H and I) of *Eng* expression were all observed (see Tab. 2 for counts). Dotted line in (E) indicates the dorsal midline. st, stomodeum.

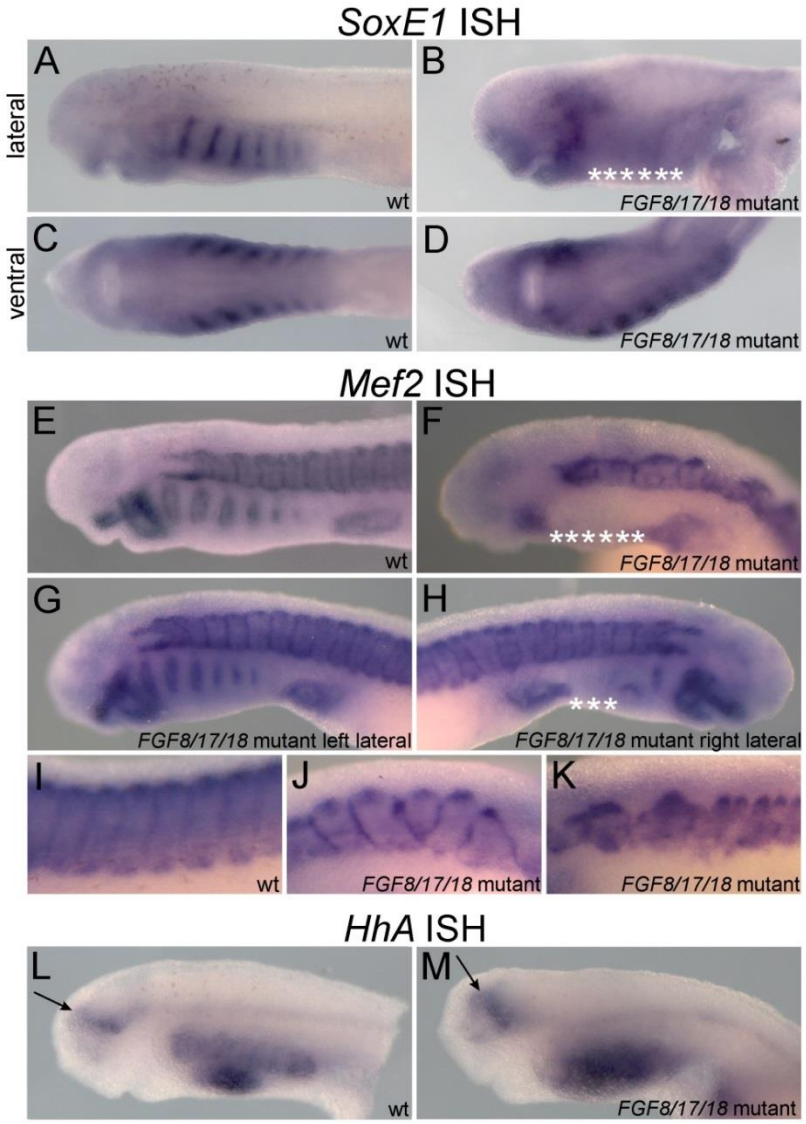

Fig. 5. *SoxE1, Mef2,* and *HhA* ISH on *FGF8/17/18* CRISPR/Cas9-mediated mutants. Anterior to left in all panels except H, where anterior is to right. (A-D) *SoxE1* ISH on wt (A and B) and *FGF8/17/18* CRISPR/Cas9-mediated mutant larvae. (B) A unilateral disruption of *FGF8/17/18* resulted in the loss of *SoxE1* expression in the posteriormost 6 pharyngeal arches (white stars). (E-K) *Mef2* ISH on wt (A) and *FGF8/17/18* CRISPR/Cas9-mediated mutants (F-K). (F) Some highly affected mutant larvae showed a complete loss of pharyngeal arch *SoxE1* expression (white stars). (G and H) Some moderately affected mutant larvae showed a unilateral disruption of *Mef2* expression in pharyngeal pouch mesoderm (white stars in H). (I-K) *Mef2* also revealed somite shape and spacing, which was sometimes slightly disrupted (J) or highly disrupted (K). (L and M) *HhA* ISH on wt (L) and a presumed *FGF8/17/18* CRISPR-mediated mutant larva (M); both larvae show proper expression of *HhA* in the ZLI (arrow in each), indicating that this organization center has expectedly remained unperturbed.

embryos with reduced heads (mutant class 3) injected that were with Cas9 mRNA + *FGF8/17/18* gRNA 2 alone were pooled and genotyped. Of the 12 alleles observed, 6 were wild type and 6 contained a deletion (Fig. S2B and C).

Negative controls

To ensure that our pigment and FGF-related phenotypes were specific to these perturbations rather than being a result of non-specific Cas9 or gRNA activity, a negative control gRNA was synthesized against *Xenopus laevis tyrosinase* and injected at 400 pg per embryo with 1 ng of Cas9 mRNA (n=231). 5.6% (13/231) of these negative control lamprey showed non-specific, inconsistent developmental defects (such as a severe delay, or bending along the antero-posterior axis), supporting the specificity of the head-reduction phenotype of *FGF8/17/18* mutants. Similarly, no control-injected embryo showed a reduction in pigmentation, ruling out the possibility that these levels of Cas9 mRNA or gRNA inherently disrupt melanin production.

Discussion

Sea lampreys take 7-15 years to mature, can grow to almost 1 meter long, and require large, live prey as adults. As a result, the establishment of mutant or transgenic lines of this species is a practical impossibility. Here we present a proof-of-principle application of CRISPR/Cas9-mediated mutagenesis in the sea lamprey. Using our optimized parameters, we found that a majority of injected "F0" embryos develop into complete or partial mutants with reproducible phenotypes and minimal confounding side effects. Remarkably, this efficiency approached 100% when using purified Cas9 protein in the injection mixture. The ability to generate large numbers of genotype-verifiable mutant F0 larvae from the embryos of wild-caught adults opens exciting new possibilities for studying the genetic bases of development in this phylogenetically important species.

Our study suggests the sea lamprey has a combination of features that make it well-suited to CRISPR/Cas9-based genetic manipulation. Like amphibian embryos, sea lamprey embryos are easy to inject, and can tolerate large quantities of synthetic RNA and protein. Having more Cas9 protein and gRNA per cell not only increases mutagenesis efficiency, but opens up the possibility of mutagenizing multiple genes simultaneously. We also found that the nucleotide composition of the lamprey genome makes it relatively easy to design multiple, high-quality gRNAs for a given target exon. The most effective CRISPR gRNAs have a GC content of over 50% (Doench et al., 2014; Gagnon et al., 2014), and all target sites must incorporate a 3' PAM signal (5' NGG 3'). Because of this, genes that are very AT rich can lack appropriate target sites near PAM motifs. Sea lamprey exons have an average GC content of 61% (Smith et al., 2013). As a result, a typical lamprey exon will have several potential CRISPR/Cas9

target sites near PAM signals. This permits multiple gRNAs to be designed against a single target exon and coinjected, mitigating potential variation in the effectiveness of individual gRNAs. Having multiple gRNAs per gene also provides opportunities to control for off-target phenotypes. For example, if two or more different gRNAs (or gRNA cocktails) targeting the same gene have the same effect on development, it can be safely assumed the phenotype is specific to the gene, even before genotyping.

While the lamprey genome lends itself to CRISPR/Cas9-mediated manipulations, we did encounter some challenges with the current publicly-available assemblies. Due to large sequence gaps, many sea lamprey genes identified as cDNAs, or through transcriptome sequencing, are not present in the current version of the genome. Because the exon structure is unknown, CRISPR target sites for these genes run the risk of spanning exons, and thus not being contiguously represented in the genome (which is required for recognition by the Cas9/gRNA complex). Fortunately, exon structure tends to be strongly conserved across animals (Rogozin et al., 2003), thus the risk of spanning exons can be reduced by focusing on evolutionarily conserved exons. The incomplete genome also makes it impossible to scan for all potential off-targets, as is becoming standard practice for other model organisms. These issues should be resolved with future versions of the lamprey genome assembly.

A perennial challenge for understanding the function of developmental regulatory genes is that many are highly pleiotropic. In genetic model systems, strategies for temporally or spatially restricting expression of Cas9 enzyme or gRNA have been proposed as a way to circumvent this problem. We targeted the highly pleiotropic gene *FGF8/17/18* to get a sense for how this issue might affect the creation and scoring of sea lamprey mutants. We found that while a large percentage of CRISPR/Cas9-mediated *FGF8/17/18* mutants appeared to be near-complete nulls, many others had milder phenotypes consistent with mosaics or hypomorphs, frequently displaying unilateral disruptions in FGF-dependent processes. This difference in penetrance corresponded well to the intensity of lineage tracer fluorescence in the mutant embryos (Fig. 3), as reported in axolotl (Flowers et al., 2014). Going forward, sorting mutants according to lineage tracer intensity and distribution should allow workers to separate and score complete versus partial mutant embryos. In addition, it should be possible to intentionally restrict CRISPR/Cas9 activity to particular embryonic domains by injecting single blastomeres at later cleavage stages (i.e. 8 or 16 cell stage). One strategy for this would be to inject Cas9 mRNA or protein into zygotes, followed by gRNA injection at later cleavage stages. Though there is no modern developmental fate map for lamprey, lamprey embryos are easy to inject, and thousands can be injected by a single worker each hour. Coinjections of gRNA with a lineage tracer should allow selection of embryos and larvae with the gRNA restricted to the desired tissue, facilitating the study of highly pleiotropic developmental regulators.

In this study, a pair of gRNAs designed against FGF8/17/18 showed different levels of activity, while the pair designed against *Tyr* did not. Variation in gRNA efficiency is well-documented (Hsu et al., 2013), and likely has to do with the specific binding properties of each RNA. Individual genotyping of *Tyr* mutants revealed that there was likely a correlation between the number of mutated *Tyr* sequences and the level of albinism observed in a given individual (Fig. 2), though a true test of this would require much more genotyping of many more individuals. Furthermore, *Tyr* disruptions using gRNA 1 yielded up to 99% albino individuals when used in combination with the protein (see Tab. 3), suggesting that the targeted region of Tyrosinase is catalytically important, not requiring a frameshift to hinder melanin synthesis. If a frameshift were required, we would at best still expect ~1/3 of all pigment cells in a given batch of larvae to retain proper melanin deposition.

CHAPTER VI: Beyond the jaw joint: Endothelins patterned neural crest cells in stem vertebrates

Jawed vertebrates (gnathostomes) are the most species-rich deuterostome group and account for ~97% of all living chordate species (Nelson et al., 2016). Developmentally, the jaws form from neural crest cells (NCCs) that migrate into the oral region (reviewed by (Santagati and Rijli, 2003)). Both Endothelin receptor (Ednr) and ligand (Edn) loss-of-function leads to a loss of the primary jaw joint in all gnathostomes examined to date, all of which are bony vertebrates (osteichthyans) (Alexander et al., 2011; Charite et al., 2001; Miller et al., 2000; Miller et al., 2003; Nair et al., 2007; Sato et al., 2008b; Tavares et al., 2012). During head skeleton development, Edn1 is secreted from pharyngeal epithelia and mesoderm and bound by the endothelin receptor Ednra, which is expressed by pre-skeletal NCCs as they invade and proliferate in the pharynx. In zebrafish and mouse, these signals regulate the combinatorial expression of *dlx* and *hand* transcription factors in oral region and pharyngeal arches (PAs) which, in turn, activate the morphogenetic and differentiation programs that shape skeletal elements and position the jaw and other joints. Specifically, Edn1 signaling is needed for *dlx5/6* and *hand* expression in ventral PA NCCs, and also dorsal to this region it similarly drives *dlx5/6* expression in the intermediate PA NCCs (Alexander et al., 2011; Charite et al., 2001; Miller et al., 2000; Miller et al., 2003; Nair et al., 2007; Sato et al., 2008b; Tavares et al., 2012), but not the dorsalmost PA domains. Importantly, intermediate domain morphologies and gene expression patterns have been shown to be more sensitive to Edn1 or Ednra perturbations than the ventral domains (Alexander et al., 2011; Tavares et al., 2012). In addition to oropharyngeal skeleton patterning, Endothelin signals have multiple other early roles in NCC development. Ednra mediates the migration and differentiation of ectomesechymal NCC in the cranial region, including precursors of the head skeleton and heart (Bonano et al., 2008; Clouthier et al., 1998; Yanagisawa et al., 1998). Meanwhile, Edn3 signaling through Ednrb receptors is critical for the migration, proliferation and differentiation of NCC-derived pigment cells and certain peripheral nervous system components (Adachi et al., 1994b; Baynash et al., 1994; Kawasaki-Nishihara et al., 2011; Krauss et al., 2014; Kusafuka and Puri, 1997; Lahav et al., 1998; RM et al., 2006; Sakai et al., 2000; Shin et al., 1999). Gnathostome Edn2 and Edn4 have no known function in NCC development, though it has been shown that Edn2 is capable of binding mouse and *Xenopus* Ednra and Ednrb-type receptors (Karne et al., 1993; Sakamoto et al., 1993). Intriguingly, Endothelin signaling appears to be absent in invertebrates, suggesting the recruitment of this signaling pathway by NCCs was a key factor not only in jaw evolution, but in NCC evolution overall (Braasch and Schartl, 2014; Braasch et al., 2009; Martinez-Morales et al., 2007; Square et al., 2016a).

To better understand when and how the various roles of Endothelin signaling in NCC development arose, we have been analyzing the expression and function of *edn*s and *ednr*s in the modern jawless (cyclostome) vertebrate, the Sea Lamprey (*Petromyzon marinus*) (Square et al., 2016a). Cyclostomes and gnathostomes diverged shortly after the appearance of vertebrates around 500-550 million years ago(Donoghue and Keating, 2014; Kuraku and Kuratani, 2006). While the gnathostome clade contains many well-developed model systems, the only cyclostomes amenable to developmental genetic perturbations are lampreys. While all living vertebrate lineages display some level of specialization in development, any characters shared by lamprey and gnathostomes were likely present in the vertebrate common ancestor (reviewed by (Square et al., 2016b)). Thus we used CRISPR/Cas9-driven mutagenesis to disrupt both Ednrs and two Edns expressed specifically during *P. marinus* embryogenesis, and generated the first completely Endothelin signaling-deficient vertebrate. To help with side-by-side comparisons and add phylogenetic depth to our understanding of gnathostome Endothelin function, we performed some parallel experiments in the African Clawed Frog (*Xenopus laevis*). Our data provide strong support for a scenario wherein Endothelin signaling gained a new function in NCC development early in vertebrate evolution. This cooption event was then followed by duplication and specialization of two functionally distinct Ednr signaling pathways with separate and essential functions in the development of ectomesenchymal (Ednra-sustaining) versus non-ectomesenchymal (Ednrb-sustaining) NCC lineages. We posit that this early specialization was critical in initially achieving and/or maintaining the developmental distinction between these two major NCC lineages through the selective regulation of skeletogenic *SoxE* genes. Finally, we find that lamprey also uses Endothelin signaling to achieve nested *dlx* expression in the pharynx, though uniquely, both lamprey Ednrs work synergistically in this process. Furthermore, unlike zebrafish, *X. laevis*, and mouse, lamprey *hand* expression is largely unaffected by Endothelin disruption, even when all endothelin signaling is perturbed. Based on our findings, we speculate that new roles for Endothelins in NCCs during stem vertebrate evolution may have facilitated finer-scale patterning and/or developmental modularity apparent in all NCCs, including head skeleton precursors.

Cas9-driven mutagenesis of Endothelin pathway genes

To discern the functions of Edn signaling during lamprey development, we disrupted *P. marinus ednA, ednE, ednra,* and *ednrb* using Cas9-mediated mutagenesis (Fig. 1; see methods) and analyzed injected (F0) embryos and larvae. We targeted each gene with at least two different single guide RNAs (sgRNAs) designed to recognize unique target loci within each coding region (see controls and methods sections). We used this same strategy to disrupt *edn1, edn3,* and *ednra* in *Xenopus laevis*. Because *X. laevis* has an allotetraploid genome (Session et al., 2016), it is typically necessary to target both the

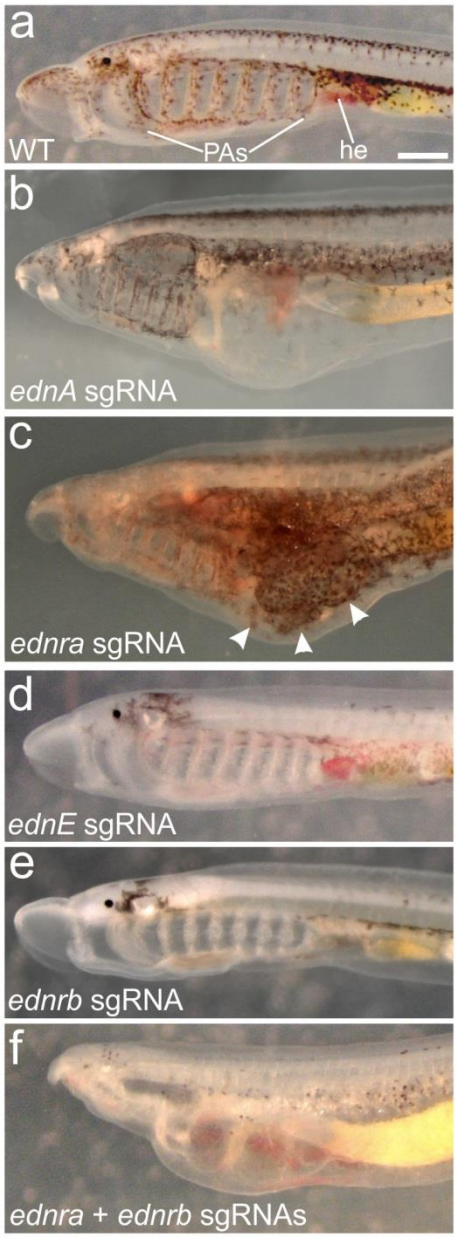

Fig. 1. Endothelin ligand and receptor mutant *P. marinus*. a-e, Affected individuals from each type of gene perturbation at st. 30. All photos are left lateral views. Scale bar in A represents 0.5 mm and applies to all panels. Targeted gene is indicated in each panel. Arrowheads in c indicate ectopic melanocytes. he, heart; PAs, pharyngeal arches.

"Long" and "Short" homoeologs to disrupt protein function (Wang et al., 2015); e.g. to disrupt Ednra, we targeted both *ednra.L* (*ednra-a*) and *ednra.S* (*ednra-b*) simultaneously (see control and methods sections). Despite the mosaic nature of the indel mutations created in these Cas9-injected F0 embryos, the consistency of the phenotypes observed and our genotyping data suggest that we regularly achieve a high percentage of null- or near-null individuals (See Tab. S1 for phenotype frequencies).

Disruption of *ednra*, *ednA*, and *edn1*

In *P. marinus*, both *ednA* and *ednra* mutants consistently developed a heart edema starting at st. 24-25. By st. 26, the future head skeleton appeared slightly smaller and condensed. Both of these features became more pronounced as development proceeded (Fig. S2). We presume that the presence of this heart edema was responsible for the head skeleton being stretched ventrally. These aspects of the *ednA* and *ednra* phenotypes were highly consistent, allowing us to track them through development and diagnose specimens as affected for ISH, IHC, and alcian blue staining. Occasionally, the upper lip was wholly absent in the most severely affected *ednra* mutants. *X. laevis edn1* and *ednra* also produced similar head skeleton phenotypes, however this was not apparent until stage 40-41 (the beginning of cartilage differentiation). Consistent with the larval onset of these phenotypes, we observed no downregulation or disruption of *X. laevis sox8.L* or *sox9.S* during cranial NCC migration st. 25, nor *P. marinus SoxE2* in at st. 22 (not shown; see Tab. S2 for ISH counts).

Alcian blue staining in both species revealed that cartilages in the pre-oral region, PA1, and PA2 (the mandibular and hyoid arches) more considerably more affected in both shape and size when *ednra, ednA*, or, *edn1* were disrupted (Fig 2a-d, Fig. 3a-c). Some small regions of cartilage, or occasionally entire posterior elements were absent in *P. marinus*, though the presence of any heterotopic PA morphologies was not clear (Fig. S1). Consistent with a reduction in chondrogenesis, we found that the *P. marinus* aggrecan homologs *lectican1* (Fig. S3) and *lectican2* (Fig. 2e-g) exhibited diminished, sometimes patchy expression in postmigratory pharyngeal NCC. We next looked at lamprey *SoxE2* and *X. laevis sox9.S*, putative upstream regulators of chondrogenesis in lamprey and gnathostomes, respectively. We detected reduced transcription of both genes (Fig. 2h-j and Fig. 3d and e), suggesting that Ednra signaling is necessary for the maintenance of chondrogenic *soxE* transcription in both species. Similarly, lamprey *FGFRa* and *FGFRb* were both downregulated in *ednra* mutant lamprey, including both nascent heart and PA tissues (Fig. S3). In contrast to *soxE* and *FGFR* genes, general markers of ectomesenchymal NCC specification such as *TwistA*, *myc*, and *ID,* in lamprey (Fig. S3), and *sox8* in *X. laevis* (not shown), were expressed normally (see Tab. S2 for counts).

In addition to being hypomorphic, we also observed skeletal patterning defects in the head skeleton of *X. laevis Ednra* mutants. As in other bony vertebrates, intermediate and ventral components of

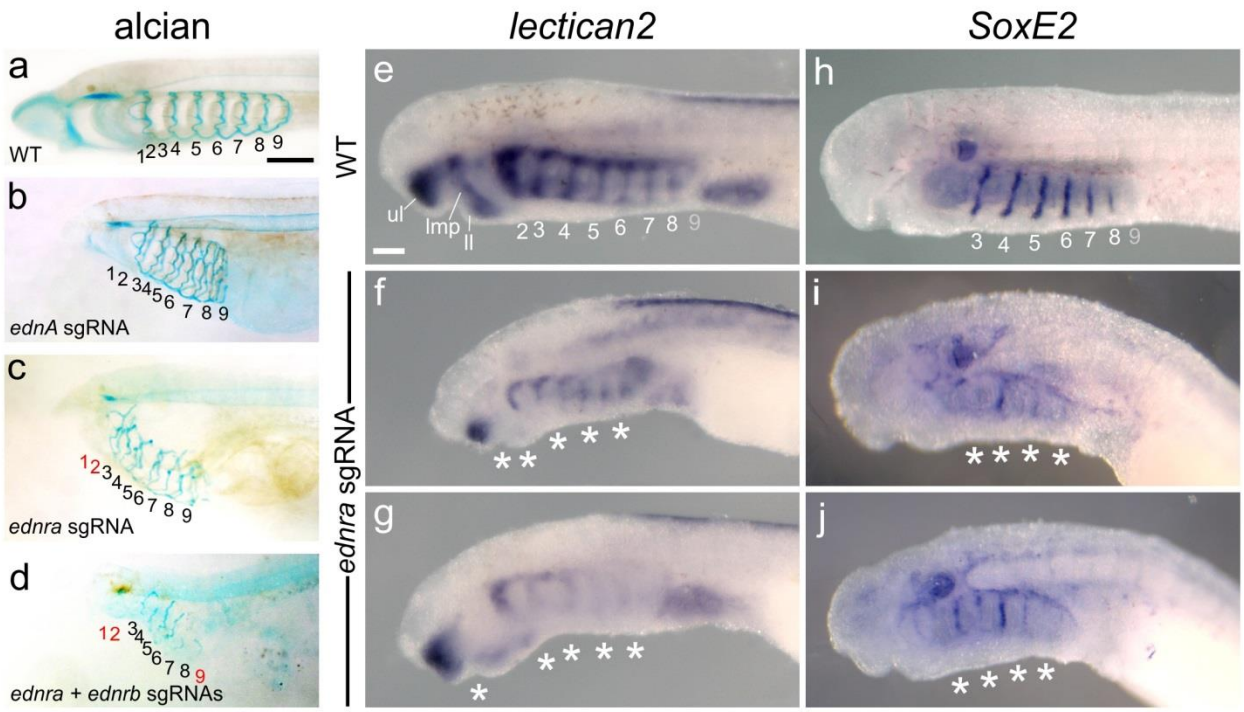

Fig. 2. Head skeleton defects in *ednA* and *ednra* mutant *P. marinus*. a-d, Alcian blue staining reveals truncated or missing head skeletal elements, especially in certain PA1 and PA2 derivatives (red numbers indicate alcian-negative PAs). Scale bar in A represents 500 µm and applies to b-d. **e-j**, ISH for *SoxE2* (e-g) and *Lec2* (h-j) in *ednra* sgRNA-injected lampreys reveals downregulation in the PAs at st. 26.5 (white stars). Scale bar in e represents 100 µm and applies to E-I. PAs are numbered in a-d and WT ISH panels. ll, lower lip; lmp, lateral mouth plate; ul, upper lip.

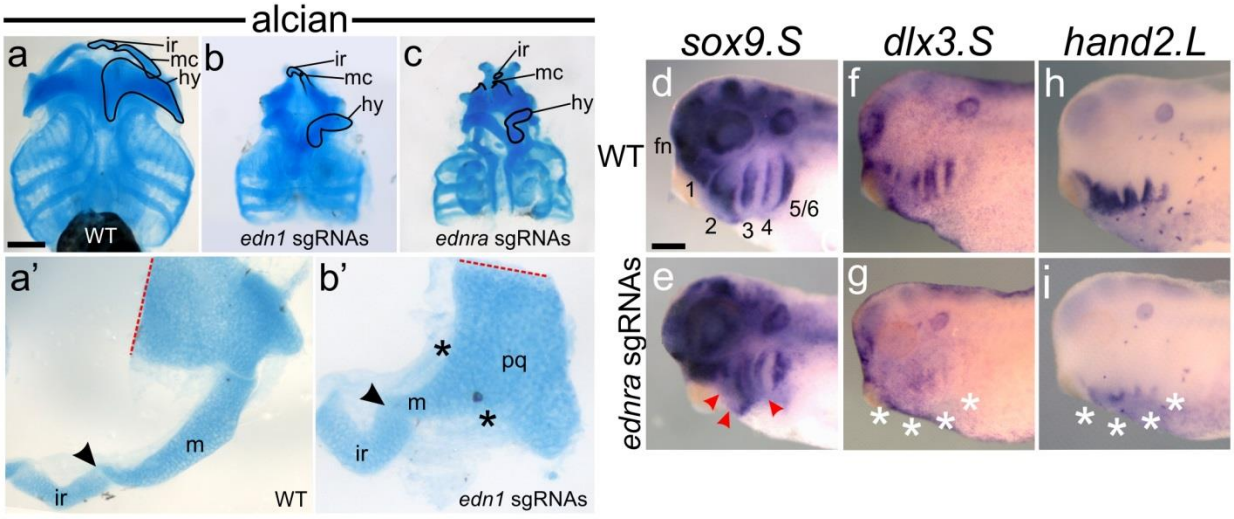

Fig. 3. Head skeleton defects in *edn1.L/S* and *ednra.L/S* mutant *X. laevis*. a-c, alcian blue staining reveals truncated head skeletal elements, especially in certain PA1 and PA2 derivatives (outlined). a-c are shown in ventral view with anterior facing upwards. Scale bar in a represents 500 µm and applies to b and c. a' and b' show further dissected (red dotted lines) and enlarged oral cartilages from panels a and b, respectively, and are not shown to scale with each other. Arrowheads in a' and b' indicate the remaining intramandibular joint. Arrow in a' indicates the primary jaw joint. Stars in b' indicate the loss of the primary jaw joint. **d-f**, *sox9.S* (d and e), *dlx3.S* (f and g), and *hand2.L* (h and i) ISH in *ednra.L/S* mutants reveals disrupted expression in post-migratory cranial NCCs (red arrowheads and white stars). PAs are numbered in panel d. Scale bar in d represents 100 µm and applies to e-i. fn, frotonasal prominence; hy, hyoid; ir, infrorostral; mc, Meckel's cartilage; pq, palatoquadrate.

the pharyngeal skeleton were most affected, including a loss of the primary jaw joint, and a severe reduction of the hyoid (Fig. 3a-c). Consistent with this, *dlx3.S* and *hand2.L* (Fig. 3 f-i) were disrupted in ~50% of randomly chosen injected individuals at st. 33 (mirroring the penetrance of the assay; see Tab S1 for ISH counts), prior to the gross morphological onset of the phenotype (Fig. 3). We detected no defects in *msx1.S* expression (n=0/15; not shown). Interestingly, the anuran-specific intramandibular joint was still recognizable in most cases (arrowheads in Fig. 3b'). In the *P. marinus* branchial skeleton, we found slightly reduced *DlxA, -B, -C,* and *-D* expression in the intermediate PAs of the most severely affected *ednra* mutants (Fig. 4a-i), though this affect was generally much weaker than is observed in bony vertebrate models, wherein a complete or near-complete absence of these gene transcripts is typical in the absence of *edn1* or *ednra* prior to any gross malformation in the nascent head skeleton (e.g. Fig. 3f-i). By contrast, *Alx, Hand,* and *MsxB* displayed normal overall expression patterns in the *P. marinus* branchial skeleton, with no DV shifts in expression (Fig. 4j-l, Fig. S3). Correlating with the more severe effect on anterior cartilages as seen by alcian blue staining, lamprey PA1, PA2, and pre-oral ectomesenchyme was most severely affected transcriptionally as well, frequently lacking the expression of most *Dlx* genes, *Alx, Hand, MsxB,* and *SoxE1* (Fig. 4f-i and Fig. S3).

While we observed slight defects in skeletal patterning genes in lamprey *ednra* mutants (relative to bony vertebrates), *ednra* disruption led to improper migration and dramatically increased numbers of melanocytes, with ectopic melanocytes populating the cardiac region and fin folds (arrowheads in Fig. 1c and Fig. S4). This effect appears unique to lamprey, as Ednra disruption does not cause excess pigment cell production in *X. laevis* or any other bony vertebrate thus addressed. Both *ednra* sgRNAs also occasionally caused excess pigmentation without any heart edema or reduced head skeleton (in ~8% of injected individuals; Fig. S4), though we never observed any individual that had a reduced head skeleton without aberrant pigment cells. Neither ectopic melanocytes nor novel melanocyte migration patterns were ever observed in any *ednA* mutagenesis experiment (Fig. S5). Given that *ednra* is not expressed in WT pigment cells, and *ednra* mutants have less skeletal tissue, our results strongly suggest that the excess pigment cells in *ednra* mutants are preusmptive skeletogenic NCCs that have taken on a pigment cell fate. Interestingly, *P. marinus ednrb* is temporarily coexpressed with *ednra* throughout all postmigratory pharyngeal NCCs (Square et al., 2016a), an expression domain not seen in gnathostomes. ISH for *ednrb* at mid-pharyngula stages revealed that its expression is completely eliminated from the pharynx in *ednra* mutants during late NCC migration, and thereafter an excess of *ednrb*-positive cells begin aggregating over the future heart and at the top of the pharynx (Fig. S3), indicating that these genes somehow interact in the future head skeleton. Regardless of mechanism underlying these ectopic pigment cells, these results are consistent with an ancient function for of vertebrate Ednra signaling in specifying and/or reinforcing ectomesenchymal NCC fate.

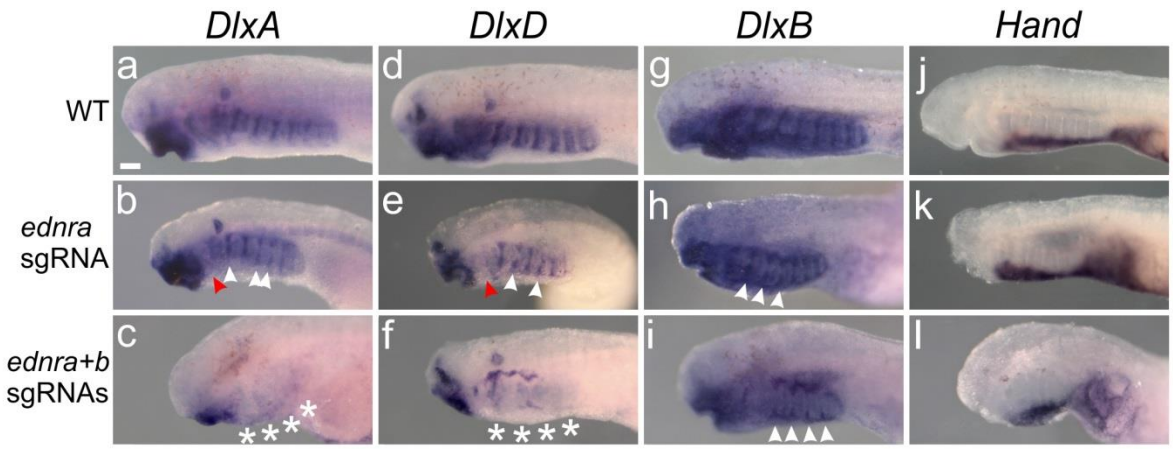

Fig. 4. *DlxA, -D, -B,* and *Hand* expression in sgRNA injected lampreys. a-l, left lateral views showing ISHs of *DlxA* (a-c), *DlxD* (d-f), *DlxB* (g-i), and *Hand* (j-l), reveal disrupted expression in sgRNA injected lampreys. Red arrowheads indicate missing PA1 expression in *ednra* mutants. White arrowheads indicate PAs missing only intermediate expression. White stars indicate highly disrupted pharyngeal expression in *ednra+b* sgRNA injected lampreys. Scale bar in a represents 100 µm and applies to all panels.

Disruption of *ednrb*, *ednE*, and *edn3*

Previous work on the Edn3/Ednrb signaling in gnathostomes has shown it is necessary for normal NCC-derived pigment cell proliferation, differentiation and migration (Kawasaki-Nishihara et al., 2011; Krauss et al., 2014; RM et al., 2006). Similarly, Cas9-driven mutagenesis of *P. marinus ednrb* or *ednE,* and *X. laevis edn3* all consistently produced a severely (>50%) reduced pigmentation phenotype, together with pigment cell migration defects (Fig. 1, Fig. S6-8; see Tab. S1 for phenotype counts). In *P. marinus,* these defects included the accumulation of melanocytes around the otic placode starting at st. 25, which is the region where they first emerge in WT larvae. Despite the extensive expression of *ednrb* in migratory pre-skeletal NCCs (see Fig. S3), no defects in head skeleton formation were observed during *ednrb* mutagenesis. Concordantly, *DlxA, DlxB,* and *DlxD* in the nascent head skeleton at st. 26.5 were completely normal (not shown). In *X. laevis,* unilateral injection (one of two blastomeres) of *edn3* sgRNA designed to target both *edn3.L* and *-S* led to a unilateral disruption in pigment migration. Injection at the one cell stage led to a higher rate of overall pigment loss, including some completely leucistic larvae (Fig. S8; n=11/71). These leucistic larvae were lacking both NCC-derived melanophores on the body, as well as iridophores overlying the gut (Fig. S8).

In placental mammals, Ednrb(1) signaling is also required for proper formation and migration of some, but not all NCCs that generate the peripheral nervous system (PNS), namely enteric neurons of the esophagus and gut (RM et al., 2006; Sanchez-Mejias et al., 2010; Stanchina et al., 2006). We thus looked at the distribution of Neurofilament protein, a marker of neuronal processes, and *Soxb1b*, *Jagged*, *Gremlin*, and *Phox* mRNA, which mark various subsets of the PNS in *P. marinus*. Neurofilament IHC revealed the complete absence of a population of ganglia above the heart at st. 27 (Fig. 5), in the region of the pronephros, surrounding the esophagus (red arrows in Fig. 5D). By contrast, we observed no defects in gene expression within, or differentiation of cranial ganglia (which may not form from NCCs in lamprey), dorsal root ganglia (DRGs), or the putative *Phox*-positive Schwann Cell Precursors (SCPs) overlying the gut (Green et al., 2017) (Fig. S3, some data not pictured; see Tab. S2). This suggests that, like in bony vertebrates, despite Endothelin receptors being expressed throughout all or most of the PNS, they are critical for the normal development of only a small subset of these neurons and glia. Given that recent evidence suggests that enteric neurons are a jawed vertebrate synapomorphy (Green et al., 2017), the Endothelin-independent NCC subtypes (DRGs and SCPs) might represent the most ancient type of neural crest derivative, predating the cooption of Endothelin signaling by NCCs. Alternatively, these PNS derivatives may simply have a more robust developmental program, with a higher level of redundancy that negates Endothelin signaling as a necessity. In any case, in both lamprey and bony vertebrates, only a

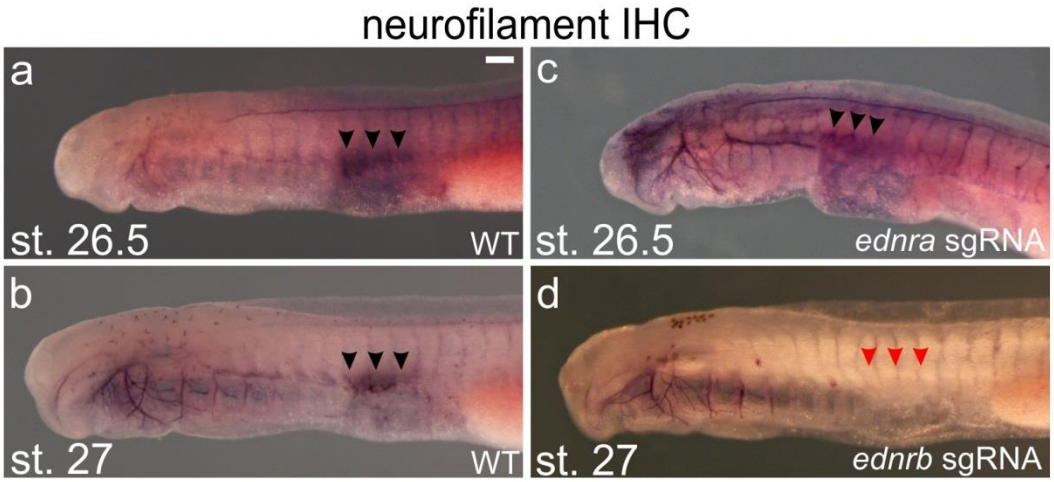

Fig. 5. Neurofilament staining reveals a missing neuronal population in *ednrb* mutants. a-d, left lateral views showing neurofilament IHC in WT and sgRNA injected lampreys. Stage and assay are indicated in each panel. Scale bar in a represents 100 µm and applies to all panels. Black arrowheads indicate peripheral nervous system derivatives near the esophagus and pronephros present in WT from st 26-27 (a and b) and *ednra* mutants (c) at st. 26.5. Red arrowheads in d indicate the lack of this neuronal cluster in *ednrb* mutants.

small subset of the PNS relies on Endothelin signaling in its proper formation; thus we infer this condition was also present in the vertebrate common ancestor.

P. marinus combined *ednra + ednrb* disruption

Given their brief expression overlap in NCCs (Square et al., 2016a), to discern the level of Ednr redundancy and/or synergy in *P. marinus* we targeted *ednra* and *ednrb* simultaneously using three different pairs of sgRNAs. All three of these assays yielded a range of different phenotypes and combinations of phenotypes all characteristic of the single receptor perturbations. This result would be expected given the imperfect penetrance of Cas9-mediated mutagenesis occurring independently at two different loci with different developmental effects. Despite this range in phenotypes, we did find that the most consistent phenotype was combinatorial: a reduced head skeleton, an enlarged heart, and a very small number of pigment cells (Fig. 1F and Fig. S9; see Tab. S1 for counts). Notably, the onset of the head skeleton reduction was at the same stage as in the *ednra* perturbations (24-25), but was perhaps slightly more severe than in the *ednra*-only mutants by st. 26+ (compare Fig. 1c and f). The most severely affected individuals were missing their upper lip, as in *ednra* mutagenesis. Genotyping of three individuals with a strong phenotype revealed no more than 17% WT alleles between *ednra* and *ednrb*, with one individual (Fig. 1F) returning 0% WT alleles (Fig. S9). ISH at st. 26.5 revealed that *SoxE2* expression was disrupted throughout the head, but not *Hand* and *MsxB* in the ventral and dorsalmost domains (save PA1 derivatives) just as in the *ednra* single perturbations. Strikingly, in contrast to the *ednra*-only perturbations, *DlxA, -B, -C,* and *-D* expression were severely reduced at st. 26.5 in the intermediate domain of pharyngeal skeletal precursors, with *DlxA* being wholly undetectable in the future branchial skeleton in many cases (Fig. 4). Alcian blue (n=6) and Toluidene blue (n=2) staining on the most highly affected specimens at st. 30 confirmed the presence of a highly truncated, but never completely absent head skeleton, wherein some or most cartilage elements were abbreviated or completely missing.

Genotyping individuals

To verify that affected phenotypes in *P. marinus* were a predictor of a high proportion of mutagenized target loci, we genotyped three or more individuals with severe phenotypes from each perturbation (see Fig. S4, S5, S6, S7, and S9). In most cases, WT alleles accounted for 0-15% of the alleles we observed; in only three affected individuals did WT alleles exceed 20% of sampled loci, and never accounted for more than 33%. In *X. laevis,* we genotyped at least four specimens from each of the *edn1* and *ednra* experiments (Fig. S10), including some *sox9.S* ISH stained larvae (both affected and

unaffected individuals). All affected larvae displayed <25% WT loci. Of the three unaffected (but injected) *X. laevis* individuals we genotyped from the *ednra* assays, one had 100% WT *ednra.L* and *-S* loci, while the two others showed moderate or even high, but not complete levels of indels (37% and 27% WT loci). Considering the tetraploid nature of the *X. laevis* genome, it seems logical that the mutation rate in this species might be a particularly poor predictor of an effect given that we lack information regarding whether or not these indel alleles are combined in different cell types. For *edn3*, we genotyped a single, leucistic individual that returned 7% WT alleles between *edn3.L* and *edn3.S* (Fig. S8).

Of all 35 specimens' loci thoroughly genotyped for this work (nine or more allele sequences returned), the smallest number of unique indel alleles ever observed was four, which occurred six times between *P. marinus* and *X. laevis*. Most frequently, we found six or more alleles at a given locus in a single specimen (considering the homoeologous "L" and "S" *X. laevis* loci separately), which indicates that biallelic Cas9-driven mutagenesis in the pre-mitotic zygote may be rare, or biochemically implausible in these two species. Notably, when insertions of DNA fragments were discovered, these motifs were always fully or mostly accounted for on the reverse and/or forward strand very close to the target site/lesion (see alleles in Fig. S4, S6, S7, S9, and S10). This insertion pattern can also be seen in other studies (Flowers et al., 2014; Square et al., 2015b; Wang et al., 2015).

CRISPR/Cas9 controls

To control for off-target effects, each *P. marinus* Endothelin gene was targeted via at least two unique target sites, as previously described (Square et al., 2015b). For *ednA, ednE, ednra,* and *ednrb*, all sgRNAs targeting the same gene consistently produced the same mutant phenotype, though usually with different efficiencies (see Tab S1). Observing the same phenotype by targeting the same gene in two locations mitigates the possibility of off-target effects of a given sgRNA. To further test for sgRNA specificity in *P. marinus,* and to ensure that sgRNA or Cas9 injections do not inherently cause any phenotype we address here, we designed a negative control sgRNA that bridges an exon/exon boundary in the *P. marinus ednrb* gene transcript (see Tab. S1). This sgRNA should target a locus that does not exist within the *ednrb* coding sequence, or elsewhere in the *P. marinus* genome to our knowledge. Negative control sgRNA was injected as all others (see methods) with either cas9 mRNA or Cas9 protein (see Tab. S1), and resulted in neither pigment reduction nor any other consistent, discernable larval phenotype.

For *X. laevis ednra* and *edn1* experiments, sgRNAs targeting both homoeologs were coinjected to strongly disrupt gene function, as previously described (Wang et al., 2015). As negative controls for the *edn1* and *ednra* experiments, we targeted the *-L* and *-S* loci separately (see methods for injection specifics). These injections caused no discernible head skeleton malformations at st. 48 (see Tab. S1). For

edn3, a single sgRNA was designed to target both loci, confirming the morpholino-based knock down phenotype previously published in *X. laevis* (Kawasaki-Nishihara et al., 2011).

Discussion

Outside of vertebrates, the only known rudiment of the endothelin pathway is a single *ednr-like* gene in amphioxus, which does not appear to be expressed at larval stages (Yu et al., 2008a), and is missing an amino acid residue critical for binding Edns. Furthermore, there are no known endothelin ligands or Endothelin Converting Enzymes (ECEs) in any living invertebrate (Braasch and Schartl, 2014; Braasch et al., 2009; Martinez-Morales et al., 2007; Square et al., 2016a). Thus the initial appearance of *bona fide* Endothelin signaling occurred in early stem vertebrates after the evolution of the first *ECE* and *edn* genes. This was followed by one or more duplication events for each of these genes, as well as *ednr,* in stem vertebrates, likely during whole-genome duplication(s) (Braasch and Schartl, 2014; Smith and Keinath, 2015). Previous work showed that *P. marinus* shares extensive gnathostome-like expression of DV patterning genes in its nascent head skeleton (Cerny et al., 2010). Similarly, they possess a gnathostome-like complement of Edns and Ednrs in their genome, which are expressed in and around different NCC derivatives. Here, we confirm grossly similar specialization in Endothelin pathway function between bony vertebrates and *P. marinus*. Both groups rely on Ednra for proper heart formation, and head skeleton proliferation via skeletogenic *soxE* maintenance, while Ednrb signaling is critical for the development of pigment cells and some neuronal derivatives. This suggests the evolution of distinct Ednra and Ednrb signaling pathways was a critical step in the segregation of ectomesenchymal (heart and skeleton) and non-ectomesenchymal (pigment and nervous system) NCC lineages in the first vertebrates. Beyond broad-scale NCC fate and migration, the use of Endothelin signaling in head skeleton patterning is somewhat divergent between these two lineages. In *P. marinus*, Ednra and Ednrb work synergistically to activate nested *Dlx,* but not *Hand* transcription throughout the pharynx, while in all bony vertebrates thus addressed Ednra functions alone to activate both nested *dlx* and *hand* transcription. This suggests that the evolution of Ednra-dependent jaw joint and facial bone development made use of already-present Endothelin-driven Dlx expression in the head skeleton, which predated these gnathostome morphological innovations by 75 million years or more (Fig. 6). Barring extreme convergence, this means that the head skeleton "pre-pattern" previously inferred to have existed in the vertebrate common ancestor (Cerny et al., 2010) was at least partially driven by Endothelin signaling.

Thus, phylogenetically coincident with the evolution of *bona fide* Endothelin signaling and the diversification of these ligands and receptors, NCCs began utilizing these pathways in the development of most NCC subtypes during stem vertebrate evolution. This includes driving *soxE* and *dlx* transcription in post-migratory NCCs. The relative timing of these events is not clear, though post-migratory

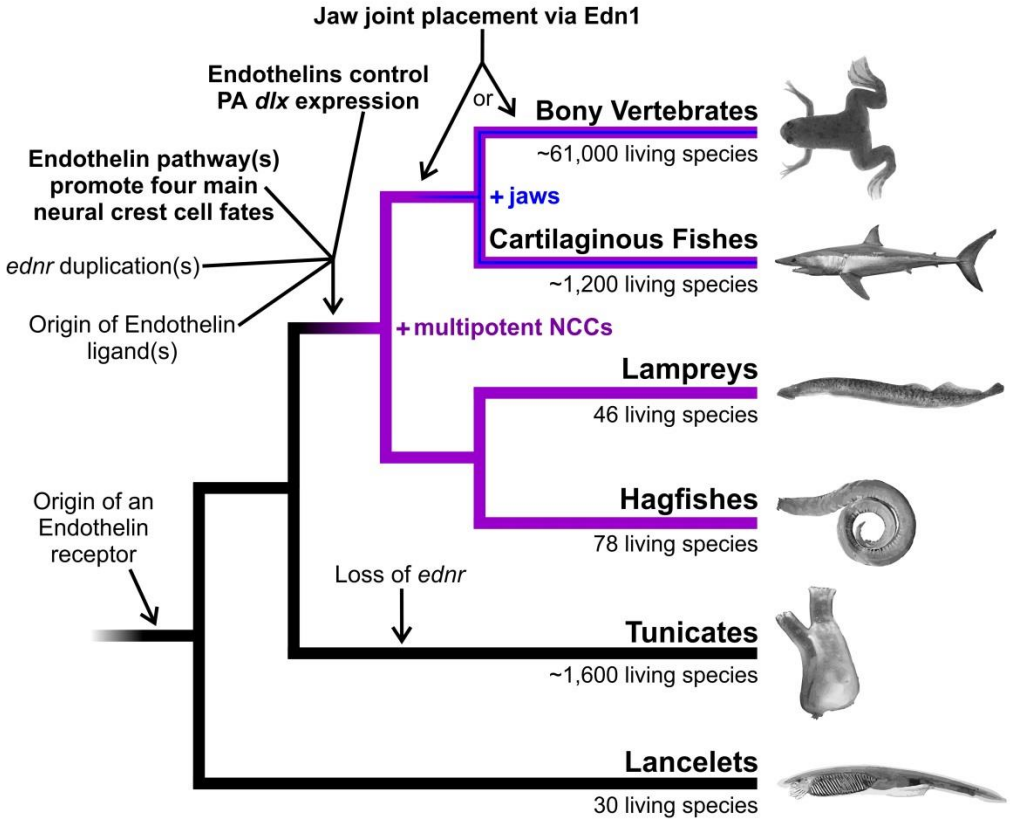

Fig. 6. Endothelin pathway evolution in chordates. A summary tree depicting the inferred phylogenetic position of various events in Endothelin pathway evolution. In stem chordates, the first *ednr* evolved, which was apparently lost in stem tunicates. In stem vertebrates first occurred Endothelin ligand(s) and Endothelin Converting Enzyme(s), *ednr* duplication(s), the specialization of these pathways into different NCC fates, and their driving *dlx* expression in the pharynx. These processes that occurred in stem vertebrates coincide with the appearance of multipotent NCCs (purple). Later, either in stem gnathostomes or stem bony vertebrates (osteichythyians), the Edn-driven *dlx* expression was coopted for jaw joint placement in the head skeleton.

ectomesenchyme that sustains *ednra* expression vs. post-migratory non-ectomesenchyme that sustains only *ednrb* expression appears to be one of the most deeply conserved differences in the specialization of these receptors (see (Square et al., 2016a) for a table comparing *ednr* expression domains in vertebrates). Cartilaginous fish (Chondrichythian) Endothelin functional assays are needed to determine if Edn1/Ednra-dependent jaw joint placement is a synapomorphy of jawed vertebrates as a whole, or bony fishes only, however using the inferences generated here we predict that Chondrichythyian *dlx* expression in the pharynx will also be driven by some form of Endothelin signaling.

Methods

Lamprey husbandry, Cas9 manipulations, and assays

P. marinus fertilizations, husbandry, embryo injections, design and preparation of all sgRNAs, preparation of Cas9 mRNA, and preparation of Cas9 protein were as described previously (Square et al., 2015b), with some slight modification. Lamprey zygotes were injected with a 5 nL droplet containing 400 pg of sgRNA, and either 800 pg of Cas9 protein (a 2:1 ratio of protein:sgRNA by mass) or 1 ng of Cas9 mRNA, 5 mg/mL lysinated rhodamine dextran, and nuclease free water. For the combined *ednra* + *ednrb* experiments, 200 pg of each sgRNA was used with .8 ug of Cas9 protein. Genotyping was carried out as described previously (Square et al., 2015b) to confirm that animals with severe phenotypes generated from each sgRNA were reflective of successful mutagenesis. A list of target sites and genotyping primers can be found in Tab. S1. Lampreys were staged according to (Tahara, 1988). As previously noted, all injected animals (including negative controls and otherwise unaffected individuals) displayed a slight developmental delay (~10%, i.e. 10 day old animals appeared 9 days old); due to this, developmental events such as somite segregation and yolk absorption were used to stage the embryos and larvae, and injected animals were usually fixed 1-3 days later than their WT siblings. All injected embryos were sorted for fluorescence at 4-6 days post fertilization; individuals lacking fluorescence were discarded. IHC was as described previously (McCauley and Bronner-Fraser, 2002), with the addition of 1% dimethyl sulfoxide (DMSO) to the phosphate buffer solution prior to the blocking step. For IHC, the WT embryos shown were stained in parallel (in the same tubes). All ISH and alcian blue cartilage staining was carried out as described previously (Cerny et al., 2010; Square et al., 2015a). Control ISHs on WTs were always performed in parallel (in the same tubes, with tails cut off for identification) to ensure proper signal development. See supplementary methods for drug treatment information. All lamprey manipulations were in accordance with CU-Boulder IACUC protocol #2392.

Xenopus husbandry, manipulations, and assays

X. laevis fertilizations and husbandry were performed according to standard methods. As reported previously (Wang et al., 2015), both the "long" (-*L*) and "short" (-*S*) alleles of the allotetraploid *X. laevis* were simultaneously targeted to make null animals for each *edn1*, *edn3*, and *ednra*. Zygotes or two-cell embryos were injected with a 5 nL droplet containing 800 pg of total sgRNA (a single sgRNA targeting both *edn3.L* (*edn3-a*) and *edn3.S* (-*b*), or 400 pg of each of two sgRNAs in for *edn1* and *ednra*), and either 1 ng of Cas9 mRNA, or 1.6 ng of Cas9 protein (a 2:1 protein:sgRNA ratio, as suggested by the manufacturer). All staging was according to (Nieuwkoop and Faber, 1956). A list of target sites and genotyping primers can be found in Tab. S1. All ISH and alcian blue cartilage staining was carried out as described previously (Square et al., 2015a). Control ISHs on WTs were always performed in parallel (in the same tubes, with tails cut off for identification) to ensure proper signal development. All *Xenopus* manipulations were in accordance with CU-Boulder IACUC protocol #2392.

Bibliography

Abassi, Z.A., Golomb, E., Bridenbaugh, R., Keiser, H.R., 1993. METABOLISM OF ENDOTHELIN-1 AND BIG ENDOTHELIN-1 BY RECOMBINANT NEUTRAL ENDOPEPTIDASE EC.3.4.24.11. British Journal of Pharmacology 109, 1024-1028.

Abitua, P.B., Wagner, E., Navarrete, I.A., Levine, M., 2012. Identification of a rudimentary neural crest in a non-vertebrate chordate. Nature 492, 104-+.

Abu-Issa, R., Smyth, G., Smoak, I., Yamamura, K., Meyers, E.N., 2002. Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. Development 129, 4613-4625.

Abzhanov, A., Tabin, C.J., 2004. Shh and Fgf8 act synergistically to drive cartilage outgrowth during cranial development. Dev Biol (N Y 1985) 273, 134-148.

Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A., Brulet, P., 1995. FOREBRAIN AND MIDBRAIN REGIONS ARE DELETED IN OTX2(-/-) MUTANTS DUE TO A DEFECTIVE ANTERIOR NEUROECTODERM SPECIFICATION DURING GASTRULATION. Development 121, 3279-3290.

Adachi, M., Furuichi, Y., Miyamoto, C., 1994a. IDENTIFICATION OF A LIGAND-BINDING SITE OF THE HUMAN ENDOTHELIN-A RECEPTOR AND SPECIFIC REGIONS REQUIRED FOR LIGAND SELECTIVITY. European Journal of Biochemistry 220, 37-43.

Adachi, M., Furuichi, Y., Miyamoto, C., 1994b. IDENTIFICATION OF SPECIFIC REGIONS OF THE HUMAN ENDOTHELIN-B RECEPTOR REQUIRED FOR HIGH-AFFINITY BINDING WITH ENDOTHELIN-3. Biochimica Et Biophysica Acta-Molecular Cell Research 1223, 202-208.

Akimenko, M.A., Ekker, M., Wegner, J., Lin, W., Westerfield, M., 1994. COMBINATORIAL EXPRESSION OF 3 ZEBRAFISH GENES RELATED TO DISTAL-LESS - PART OF A HOMEOBOX GENE CODE FOR THE HEAD. Journal of Neuroscience 14, 3475-3486.

Alexander, C., Zuniga, E., Blitz, I.L., Wada, N., Le Pabic, P., Javidan, Y., Zhang, T.L., Cho, K.W., Crump, J.G., Schilling, T.F., 2011. Combinatorial roles for BMPs and Endothelin 1 in patterning the dorsal-ventral axis of the craniofacial skeleton. Development 138, 5135-5146.

Antonopoulou, I., Mavrogiannis, L.A., Wilkie, A.O.M., Morriss-Kay, G.M., 2004. Alx4 and Msx2 play phenotypically similar and additive roles in skull vault differentiation. Journal of Anatomy 204, 487-499.

Arduini, B.L., Bosse, K.M., Henion, P.D., 2009. Genetic ablation of neural crest cell diversification. Development 136, 1987-1994.

Aybar, M.J., Nieto, M.A., Mayor, R., 2003. Snail precedes Slug in the genetic cascade required for the specification and migration of the Xenopus neural crest. Development 130, 483-494.

Barlow, A.J., Bogardi, J.P., Ladher, R., Francis-West, P.H., 1999. Expression of chick Barx-1 and its differential-regulation by FGF-8 and BMP signaling in the maxillary primordia. Developmental Dynamics 214, 291-302.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709-1712.

Basch, M.L., Bronner-Fraser, M., Garcia-Castro, M.I., 2006. Specification of the neural crest occurs during gastrulation and requires Pax7. Nature 441, 218-222.

Bassett, A.R., Tibbit, C., Ponting, C.P., Liu, J.L., 2013. Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. Cell Rep 4, 220-228.

Baynash, A.G., Hosoda, K., Giaid, A., Richardson, J.A., Emoto, N., Hammer, R.E., Yanagisawa, M., 1994. INTERACTION OF ENDOTHELIN-3 WITH ENDOTHELIN-B RECEPTOR IS ESSENTIAL FOR DEVELOPMENT OF EPIDERMAL MELANOCYTES AND ENTERIC NEURONS. Cell 79, 1277-1285.

Beamish, F.W.H., 1980. Biology of the North-American Anadromous Sea Lamprey, Petromyzon-Marinus. Can J Fish Aquat Sci 37, 1924-1943.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Nekrasov, V., 2013. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9.

Beverdam, A., Brouwer, A., Reijnen, M., Korving, J., Meijlink, F., 2001. Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice. Development 128, 3975-3986.

Beverdam, A., Meijlink, F., 2001. Expression patterns of group-I aristaless-related genes during craniofacial and limb development. Mechanisms of Development 107, 163-167.

Beverdam, A., Merlo, G.R., Paleari, L., Mantero, S., Genova, F., Barbieri, O., Janvier, P., Levi, G., 2002. Jaw transformation with gain of symmetry after Dlx5/Dlx6 inactivation: Mirror of the past? Genesis 34, 221-227.

Bianchetti, L., Oudet, C., Poch, O., 2002. M13 endopeptidases: New conserved motifs correlated with structure, and simultaneous phylogenetic occurrence of PHEX and the bony fish. Proteins-Structure Function and Bioinformatics 47, 481-488.

Bland, N.D., Pinney, J.W., Thomas, J.E., Turner, A.J., Isaac, R.E., 2008. Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships. Bmc Evolutionary Biology 8.

Blitz, I.L., Biesinger, J., Xie, X.H., Cho, K.W.Y., 2013. Biallelic Genome Modification in F-0 Xenopus tropicalis Embryos Using the CRISPR/Cas System. Genesis 51, 827-834.

Bonano, M., Tribulo, C., De Calisto, J., Marchant, L., Sanchez, S.S., Mayor, R., Aybar, M.J., 2008. A new role for the Endothelin-1/Endothelin-A receptor signaling during early neural crest specification. Developmental Biology 323, 114-129.

Borday-Birraux, V., Van der Heyden, C., Debiais-Thibaud, M., Verreijdt, L., Stock, D.W., Huysseune, A., Sire, J.Y., 2006. Expression of Dlx genes during the development of the zebrafish pharyngeal dentition: evolutionary implications. Evolution & Development 8, 130-141.

Bowes, J.B., Snyder, K.A., Segerdell, E., Gibb, R., Jarabek, C., Noumen, E., Pollet, N., Vize, P.D., 2008. Xenbase: a Xenopus biology and genomics resource. Nucleic Acids Research 36, D761-D767.

Braasch, I., Schartl, M., 2014. Evolution of endothelin receptors in vertebrates. General and Comparative Endocrinology 209, 21-34.

Braasch, I., Volff, J.-N., Schartl, M., 2009. The Endothelin System: Evolution of Vertebrate-Specific Ligand-Receptor Interactions by Three Rounds of Genome Duplication. Molecular Biology and Evolution 26, 783-799.

Brazeau, M.D., Friedman, M., 2015. The origin and early phylogenetic history of jawed vertebrates. Nature 520, 490-497.

Brown, S.T., Wang, J.M., Groves, A.K., 2005. Dlx gene expression during chick inner ear development. Journal of Comparative Neurology 483, 48-65.

Bryant, S., Herdy, J., Amemiya, C., Smith, J., 2016. Characterization of Somatically-Eliminated Genes During Development of the Sea Lamprey (Petromyzon marinus). Mol Biol Evol.

Campo-Paysaa, F., Jandzik, D., Takio-Ogawa, Y., Cattell, M.V., Neef, H.C., Langeland, J.A., Kuratani, S., Medeiros, D.M., Mazan, S., Kuraku, S., Laudet, V., Schubert, M., 2015. Evolution of retinoic acid receptors in chordates: insights from three lamprey species, Lampetra fluviatilis, Petromyzon marinus, and Lethenteron japonicum. Evodevo 6.

Cattell, M., Lai, S., Cerny, R., Medeiros, D.M., 2011. A New Mechanistic Scenario for the Origin and Evolution of Vertebrate Cartilage. Plos One 6.

Cerny, R., Cattell, M., Sauka-Spengler, T., Bronner-Fraser, M., Yu, F., Medeiros, D.M., 2010. Evidence for the prepattern/cooption model of vertebrate jaw evolution. Proceedings of the National Academy of Sciences of the United States of America 107, 17262-17267.

Cerny, R., Meulemans, D., Berger, J., Wilsch-Brauninger, M., Kurth, T., Bronner-Fraser, M., Epperlein, H.H., 2004. Combined intrinsic and extrinsic influences pattern cranial neural crest migration and pharyngeal arch morphogenesis in axolotl. Developmental Biology 266, 252-269.

Charite, J., McFadden, D.G., Merlo, G., Levi, G., Clouthier, D.E., Yanagisawa, M., Richardson, J.A., Olson, E.N., 2001. Role of Dlx6 in regulation of an endothelin-1-dependent, dHAND branchial arch enhancer. Genes & Development 15, 3039-3049.

Clouthier, D.E., Garcia, E., Schilling, T.F., 2010. Regulation of Facial Morphogenesis by Endothelin Signaling: Insights From Mice and Fish. American Journal of Medical Genetics Part A 152A, 2962-2973.

Clouthier, D.E., Hosoda, K., Richardson, J.A., Williams, S.C., Yanagisawa, H., Kuwaki, T., Kumada, M., Hammer, R.E., Yanagisawa, M., 1998. Cranial and cardiac neural crest defects in endothelin-A receptor-deficient mice. Development 125, 813-824.

Compagnucci, C., Debiais-Thibaud, M., Coolen, M., Fish, J., Griffin, J.N., Bertocchini, F., Minoux, M., Rijli, F.M., Borday-Birraux, V., Casane, D., Mazan, S., Depew, M.J., 2013. Pattern and polarity in the development and evolution of the gnathostome jaw: Both conservation and heterotopy in the branchial arches of the shark, Scyliorhinus canicula. Developmental Biology 377, 428-448.

Cong, L., Ran, F.A., Cox, D., Lin, S.L., Barretto, R., Habib, N., Hsu, P.D., Wu, X.B., Jiang, W.Y., Marraffini, L.A., Zhang, F., 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339, 819-823.

Creuzet, S., Couly, G., Vincent, C., Le Douarin, N.M., 2002. Negative effect of Hox gene expression on the development of the neural crest-derived facial skeleton. Development 129, 4301-4313.

Crump, J.G., Maves, L., Lawson, N.D., Weinstein, B.M., Kimmel, C.B., 2004a. An essential role for Fgfs in endodermal pouch formation influences later craniofacial skeletal patterning. Development 131, 5703-5716.

Crump, J.G., Maves, L., Lawson, N.D., Weinstein, B.M., Kimmel, C.B., 2004b. An essential role for Fgfs in endodermal pouch formation influences later craniofacial skeletal patterning. Development 131, 5703-5716.

De Robertis, E.M., Kuroda, H., 2004. Dorsal-ventral patterning and neural induction in Xenopus embryos. Annual Review of Cell and Developmental Biology 20, 285-308.

Debiais-Thibaud, M., Metcalfe, C.J., Pollack, J., Germon, I., Ekker, M., Depew, M., Laurenti, P., Borday-Birraux, V., Casane, D., 2013. Heterogeneous Conservation of Dlx Paralog Co-Expression in Jawed Vertebrates. Plos One 8.

Dee, C.T., Szymoniuk, C.R., Mills, P.E.D., Takahashi, T., 2013. Defective neural crest migration revealed by a Zebrafish model of Alx1-related frontonasal dysplasia. Human Molecular Genetics 22, 239-251.

Delarbre, C., Gallut, C., Barriel, V., Janvier, P., Gachelin, G., 2002. Complete mitochondrial DNA of the hagfish, Eptatretus burgeri: The comparative analysis of mitochondrial DNA sequences strongly supports the cyclostome monophyly. Molecular Phylogenetics and Evolution 22, 184-192.

Depew, M.J., Lufkin, T., Rubenstein, J.L.R., 2002. Specification of jaw subdivisions by Dix genes. Science 298, 381-385.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E., 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32, 1262-U1130.

Donoghue, P.C.J., Keating, J.N., 2014. EARLY VERTEBRATE EVOLUTION. Palaeontology 57, 879-893.

Donoghue, P.C.J., Purnell, M.A., 2009. Distinguishing heat from light in debate over controversial fossils. Bioessays 31, 178-189.

Dupin, E., Le Douarin, N.M., 2014. The Neural Crest, A Multifaceted Structure of the Vertebrates. Birth Defects Research Part C-Embryo Today-Reviews 102, 187-209.

Dutton, K.A., Pauliny, A., Lopes, S.S., Elworthy, S., Carney, T.J., Rauch, J., Geisler, R., Haffter, P., Kelsh, R.N., 2001. Zebrafish colourless encodes sox10 and specifies non-ectomesenchymal neural crest fates. Development 128, 4113-4125.

Ellies, D.L., Langille, R.M., Martin, C.C., Akimenko, M.A., Ekker, M., 1997. Specific craniofacial cartilage dysmorphogenesis coincides with a loss of dlx gene expression in retinoic acid-treated zebrafish embryos. Mechanisms of Development 61, 23-36.

Emoto, N., Yanagisawa, M., 1995. ENDOTHELIN-CONVERTING ENZYME-2 IS A MEMBRANE-BOUND, PHOSPHORAMIDON-SENSITIVE METALLOPROTEASE WITH ACIDIC PH OPTIMUM. Journal of Biological Chemistry 270, 15262-15268.

Firulli, A.B., 2003. A HANDful of questions: the molecular biology of the heart and neural crest derivatives (HAND)-subclass of basic helix-loop-helix transcription factors. Gene 312, 27-40.

Fish, J.L., Sklar, R.S., Woronowicz, K.C., Schneider, R.A., 2014. Multiple developmental mechanisms regulate species-specific jaw size. Development 141, 674-684.

Fish, J.L., Villmoare, B., Koebernick, K., Compagnucci, C., Britanova, O., Tarabykin, V., Depew, M.J., 2011. Satb2, modularity, and the evolvability of the vertebrate jaw. Evolution & Development 13, 549-564.

Flowers, G.P., Timberlake, A.T., Mclean, K.C., Monaghan, J.R., Crews, C.M., 2014. Highly efficient targeted mutagenesis in axolotl using Cas9 RNA-guided nuclease. Development 141, 2165-2171.

Fujimi, T.J., Mikoshiba, K., Aruga, J., 2006. Xenopus Zic4: Conservation and diversification of expression profiles and protein function among the Xenopus Zic family. Developmental Dynamics 235, 3379-3386.

Fujimoto, S., Oisi, Y., Kuraku, S., Ota, K.G., Kuratani, S., 2013. Non-parsimonious evolution of hagfish Dlx genes. Bmc Evolutionary Biology 13.

Gagnon, J.A., Valen, E., Thyme, S.B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T.G., Zimmerman, S., Richter, C., Schier, A.F., 2014. Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. PLoS ONE 9.

Gans, C., Northcutt, R.G., 1983. NEURAL CREST AND THE ORIGIN OF VERTEBRATES - A NEW HEAD. Science 220, 268-273.

Garnett, A.T., Square, T.A., Medeiros, D.M., 2012. BMP, Wnt and FGF signals are integrated through evolutionarily conserved enhancers to achieve robust expression of Pax3 and Zic genes at the zebrafish neural plate border. Development 139, 4220-4231.

Gaunt, S.J., Blum, M., Derobertis, E.M., 1993. EXPRESSION OF THE MOUSE GOOSECOID GENE DURING MID-EMBRYOGENESIS MAY MARK MESENCHYMAL CELL LINEAGES IN THE DEVELOPING HEAD, LIMBS AND BODY WALL. Development 117, 769-778.

Gess, R.W., Coates, M.I., Rubidge, B.S., 2006. A lamprey from the Devonian period of South Africa. Nature 443, 981-984.

Gillis, J.A., Modrell, M.S., Baker, C.V.H., 2013. Developmental evidence for serial homology of the vertebrate jaw and gill arch skeleton. Nature Communications 4.

Gordon, C.T., Brinas, I.M.L., Rodda, F.A., Bendall, A.J., Farlie, P.G., 2010. Role of Dlx genes in craniofacial morphogenesis: Dlx2 influences skeletal patterning by inducing ectomesenchymal aggregation in ovo. Evolution & Development 12, 459-473.

Green, S.A., Simoes-Costa, M., Bronner, M.E., 2015. Evolution of vertebrates as viewed from the crest. Nature 520, 474-482.

Green, S.A., Uy, B.R., Bronner, M.E., 2017. Ancient evolutionary origin of vertebrate enteric neurons from trunk-derived neural crest. Nature 544, 88-+.

Grinblat, Y., Sive, H., 2001. zic gene expression marks anteroposterior pattern in the presumptive neurectoderm of the zebrafish gastrula. Developmental Dynamics 222, 688-693.

Guo, X.G., Zhang, T.J., Hu, Z., Zhang, Y.Q., Shi, Z.Y., Wang, Q.H., Cui, Y., Wang, F.Q., Zhao, H., Chen, Y.L., 2014. Efficient RNA/Cas9-mediated genome editing in Xenopus tropicalis. Development 141, 707-714.

Hall, B.K., 2000. The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. Evolution & Development 2, 3-5.

Hauptmann, G., Gerster, T., 2000. Combinatorial expression of zebrafish Brn-1- and Brn-2-related POU genes in the embryonic brain, pronephric primordium, and pharyngeal arches. Developmental Dynamics 218, 345-358.

Heimberg, A.M., Cowper-Sallari, R., Semon, M., Donoghue, P.C.J., Peterson, K.J., 2010. microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. Proceedings of the National Academy of Sciences of the United States of America 107, 19379-19383.

Hernandez-Vega, A., Minguillon, C., 2011. The Prx1 Limb Enhancers: Targeted Gene Expression in Developing Zebrafish Pectoral Fins. Developmental Dynamics 240, 1977-1988.

Hildebrand, M., Goslow, G.E., 2001. Analysis of vertebrate structure, 5th ed. John Wiley, New York.

Hirata, Y., Yoshimi, H., Takata, S., Watanabe, T.X., Kumagai, S., Nakajima, K., Sakakibara, S., 1988. CELLULAR MECHANISM OF ACTION BY A NOVEL VASOCONSTRICTOR ENDOTHELIN IN CULTURED RAT VASCULAR SMOOTH-MUSCLE CELLS. Biochemical and Biophysical Research Communications 154, 868-875.

Hohimer, A.R., Giraud, G.D., Thornburg, K.L., Jeffrey, J., Cartwright, E.J., Sharpe, P.T., 1993. EXPRESSION OF HOMEOBOX GENES MSX-1 AND MSX-2 IN EMBRYONIC MOUSE HEART. Circulation 88, 282-282.

Holland, L.Z., Holland, N.D., 1998. Developmental gene expression in amphioxus: New insights into the evolutionary origin of vertebrate brain regions, neural crest, and rostrocaudal segmentation. American Zoologist 38, 647-658.

Howard-Ashby, M., Materna, S.C., Brown, C.T., Chen, L., Cameron, R.A., Davidson, E.H., 2006. Identification and characterization of homeobox transcription factor genes in Strongylocentrotus purpuratus, and their expression in embryonic development. Developmental Biology 300, 74-89.

Hunt, P., Gulisano, M., Cook, M., Sham, M.H., Faiella, A., Wilkinson, D., Boncinelli, E., Krumlauf, R., 1991a. A DISTINCT HOX CODE FOR THE BRANCHIAL REGION OF THE VERTEBRATE HEAD. Nature 353, 861-864.

Hunt, P., Whiting, J., Muchamore, I., Marshall, H., Krumlauf, R., 1991b. HOMEOBOX GENES AND MODELS FOR PATTERNING THE HINDBRAIN AND BRANCHIAL ARCHES. Development, 187-&.

Hunt, P., Wilkinson, D., Krumlauf, R., 1991c. PATTERNING THE VERTEBRATE HEAD - MURINE HOX-2 GENES MARK DISTINCT SUBPOPULATIONS OF PREMIGRATORY AND MIGRATING CRANIAL NEURAL CREST. Development 112, 43-50.

Hwang, W.Y., Fu, Y.F., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R.J., Joung, J.K., 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31, 227-229.

Ikmi, A., McKinney, S.A., Delventhal, K.M., Gibson, M.C., 2014. TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. Nature Communications 5.

Imai, K.S., Stolfi, A., Levine, M., Satou, Y., 2009. Gene regulatory networks underlying the compartmentalization of the Ciona central nervous system. Development 136, 285-293.

Itoh, Y., Yanagisawa, M., Ohkubo, S., Kimura, C., Kosaka, T., Inoue, A., Ishida, N., Mitsui, Y., Onda, H., Fujino, M., Masaki, T., 1988. CLONING AND SEQUENCE-ANALYSIS OF CDNA-ENCODING THE PRECURSOR OF A HUMAN ENDOTHELIUM-DERIVED VASOCONSTRICTOR PEPTIDE, ENDOTHELIN - IDENTITY OF HUMAN AND PORCINE ENDOTHELIN. Febs Letters 231, 440-444.

Jandzik, D., Garnett, A.T., Square, T.A., Cattell, M.V., Yu, J.-K., Medeiros, D.M., 2015. Evolution of the new vertebrate head by co-option of an ancient chordate skeletal tissue. Nature 518, 534-537.

Jandzik, D., Hawkins, M.B., Cattell, M.V., Cerny, R., Square, T.A., Medeiros, D.M., 2014a. Roles for FGF in lamprey pharyngeal pouch formation and skeletogenesis highlight ancestral functions in the vertebrate head. Development 141, 629-638.

Jandzik, D., Hawkins, M.B., Cattell, M.V., Cerny, R., Square, T.A., Medeiros, D.M., 2014b. 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. Clarendon Press, New York.

Jeong, J., Li, X., McEvilly, R.J., Rosenfeld, M.G., Lufkin, T., Rubenstein, J.L.R., 2008. Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. Development 135, 2905-2916.

Johnson, G.D., Swenson, H.R., Ramage, R., Ahn, K., 2002. Mapping the active site of endothelin-converting enzyme-1 through subsite specificity and mutagenesis studies: A comparison with neprilysin. Archives of Biochemistry and Biophysics 398, 240-248.

Kague, E., Gallagher, M., Burke, S., Parsons, M., Franz-Odendaal, T., Fisher, S., 2012. Skeletogenic Fate of Zebrafish Cranial and Trunk Neural Crest. Plos One 7.

Karne, S., Jayawickreme, C.K., Lerner, M.R., 1993. CLONING AND CHARACTERIZATION OF AN ENDOTHELIN-3 SPECIFIC RECEPTOR (ET(C) RECEPTOR) FROM XENOPUS-LAEVIS DERMAL MELANOPHORES. Journal of Biological Chemistry 268, 19126-19133.

Kawasaki-Nishihara, A., Nishihara, D., Nakamura, H., Yamamoto, H., 2011. ET3/Ednrb2 Signaling Is Critically Involved in Regulating Melanophore Migration in Xenopus. Developmental Dynamics 240, 1454-1466.

Kee, Y., Bronner-Fraser, M., 2005. To proliferate or to die: role of Id3 in cell cycle progression and survival of neural crest progenitors. Genes & Development 19, 744-755.

Kempf, H., Linares, C., Corvol, P., Gasc, J.M., 1998. Pharmacological inactivation of the endothelin type A receptor in the early chick embryo: a model of mispatterning of the branchial arch derivatives. Development 125, 4931-4941.

Khimji, A.K., Rockey, D.C., 2010. Endothelin-Biology and disease. Cellular Signalling 22, 1615-1625.

Kido, T., Sawamura, T., Masaki, T., 1998. The processing pathway of endothelin-1 production. Journal of Cardiovascular Pharmacology 31, S13-S15.

Kimmel, C.B., Miller, C.T., Kruze, G., Ullmann, B., BreMiller, R.A., Larison, K.D., Snyder, H.C., 1998. The shaping of pharyngeal cartilages during early development of the zebrafish. Developmental Biology 203, 245-263.

Kos, R., Reedy, M.V., Johnson, R.L., Erickson, C.A., 2001. The winged-helix transcription factor FoxD3 is important for establishing the neural crest lineage and repressing melanogenesis in avian embryos. Development 128, 1467-1479.

Kralovic, M., Horacek, I., Cerny, R., 2010. Mouth development in the Senegal bichir Polypterus senegalus does not involve the oropharyngeal membrane: possible implications for the ecto-endoderm boundary and tooth initiation. Journal of Applied Ichthyology 26, 179-182.

Krauss, J., Frohnhoefer, H.G., Walderich, B., Maischein, H.-M., Weiler, C., Irion, U., Nuesslein-Volhard, C., 2014. Endothelin signalling in iridophore development and stripe pattern formation of zebrafish. Biology Open 3, 503-509.

Kudoh, T., Tsang, M., Hukriede, N.A., Chen, X.F., Dedekian, M., Clarke, C.J., Kiang, A., Schultz, S., Epstein, J.A., Toyama, R., Dawid, I.B., 2001. A gene expression screen in zebrafish embryogenesis. Genome Research 11, 1979-1987.

Kuraku, S., 2013. Impact of asymmetric gene repertoire between cyclostomes and gnathostomes. Seminars in Cell & Developmental Biology 24, 119-127.

Kuraku, S., Kuratani, S., 2006. Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. Zoological Science 23, 1053-1064.

Kuraku, S., Meyer, A., Kuratani, S., 2009. Timing of Genome Duplications Relative to the Origin of the Vertebrates: Did Cyclostomes Diverge before or after? Molecular Biology and Evolution 26, 47-59.

Kuraku, S., Takio, Y., Sugahara, F., Takechi, M., Kuratani, S., 2010. Evolution of oropharyngeal patterning mechanisms involving Dlx and endothelins in vertebrates. Developmental Biology 341, 315-323.

Kuratani, S., 2004. Evolution of the vertebrate jaw: comparative embryology and molecular developmental biology reveal the factors behind evolutionary novelty. Journal of Anatomy 205, 335-347.

Kuratani, S., Adachi, N., Wada, N., Oisi, Y., Sugahara, F., 2013. Developmental and evolutionary significance of the mandibular arch and prechordal/premandibular cranium in vertebrates: revising the heterotopy scenario of gnathostome jaw evolution. Journal of Anatomy 222, 41-55.

Kuratani, S., Horigome, N., Ueki, T., Aizawa, S., Hirano, S., 1998a. Stereotyped axonal bundle formation and neuromeric patterns in embryos of a cyclostome, Lampetra japonica. Journal of Comparative Neurology 391, 99-114.

Kuratani, S., Nobusada, Y., Horigome, N., Shigetani, Y., 2001. Embryology of the lamprey and evolution of the vertebrate jaw: insights from molecular and developmental perspectives. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 356, 1615-1632.

Kuratani, S., Oisi, Y., Ota, K.G., 2016. Evolution of the Vertebrate Cranium: Viewed from Hagfish Developmental Studies. Zoological Science 33, 229-238.

Kuratani, S., Ueki, T., Hirano, S., Aizawa, S., 1998b. Rostral truncation of a cyclostome, Lampetra japonica, induced by all-trans retinoic acid defines the head/trunk interface of the vertebrate body. Developmental Dynamics 211, 35-51.

Kusafuka, T., Puri, P., 1997. Mutations of the endothelin-B receptor and endothelin-3 genes in Hirschsprung's disease. Pediatric surgery international 12, 19-23.

Kusakabe, R., Kuraku, S., Kuratani, S., 2011. Expression and interaction of muscle-related genes in the lamprey imply the evolutionary scenario for vertebrate skeletal muscle, in association with the acquisition of the neck and fins. Developmental Biology 350, 217-227.

Lahav, R., Dupin, E., Lecoin, L., Glavieux, C., Champeval, D., Ziller, C., Le Douarin, N.M., 1998. Endothelin 3 selectively promotes survival and proliferation of neural crest-derived glial and melanocytic precursors in vitro. Proc Natl Acad Sci U S A 95, 14214-14219.

Lakiza, O., Miller, S., Bunce, A., Lee, E.M.J., McCauley, D.W., 2011. SoxE gene duplication and development of the lamprey branchial skeleton: Insights into development and evolution of the neural crest. Dev Biol (N Y 1985) 359, 149-161.

Le Douarin, N.M., Dupin, E., 2003. Multipotentiality of the neural crest. Current Opinion in Genetics & Development 13, 529-536.

Le Pabic, P., Ng, C., Schilling, T.F., 2014. Fat-Dachsous Signaling Coordinates Cartilage Differentiation and Polarity during Craniofacial Development. Plos Genetics 10.

Lecoin, L., Sakurai, T., Ngo, M.T., Abe, Y., Yanagisawa, M., Le Douarin, N.M., 1998. Cloning and characterization of a novel endothelin receptor subtype in the avian class. Proceedings of the National Academy of Sciences of the United States of America 95, 3024-3029.

Lee, H.Y., Kleber, M., Hari, L., Brault, V., Suter, U., Taketo, M.M., Kemler, R., Sommer, L., 2004. Instructive role of Wnt/beta-catenin in sensory fate specification in neural crest stem cells. Science 303, 1020-1023.

Leoyklang, P., Suphapeetiporn, K., Siriwan, P., Desudchit, T., Chaowanapanja, P., Gahl, W.A., Shotelersuk, V., 2007. Heterozygous nonsense mutation SATB2 associated with cleft palate, osteoporosis, and cognitive defects. Human Mutation 28, 732-738.

Lister, J.A., Cooper, C., Nguyen, K., Modrell, M., Grant, K., Raible, D.W., 2006. Zebrafish Foxd3 is required for development of a subset of neural crest derivatives. Developmental Biology 290, 92-104.

Liu, K.J., Harland, R.A., 2003. Cloning and characterization of Xenopus Id4 reveals differing roles for Id genes. Developmental Biology 264, 339-351.

Lo, T.W., Pickle, C.S., Lin, S., Ralston, E.J., Gurling, M., Schartner, C.M., Bian, Q., Doudna, J.A., Meyer, B.J., 2013. Precise and Heritable Genome Editing in Evolutionarily Diverse Nematodes Using TALENs and CRISPR/Cas9 to Engineer Insertions and Deletions. Genetics 195, 331-+.

Lyon, R.S., Davis, A., Scemama, J.L., 2013. Spatio-temporal expression patterns of anterior Hox genes during Nile tilapia (Oreochromis niloticus) embryonic development. Gene Expression Patterns 13, 104-108.

MacDonald, R.B., Debiais-Thibaud, M., Ekker, M., 2010. Regulation of Dlx gene expression in the zebrafish pharyngeal arches: from conserved enhancer sequences to conserved activity. Journal of Applied Ichthyology 26, 187-191.

Mali, P., Esvelt, K.M., Church, G.M., 2013. Cas9 as a versatile tool for engineering biology. Nat Methods 10, 957-963.

Mallarino, R., Campas, O., Fritz, J.A., Burns, K.J., Weeks, O.G., Brenner, M.P., Abzhanov, A., 2012. Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. Proceedings of the National Academy of Sciences of the United States of America 109, 16222-16227.

Mallatt, J., Chen, J.Y., 2003. Fossil sister group of craniates: Predicted and found. Journal of Morphology 258, 1-31.

Marchant, L., Linker, C., Ruiz, P., Guerrero, N., Mayor, R., 1998. The inductive properties of mesoderm suggest that the neural crest cells are specified by a BMP gradient. Developmental Biology 198, 319-329.

Martinez-Morales, J.R., Henrich, T., Ramialison, M., Wittbrodt, J., 2007. New genes in the evolution of the neural crest differentiation program. Genome Biology 8.

Mason, I., Chambers, D., Shamim, H., Walshe, J., Irving, C., 2000. Regulation and function of FGF8 in patterning of midbrain and anterior hindbrain. Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire 78, 577-584.

Matsuo, I., Kuratani, S., Kimura, C., Takeda, N., Aizawa, S., 1995. MOUSE OTX2 FUNCTIONS IN THE FORMATION AND PATTERNING OF ROSTRAL HEAD. Genes & Development 9, 2646-2658.

McCauley, D.W., Bronner-Fraser, M., 2002. Conservation of Pax gene expression in ectodermal placodes of the lamprey. Gene 287, 129-139.

McCauley, D.W., Bronner-Fraser, M., 2003. Neural crest contributions to the lamprey head. Development 130, 2317-2327.

McCauley, D.W., Bronner-Fraser, M., 2004. Conservation and divergence of BMP2/4 genes in the lamprey: expression and phylogenetic analysis suggest a single ancestral vertebrate gene. Evolution & Development 6, 411-422.

McCauley, D.W., Bronner-Fraser, M., 2006. Importance of SoxE in neural crest development and the evolution of the pharynx. Nature 441, 750-752.

McCauley, D.W., Kuratani, S., 2008. Cyclostome studies in the context of vertebrate evolution. Zoolog Sci 25, 953-954.

McCauley, D.W., Lee, E., Yuan, T., Nguyen, K., 2014. Differential Activity of SoxE Transcription Factors in Neural Crest Development and Evolution. Integrative and Comparative Biology 54, E136-E136.

McGonnell, I.M., Graham, A., Richardson, J., Fish, J.L., Depew, M.J., Dee, C.T., Holland, P.W.H., Takahashi, T., 2011. Evolution of the Alx homeobox gene family: parallel retention and independent loss of the vertebrate Alx3 gene. Evolution & Development 13, 343-351.

Medeiros, D.M., Crump, J.G., 2012. New perspectives on pharyngeal dorsoventral patterning in development and evolution of the vertebrate jaw. Developmental Biology 371, 121-135.

Mesbah, K., Harrelson, Z., Theveniau-Ruissy, M., Papaioannou, V.E., Kelly, R.G., 2008. Tbx3 is required for outflow tract development. Circulation Research 103, 743-750.

Meulemans, D., Bronner-Fraser, M., 2004. Gene-regulatory interactions in neural crest evolution and development. Developmental Cell 7, 291-299.

Meulemans, D., Bronner-Fraser, M., 2005. Central role of gene cooption in neural crest evolution. Journal of Experimental Zoology Part B-Molecular and Developmental Evolution 304B, 298-303.

Milet, C., Monsoro-Burq, A.H., 2012. Neural crest induction at the neural plate border in vertebrates. Developmental Biology 366, 22-33.

Miller, C.T., Schilling, T.F., Lee, K.H., Parker, J., Kimmel, C.B., 2000. sucker encodes a zebrafish Endothelin-1 required for ventral pharyngeal arch development. Development 127, 3815-3828.

Miller, C.T., Swartz, M.E., Khuu, P.A., Walker, M.B., Eberhart, J.K., Kimmel, C.B., 2007. mef2ca is required in cranial neural crest to effect Endothelin1 signaling in zebrafish. Developmental Biology 308, 144-157.

Miller, C.T., Yelon, D., Stainier, D.Y.R., Kimmel, C.B., 2003. Two endothelin 1 effectors, hand2 and bapx1, pattern ventral pharyngeal cartilage and the jaw joint. Development 130, 1353-1365.

Minoux, M., Antonarakis, G.S., Kmita, M., Duboule, D., Rijli, F.M., 2009. Rostral and caudal pharyngeal arches share a common neural crest ground pattern. Development 136, 637-645.

Minoux, M., Rijli, F.M., 2010. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. Development 137, 2605-2621.

Morita, T., Nitta, H., Kiyama, Y., Mori, H., Mishina, M., 1995. DIFFERENTIAL EXPRESSION OF 2 ZEBRAFISH EMX HOMEOPROTEIN MESSENGER-RNAS IN THE DEVELOPING BRAIN. Neuroscience Letters 198, 131-134.

Morris, S.C., Caron, J.-B., 2014. A primitive fish from the Cambrian of North America. Nature 512, 419-U413.

Morrison, S.L., Campbell, C.K., Wright, G.M., 2000. Chondrogenesis of the branchial skeleton in embryonic sea lamprey, Petromyzon marinus. Anatomical Record 260, 252-267.

Nair, S., Li, W., Cornell, R., Schilling, T.F., 2007. Requirements for Endothelin type-A receptors and Endothelin-1 signaling in the facial ectoderm for the patterning of skeletogenic neural crest cells in zebrafish. Development 134, 335-345.

Nakayama, T., Fish, M.B., Fisher, M., Oomen-Hajagos, J., Thomsen, G.H., Grainger, R.M., 2013. Simple and Efficient CRISPR/Cas9-Mediated Targeted Mutagenesis in Xenopus tropicalis. Genesis 51, 835-843.

Nandadasa, S., Tao, Q., Menon, N.R., Heasman, J., Wylie, C., 2009. N- and E-cadherins in Xenopus are specifically required in the neural and non-neural ectoderm, respectively, for F-actin assembly and morphogenetic movements. Development 136, 1327-1338.

Nataf, V., Grapin-Botton, A., Champeval, D., Amemiya, A., Yanagisawa, M., Le Douarin, N.M., 1998. The expression patterns of endothelin-A receptor and endothelin 1 in the avian embryo. Mechanisms of Development 75, 145-149.

Nataf, V., Lecoin, L., Eichmann, A., LeDouarin, N.M., 1996. Endothelin-B receptor is expressed by neural crest cells in the avian embryo. Proceedings of the National Academy of Sciences of the United States of America 93, 9645-9650.

Nelson, J.S., Grande, T.C., Wilson, M.V.H., 2016. Fishes of the World.

Newman, C.S., Krieg, P.A., 1999. The Xenopus bagpipe-related homeobox gene zampogna is expressed in the pharyngeal endoderm and the visceral musculature of the midgut. Development Genes and Evolution 209, 132-134.

Nichols, J.T., Pan, L., Moens, C.B., Kimmel, C.B., 2013. barx1 represses joints and promotes cartilage in the craniofacial skeleton. Development 140, 2765-2775.

Nieuwkoop, P.D., Faber, J., 1956. Normal table of Xenopus laevis (Daudin). A systematical and chro-nological survey of the development from the fertilized egg till the end of the metamorphosis. Normal table of Xenopus laevis (Daudin). A systematical and chro-nological survey of the development from the fertilized egg till the end of the metamorphosis., 243-pp. vi + 243.

Nikitina, N., Bronner-Fraser, M., Sauka-Spengler, T., 2009. Culturing lamprey embryos. Cold Spring Harb Protoc 2009, pdb prot5122.

Nikitina, N., Sauka-Spengler, T., Bronner-Fraser, M., 2008. Dissecting early regulatory relationships in the lamprey neural crest gene network. Proceedings of the National Academy of Sciences of the United States of America 105, 20083-20088.

Nikitina, N., Tong, L., Bronner, M.E., 2011. Ancestral Network Module Regulating prdm1 Expression in the Lamprey Neural Plate Border. Dev Dyn 240, 2265-2271.

Nishida, H., Stach, T., 2014. Cell Lineages and Fate Maps in Tunicates: Conservation and Modification. Zoological Science 31, 645-652.

Noden, D.M., 1978. CONTROL OF AVIAN CEPHALIC NEURAL CREST CYTODIFFERENTIATION .1. SKELETAL AND CONNECTIVE TISSUES. Developmental Biology 67, 296-312.

Northcutt, R.G., Gans, C., 1983. THE GENESIS OF NEURAL CREST AND EPIDERMAL PLACODES - A REINTERPRETATION OF VERTEBRATE ORIGINS. Quarterly Review of Biology 58, 1-28.

Oisi, Y., Ota, K.G., Fujimoto, S., Kuratani, S., 2013. Development of the Chondrocranium in Hagfishes, with Special Reference to the Early Evolution of Vertebrates. Zoological Science 30, 944-961.

Ono, H., Kozmik, Z., Yu, J.-K., Wada, H., 2014. A novel N-terminal motif is responsible for the evolution of neural crest-specific gene-regulatory activity in vertebrate FoxD3. Developmental Biology 385, 396-404.

Ota, K.G., Kuraku, S., Kuratani, S., 2007. Hagfish embryology with reference to the evolution of the neural crest. Nature 446, 672-675.

Pancer, Z., Amemiya, C.T., Ehrhardt, G.R.A., Ceitlin, J., Gartland, G.L., Cooper, M.D., 2004. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature 430, 174-180.

Parker, H.J., Bronner, M.E., Krumlauf, R., 2014. A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. Nature 514, 490-+.

Pasqualetti, M., Ori, M., Nardi, I., Rijli, F.M., 2000. Ectopic Hoxa2 induction after neural crest migration results in homeosis of jaw elements in Xenopus. Development 127, 5367-5378.

Patthey, C., Gunhaga, L., Edlund, T., 2008. Early Development of the Central and Peripheral Nervous Systems Is Coordinated by Wnt and BMP Signals. Plos One 3.

Piavis, G.W., 1971. The Biology of Lampreys. Academic Press, London.

Piekarski, N., Gross, J.B., Hanken, J., 2014. Evolutionary innovation and conservation in the embryonic derivation of the vertebrate skull. Nature Communications 5.

Pla, P., Larue, L., 2003. Involvement of endothelin receptors in normal and pathological development of neural crest cells. International Journal of Developmental Biology 47, 315-325.

Pugener, L.A., Maglia, A.M., Trueb, L., 2003. Revisiting the contribution of larval characters to an analysis of phylogenetic relationships of basal anurans. Zoological Journal of the Linnean Society 139, 129-155.

Rawlings, N.D., Barrett, A.J., 1995. EVOLUTIONARY FAMILIES OF METALLOPEPTIDASES. Proteolytic Enzymes: Aspartic and Metallo Peptidases 248, 183-228.

Reifers, F., Bohli, H., Walsh, E.C., Crossley, P.H., Stainier, D.Y.R., Brand, M., 1998. Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis. Development 125, 2381-2395.

Renz, A.J., Gunter, H.M., Fischer, J.M.F., Qiu, H., Meyer, A., Kuraku, S., 2011. Ancestral and derived attributes of the dlx gene repertoire, cluster structure and expression patterns in an African cichlid fish. Evodevo 2.

Ribeiro, I., Kawakami, Y., Buscher, D., Raya, A., Rodriguez-Leon, J., Morita, M., Esteban, C.R., Belmonte, J.C.I., 2007. Tbx2 and Tbx3 Regulate the Dynamics of Cell Proliferation during Heart Remodeling. Plos One 2.

RM, H., L, H., LL, B., SK, L., EM, S.-S., A, I., J, C., WJ, P., 2006. Genetic evidence does not support direct regulation of \ EDNRB\ by \ SOX10\ in migratory neural crest and the melanocyte lineage. Mechanisms of Development 123, 124 - 134.

Rogozin, I.B., Wolf, Y.I., Sorokin, A.V., Mirkin, B.G., Koonin, E.V., 2003. Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. Curr. Biol. 13, 1512-1517.

Rose, C.S., 2014. The importance of cartilage to amphibian development and evolution. International Journal of Developmental Biology 58, 917-927.

Ruest, L.B., Hammer, R.E., Yanagisawa, M., Clouthier, D.E., 2003. Dlx5/6-enhancer directed expression of Cre recombinase in the pharyngeal arches and brain. Genesis 37, 188-194.

Sakai, T., Nirasawa, Y., Itoh, Y., Wakizaka, A., 2000. Japanese patients with sporadic Hirschsprung: mutation analysis of the receptor tyrosine kinase proto-oncogene, endothelin-B receptor, endothelin-3, glial cell line-derived neurotrophic factor and neurturin genes: a comparison with similar studies. European Journal of Pediatrics 159, 160-167.

Sakamoto, A., Yanagisawa, M., Sawamura, T., Enoki, T., Ohtani, T., Sakurai, T., Nakao, K., Toyooka, T., Masaki, T., 1993. DISTINCT SUBDOMAINS OF HUMAN ENDOTHELIN RECEPTORS DETERMINE THEIR SELECTIVITY TO ENDOTHELIN(A)-SELECTIVE ANTAGONIST AND ENDOTHELIN(B)-SELECTIVE AGONISTS. Journal of Biological Chemistry 268, 8547-8553.

Sanchez-Mejias, A., Fernandez, R.M., Lopez-Alonso, M., Antinolo, G., Borrego, S., 2010. New roles of EDNRB and EDN3 in the pathogenesis of Hirschsprung disease. Genetics in Medicine 12, 39-43.

Sansom, R.S., Freedman, K., Gabbott, S.E., Aldridge, R.J., Purnell, M.A., 2010. TAPHONOMY AND AFFINITY OF AN ENIGMATIC SILURIAN VERTEBRATE, JAMOYTIUS KERWOODI WHITE. Palaeontology 53, 1393-1409.

Santagati, F., Minoux, M., Ren, S.Y., Rijli, F.M., 2005. Temporal requirement of Hoxa2 in cranial neural crest skeletal morphogenesis. Development 132, 4927-4936.

Santagati, F., Rijli, F.M., 2003. Cranial neural crest and the building of the vertebrate head. Nature Reviews Neuroscience 4, 806-818.

Sato, T., Kawamura, Y., Asai, R., Amano, T., Uchijima, Y., Dettlaff-Swiercz, D.A., Offermanns, S., Kurihara, Y., Kurihara, H., 2008a. Recombinase-mediated cassette exchange reveals the selective use of G(q)/G(11)-dependent and -independent endothelin 1/endothelin type A receptor signaling in pharyngeal arch development. Development 135, 755-765.

Sato, T., Kurihara, Y., Asai, R., Kawamura, Y., Tonami, K., Uchijima, Y., Heude, E., Ekker, M., Levi, G., Kurihara, H., 2008b. An endothelin-1 switch specifies maxillomandibular identity. Proceedings of the National Academy of Sciences of the United States of America 105, 18806-18811.

Sauka-Spengler, T., Meulemans, D., Jones, M., Bronner-Fraser, M., 2007. Ancient evolutionary origin of the neural crest gene regulatory network. Developmental Cell 13, 405-420.

Schneider, R.A., Helms, J.A., 2003. The cellular and molecular origins of beak morphology. Science 299, 565-568.

Schultemerker, S., Hammerschmidt, M., Beuchle, D., Cho, K.W., Derobertis, E.M., Nussleinvolhard, C., 1994. EXPRESSION OF ZEBRAFISH GOOSECOID AND NO TAIL GENE-PRODUCTS IN WILD-TYPE AND MUTANT NO TAIL EMBRYOS. Development 120, 843-852.

Schwend, T., Ahlgren, S.C., 2009. Zebrafish con/disp1 reveals multiple spatiotemporal requirements for Hedgehog-signaling in craniofacial development. Bmc Developmental Biology 9.

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

Settle, S.H., Rountree, R.B., Sinha, A., Thacker, A., Higgins, K., Kingsley, D.M., 2003. Multiple joint and skeletal patterning defects caused by single and double mutations in the mouse Gdf6 and Gdf5 genes. Developmental Biology 254, 116-130.

Sheehan-Rooney, K., Palinkasova, B., Eberhart, J.K., Dixon, M.J., 2010. A Cross-Species Analysis of Satb2 Expression Suggests Deep Conservation Across Vertebrate Lineages. Developmental Dynamics 239, 3481-3491.

Shigetani, Y., Sugahara, F., Kawakami, Y., Murakami, Y., Hirano, S., Kuratani, S., 2002. Heterotopic shift of epithelial-mesenchymal interactions in vertebrate jaw evolution. Science 296, 1316-1319.

Shin, M.K., Levorse, J.M., Ingram, R.S., Tilghman, S.M., 1999. The temporal requirement for endothelin receptor-B signalling during neural crest development. Nature 402, 496-501.

Shu, D.G., Luo, H.L., Morris, S.C., Zhang, X.L., Hu, S.X., Chen, L., Han, J., Zhu, M., Li, Y., Chen, L.Z., 1999. Lower Cambrian vertebrates from South China. Nature 402, 42-46.

Shu, D.G., Morris, S.C., Han, J., Zhang, Z.F., Yasui, K., Janvier, P., Chen, L., Zhang, X.L., Liu, J.N., Li, Y., Liu, H.Q., 2003. Head and backbone of the Early Cambrian vertebrate Haikouichthys. Nature 421, 526-529.

Shubin, N., Tabin, C., Carroll, S., 2009. Deep homology and the origins of evolutionary novelty. Nature 457, 818-823.

Simoes-Costa, M., Bronner, M.E., 2015. Establishing neural crest identity: a gene regulatory recipe. Development 142, 242-257.

Smith, J., Morgan, J.R., Zottoli, S.J., Smith, P.J., Buxbaum, J.D., Bloom, O.E., 2011. Regeneration in the Era of Functional Genomics and Gene Network Analysis. Biol. Bull. 221, 18-34.

Smith, J.J., Keinath, M.C., 2015. The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. Genome Research 25, 1081-1090.

Smith, J.J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M.S., Yandell, M.D., Manousaki, T., Meyer, A., Bloom, O.E., Morgan, J.R., Buxbaum, J.D., Sachidanandam, R., Sims, C., Garruss, A.S., Cook, M., Krumlauf, R., Wiedemann, L.M., Sower, S.A., Decatur, W.A., Hall, J.A., Amemiya, C.T., Saha, N.R., Buckley, K.M., Rast, J.P., Das, S., Hirano, M., McCurley, N., Guo, P., Rohner, N., Tabin, C.J., Piccinelli, P., Elgar, G., Ruffier, M., Aken, B.L., Searle, S.M.J., Muffato, M., Pignatelli, M., Herrero, J., Jones, M., Brown, C.T., Chung-Davidson, Y.W., Nanlohy, K.G., Libants, S.V., Yeh, C.Y., McCauley, D.W., Langeland, J.A., Pancer, Z., Fritzsch, B., de Jong, P.J., Zhu, B.L., Fulton, L.L., Theising, B., Flicek, P., Bronner, M.E., Warren, W.C., Clifton, S.W., Wilson, R.K., Li, W.M., 2013. Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. Nature Genet. 45, 415-421.

Sower, S.A., Freamat, M., Kavanaugh, S.I., 2009. The origins of the vertebrate hypothalamic-pituitary-gonadal (HPG) and hypothalamic-pituitary-thyroid (HPT) endocrine systems: New insights from lampreys. Gen. Comp. Endocrinol. 161, 20-29.

Sperber, S.M., Dawid, I.B., 2008. barx1 is necessary for ectomesenchyme proliferation and osteochondroprogenitor condensation in the zebrafish pharyngeal arches. Developmental Biology 321, 101-110.

Square, T., Jandzik, D., Cattell, M., Coe, A., Doherty, J., Medeiros, D.M., 2015a. A gene expression map of the larval Xenopus laevis head reveals developmental changes underlying the evolution of new skeletal elements. Developmental Biology 397, 293-304.

Square, T., Jandzik, D., Cattell, M., Hansen, A., Medeiros, D.M., 2016a. Embryonic expression of endothelins and their receptors in lamprey and frog reveals stem vertebrate origins of complex Endothelin signaling. Scientific Reports 6.

Square, T., Jandzik, D., Romasek, M., Cerny, R., Medeiros, D.M., 2016b. The origin and diversification of the developmental mechanisms thatpattern the vertebrate head skeleton. Dev Biol.

Square, T., Romasek, M., Jandzik, D., Cattell, M.V., Klymkowsky, M., Medeiros, D.M., 2015b. CRISPR/Cas9-mediated mutagenesis in the sea lamprey Petromyzon marinus: a powerful tool for understanding ancestral gene functions in vertebrates. Development 142, 4180-4187.

Stanchina, L., Baral, V., Robert, F., Pingault, V., Lemort, N., Pachnis, V., Goossens, M., Bondurand, N., 2006. Interactions between Sox10, Edn3 and Ednrb during enteric nervous system and melanocyte development. Developmental Biology 295, 232-249.

Stock, D.W., 2005. The Dlx gene complement of the leopard shark, Triakis semifasciata, resembles that of mammals: Implications for genomic and morphological evolution of jawed vertebrates. Genetics 169, 807-817.

Stock, D.W., Ellies, D.L., Zhao, Z.Y., Ekker, M., Ruddle, F.H., Weiss, K.M., 1996. The evolution of the vertebrate Dlx gene family. Proceedings of the National Academy of Sciences of the United States of America 93, 10858-10863.

Stock, D.W., Whitt, G.S., 1992. EVIDENCE FROM 18S RIBOSOMAL-RNA SEQUENCES THAT LAMPREYS AND HAGFISHES FORM A NATURAL GROUP. Science 257, 787-789.

Stolfi, A., Ryan, K., Meinertzhagen, I.A., Christiaen, L., 2015. Migratory neuronal progenitors arise from the neural plate borders in tunicates. Nature 527, 371-+.

Storm, E.E., Kingsley, D.M., 1999. GDF5 coordinates bone and joint formation during digit development. Developmental Biology 209, 11-27.

Sugahara, F., Aota, S., Kuraku, S., Murakami, Y., Takio-Ogawa, Y., Hirano, S., Kuratani, S., 2011. Involvement of Hedgehog and FGF signalling in the lamprey telencephalon: evolution of regionalization and dorsoventral patterning of the vertebrate forebrain. Development 138, 1217-1226.

Sumiyama, K., Irvine, S.Q., Stock, D.W., Weiss, K.M., Kawasaki, K., Shimizu, N., Shashikant, C.S., Miller, W., Ruddle, F.H., 2002. Genomic structure and functional control of the Dlx3-7 bigene cluster. Proceedings of the National Academy of Sciences of the United States of America 99, 780-785.

Sun, X., Meyers, E.N., Lewandoski, M., Martin, G.R., 1999. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev. 13, 1834-1846.

Suzuki, T., Kusakabe, M., Nakayama, K., Nishida, E., 2012. The protein kinase MLTK regulates chondrogenesis by inducing the transcription factor Sox6. Development 139, 2988-2998.

Svensson, M.E., Haas, A., 2005. Evolutionary innovation in the vertebrate jaw: a derived morphology in anuran tadpoles and its possible developmental origin. Bioessays 27, 526-532.

Swartz, M.E., Sheehan-Rooney, K., Dixon, M.J., Eberhart, J.K., 2011. Examination of a Palatogenic Gene Program in Zebrafish. Developmental Dynamics 240, 2204-2220.

Tahara, Y., 1988. NORMAL STAGES OF DEVELOPMENT IN THE LAMPREY, LAMPETRA-REISSNERI (DYBOWSKI). Zoological Science 5, 109-118.

Takechi, M., Adachi, N., Hirai, T., Kuratani, S., Kuraku, S., 2013. The Dlx genes as clues to vertebrate genomics and craniofacial evolution. Seminars in Cell & Developmental Biology 24, 110-118.

Takio, Y., Kuraku, S., Murakami, Y., Pasqualetti, M., Rijli, F.M., Narita, Y., Kuratani, S., Kusakabe, R., 2007. Hox gene expression patterns in Lethenteron japonicum embryos - Insights into the evolution of the vertebrate Hox code. Developmental Biology 308, 606-620.

Talbot, J.C., Johnson, S.L., Kimmel, C.B., 2010. hand2 and Dlx genes specify dorsal, intermediate and ventral domains within zebrafish pharyngeal arches. Development 137, 2506-2516.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Tarazona, O.A., Slota, L.A., Lopez, D.H., Zhang, G.J., Cohn, M.J., 2016. The genetic program for cartilage development has deep homology within Bilateria. Nature 533, 86-+.

Tavares, A.L.P., Garcia, E.L., Kuhn, K., Woods, C.M., Williams, T., Clouthier, D.E., 2012. Ectodermal-derived Endothelin1 is required for patterning the distal and intermediate domains of the mouse mandibular arch. Developmental Biology 371, 47-56.

ten Berge, D., Brouwer, A., Korving, J., Martin, J.F., Meijlink, F., 1998. Prx1 and Prx2 in skeletogenesis: roles in the craniofacial region, inner ear and limbs. Development 125, 3831-3842.

Theveneau, E., Mayor, R., 2012. Neural crest delamination and migration: From epithelium-to-mesenchyme transition to collective cell migration. Developmental Biology 366, 34-54.

Thisse, B., Heyer, V., Lux, A., Alunni, V., Degrave, A., Seiliez, I., Kirchner, J., Parkhill, J.P., Thisse, C., 2004. Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. Zebrafish:2nd Edition Genetics Genomics and Informatics 77, 505-519.

Thomas, T., Kurihara, H., Yamagishi, H., Kurihara, Y., Yazaki, Y., Olson, E.N., Srivastava, D., 1998. A signaling cascade involving endothelin-1, dHAND and Msx1 regulates development of neural-crest-derived branchial arch mesenchyme. Development 125, 3005-3014.

Tomobe, Y., Miyauchi, T., Saito, A., Yanagisawa, M., Kimura, S., Goto, K., Masaki, T., 1988. EFFECTS OF ENDOTHELIN ON THE RENAL-ARTERY FROM SPONTANEOUSLY HYPERTENSIVE AND WISTAR KYOTO RATS. European Journal of Pharmacology 152, 373-374.

Tomsa, J.M., Langeland, J.A., 1999. Otx expression during lamprey embryogenesis provides insights into the evolution of the vertebrate head and jaw. Developmental Biology 207, 26-37.

Trueb, L., Hanken, J., 1992. SKELETAL DEVELOPMENT IN XENOPUS-LAEVIS (ANURA, PIPIDAE). Journal of Morphology 214, 1-41.

Tucker, A.S., Watson, R.P., Lettice, L.A., Yamada, G., Hill, R.E., 2004. Bapx1 regulates patterning in the middle ear: altered regulatory role in the transition from the proximal jaw during vertebrate evolution. Development 131, 1235-1245.

Valdenaire, O., Schweizer, A., 2000. Endothelin-converting enzyme-like I (ECELI; 'XCE'): a putative metallopeptidstse crucially involved in the nervous control of respiration. Biochemical Society Transactions 28, 426-430.

Van Otterloo, E., Cornell, R.A., Medeiros, D.M., Garnett, A.T., 2013. Gene regulatory evolution and the origin of macroevolutionary novelties: Insights from the neural crest. Genesis 51, 457-470.

Van Otterloo, E., Li, W., Garnett, A., Cattell, M., Medeiros, D.M., Cornell, R.A., 2012. Novel Tfap2-mediated control of soxE expression facilitated the evolutionary emergence of the neural crest. Development 139, 720-730.

Verzi, M.P., Agarwal, P., Brown, C., McCulley, D.J., Schwarz, J.J., Black, B.L., 2007. The transcription factor MEF2C is required for craniofacial development. Developmental Cell 12, 645-652.

Wada, H., Makabe, K., 2006. Genome duplications of early vertebrates as a possible chronicle of the evolutionary history of the neural crest. International Journal of Biological Sciences 2, 133-141.

Wang, F.Q., Shi, Z.Y., Cui, Y., Guo, X.G., Shi, Y.B., Chen, Y.L., 2015. Targeted gene disruption in Xenopus laevis using CRISPR/Cas9. Cell Biosci 5.

Wang, H.Y., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 153, 910-918.

Wilson, J., Tucker, A.S., 2004. Fgf and Bmp signals repress the expression of Bapx1 in the mandibular mesenchyme and control the position of the developing jaw joint. Developmental Biology 266, 138-150.

Witschi, E., 1949. THE LARVAL EAR OF THE FROG AND ITS TRANSFORMATION DURING METAMORPHOSIS. Zeitschrift Fur Naturforschung Section B-a Journal of Chemical Sciences 4, 230-242.

Yanagisawa, H., Clouthier, D.E., Richardson, J.A., Charite, J., Olson, E.N., 2003. Targeted deletion of a branchial arch-specific enhancer reveals a role of dHAND in craniofacial development. Development 130, 1069-1078.

Yanagisawa, H., Hammer, R.E., Richardson, J.A., Emoto, N., Williams, S.C., Takeda, S., Clouthier, D.E., Yanagisawa, M., 2000. Disruption of ECE-1 and ECE-2 reveals a role for endothelin-converting enzyme-2 in murine cardiac development. Journal of Clinical Investigation 105, 1373-1382.

Yanagisawa, H., Hammer, R.E., Richardson, J.A., Williams, S.C., Clouthier, D.E., Yanagisawa, M., 1998. Role of endothelin-1/endothelin-A receptor-mediated signaling pathway in the aortic arch patterning in mice. Journal of Clinical Investigation 102, 22-33.

Yanagisawa, M., 1994. THE ENDOTHELIN SYSTEM - A NEW TARGET FOR THERAPEUTIC INTERVENTION. Circulation 89, 1320-1322.

Yanagisawa, M., Kurihara, H., Kimura, S., Goto, K., Masaki, T., 1988. ENDOTHELIN - A NOVEL POTENT VASOCONSTRICTOR PEPTIDE PRODUCED BY VASCULAR ENDOTHELIAL-CELLS. Faseb Journal 2, A393-A393.

Yu, J.-K., Wang, M.-C., Shin, T.I., Kohara, Y., Holland, L.Z., Satoh, N., Satou, Y., 2008a. A cDNA resource for the cephalochordate amphioxus Branchiostoma floridae. Development Genes and Evolution 218, 723-727.

Yu, J.K., Meulemans, D., McKeown, S.J., Bronner-Fraser, M., 2008b. Insights from the amphioxus genome on the origin of vertebrate neural crest. Genome Research 18, 1127-1132.

Yu, J.K.S., Holland, L.Z., Cold Spring Harbor Laboratory, P., 2010. Cephalochordates (Amphioxus or Lancelets): A Model for Understanding the Evolution of Chordate Characters. Emerging Model Organisms: Laboratory Manual, Vol 2, 335-363.

Zerucha, T., Stuhmer, T., Hatch, G., Park, B.K., Long, Q.M., Yu, G.Y., Gambarotta, A., Schultz, J.R., Rubenstein, J.L.R., Ekker, M., 2000. A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. Journal of Neuroscience 20, 709-721.

Zhang, G.J., Miyamoto, M.M., Cohn, M.J., 2006. Lamprey type II collagen and Sox9 reveal an ancient origin of the vertebrate collagenous skeleton. Proceedings of the National Academy of Sciences of the United States of America 103, 3180-3185.

Zhang, P., Liang, D., Mao, R.L., Hillis, D.M., Wake, D.B., Cannatella, D.C., 2013. Efficient Sequencing of Anuran mtDNAs and a Mitogenomic Exploration of the Phylogeny and Evolution of Frogs. Molecular Biology and Evolution 30, 1899-1915.

Zhang, S.W., Li, J.J., Lea, R., Vleminckx, K., Amaya, E., 2014. Fezf2 promotes neuronal differentiation through localised activation of Wnt/beta-catenin signalling during forebrain development. Development 141, 4794-4805.

Zhao, Q., Eberspaecher, H., Lefebvre, V., deCrombrugghe, B., 1997. Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis. Developmental Dynamics 209, 377-386.

Zhao, Z.Y., Stock, D.W., Buchanan, A.V., Weiss, K.M., 2000. Expression of Dlx genes during the development of the murine dentition. Development Genes and Evolution 210, 270-275.

Zirzow, S., Ludtke, T.H.W., Brons, J.F., Petry, M., Christoffels, V.M., Kispert, A., 2009. Expression and requirement of T-box transcription factors Tbx2 and Tbx3 during secondary palate development in the mouse. Developmental Biology 336, 145-155.

Appendix

Contents

Methods for culturing Florida Amphioxus (*Branchiostoma floridae*)

In order for us to have a tractable, relevant outgroup for developmental assays of vertebrate evolution, being able to culture cephalochordates in the lab becomes pertinent. Despite their modernly accepted position as the sister group to Olfactores (vertebrates + urochordates), the body plan of an adult amphioxus is widely held as less derived than that of urochordates Thus *B. floridae* represents a meaningful system to use as an outgroup despite their not being the closest sister group to vertebrates. To these ends, I visited the lab of Dr. Jr Kai Yu in order to experience firsthand how they have been so successful in culturing the Florida amphioxus (*Branchiostoma floridae*). This lab has successfully raised an F3 generation of this species, which to our knowledge was the only case of this in the world at the time of my visit. Success raising amphioxus is contingent on four main practices (arranged into sections within this chapter):

1) Culturing algae
    1.1) Making Seawater
    1.2) Algae inoculations
2) Animal husbandry
    2.1) Larval husbandry
    2.2) Adult husbandry

Previous to this visit, our lab had little or modest success with all of the above items, but our inability to feed the adults properly meant we were unable to ripen and spawn our own *B. floridae*, a critical step that is crucial to the efficacy of this developmental system. During my visit I was able to participate in all of the above items daily, except for item 3) which only occurred one time upon my arrival. During my time there I took extensive notes on how these methods are executed in order to generate a protocol to be disseminated to anyone interested in working with *B. floridae*.

1) Culturing algae

The following steps are applicable to most algal genera, however *Rhodomonas* specifically tends to be quite variable in different environments, and requires some special practices (mentioned when appropriate below). For all other cultures, all of the following steps can be carried out at room temperature (RT), at about 3000 to 3500 lux. When choosing light bulbs, It is best to use between 25%

and 50% of bubs designed specifically for growing plants (which have high emission in the 500-600nm range), while the rest should be standard "soft-white" fluorescent bulbs. Avoid bulbs with an orange or red hue, as these emit light that is mostly unusable to plants.

Rhodomonas requires special conditions. In our hands, cultures of this genera do best at 19 degrees C (the cultures do not crash as quickly), though we do grow 10L cultures at RT, these typically are only useable (alive) at or near saturation for around four days. We also shade our cultures down to between $^1/_3$ and ½ of the amount of light input as the other genera receive, as they seem to do better in lower light (for small cultures, we use a paper 'veil,' for large cultures we block ½ of the light bulbs with a black piece of foam). This genera also does best with much higher aeration (for small cultures, we use a stir rig on a medium to high setting to create a small vortex in the solution), and its highest density is still far below the other strains outlined here. Although this strain tends to be more difficult to culture, it has a high nutritional value, and it is therefore a very healthy food source (especially for larvae). At the time of this visit, Linda Holland's lab has had success growing their larvae without *Rhodomonas*, using *Isochrysis* as a substitute, while Dr. Jr Kai Yu's lab never feeds larvae without adding *Rhodomonas*. We feed this stain to both larvae and adults, and we are currently collecting data on growth rates (of both larvae and adult gonads) to determine the level of its value.

1.1) Seawater

All seawater should have salt concentration between 29 and 32 parts per thousand (ppt). If water from the ocean is available, be sure to filter it with at least a .22 micron filter and adjust salinity accordingly. If using artificial sea water (ASW), either use a .22 micron filter, or chemically sterilize the container which you will use for the algae, and keep the salt and DI H2O free of contaminants before/while adding. Chemical sterilization with bleach is an option, however the sodium thiosulfate used to inactivate the bleach can alter algal growth (and potentially the ASW buffering mechanism), so we do not use this procedure. Autoclaving seawater can also destroy its buffering system. It is best to allow large volumes of ASW to sit overnight on medium aeration before using them – this will completely dissolve the salt (especially important for amphioxus). We use Crystal Sea Marinemix (Aquaneering, San Diego, CA) as our sea salt, a very high-quality variety of aquarium salt. This tends to leave behind a precipitate in larger volume carboys, which is easily dissolved away with a 15 mM solution of HCl. Certain brands of salt are best avoided for algae and amphioxus: namely Instant Ocean (InstantOcean, Blacksburg, VA) which contains some sort of chelator (likely EDTA) to assist in its dissolution.

1.2) Algal medium and growth

We use F/2 medium for all algal strains as defined by Bigelow Laboratory for Ocean Sciences. As recommended, silica is omitted whenever we are not growing diatoms. To create F/2 medium plates, mix nine grams of Bacto-Agar with 1L ASW, and heat until mostly dissolved (to avoid autoclaving, we use a microwave). Once cool, add F/2 ingredients, mix, pour into plates (preferably at least 1 cm thick), and let settle.

Here I outline how to grow up to a 20L culture of algae. This is enough to feed ~200 animals for 10 days. It is advised that this protocol be started from the beginning every 6 months to ensure healthy algae strains; the .5L or 2L stages can be used as a 'reservoir' to inoculate larger amounts more frequently.

To create a plate of algae, add 1 mL of liquid F/2 medium, spread, and let soak for 3-5 minutes. With your alga of choice, pick a colony from a plate, break it up with a spreader, or add one mL of liquid algae culture, and spread. Place face-up at around 3000 lux until the pool of liquid is soaked in/evaporated (usually two days), and flip upside down after this occurs. Algal growth should be apparent within 4 to 5 days after inoculation.

Using a sterilized loop, pick an algae colony and whisk into 50 mL of sterilized F/2 medium. Let sit in beaker (no aeration), covered with foil (or the like), swirling approx. daily until absorption reaches the minimum saturation for your given genera specified in table 1 below (these respective absorptions will be referred to as 'high density' for the remainder of this protocol). Once one is familiar with the color/density of these genera, it is possible to 'eyeball' their density before moving on to the next step (cultures top out and will remain stable for at least a few days).

Once the 50 mL culture reaches high density, use all of it to inoculate a .5L culture in a flask (a 1:10 dilution) covered with foil (or the like). Aerate lightly or use a stir bar on a medium to low setting to increase the evaporation/dissolution of gasses into the culture, but do not use so much aeration that the solution evaporates quickly. If the solution evaporates noticeably, add sterilized DI water to bring it back up to its original volume.

Once the .5L culture reaches high density, use 200 mL to inoculate a 2L culture (again, a 1:10 dilution). Aerate such that bubbles are constantly flowing, and the algae is circulated lightly (analogous to a simmer/light boil). When this 2L culture reaches high density, you can use it to inoculate a large carboy (15-20L), again using a 1/10 volume to inoculate. Aerate such that the solution is well circulated (analogous to a boil). Use as needed to feed animals (below). If a larger volume culture is desired, simply use a 1:10 inoculation volume, and aerate such that the solution is well-circulated.

Note that after the algal culture reaches its maximum density, it will die within the next 3 to 7 days (this depends on strain, temperature, light, etc.). Adult amphioxus can eat dead algae, but we don't

recommend feeding them dead algae regularly (or at all for larvae) as the dead algae will pollute the water and can lead to sudden and widespread death of the animals.

Table 1. Absorption at 740 nm of various algal genera at high density.

Genera	Optical density at 750 nm for a "high density" culture
Isochrysis	.115
Tetraselmis	.125
Rhodomonas	.045

Note: These numbers are referr to as "high density" throughout this protocol. These measurments are taken in a spectrophotometer set to 750 nm, which measures diffracted light. These are not the absorption measurments at highest saturation – only those necessary to continue to the next step.

2) Amphioxus husbandry

Amphioxus are filter feeders. With this in mind, the amount of algae added per animal, as well as the density of the algae, are both important factors. Changing the amount of water the animals are housed in can drastically increase or decrease how well they are able to feed on a given amount of algae.

2.1) Larval Husbandry

Culturing larvae at RT is sufficient, but faster development (up to about 50% faster) has been observed at 30° C (Yu et al., 2010). The seawater used to keep amphioxus is the same as outlined in section 1.1 above, but with the addition of antibiotics. Adding antibiotics to seawater is recommended for amphioxus larvae to prevent bacterial and fungal growth via dead or dying algae. This prevents a biofilm from forming at the surface of the water (which greatly reduces the dissolution of oxygen into the water), and reduces the amount of effort needed to keep animals healthy. If desired, add .1mg/mL streptomycin, and .5 mg/mL penicillin G. If no antibiotics are to be used, daily water changes are suggested to avoid suffocation and widespread death.

Since larvae swim towards light until they are a few weeks old, it is best to keep them in the dark, or with lighting beneath them to avoid their getting caught in the air/water interface and drying out (this is especially important for animals less than 20 days old). Animals can be kept in small bowls or petri

dishes, initially at a maximum of 5 animals per mL of sea water. We typically keep 500 – 1000 young larvae in around .5L of ASW (1-2 animals per mL), though this volume changes between water changes (i.e. they start in 500 mL, but the volume increases with feeding every day until a water change, when they return to .5L again – see below for details on water changes). At 2-3 weeks of age, the volume of water per animal should be increased by 100-200%. At this point, we typically keep 500-1000 animals in 1 to 2 liters of ASW.

It is best to change the water and dish your larvae are kept in once every 2 days until one week of age, and once every 3-4 days until they are 40 days old (past 40 days we will consider them juveniles, although we have observed metamorphosis being delayed until as late as 70 days). The dead, uneaten algae will pollute the water rapidly if left alone. For small amounts of larvae, swirling them in a petri dish can concentrate them in the middle for removal with a pipette/eyedropper, but for larger amounts, a filter made from a 50 mL tube and 20 micron mesh is suggested to concentrate them (later, you can use larger sizes of mesh when the animals are larger – by 2 weeks of age, you can switch to 70 micron mesh). When changing the water and dish, removal of deformed animals, unfertilized eggs, dead algae, and feces as much as possible is recommended. Swirling the dish also aids in this. To remove deformed animals, swirl everything to the center of the dish, apply light to the edges of the dish (which healthy animals will swim towards), and wait 5-10 minutes for the healthy animals to swim to the periphery. Remove the middle contents, (which will still have some healthy animals), place these into a new dish, and repeat as necessary to select for the deformed and/or dead animals for their disposal. Transferring some waste when doing a water change is not detrimental to their health, but removing as much as possible favors healthy development.

Once the majority of the larvae reach the seven gill slit stage (2-3 weeks of age), they can be put on sterilized sand. We have seen success using both commercially available silica sand, and fine grain beach sand collected from the ocean. Whether or not sand significantly aids in amphioxus feeding is not yet known (successful spawnings have occurred without the use of sand).

2.1.1) Feeding Larvae

Amphioxus larvae need to eat live algae starting from the time their digestive systems begin working (about 36h at RT, about 28h at 30 degrees C). If their feeding is disrupted long enough (as little as one day for very young larvae), their guts shut down, and the animals will be destined for death within a few weeks. Therefore, it is safest to constantly give them small amounts of food from the time their mouths form until metamorphosis.

The best feeding strategy is to check your larvae periodically to see how much food is in their guts, and how much uneaten food is in their water, and modulate their feeding based on these criteria.

With this in mind, the following feeding schedule is approximate; if followed stringently, animals will likely be fed successfully, but monitoring every colony is the only way to ensure adequate but not excessive feeding. After feeding, it is useful to observe the contents of their stomachs under a microscope (a light source beneath the animals greatly aids in this). Their ilium (where the spinning contents of their hindgut are housed) should always have food in it, and it is best if there is also some food in their gut directly anterior to this (not spinning, and larger in diameter). If they are given more food than needed, their water can become easily polluted. If there is ever uneaten food 24 hours after a feeding, delay the next feeding and/or delay the next increase in feeding (see below). It may also be appropriate to do an immediate water change to remove the uneaten algae if there are very high amounts (i.e. the water is considerably brown). After one week of age, they are more resilient, and can go without eating for up to two days if necessary (although this is not encouraged). Be certain to look for the contents inside their guts, not the outer layer of pigment that builds up over time.

From the time their mouths open until one week of age, feed one mL of 20 micron filtered 1:1 *Rhodomonas*/*Isochrysis* (by cell count) per 200 animals (which we house in around 60 mL of ASW) daily. If your cultures behave as ours did, this should be about 4:1 *Rhodomonas*/*Isochrysis* by volume (again, *Rhodomonas* is the most dilute algae at saturation). This ratio is optimal for larvae, though feeding only *Rhodomonas* or *Isochrysis* seems sufficient as well. Starting at one week, increase the amount of feed by around 50% (rounding is fine) every two days until they are 40 days old. This should mean that at the 40 day mark, the animals are being fed about three mL of the above mix per animal (656 mL/200 larvae). This will obviously require a larger bowl or container, and the water they are in will be composed largely of the algae solution between water changes. This does not seem to negatively affect them.

2.2) Juvenile and Adult Husbandry

All animals older than 40 days are housed in tubs with 8-12L of ASW, with 2-10 cm of sand. Animals can be kept at RT, although preliminary data indicates that faster gonad growth in the adults occurs at 19 degrees C (Jr Kai Yu, personal communication). A full water change with sand rinsing is performed weekly (below).

2.2.1) Cleaning Tubs

To clean their enclosures, shake/swirl the tub to sift the animals from the sand, and pour them out with the water. Rinse with seawater and repeat until all animals are removed. Then rinse the sand out with DI water, stirring/agitating vigorously to remove waste. Once the rinse water no longer turns brown/has waste in it, rinse the sand with ASW two times, then replace 8-12 L of ASW into the tank, and catch the animals with a net (or the like) and put them back into the cleaned tank.

The number of animals housed in one tank described above depends on their size. For 40+ day old juveniles, up to 4000 per tank is ok. For every centimeter they grow, cut the number of animals in one tank down by about half until they reach 5 cm in length (which will be the maximum size of most individuals), such that they should be at a density of ~200 per tank by the time they reach 5 cm. This maximum density will be sufficient for the rest of their lives (which is thought to be up to 4 years). If using silica sand, the fully grown adults are large enough to be filtered out with netting on an insert in the tank, which greatly reduces the amount of time needed to do a tank cleaning.

2.2.2) Juvenile and Adult Feeding

Feed each 8-12L tub with ~200 animals in it twice daily with 1L of 20 micron filtered 3:1 *Isochrysis*:*Tetraselmis* (by cell count, which in our hands nearly corresponds to volume), spaced out by at least 6 hours. Dr. Jr Kai Yu's lab uses sometimes replaces *Isochrysis* for a diatom, *Chaetoceros*, and *Tetraselmis* for another green alga, *Chorella*. Our lab has observed gonad growth on feeding regimens consisting of 3:1 *Isochrysis*:*Tetraselmis*, as well as 3:1 *Isochrysis*:*Rhodomonas* (by volume, not cell count; unpublished data). The food:animal ratio will increase when you reduce the number of animals per tank based on their growth. Participating labs are still collecting data on the relationship between the type of food and gonad growth, though the amount of food seems to be a bigger determining factor here (so long as their diet is composed mostly of either *Chaetoceros* or *Isochrysis*). We remove water from each tank daily to avoid overflow, which also aids in keeping their water clean.

CHAPTER II

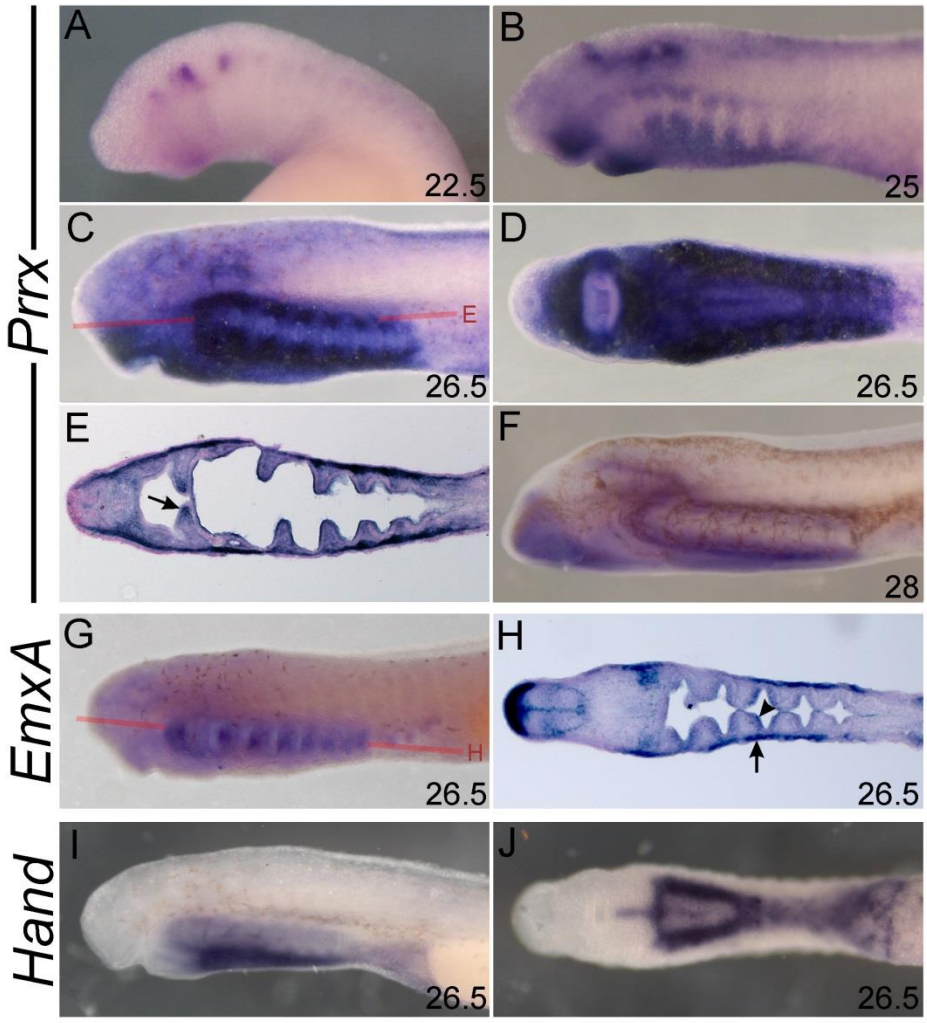

Fig. S1. Expression of *prrxA*, *emxA*, and *hand* in *P. marinus.* A, B, C, F, G, I, and J show left lateral views. D and J show ventral/oral views with anterior to left. E and H show coronal sections with anterior to left. Developmental stage(Tahara, 1988) is indicated in the bottom right corner of each panel. (A-F) *prrxA* expression marks the poles of each PA, the upper lip, and mesenchyme surrounding the otic placode. Red line in C indicates the plane of section in E. Arrow in E indicates *prrxA* expression in the future medial velar skeleton. (G and H) *emxA* expression is not found in mid-pharyngula CNCCs. Red line in G indicates the plane of section in H. Arrow in H indicates epithelial ectodermal expression of *emxA.* Arrowhead in H indicates endodermal expression of *emxA.*(I and J) *hand* expression marks ventral CNCCs. Note the medial stripe of *hand* expression in the lower lip.

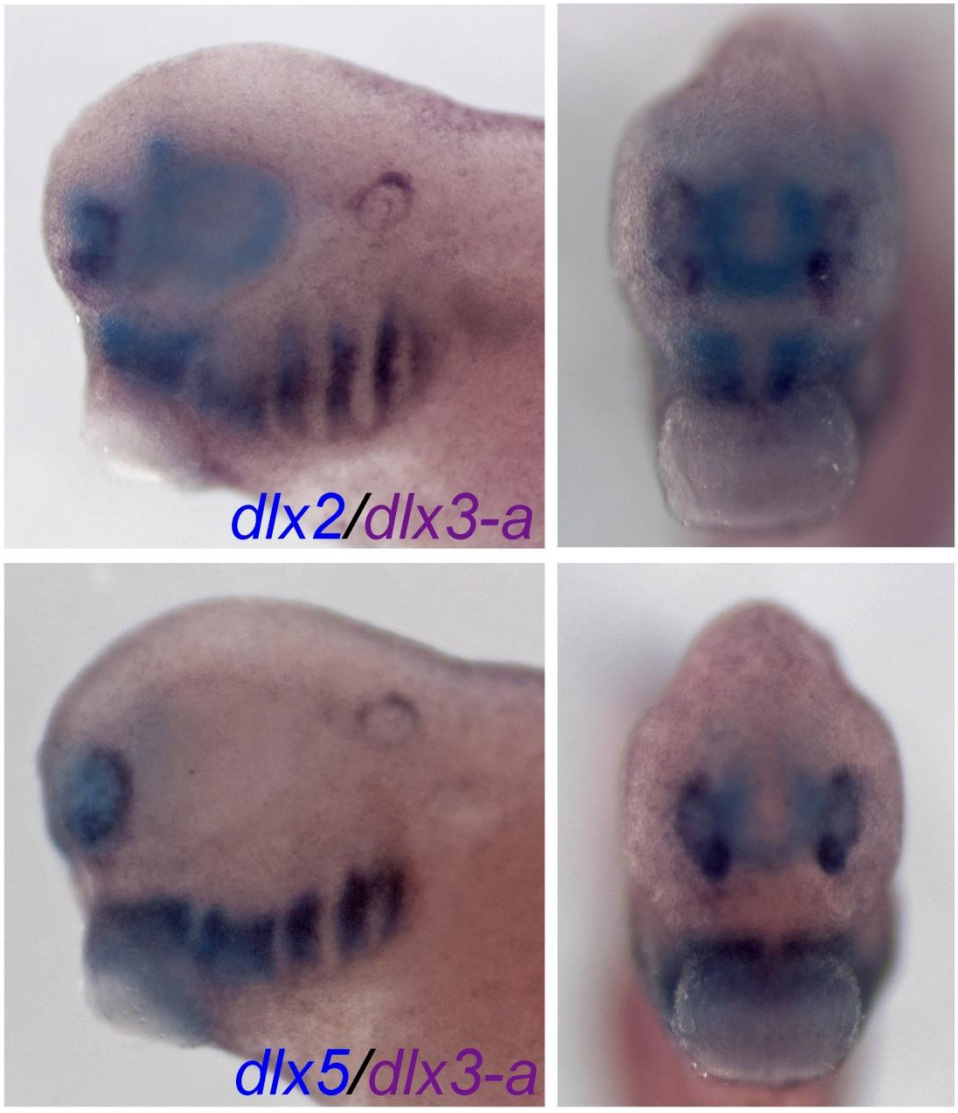

Fig. S2. Nested *dlx* expression in *Xenopus laevis* at st. 33. Left column of photos shows left lateral views, and the right column shows anterior/oral views. *dlx3-a* is shown in purple in all panels. *dlx2* is shown in purple in the top row, and *dlx5* is shown in purple in the bottom row. Note the common dorsal border of *dlx3-a* and *dlx5* (lack of blue staining above the purple).

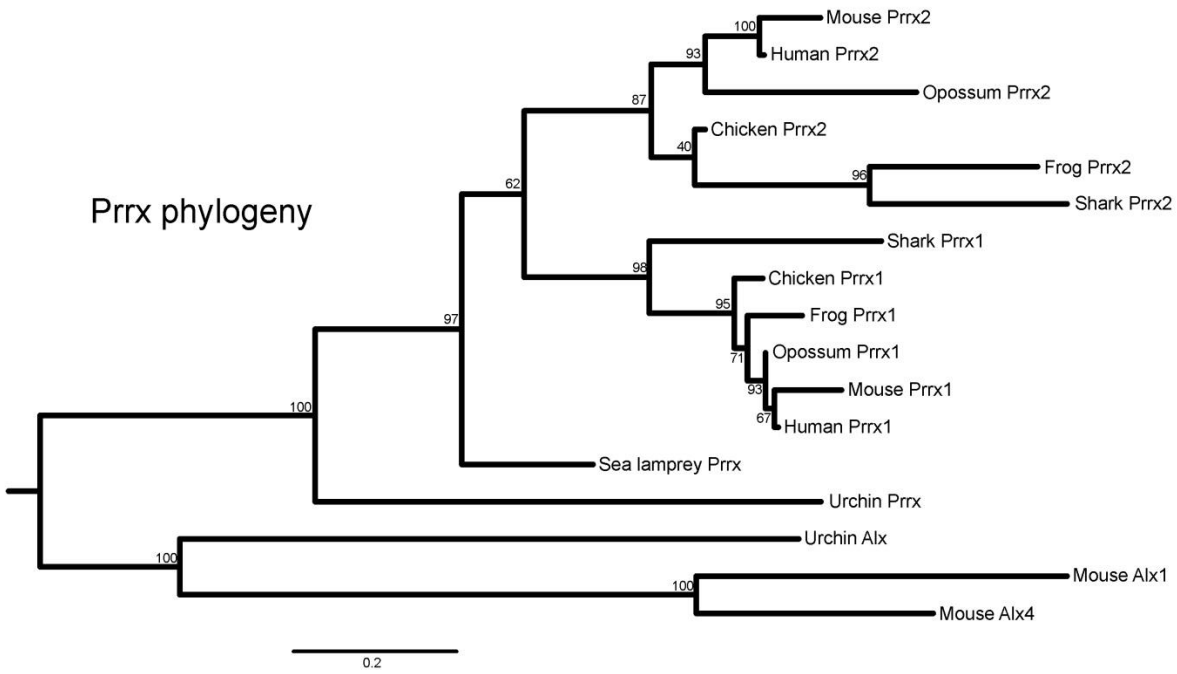

Fig. S3. A phylogeny of deuterostome Prrx proteins, with urchin and mouse Alx proteins serving as the outgroup. Bootstrap supports are shown at each node. Accession numbers are as follows (in common name alphabetical order): *Gallus gallus* (**chicken**) Prrx1: NP_001007822.1, Prrx2: NP_001280027.1; Prrx2: XP_005234079.1; *Xenopus tropicalis* (**frog**) Prrx1: XP_004913839.1, Prrx2: XP_002940803.1; *Homo sapiens* (**human**) Prrx1: NP_073207, Prrx2: NP_057391.1; *Mus musculus* (**mouse**) Alx1: NP_766141.1, Alx4: NP_031468.1, Prrx1: XP_006496766, Prrx2: NP_033142.2; *Monodelphis domestica* (**opossum**) Prrx1: XP_001364765, Prrx2: XP_007475194.1; *Petromyzon marinus* (**sea lamprey**) Prrx: KX650614; *Callorhinchus milii* (**shark**) Prrx1: XP_007893876, Prrx2: XP_007901045.1; *Stongylocentrotus purpuratus* (**urchin**) Alx: NP_999809.1, Prrx: XP_787382.1.

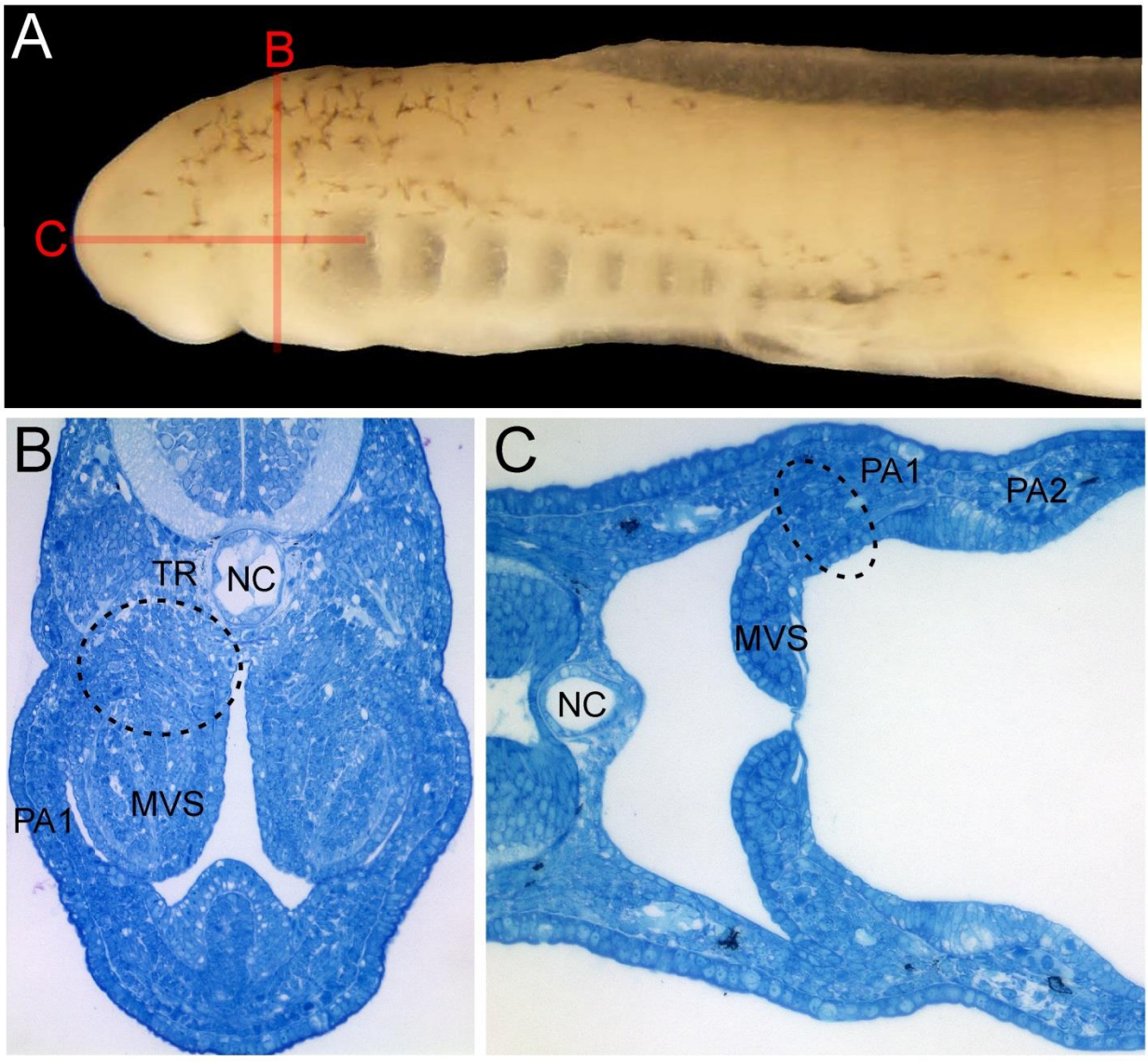

Fig. S4. The lamprey medial velar skeleton is contiguous with both PA1 and the 'trabecular' region. (A) Image of a st. 26.5 lamprey. Red lines in (A) indicate the plane of section in (B) and (C). (B) a transverse section of a st. 26.5 lamprey showing the medial velar skeleton is contiguous with the dorsal trabecular cartilages (dotted oval). (C) a coronal section of a st. 26.5 lamprey showing the medial velar skeleton is contiguous with PA1 mesenchyme. MVS, medial velar skeleton; NC, notochord; PA1, 1st pharyngeal arch; TR, trabecular.

CHAPTER III

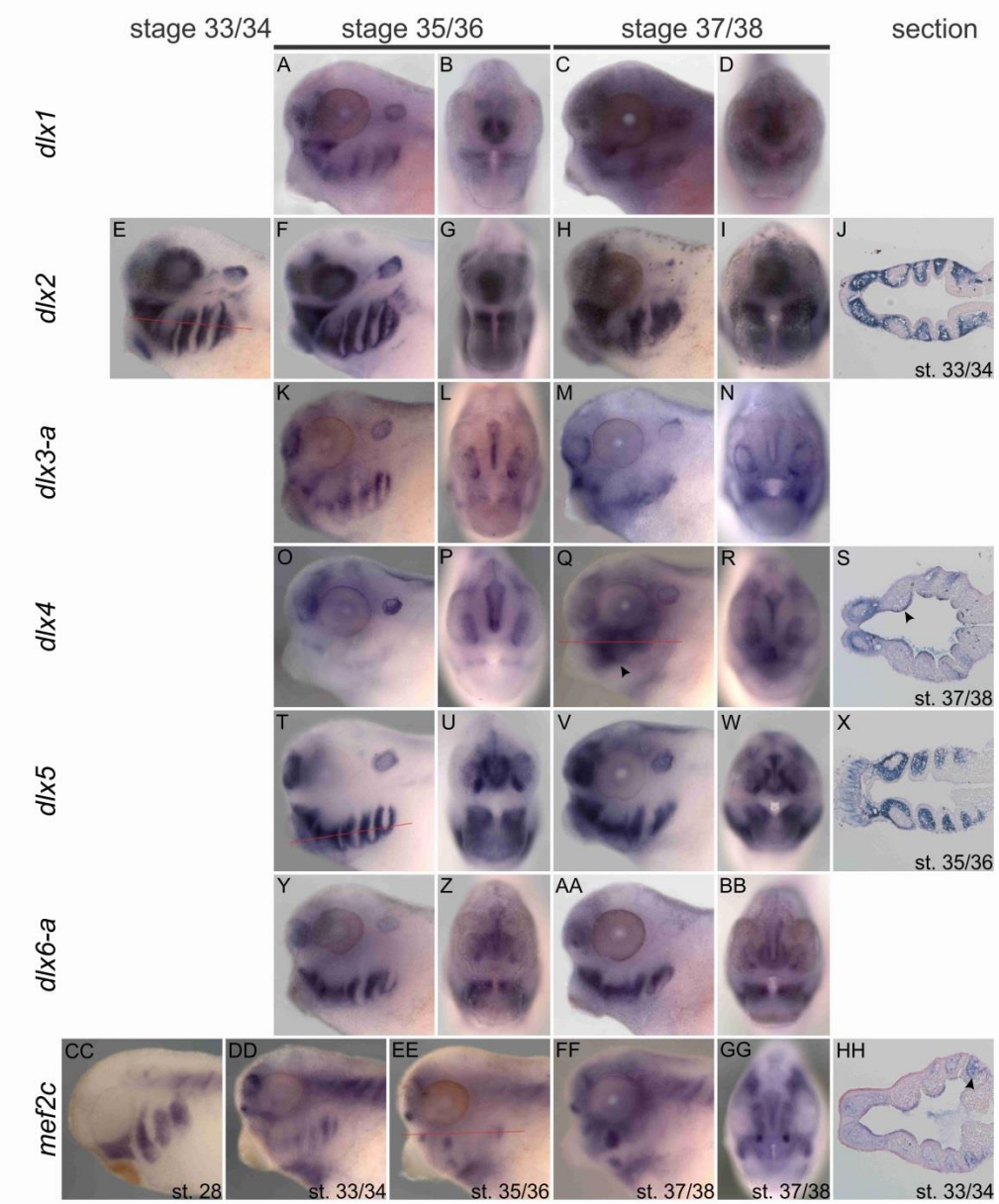

Fig. S1. Expression of *dlx* genes and *mef2c*. Gene and stage are indicated in figure. Anterior to left in all lateral views, anterior to top in all sections. For those genes with a section shown, the plane of section is shown as a red line on the corresponding wholemount panel. (Q) at stage 37/38, the *dlx4* signal in the pharynx is mostly due to riboprobe trapping (arrows in Q and S), though some expression in the posterior arch CNC remains. (HH) expression of *mef2c* is detected in pharyngeal CNC (arrows).

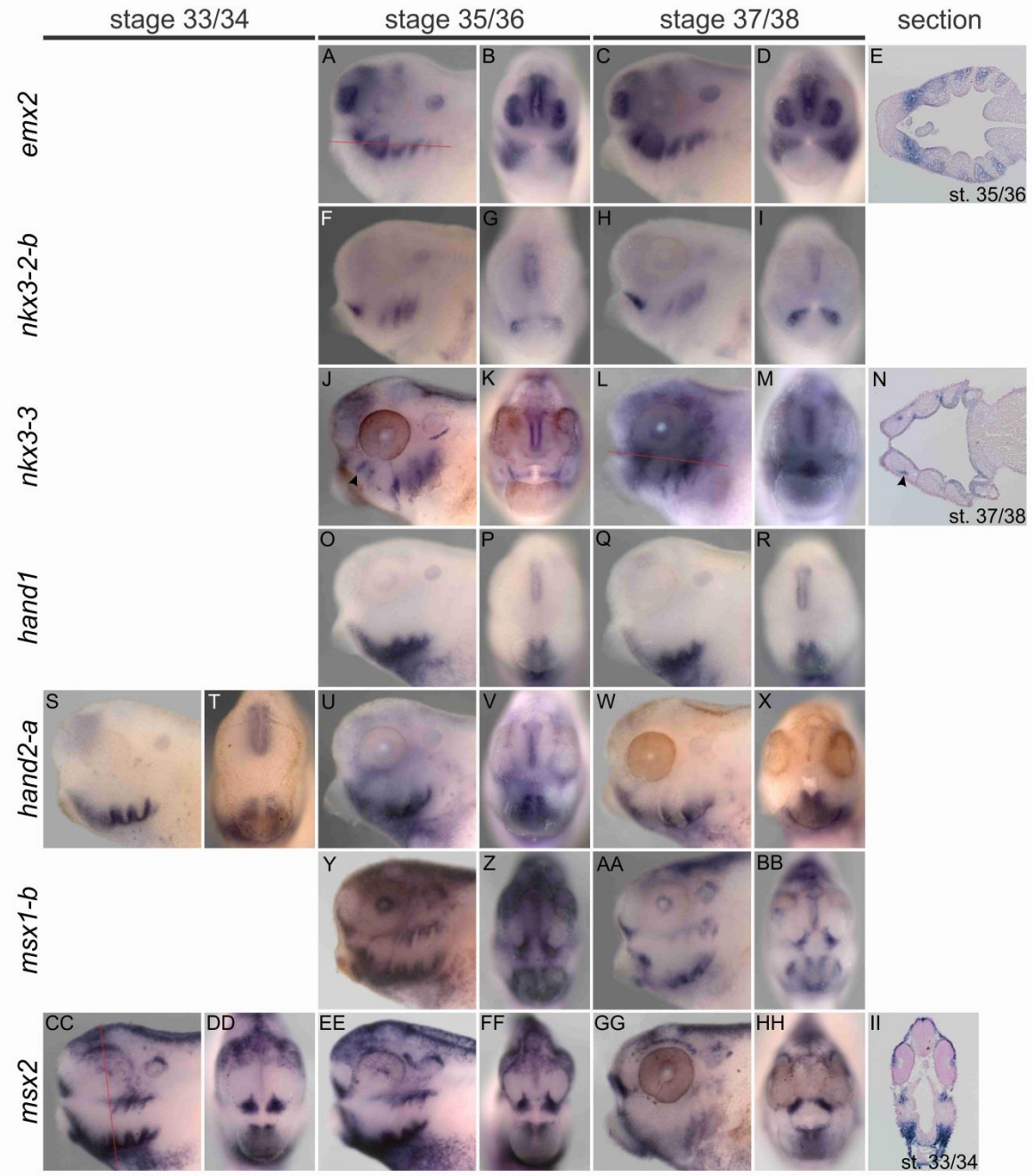

Fig. S2. Expression of *emx2*, *nkx3-2-b*, *nkx3-3*, *hand1*, *hand2-a msx1-b*, and *msx2*. Gene and stage are indicated in figure. Anterior to left in all lateral views, anterior to top in E, N, dorsal to top in II. For those genes with a section shown, the plane of section is shown as a red line on the corresponding wholemount panel. (J, N) *nkx3-3* is expressed in CNC where the primary jaw joint will eventually form (arrow).

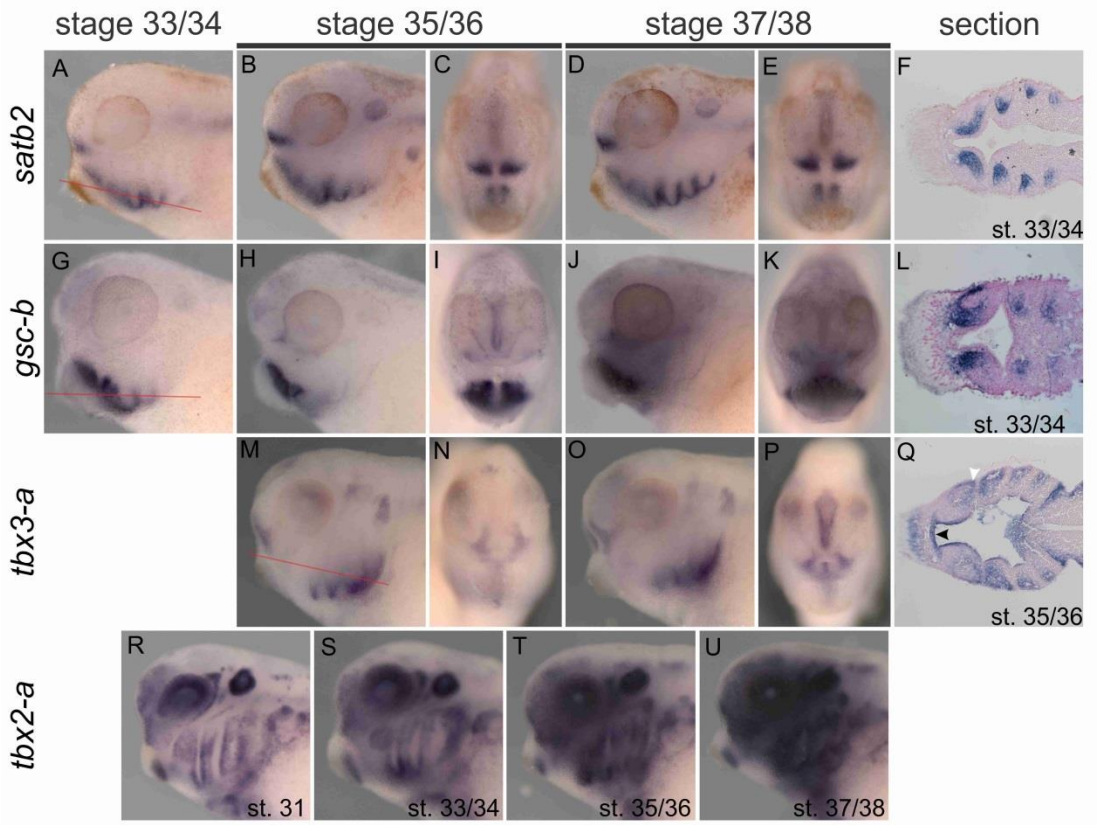

Fig. S3. Expression of *satb2*, *gsc-b*, *tbx3-a*, and *tbx2-a*. Gene and stage are indicated in figure. Anterior to left in all lateral views, anterior to top in all sections. For those genes with a section shown, the plane of section is shown as a red line on the corresponding wholemount panel. (Q) *tbx3-a* is expressed in both CNC (black arrow) and endoderm (white arrow) of the PAs.

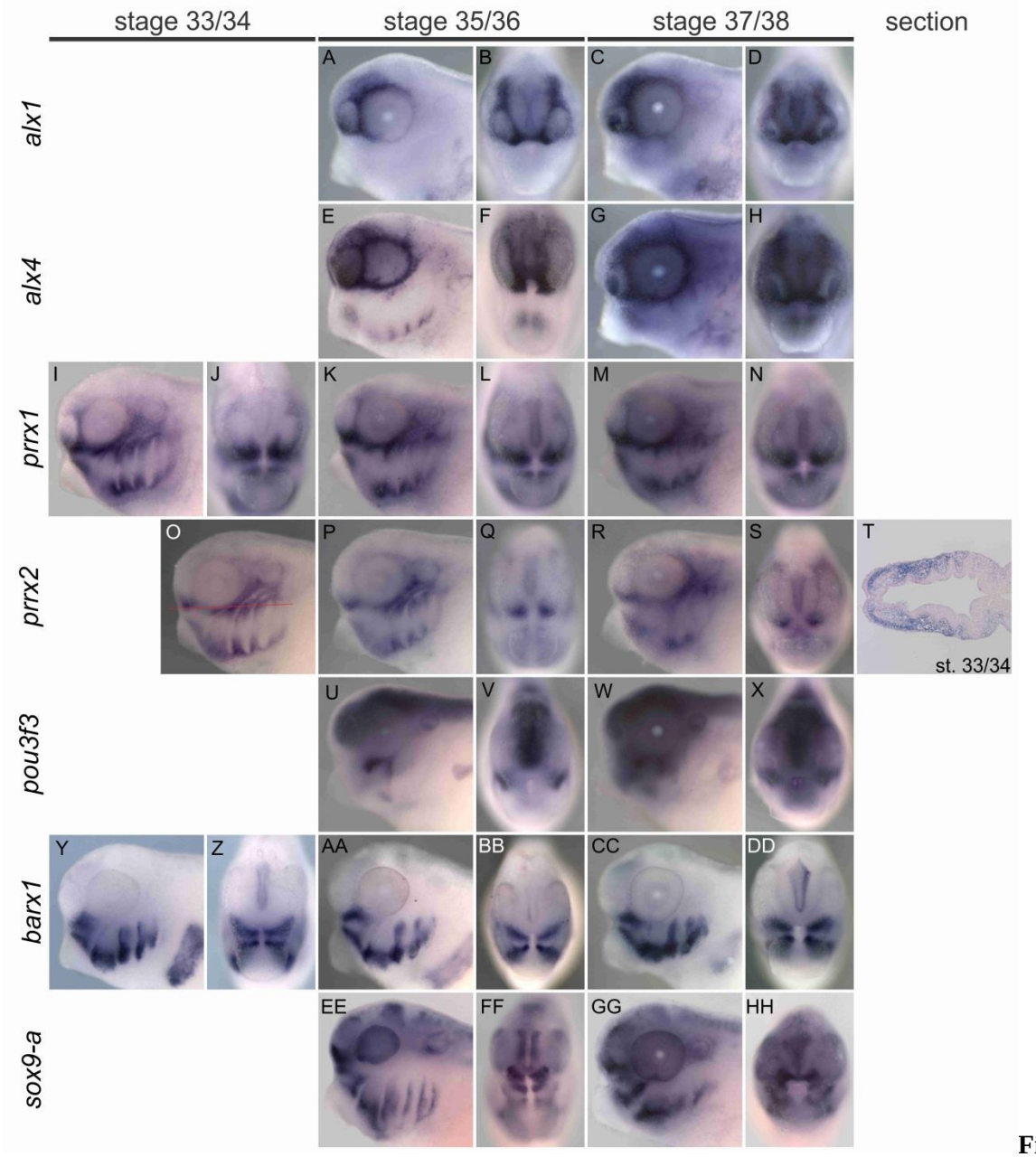

Fig. S4.

Expression of *alx1*, *alx4*, *prrx1*, *prrx2*, *pou3f3*, *barx1*, and *sox9-a*. Gene and stage are indicated in figure. Anterior to left in all lateral views. Red line in (O) shows plane of section in (T).

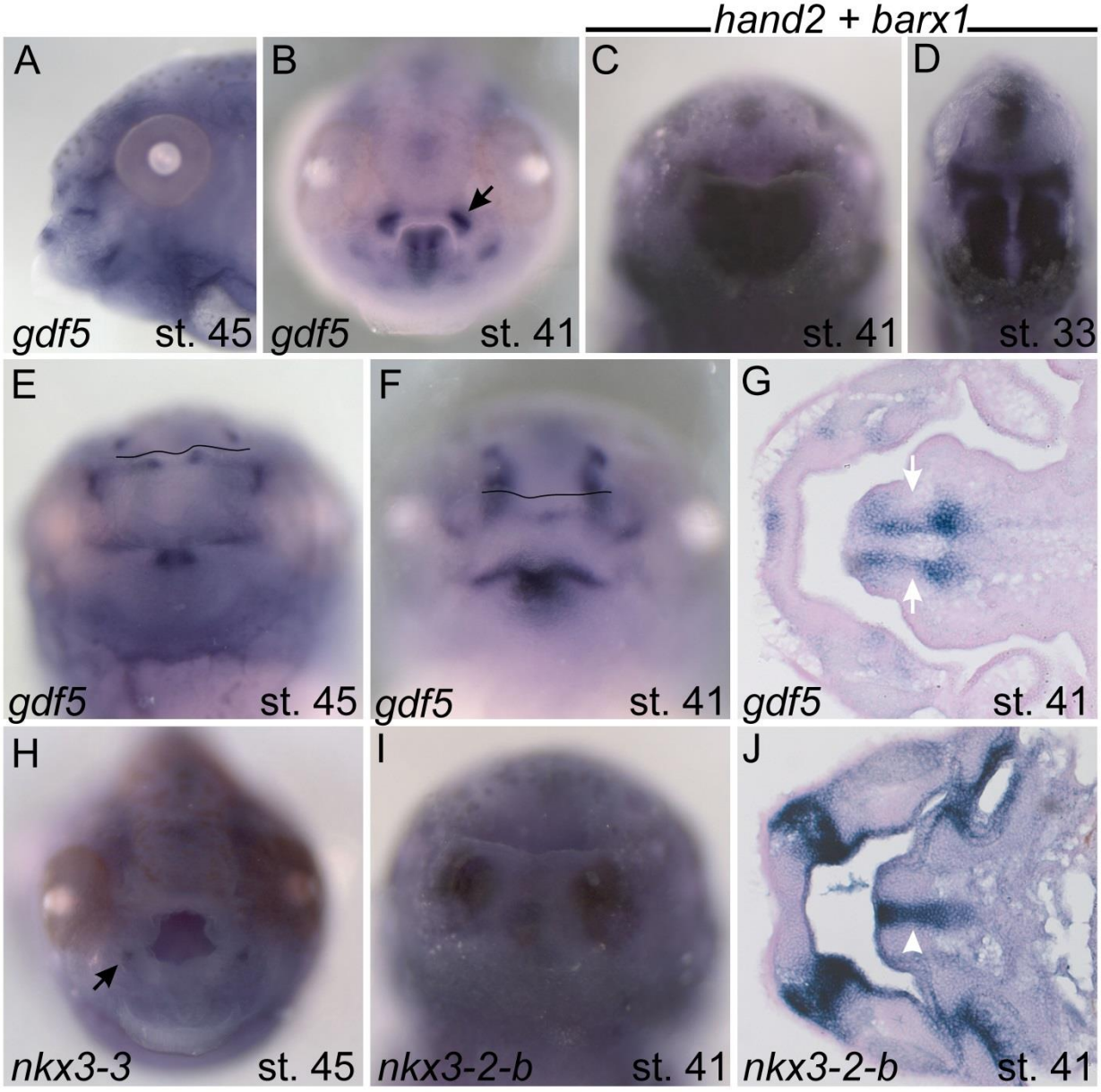

Fig. S5. Gene expression in the intramandibular joint, PA1, and PA2 derivatives. Gene and stage are indicated in figure. Lateral view with anterior to left (A), anterior views (B, D, H) and anteroventral views (C, E, F, I), and sections with anterior to left (G, K). (B) *gdf5* expression between the suprarostal and ethmoid plate (arrow). (E, F) black line shows the outline of the lower jaw. (G) *gdf5* is expressed in tissue flanking the future basihyal (white arrows). (H) *nkx3-3* is expressed in the lateral portion of the primary jaw joint (arrow). (J) *nkx3-2-b* is expressed within the future basihyal (white arrowhead).

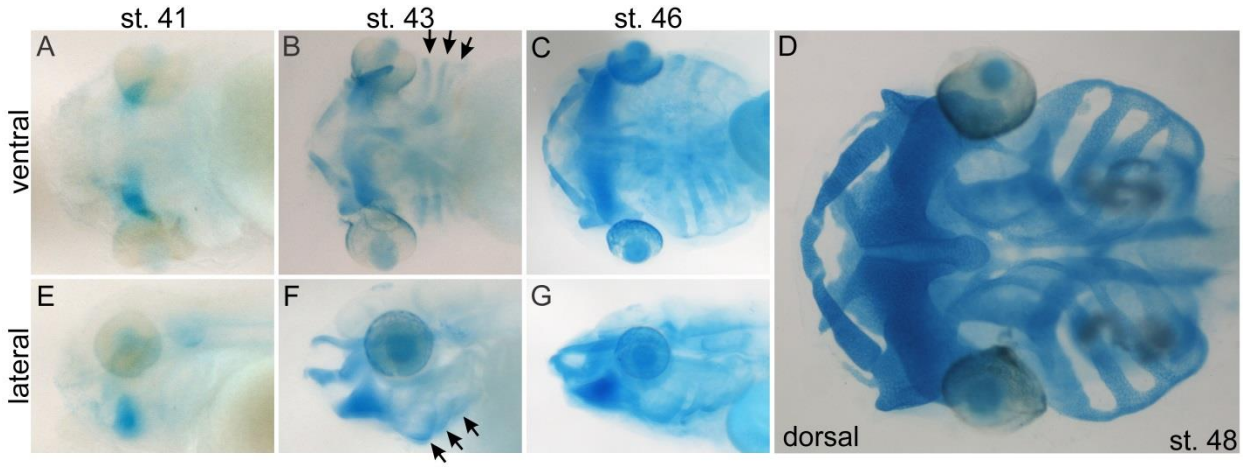

Fig. S6. Progression of cartilage differentiation in *X. laevis* as revealed by alcian blue reactivity. View is indicated in figure. (A, E) at st. 41 heavy staining in the hyoid is seen, with no reactivity in the posterior PAs. (B, F) at st. 43, staining becomes apparent in the intermediate/ventral aspect of the posterior arches (arrows). (C, D) at stage 46, the majority of all future head skeleton elements have begun glycosaminoglycan deposition. (D) a dorsal view with a ventral focal plane of at st. 48 when head skeleton differentiation is largely complete.

CHAPTER IV

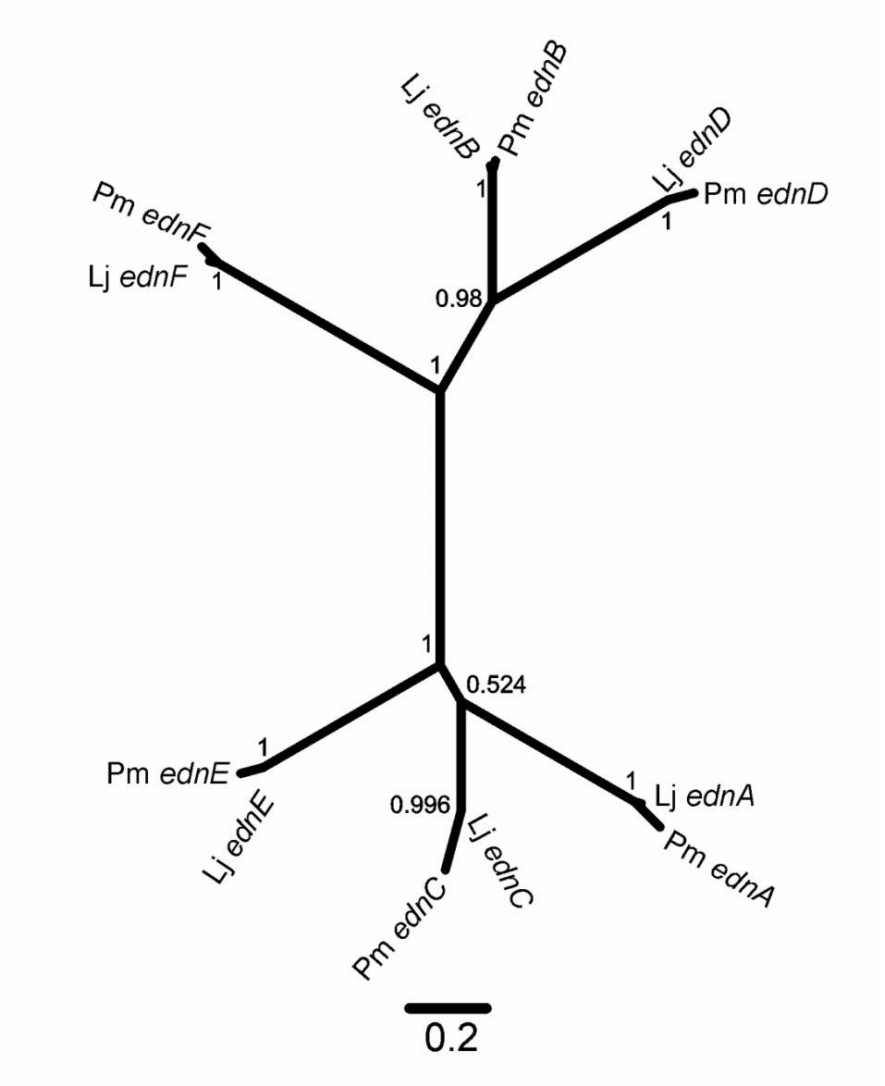

Fig. S1. An unrooted phylogenetic tree built from *edn* gene sequences in sea lamprey (*Petromyzon marinus*) and the Japanese lamprey (*Lethenteron japonicum*). Maximum likelihood analysis was used to determine lamprey ligand orthology. Bootstrap values are shown at the base of each node. Accession numbers for all sequences can be found in Tab. S2.

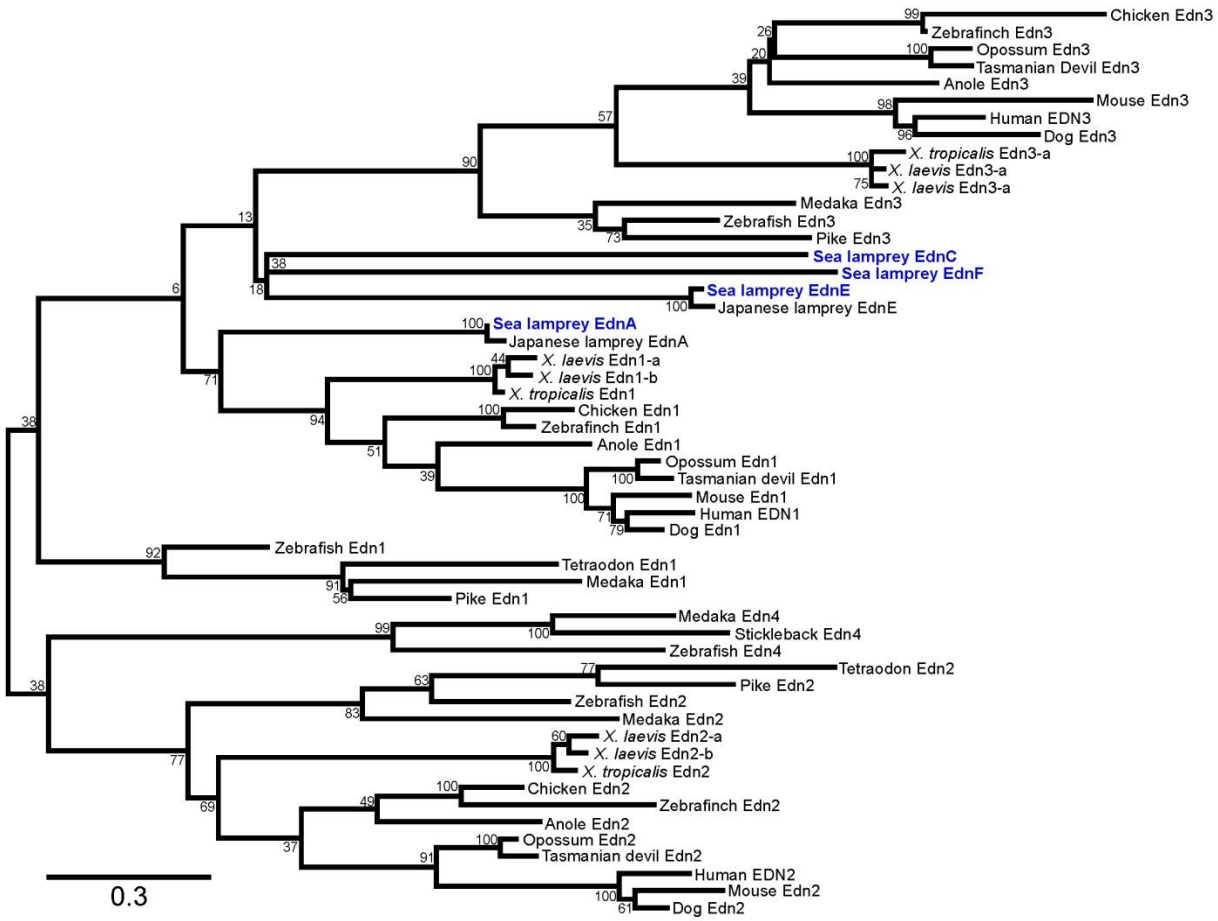

Fig. S2. A phylogenetic tree built from Edn amino acid sequences in vertebrates. Maximum likelihood analysis was used in attempt to determine the orthology of lamprey Edn ligands. This tree was midpoint rooted between the Edn1/3 clade and the Edn2/4 clade. All calculated bootstrap values are shown at the base of each node. Accession numbers for all sequences can be found in Tab. S2.

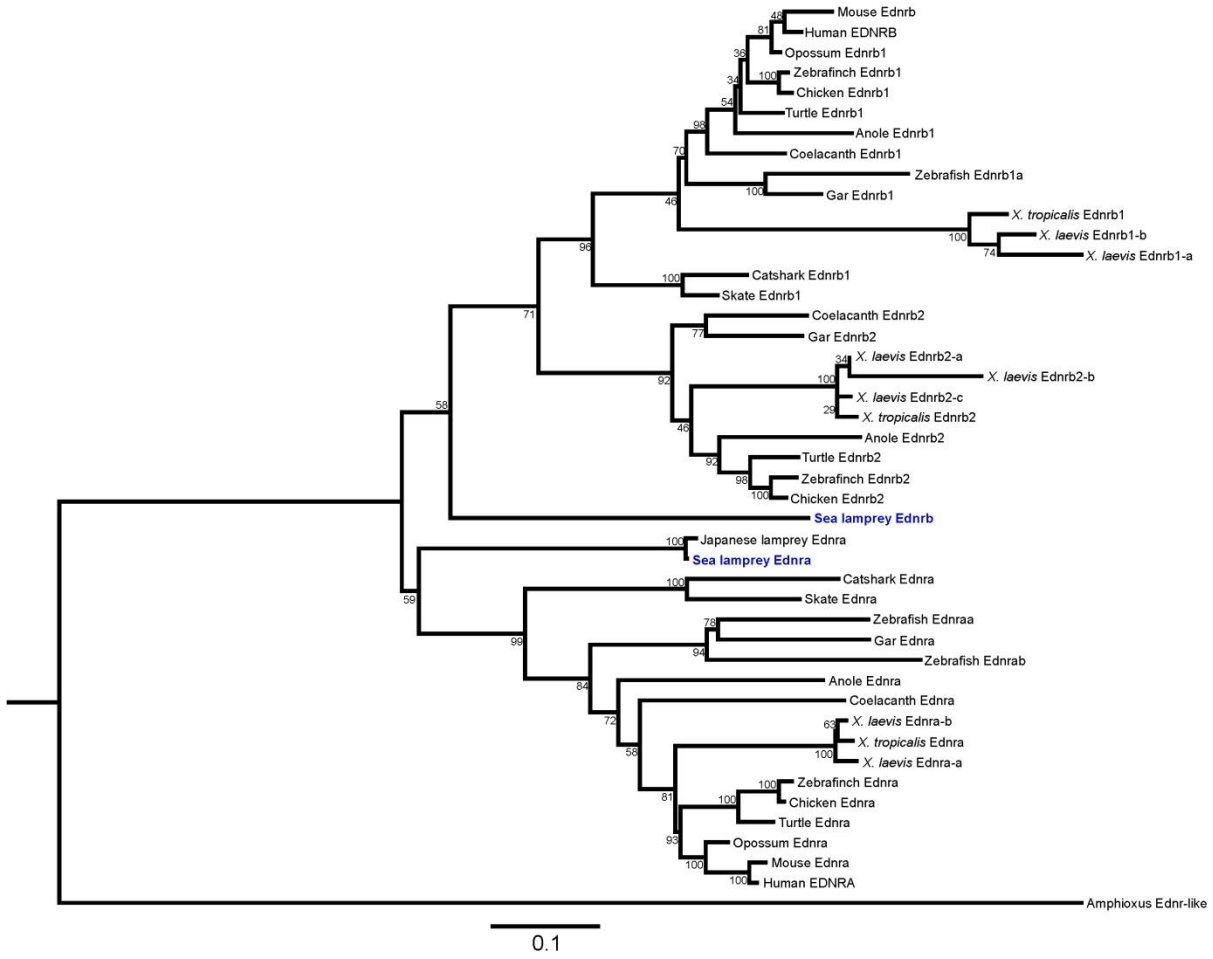

Fig. S3. A phylogenetic tree built from EdnR amino acid sequences in chordates. Maximum likelihood analysis was used to assign orthology to the lamprey EdnRs. Bootstrap values above 50 are shown at the base of each node. Amphioxus EdnR-like was selected as the outgroup. Accession numbers for all sequences can be found in Tab. S2.

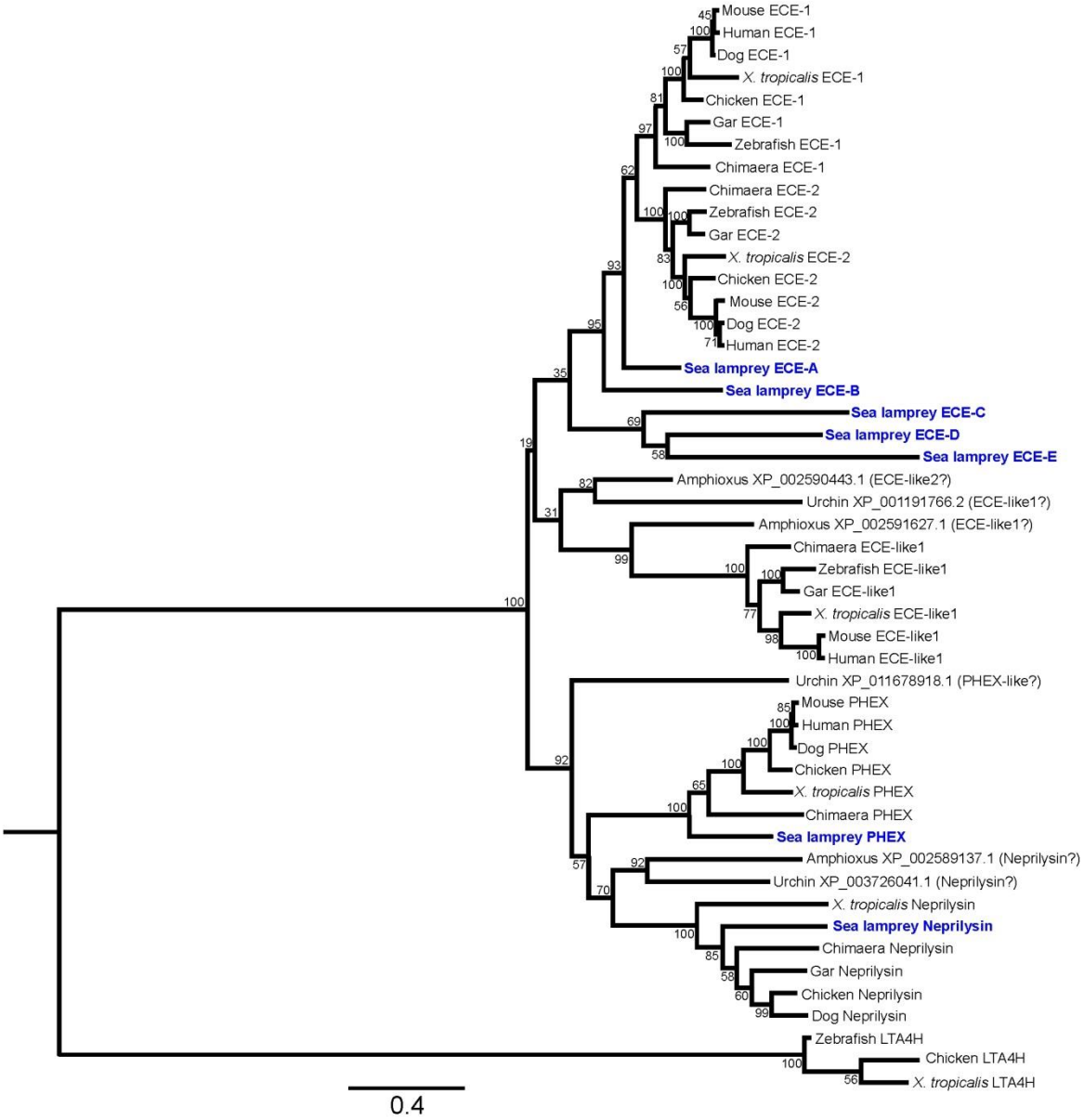

Fig. S4. A phylogenetic tree built from ECE-related metallopeptidase amino acid sequences in deuterostomes. Maximum likelihood analysis was used to assign orthology to the lamprey ECEs. All bootstrap values are shown at the base of each node. Three gnathostome LTA4H sequences were selected as the outgroup. Accession numbers for all sequences can be found in Tab. S2.

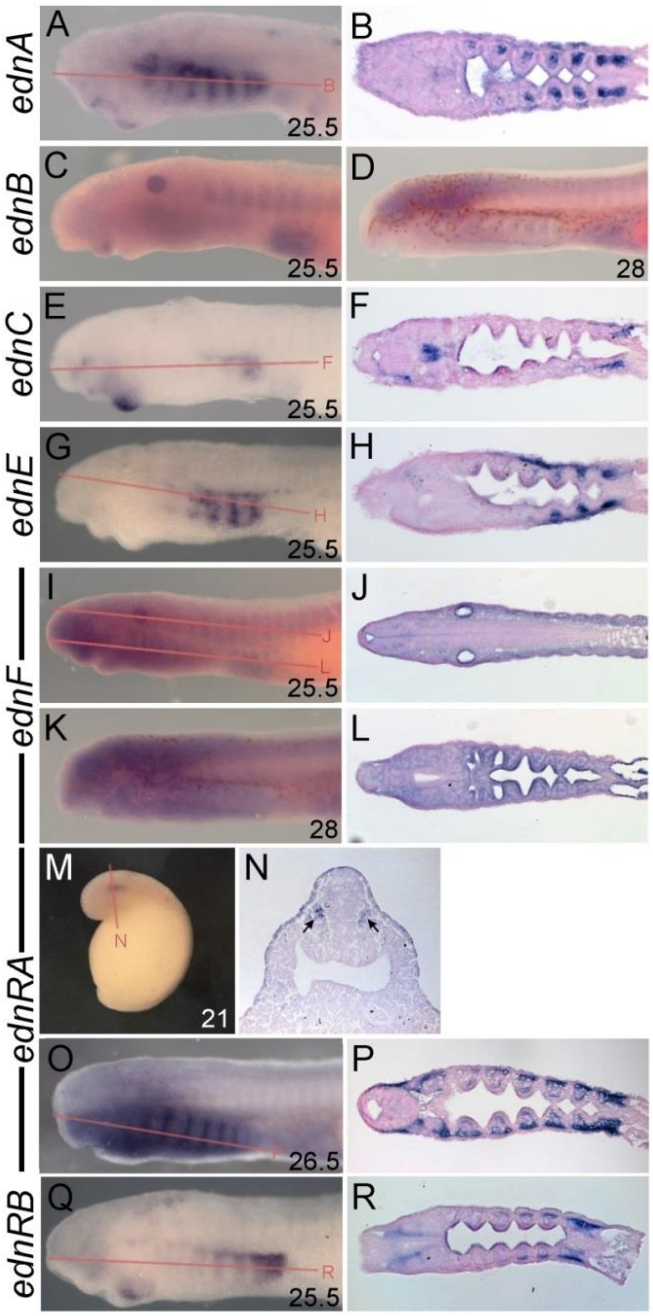

Fig. S5. Expression of *ednA-F*, *ednRA*, and *ednRB* in *P. marinus* at pharyngula stages. Left lateral views in panels A, C, E, G, I, K, M, O, and Q. Sections shown in panels B, F, H, J, L, N, P, and R correspond to the labeled red line in the wholemount ISH panel for each gene. Developmental stage(Tahara, 1988) for each specimen is indicated in the bottom right corner of the wholemount panel for each specimen.

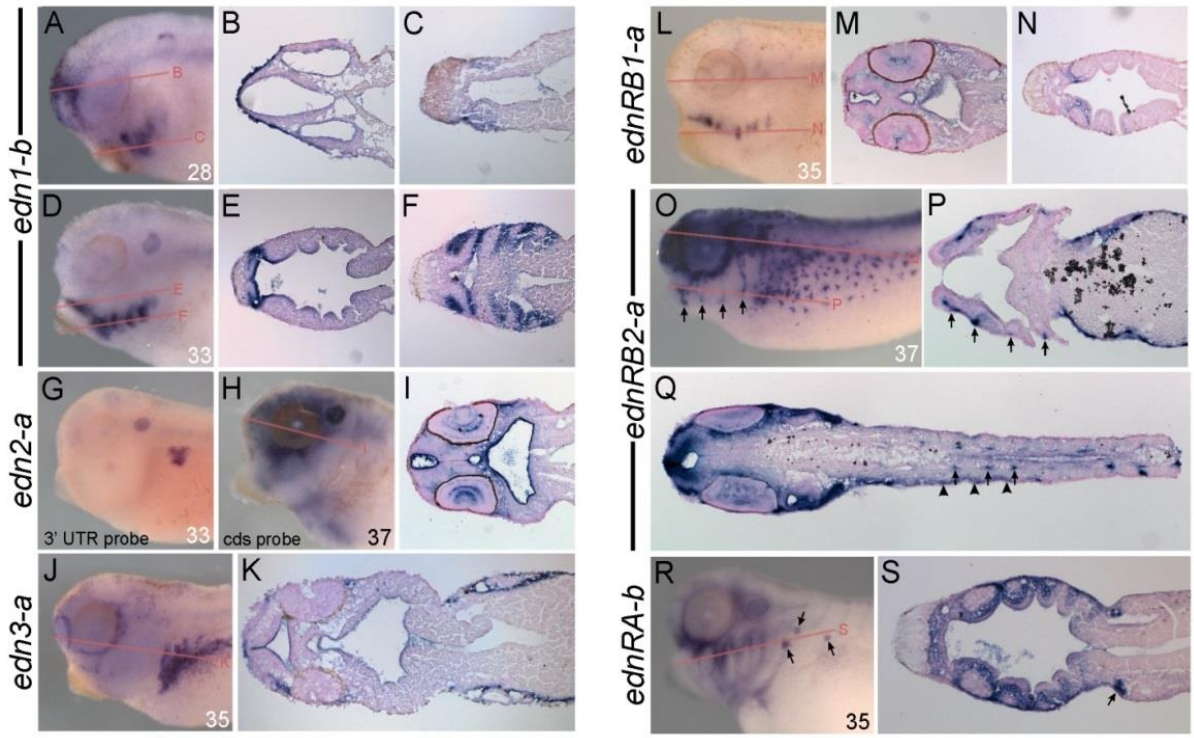

Fig. S6. Expression of *edn1-3, ednRA*, and *ednRB* in *X. laevis*. Left lateral views in panels A, D, G, H, J, L, O, and R. Sections shown in panels B, C, E, F, I, K, M, N, P, Q, and S correspond to the labeled red line in the wholemount ISH panel for each gene. Developmental stage(Nieuwkoop and Faber, 1956) for each specimen is indicated in the bottom right corner of the wholemount panel for each specimen. Arrows in O and P indicate *ednRB2-a* expression in the branchial nerves. Arrows in Q indicate expression consistent with dorsal root ganglia. Arrowheads in Q indicate expression in melanocytes. Arrows in R and S indicate *ednRA-b* expression in nephrostomes.

CHAPTER V

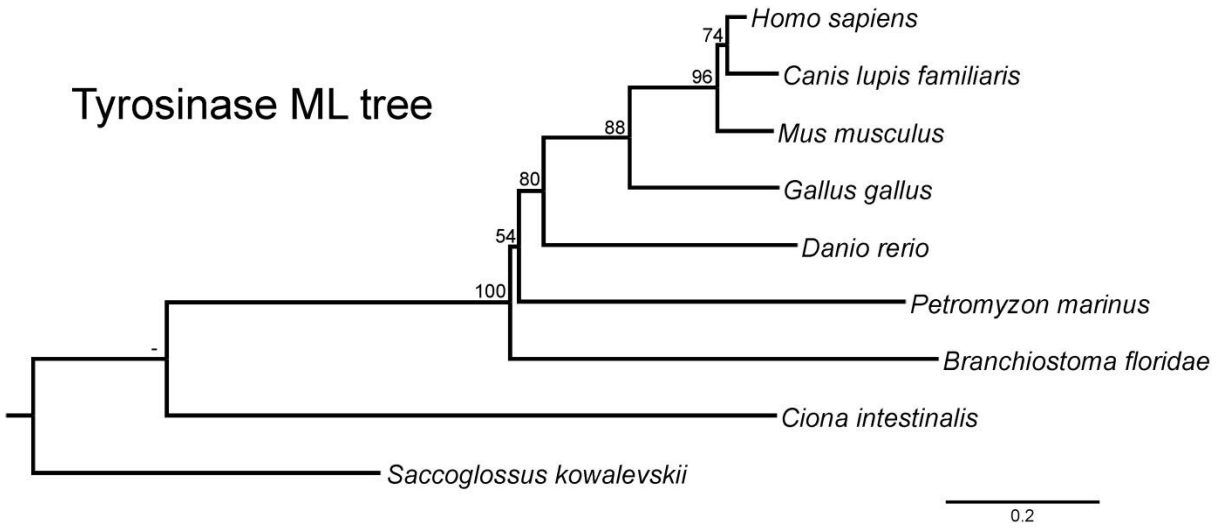

Fig. S1. Phylogenetic analysis of Tyrosinase amino acid sequences. Alignment and phylogenetic analysis of nine deuterostome Tyrosinase sequences were performed in MEGA6, using the WAG model of protein evolution for Maximum Likelihood tree construction. A 250 replicate bootstrap test was performed; the scores for this analysis are found at the nodes in the tree. Sequence accession numbers are as follows: *Branchiostoma floridae* Tyrosinase: XM_002598584, *Canis lupis familiaris* Tyrosinase: NP_001002941.1, *Ciona intestinalis* Tyrosinase: XM_002123004, *Danio rerio* tyrosinase: AAN17339.1, *Gallus gallus* Tyrosinase: NP_989491.1, *Homo sapiens* TYROSINASE: NP_000363.1, *Mus musculus*: NM_011661, *Petromyzon marinus* Tyrosinae: KR150760, *Saccoglossus kowalevskii* Tyrosinase-like: XP_006817095.1.

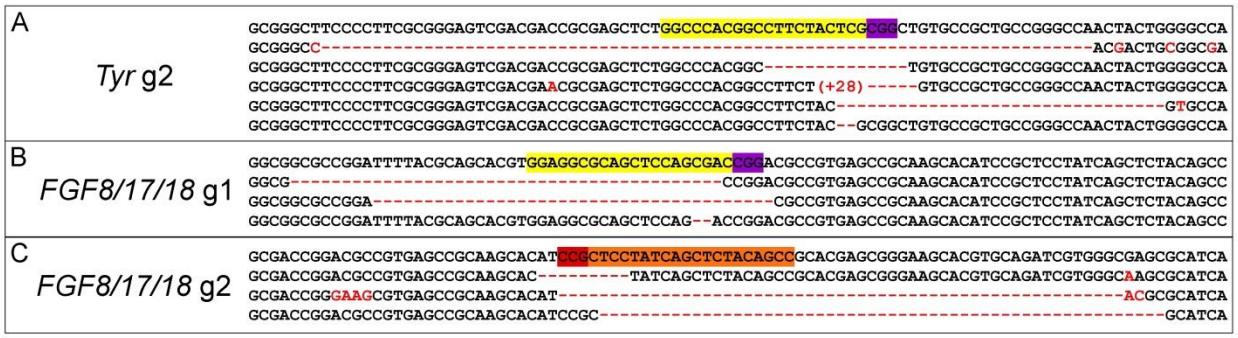

Fig. S2. Genotyped loci of targeted genes. All sequences show the sense strand in 5' to 3' orientation with respect to the corresponding gene. Forward target sites and PAMs are shown in yellow and purple, respectively, while the reverse strand target and PAM are shown in orange and red, respectively, for *FGF8/17/18* gRNA 2. Deletions relative to the WT sequence are shown as red dashes, while polymorphisms and inferred insertions are shown as red letters. (A) WT (top sequence) and 5 example mutant sequences are shown for *Tyr* gRNA 2 target site; "(+28)" indicates a 28 bp insertion. (B) WT (top sequence) and 3 mutant sequences returned for genotyping of the *FGF8/17/18* gRNA 1 target site. (C) WT (top sequence) and 3 mutant sequences returned for genotyping of the *FGF8/17/18* gRNA 2 target site.

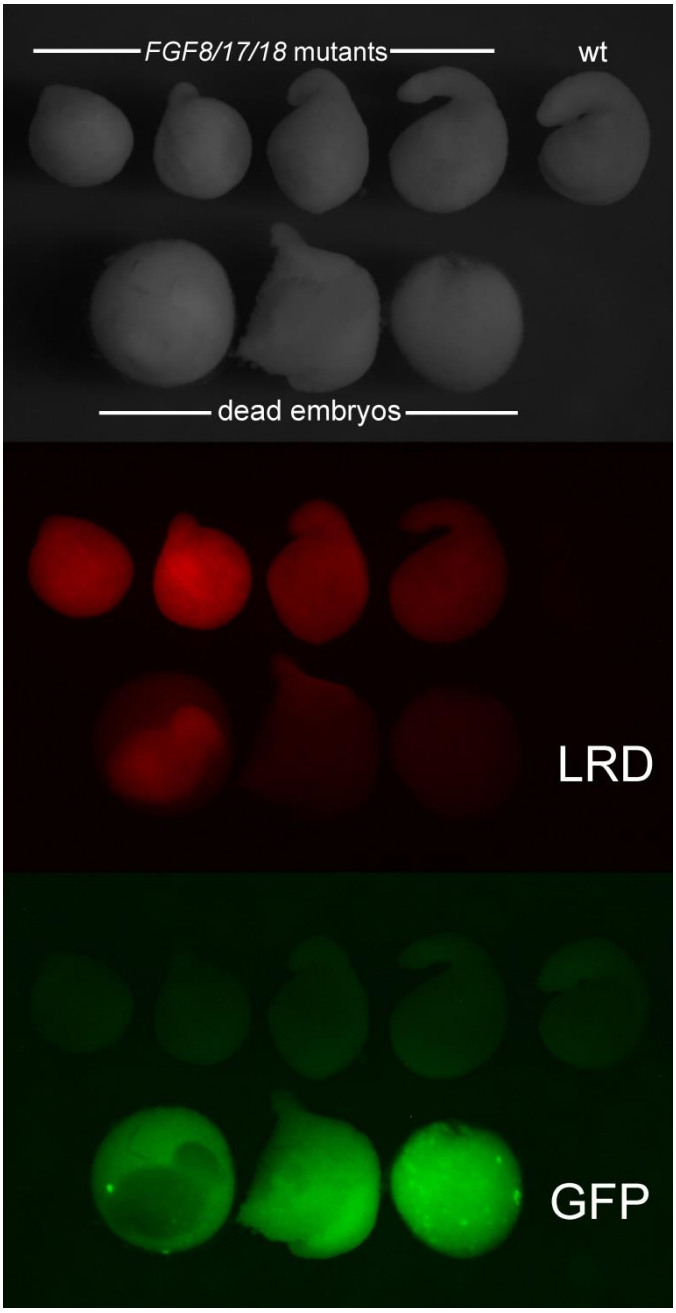

Fig. S3. Severely affected *FGF8/17/18* mutants do not display autofluorescence associated with death. Whole-mount embryos from each of 4 mutant classes, and a WT embryo shown with normal light (top), LRD fluorescence (middle), and on the GFP channel (bottom). Note the high level of autofluorescence exhibited by the dead embryos, whereas even the most affected mutants glow only as strongly as a WT embryo.

CHAPTER VI

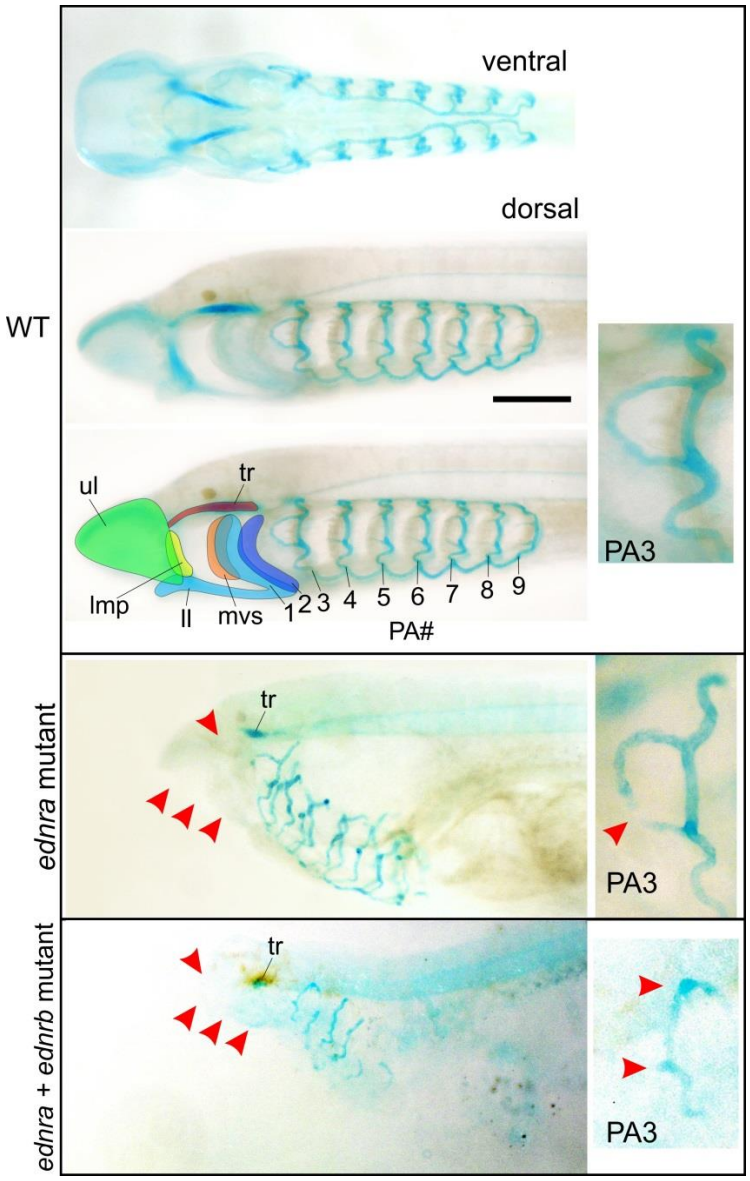

Fig. S1. Head skeleton morphology in WT and *ednra* mutant *P .marinus.* Top panels show a WT alcian blue stained st. 30 lamprey with anterior to left, which is labeled below to identify different cartilage elements. Warmer colors represent more medial structures, cooler colors represent more lateral structures. Inset shows WT PA3 from the same specimen. Below shows an *ednra* and *ednra+b* mutants, which are missing most anterior cartilages as revealed by alcian blue (red arrowheads). Inset shows PA3 from these same mutant specimens, though for *ednra* this photo was taken from a different angle to better see its shape. Scale bar represents 500 µm and applies to all images save the PA3 enlarged insets. ll, lower lip; lmp, lateral mouth plate; mvs, medial velar skeleton; PA, pharyngeal arch (numbered); tr, trabecular; ul, upper lip.

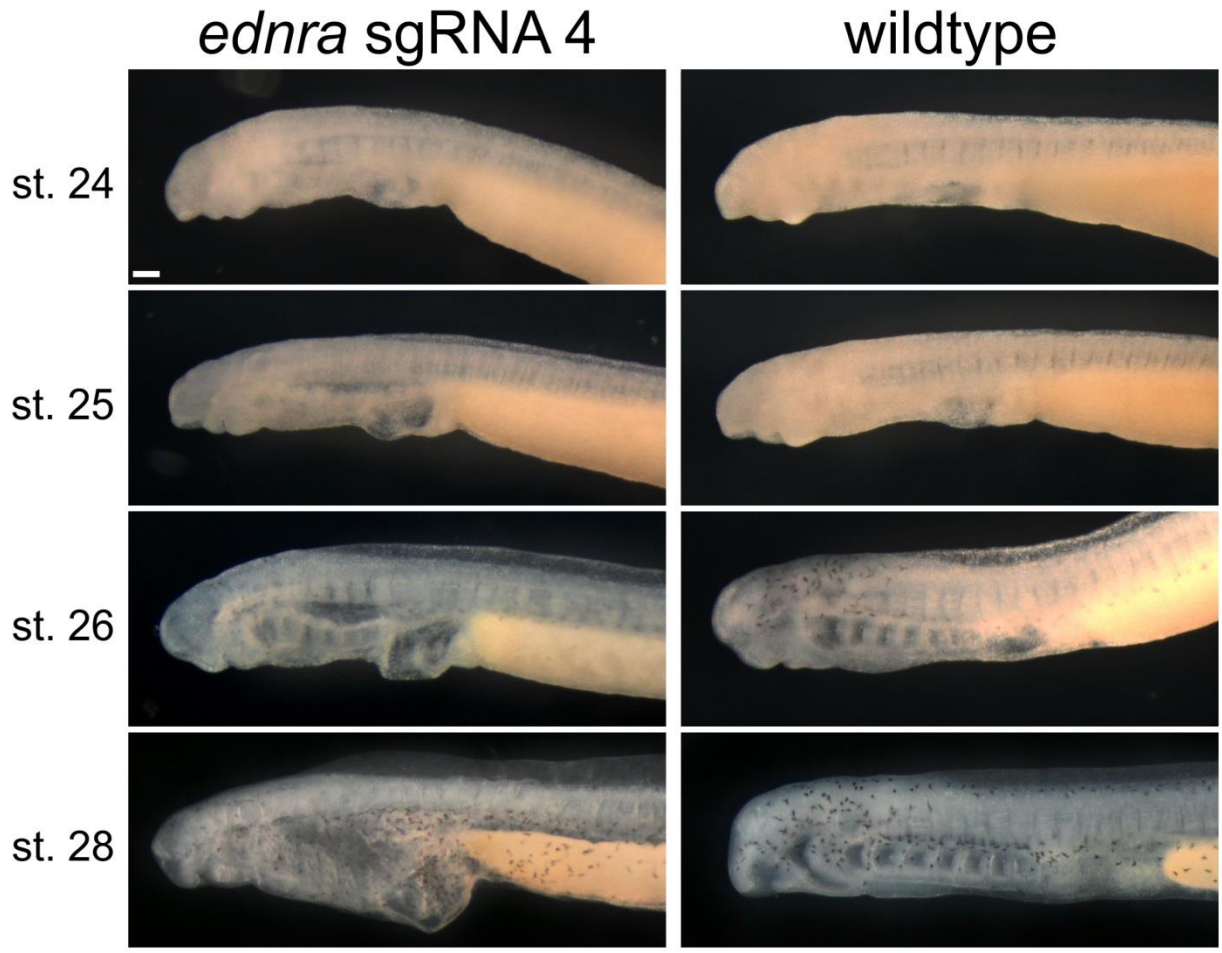

Fig. S2. *ednra* mutant development. A time series of lamprey larvae showing how the phenotype manifests through time. Anterior to left in all panels, which are all to scale with each other. This phenotype was extremely repeatable and consistent, allowing identification of severely affected individuals prior to ISH or IHC at st. 26.5. The phenotype first appears as a slight heart edema at st. 24, which grows constantly until at least st. 30. By St. 25-26, head skeleton morphology is slightly perturbed, which becomes greatly accentuated during histogenesis at st. 28+. Excess pigment cells covering the heart also appear at these later stages. Scale bar in the upper right image represents 100 µm and applies to all images.

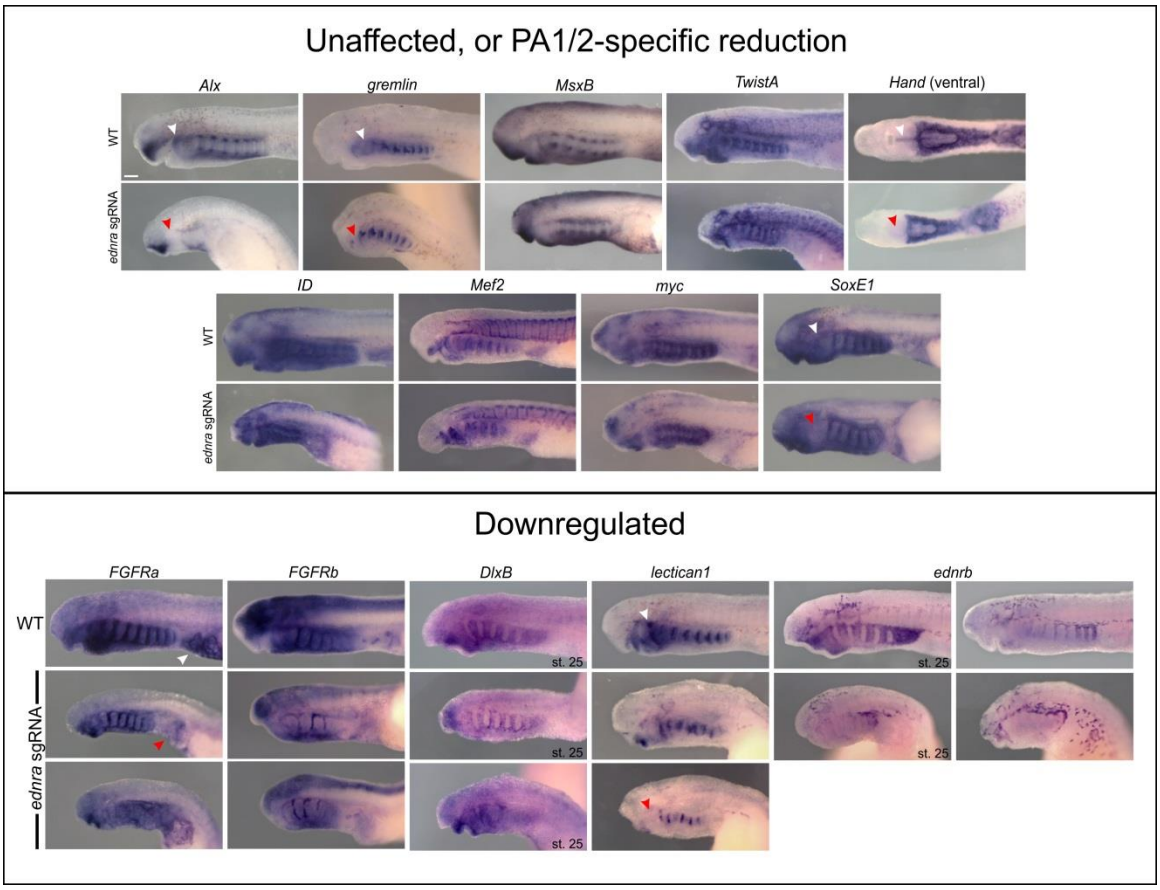

Fig. S3. Gene expression in *ednra* sgRNA injected lampreys. Anterior to left in all panels. Gene is indicated for each panel. All specimens are at st. 26.5 unless indicated otherwise. Top panels show genes mostly unaffected, though occasionally missing from PA1/2 only, bottom shows genes with stronger downregulation throughout the pharynx or heart. White arrowheads indicate WT expression in PA1 (or the heart in *FGFRa*), red arrowheads indicate missing expression in *ednra* mutants. Scale bar in the upper right image represents 100 µm and applies to all images.

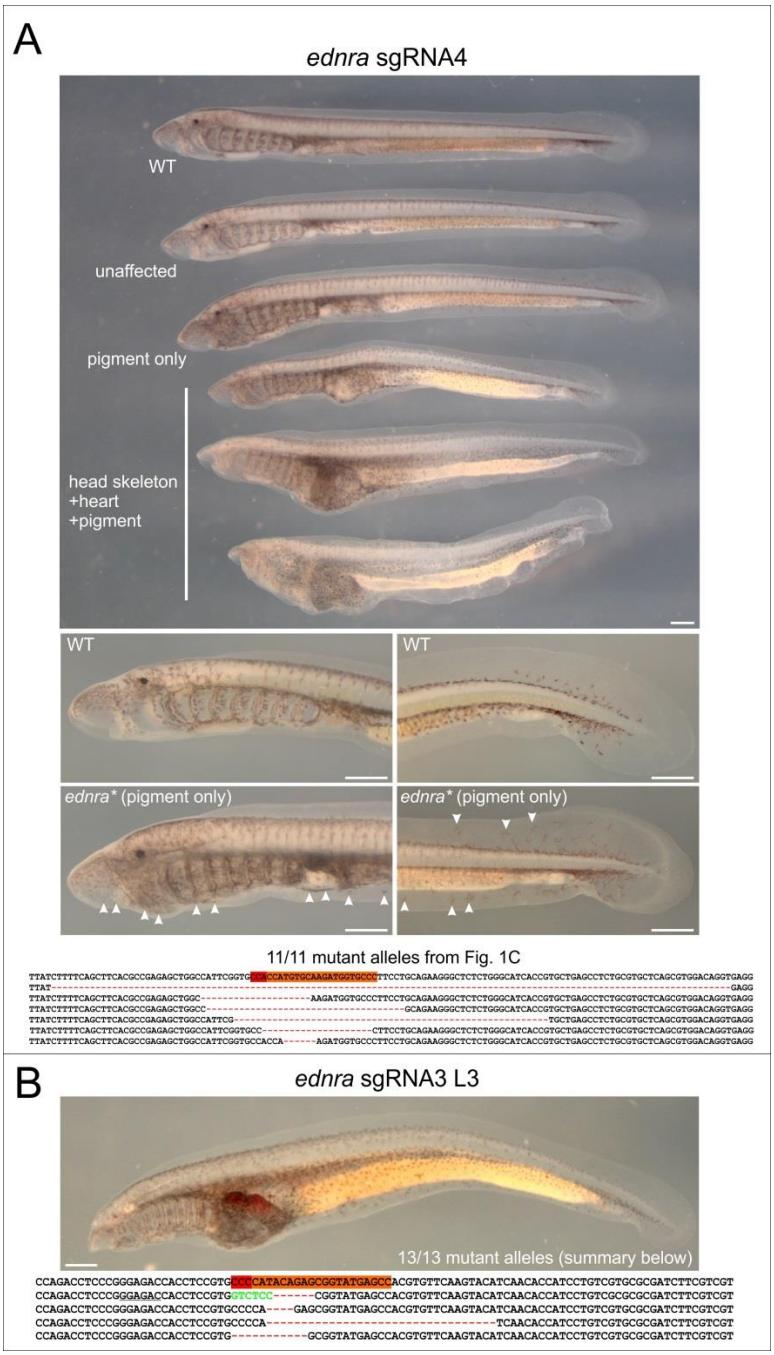

Fig. S3. Gene expression in *ednra* sgRNA injected lampreys. Anterior to left in all panels. Gene is indicated for each panel. All specimens are at st. 26.5 unless indicated otherwise. Top panels show genes mostly unaffected, though occasionally missing from PA1/2 only, bottom shows genes with stronger downregulation throughout the pharynx or heart. White arrowheads indicate WT expression in PA1 (or the heart in *FGFRa*), red arrowheads indicate missing expression in *ednra* mutants. Scale bar in the upper right image represents 100 µm and applies to all images.

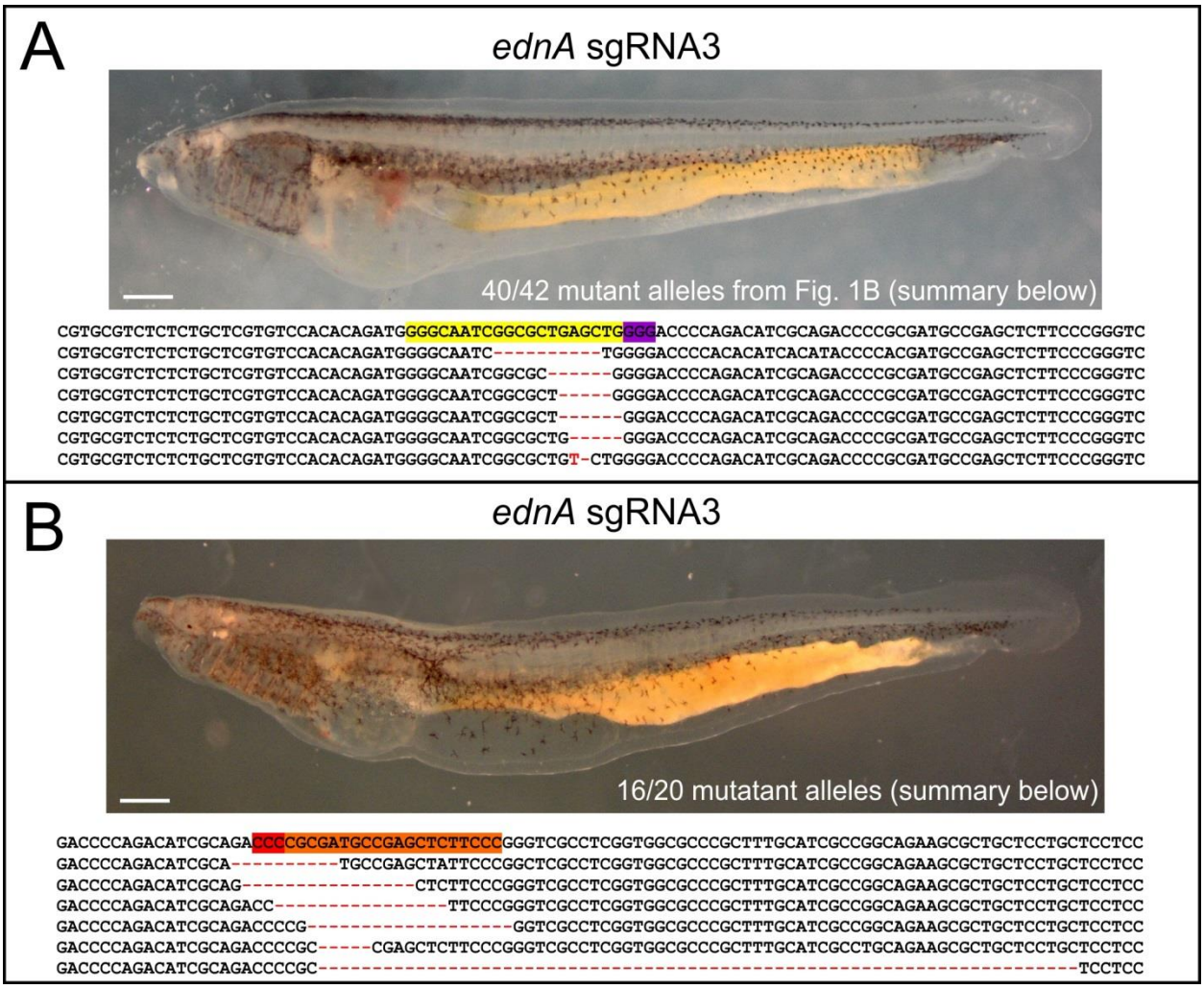

Fig. S5. *P. marinus ednA* genotyping examples. Example individuals and their loci shown from both *ednA* sgRNAs. Scale bars represent 500 µm. A maximum of 6 alleles are shown per individual. Forward target sites are shown in yellow with a purple PAM, while reverse target sites are shown in orange with a red PAM.

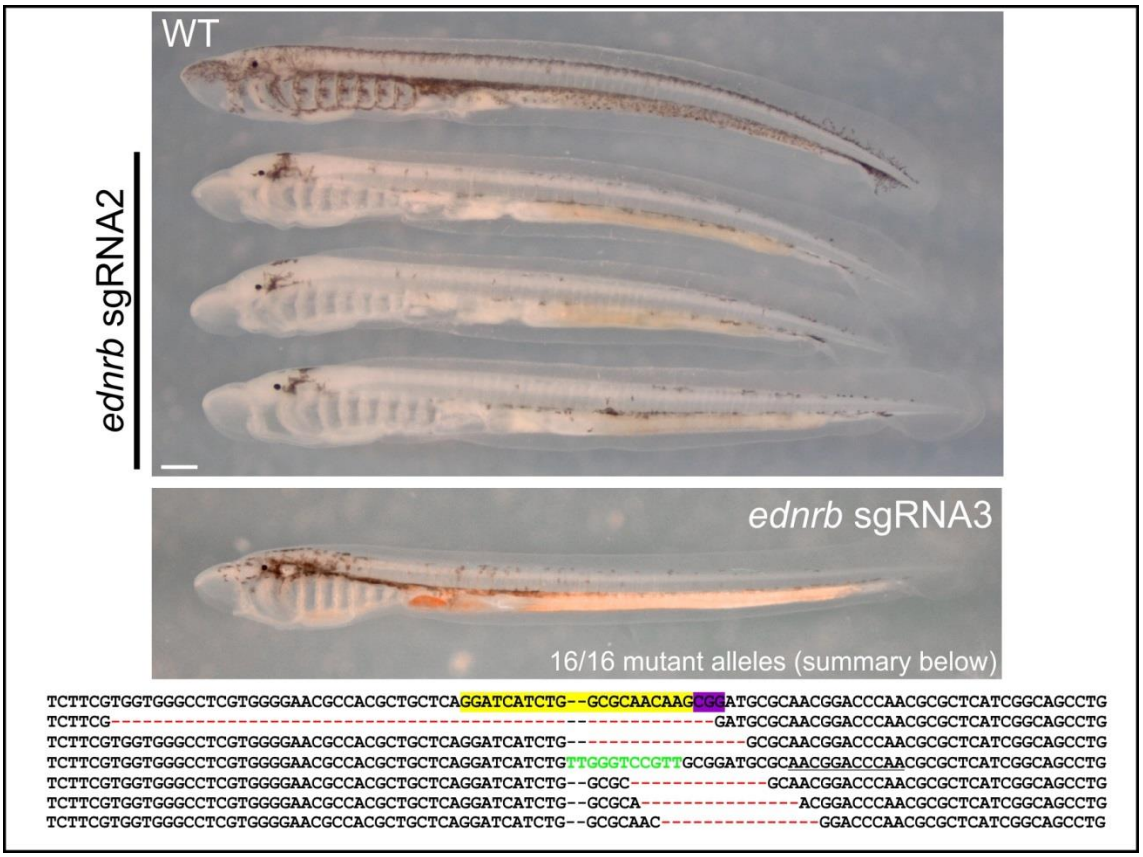

Fig. S6. *P. marinus ednrb* phenotypes and genotyping. Left lateral images of affected *ednrb* mutants. Scale bar in the first image represents 500 µm and applies to all images. 100% mutant alleles were returned for the indicated individual. Target site for sgRNA3 is shown in yellow with a purple PAM. Six example alleles are shown. An insertion from the reverse strand is shown in green, its source is underlined.

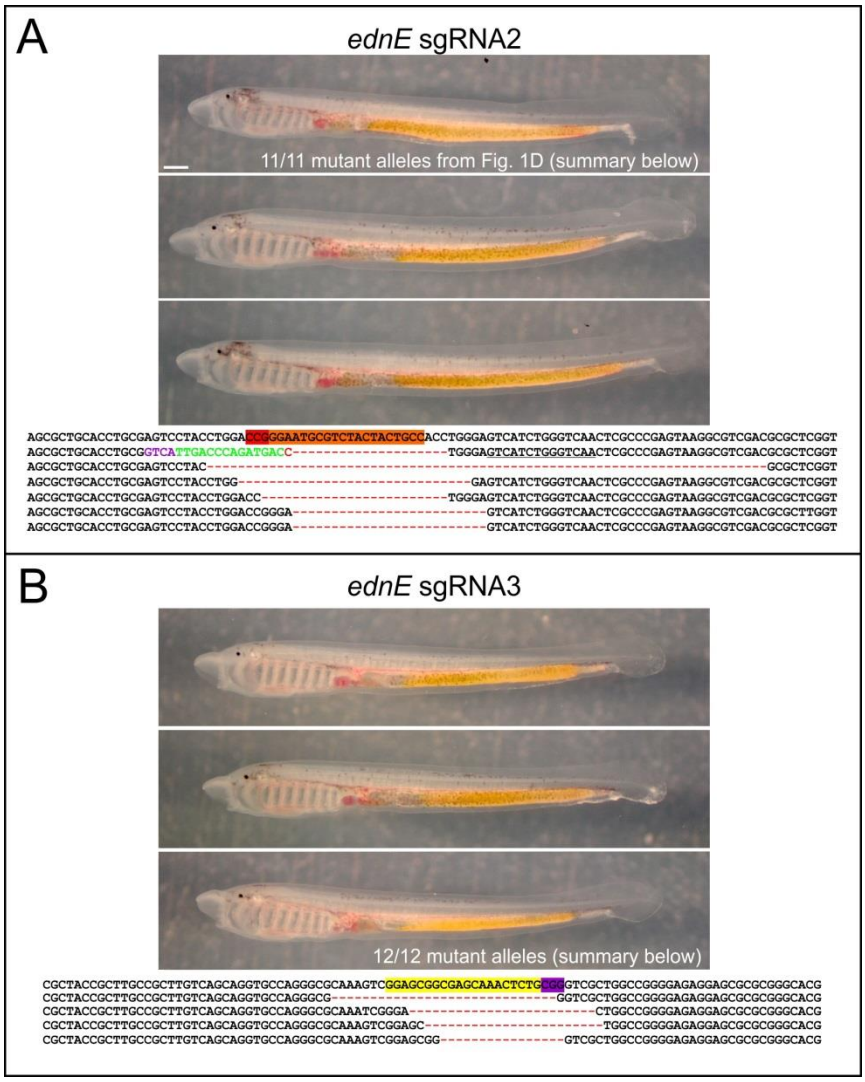

Fig. S7. *P. marinus ednE* genotyping examples. Left lateral images of genotyped individuals from *ednE* mutagenesis experiments. Scale bar in the first image represents 500 µm and applies to all images. 100% mutant alleles were returned for the indicated individuals. (A) Reverse target site is shown in orange with a red PAM. Six example alleles are shown form the indicated individual. An insertion from the reverse strand is shown in green, while an insertion from the forward strand is shown in purple; their sources are underlined. (B) Target site shown in yellow with a purple PAM. Only four types of alleles were returned from the indicated individual.

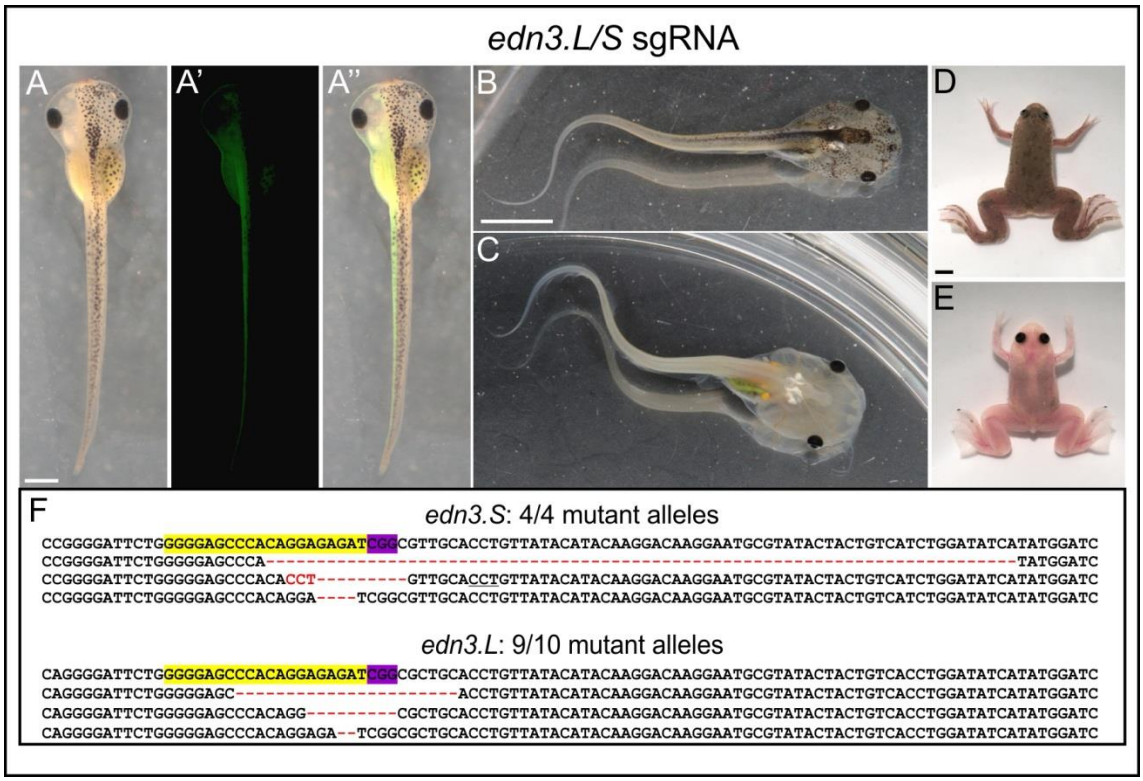

Fig. S8. *edn3.L/S* mutant *X. laevis.* (A) shows a dorsal view of a st. 41 specimen with anterior to the top. A' shows the GFP channel, A'' shows merged panels A and A'. Note unilateral pigment migration defect. Scale bar in A represents 1 mm and applies to A' and A''. (B and C) WT and *edn3* mutant *X. laevis* at st. 50. Dorsal views with anterior to the right. Note that both melanocytes throughout the body and iridophores that normally cover the gut and kidneys (green and yellow in color, respectively) are completely missing. Scale bar in B represents 5 mm and also applies to C. (D and E) WT and *edn3* mutant *X. laevis* juveniles. Scale bar in D represents 5 mm and also applies to E. (F) Genotyping of a single leucistic larva revealed 93% indel alleles. Only six alleles are shown.

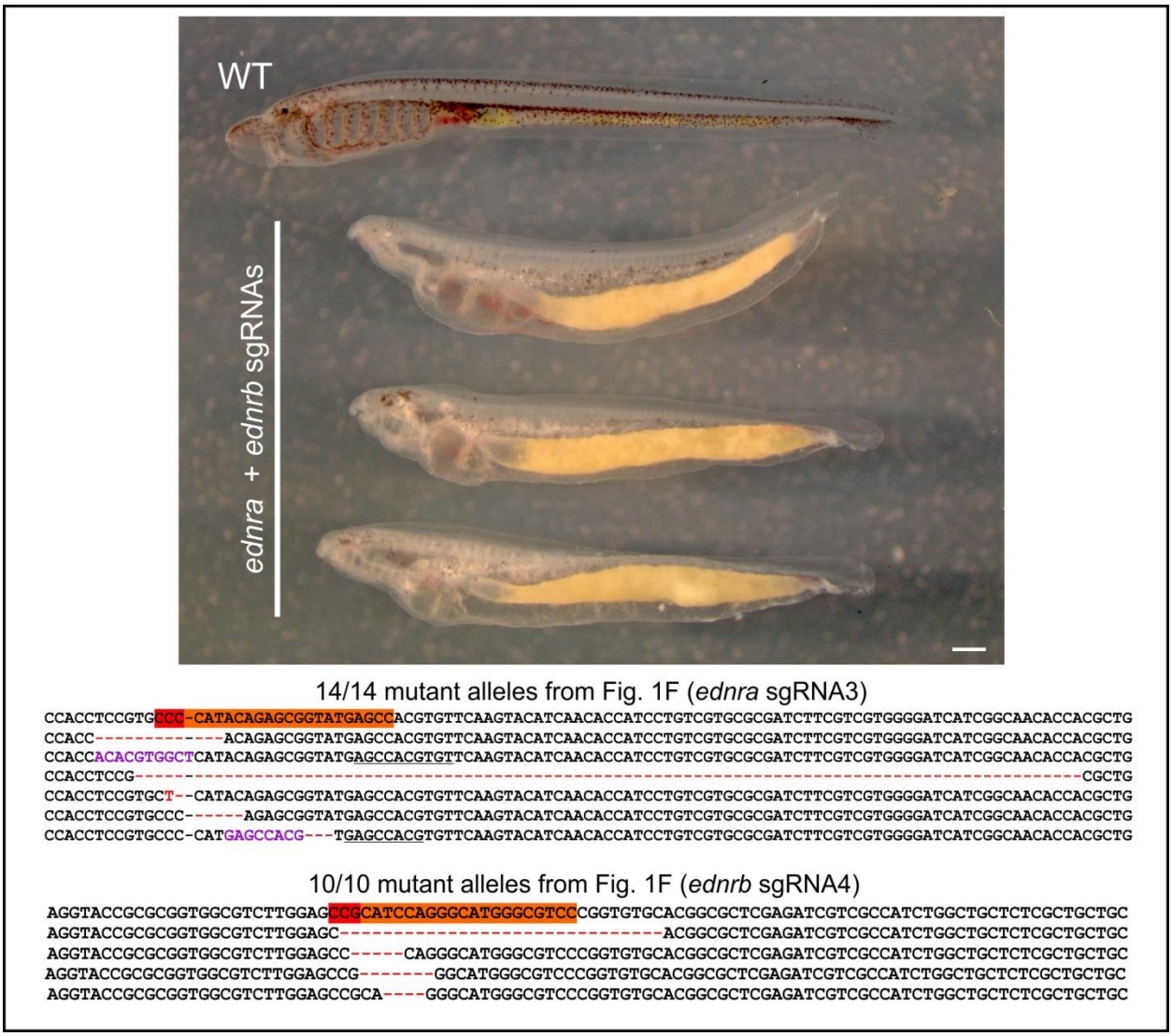

Fig. S9. *P. marinus ednra* + *ednrb* genotyping examples. Left lateral images of affected *ednra* + *ednrb* mutants. Scale bar represents 500 µm. 100% mutant alleles were returned for the individual from Fig. 1F. Target sites are shown in orange with a red PAM. For *ednra,* six example alleles are shown. Insertions from the forward strand are shown in purple, their sources are underlined. For *ednrb*, only four types of alleles were detected.

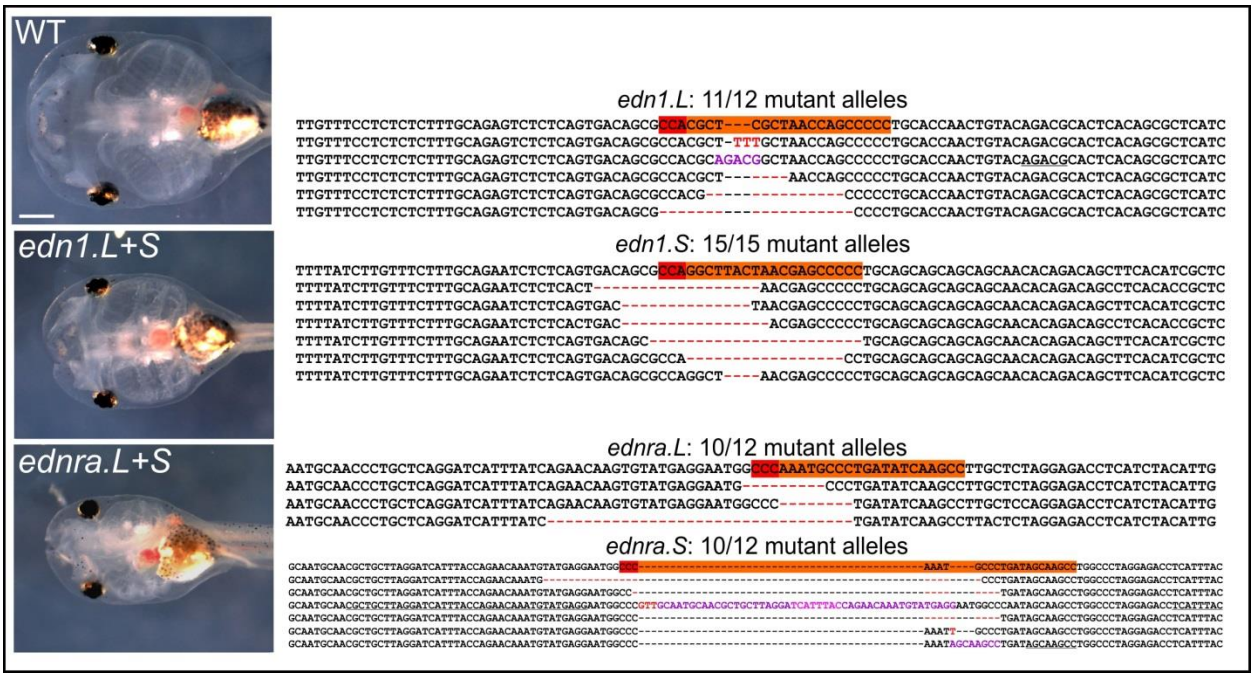

Fig. S10. *X. laevis edn1.L+S* and *ednra.L+S* genotyping examples. Ventral views with anterior to the top. Scale bar in the first image represents 500 µm and applies to all images. Insertions from the forward strand are shown in purple or pink, their source is underlined. Top: an uninjected WT sibling individual. Middle: an *edn1.L+S* mutant specimen returned 96% mutant alleles. All five types of alleles we detected are shown for *edn1.L*; six example alleles are shown for *edn1.S*. Bottom: an *ednra.L+S* mutant specimen returned 83% mutant alleles. Only four types of alleles were detected for *ednra.L*; six example alleles are shown for *ednra.S*.

CHAPTER VI: Table S1

organism	target gene(s)	target site	proportion with a strong phenotype	genotyping primers (forward, reverse)
P. marinus			(Cas9 protein counts only)	
	ednA	(g1)GGGCAATCGGCGCTGAGCTG(GGG)	22/67 (~33%)	GCACTCACAGACTCGCAC, GAGATCCGTGGGACTCAC
	ednA	(g3)(CCC)CGCGATGCCGAGCTCTTCCC	29/38 (~76%)	GCACTCACAGACTCGCAC, GAGATCCGTGGGACTCAC
	ednE	(g2)(CCG)GGAATGCGTCTACTACTGCC	70/73 (~96%)	CGACCCCCTGTATACCAG,, TAACAAGCGATCACCCTGAC
	ednE	(g3)GGAGCGGCGAGCAAACTCTG(CGG)	38/51 (~74%)	CGACCCCCTGTATACCAG, TAACAAGCGATCACCCTGAC
	ednra	(g3)(CCC)CATACAGAGCGGTATGAGCC	113/154 (~73%)	TGAACGGCCAAGGAAACCGAC, GCCCGTTGCGCATGCAC
	ednra	(g4)(CCA)CCATGTGCAAGATGGTGCCC	264/325 (~81%)	GGCCCCAACTCATTTCCTC, CCACTACGCCTCTCCAG
	ednrb	(g2)GGTACCGCGCGGTGGCGTCT(TGG)	40/52 (~77%)	n/a
	ednrb	(g3)GGATCATCTGGCGCAACAAG(CGG)	177/403 (~44%)	AAGGCGGTGAACACGGCGGTG, GACACGGGCAGCGCGATGACG
	ednrb	(g4)(CCG)CATCCAGGGCATGGGCGTCC	n/a	n/a
	ednra+ednrb	-ra SG4 (above) + -rb SG3 (above)	33/98 (~34%)	same as above for each target
	ednra+ednrb	-ra SG4 (above) + -rb SG2 (above)	54/88 (~61%)	same as above for -ra, -rb n/a
	ednra+ednrb	-ra SG4 (above) + -rb SG4 (above)	n/a	same as above for each target
	negative control	cGgGacgcAG/CTCCTGGCCG(AGG)	0/102 (0%)	n/a
X. laevis				
	edn1.L(neg. control)	(CCA)CGCTCGCTAACCAGCCCCCt	0/33 (0%)	AGACTGCTGCTATGAGCTG, ATGATGTCCAAGTGGCAG
	edn1.S (neg. control)	(CCA)GGCTTACTAACGAGCCCCCt	0/57 (0%)	TGTCATGCCAGATTGCTG, ATGATGTCCAAGTGGCAG
	edn1.L+S	(both above combined; see methods)	20/47 (~42%)	(both above)
	edn3.L/S	GGGGAGCCCACAGGAGAGAT(CGG)	31/71 (~44%)	TGTTTTCCCCAAACTCCAG, TTACTCTGGGGTGTTGATCC
	ednra.L(neg. control)	(CCC)AAATGCCCTGATATCAAGCC	0/56 (0%)	TTGTGCCCAGAAAGGAC, CTGTATCTTTGTGACTTTATAGC
	ednra.S (neg. control)	(CCC)AAATGCCCTGATAGCAAGCC	0/72 (0%)	AATCGGACTGATGCCAG, TGACTTCATACAGTAGCTTACTG
	ednra.L+S	(both above combined; see methods)	37/71 (~52%)	(both above)

CHAPTER VI: Table S2

Petromyzon marinus					
			# affected	n	notes
	*ednra**	*alx* (st. 26.5)	4	6	*4/6 missing it from PA1/2, none missing from BAs
		dlxA (st. 26.5)	7	21	
		dlxB (st. 24)	4	6	
		dlxB (st. 26.5)	2	5	
		dlxD (st. 26.5)	5	14	
		ednra (st. 26.5)	0	3	
		ednrb (st. 26.5)	8	10	*missing from skeletogenic NCCs
		ets1d (st. 26.5)	0	9	
		FGFRa (st. 26.5)	3	6	
		FGFRb (st. 26.5)	5	7	
		foxD-A (st. 26.5)	0	3	
		grem (st. 26.5)	0	4	
		hand (st. 26.5)	5	8	*only missing from lower lip, not mucocartilage ventral to the BAs
		ID (st. 26.5)	0	5	
		jagged (st. 26.5)	0	3	
		lec1 (st. 26.5)	6	8	
		lec2 (st. 26.5)	16	16	
		mef2 (st. 26.5)	0	4	*0/4 missing expression from formed pouches, though some specimens were lacking posteriormost pouches altogether
		msxB (st. 26.5)	0	12	
		myc (st. 26.5)	0	6	
		prdm (st. 26.5)	0	4	
		prrx (st. 26.5)	0	4	
		RAR1 (st. 26.5)	0	3	
		soxE1 (st. 26.5)	0	4	
		soxE2 (st. 23)	0	20	
		soxE2 (st. 26.5)	15	21	
		soxE3 (st. 26.5)	0	4	
		twistA (st. 26.5)	0	6	
	*ednrb**	*dlxA* (st. 26.5)	0	7	
		dlxB (st. 26.5)	0	5	
		dlxD (st. 26.5)	0	7	
		grem (st. 26.5)	0	8	
		jag (st. 26.5)	0	8	
		phox (st. 26.5)	0	12	
		soxb1b (st. 26.5)	0	8	

	ednra+b**	*dlxA* (st. 26.5)	20	24	
		dlxB (st. 26.5)	4	6	
		dlxC (st. 26.5)	4	7	
		dlxD (st. 26.5)	7	8	
		hand (st. 26.5)	5	9	*only missing from lower lip, not mucocartilage ventral to the BAs
		msxB (st. 26.5)	0	8	
		soxE2 (st. 26.5)	9	10	
		soxE3 (st. 26.5)	0	3	
Xenopus laevis					
	ednra.S+L**	*dlx3.S* (st. 33)	3	7	
		hand2.L (st. 33)	4	7	
		msx1.S (st. 33)	0	15	
		sox8.L (st. 25)	0	20	
		sox9.S (st. 33)	12	20	
		sox9.S (st. 25)	0	15	