Published in final edited form as:

Nature. 2019 October ; 574(7780): 675–678. doi:10.1038/s41586-019-1691-4.

Evolution of the New Head by gradual acquisition of neural crest regulatory circuits

Megan L. Martik[1], **Shashank Gandhi**[1], **Benjamin R. Uy**[1], **J. Andrew Gillis**[2,3], **Stephen A. Green**[1], **Marcos Simoes-Costa**[4], **Marianne E. Bronner**[1,*]

[1]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

[2]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, United Kingdom

[3]Marine Biological Laboratory, Woods Hole, MA, 02543, U.S.A.

[4]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

Summary:

The neural crest is a vertebrate innovation proposed to be a key component of the "New Head" that imbued vertebrates with predatory behavior. To address how evolution of this cell type impacted the vertebrate body plan, we examined the molecular circuits that control neural crest development along the anteroposterior axis of a jawless vertebrate, the sea lamprey. Gene expression analysis showed that the lamprey cranial neural crest lacks most components of an amniote cranial-specific transcriptional circuit that confers the ability to form craniofacial cartilage onto other neural crest populations[1]. Consistent with this, hierarchical clustering revealed that the transcriptional profile of the lamprey cranial crest is more similar to the amniote trunk crest. Intriguingly, analysis of the cranial neural crest in little skate and zebrafish embryos demonstrated that the cranial-specific transcriptional circuit emerged via gradual addition of network components to the neural crest of gnathostomes, which subsequently became restricted to the cephalic region. Our results indicate that the ancestral neural crest at the base of vertebrates possessed a trunk-like identity. We propose that the emergence of the cranial neural crest, by progressive assembly of a novel axial-specific regulatory circuit, allowed for the elaboration of the New Head during vertebrate evolution.

Gans and Northcutt's "New Head" hypothesis proposed that emergence of the vertebrate lineage was accompanied by advent of the neural crest (NC), an embryonic stem cell population that arises within the forming central nervous system (CNS) in all vertebrates[2,3].

---

*Corresponding author. Correspondence and request for materials should be addressed to mbronner@caltech.edu.

Author Contributions:

Project and analysis conception were designed by M.L.M., M.S.C., and M.E.B. Writing and interpretation was performed by M.L.M., S.G., B.R.U., J.A.G., S.A.G., M.S.C., M.E.B. Lamprey orthologue cloning and all *in situ* hybridization, imaging, and analysis was performed by M.L.M. Bioinformatics and chicken RNAseq was performed by S.G. Phylogenetic analysis and lamprey embryo acquisition was performed by S.A.G. Cloning of skate orthologues and skate embryo acquisition was performed by J.A.G. Lamprey embryo dissections and library preparations were performed by B.R.U. and M.S.C.

The authors declare no competing financial interests.

These cells subsequently leave the CNS, migrate to diverse locations and differentiate into many derivatives including peripheral ganglia and craniofacial skeleton[4,5]. As vertebrates evolved, NC cells contributed to morphological novelties like jaws, that enabled expansion of vertebrates.

A pan-vertebrate NC gene regulatory network (GRN), invoking sequential deployment of signaling and transcriptional events, has been proposed to underlie formation of this unique cell type. Primarily studied at cranial levels, the core of the NC GRN is largely conserved across vertebrates, including the sea lamprey, *Petromyzon marinus,* a jawless (cyclostome) vertebrate. However, differences exist in utilization of key transcription factors, like *Ets1* and *Twist,* which are deployed later in the lamprey GRN than in amniotes[6,7], suggesting regulatory differences between cyclostomes and gnathostomes. Furthermore, some NC derivatives are novelties of gnathostomes, such as jaws at cranial levels, a vagal-derived enteric nervous system, and sympathetic ganglia at trunk levels[8,9]. This raises the intriguing possibility that network differences in axial regionalization of the neural crest may have contributed to the presence of these gnathostome cell types.

In jawed vertebrates, the NC is subdivided along the body axis into cranial, vagal, and trunk populations. In contrast, lamprey lack an intermediate vagal population, suggesting there are two major subdivisions: cranial and trunk[8,5]. How axial identity in lamprey is controlled molecularly remains unknown. Avian embryos possess a “cranial crest-specific” NC GRN subcircuit with ability to drive differentiation of trunk NC into ectomesenchymal derivatives[1]. In this kernel, transcription factors *Brn3c, Lhx5,* and *Dmbx1* are expressed at the neural plate border and, in turn, activate expression of *Ets1* and *Sox8* in premigratory cranial NC (Fig 1A). In contrast to their cranial-specific expression, *Tfap2b* and *Sox10* are pan-NC genes expressed all along the body axis[10].

Here, we assessed whether this cranial subcircuit is a general feature of vertebrates by examining whether lamprey possess a homologous spatiotemporal regulatory state. Taking a candidate approach, we analyzed expression of cranial circuit orthologues in lamprey embryos at different developmental stages. In contrast to amniotes, our results show that *Brn3, Lhx5, Dmbx1,* and *Ets1* appear to be absent from lamprey premigratory or migratory NC (Fig 1B). The lack of most cranial-specific regulatory factors suggests a high degree of divergence between early regulatory states of lamprey and amniote NC. In contrast, lamprey *SoxE1* and *Tfap2* were robustly expressed in premigratory and migratory NC along the entire body axis (Fig 1B, SupFig1A-D). No *SoxE* family member was restricted in expression to the cranial NC as is Sox8 in amniotes. Of note, lamprey SoxE transcription factors are homologous to gnathostome Sox8/9/10, and there is variation in SoxE paralog usage across gnathostomes[11–13]. Consistent with the lack of restricted ‘cranial-specific’ expression, ectomesenchymal derivatives have been previously reported as present at trunk levels in the lamprey dorsal fin[9].

How then did this regulatory subcircuit evolve? Interestingly, genes from the cranial crest subcircuit are present in the genome and expressed later in pharyngeal arches populated by NC cells (Fig 1C,D, SupFig1E-L, SupFig6). An intriguing possibility is that these genes were expressed only in late NC derivatives of early vertebrates, followed by gradual co-

option of components of this regulatory program to earlier developmental stages in gnathostomes. According to this scenario, genes involved in NC differentiation in early vertebrates were co-opted to the specification program of gnathostomes at all axial levels. With subsequent regulatory modifications, they became cranially restricted, possibly endowing the cranial NC with novel morphogenetic features while the trunk NC lost ability to make cranial-like derivatives[14,15].

To explore this possibility, we examined candidate elements of the cranial NC subcircuit in the little skate *Leucoraja erinacea,* a Chondricthyan gnathostome outgroup to the bony fishes. Of the *SoxE* genes, expression of *Sox9* and *Sox10,* as well as *Tfap2b* and *Ets1,* were present at all axial levels and not restricted to cranial crest (Fig 2A, SupFig2). Since *Ets1* appears in the little skate migratory NC as in other gnathostomes, we conclude that this early node was a novelty acquired by the cranial NC GRN prior to divergence of cartilaginous and bony fishes (Fig 2A,B, SupFig2). Later, after NC cells migrate to and populate the pharyngeal arches, *SoxE, Tfap2b,* and *Ets1* were present within the arches (SupFig3). Trunk NC in the little skate produce ectomesenchymal dermal denticles, “cranial-like” derivatives, consistent with our observation that cranial subcircuit genes in the little skate are not restricted to the head but can drive differentiation of skeletogenic derivatives in the trunk[16]. In the fossil record, many stem-gnathostomes possessed extensive dermal armour, which has been retained, albeit with the dental component reduced and modified, within the gnathostome crown group (e.g. dermal denticles of chondrichthyans; dentinous scales of *Polypterus* and coelacanth). Thus, dental tissues in the post cranial dermal skeleton appears to be ancestral for gnathostomes.

Interestingly, in the teleost *Danio rerio, lhx5* and *dmbx1* are present in the early cranial circuit but absent from later pharyngeal arch derivatives (Fig 2C,D, SupFig4, SupFig5). In addition, *sox8b, sox10, tfap2a,* and *ets1* are present in premigratory and migratory crest at all axial levels, though *brn3c* is missing (Fig 2C, SupFig4). Rather than restricted to the cranial NC, many of these factors also are present in the zebrafish trunk, raising the possibility that resolution of axial level potential may have arisen within sarcopterygians. Furthermore, *in situ* analysis of pharyngeal arch derivatives in both little skate and zebrafish lend support to temporal shifts of cranial specific regulatory nodes from later NC derivatives to an early specification program. With progressive loss of nodes from late derivatives and addition to an earlier program, this suggests that regulatory modifications arose gradually throughout gnathostome evolution (SupFig 1M).

Our candidate gene approach suggests that extensive changes occurred in the NC regulatory state between jawless and jawed vertebrates. To investigate further, we conducted a comparative transcriptome analysis of cranial and trunk NC subpopulations in lamprey and chicken (Fig 3). Premigratory lamprey NC was obtained by micro-dissecting segments of cranial and trunk dorsal neural tubes at stages T21 and T23.5. For chick, premigratory NC populations were isolated using enhancers driving eGFP expression in cranial or trunk neural crest populations for FACS at stages HH9+ and HH18, respectively. After cDNA library preparation and sequencing, differential expression analysis revealed far fewer genes (1233 genes in lamprey compared with 2794 in chicken) significantly enriched in lamprey cranial versus trunk crest compared with chick cranial versus trunk (Fig 3A, 3B).

To better understand how each library correlated to the others in an unbiased fashion, we mapped each to a common reference transcriptome, created by aligning proteomes using BLAT and compiling matching sequences as a consensus alignment between species for Bowtie mapping. We next performed hierarchical clustering analysis of all known NC GRN genes (Fig 3C). Consistent with our previous *in situ* hybridization analysis (Fig 1B), *Tfap2* was enriched at all axial levels in both chicken and lamprey (Fig 3C'); *Sox8* was enriched in chicken cranial NC but was in both cranial and trunk lamprey populations (Fig 3C"). *Dmbx1* and *Ets1* were enriched in chick but not lamprey cranial datasets (Fig 3C"').

Interestingly, we found that lamprey cranial populations correlated more closely to chicken trunk than lamprey trunk libraries, suggesting that basal NC was "trunk-like" in its regulatory program (Fig 3C). These results suggest that cyclostomes possess a simpler and more trunk-like cranial crest GRN, with potentially important implications for evolution of NC subpopulations (Fig 4A). Accordingly, we speculate that the ancestral neural crest may have been relatively homogeneous and trunk-like. Throughout evolution of the vertebrate lineage, we propose that key transcription factors were progressively co-opted into an early, cranial-restricted circuit, whereas some features like skeletogenic potential were lost from the trunk.

These differences in axial-specific genes contrast with the deep conservation of the pan-NC program[6,17]. Transcription factors like *SoxEs, Tfap2,* and *Id* may be the rudiment of a larger, more complex cranial crest GRN that was expanded during early vertebrate evolution with incorporation of novel players such as *Dmbx1, Brn3c, Ets1,* and *Lhx5.* Consistent with these findings, the basal chordate Amphioxus lacks expression at the neural plate border of genes like *Dmbx, Brn3, Ets,* as well as core NC genes like *SoxE, FoxD, Tfap2,* and *Id,* although these genes are expressed in other tissues[18–21]. Our observations also show that some of these "novel" genes are expressed at later stages of NC formation, consistent with the possibility that elaboration of the GRN might have involved co-option of parts of differentiation programs to earlier portions of the network perhaps by acquisition of new regulatory elements responsible for their heterochronic shift[22]. Thus, the pan-NC program was likely the ancestral molecular recipe to make NC, with the subsequent elaboration of axial-specific regulatory programs conferring important differences in developmental potential along the body axis. Given that many key NC derivatives are gnathostome innovations, we hypothesize that gain of these derivatives may be due to gene regulatory differences associated with axial-specific regulatory programs.

Taken together, our results suggest the following scenario to explain evolution of NC subpopulations (Fig 4). We suggest that NC of early vertebrates was uniform and similar to amniote trunk populations, and that the division of NC into cranial and trunk subpopulations occurred early in vertebrate evolution (Fig 4A). Consistent with evolutionary expansion of NC cells in the vertebrate lineage, our molecular analysis of the cranial NC reveals surprising differences in lamprey compared with gnathostome counterparts (Fig 4B). Given that the Hox code was already linked to segmentation of the CNS in basal vertebrates, posteriorizing influences of Hox genes and other factors may be sufficient to account for the subtle transcriptional differences observed between these two populations[23,24]. We cannot rule out the possibility that cyclostomes lost NC subpopulations during the course of

evolution. However, the relative scarcity of cranial-specific factors in the lamprey cranial crest might suggest that the gnathostome cranial NC GRN has undergone extensive elaboration from a regulatory standpoint. Thus, we propose that regionalization of the NC, with both emergence of new subpopulations and expansion of the cranial crest GRN, played a crucial part in vertebrate evolution as a key element for driving evolution and expansion of gnathostomes.

What does this mean for the 'New Head' hypothesis? We posit that, the NC component of the New Head, rather than arising *in toto* at the base of vertebrates, underwent continued regulatory modifications, evolving gradually during the course of vertebrate evolution. Our data suggest that early vertebrates possessed a relatively simple NC that initially arose as a fairly uniform population along the body axis and lacked region-restricted regulatory programming. During gnathostome evolution, the cranial NC appears to have gained regulatory complexity that modulated differentiation capacity, gaining some individual cell fates while restricting others. We propose that co-option of distinct genes into a cranial-specific module enabled this progressive specialization of NC regulatory programs, leading to unique axial populations and morphological novelties of the gnathostome body plan.

Methods:

Animal husbandry and embryo collection

Adult sea lamprey were obtained from the US Fish and Wildlife Service and Department of the Interior. Embryos were cultured according to previously published protocols and staged according to Tahara staging methods[12,25]. All lamprey embryology work was completed in compliance with California Institute of Technology Institutional Animal Care and Use Committee (IACUC) protocol 1436. Skate eggs were obtained by the Marine Biological Laboratory (MBL) Marine Resource Center, and embryos were cultured as previously described [16]. All skate embryology work was compliant with animal protocols approved by the IACUC at the MBL. Adult zebrafish were maintained in the Beckman Institute Fish Facility at Caltech, and all animal and embryo work was compliant under approved IACUC protocol 1346. Fertilized chicken eggs were obtained from a local farm in Sylmar, CA. No statistical methods were used to predetermine sample size for analyses. For in situ hybridization, embryos were pooled from different breeding pairs (fish), brooding stocks (skates), or embryo batches (lamprey) to ensure replication of results in multiple fixed collections.

Cloning of lamprey, skate, and zebrafish orthologues

RNA was extracted from desired embryo stages using the RNAqueous-Micro Kit (Thermo Fisher Scientific), and cDNA was synthesized using a SuperScript III Reverse Transcriptase Kit (Invitrogen). The following gene specific primers used to amplify probe template sequences (accession numbers in parentheses):

PmTfap2 (MN410935): *F: 5'-GCATCGCGACAGTTGTTTGCTG-3'; R: 5'-GATGCTGTGGTGCCCTAATCC-3'*

PmSoxE2 (MN410934): *F: 5'-CGAGTCACGTGGATCTGCTGC-3'; R: 5'-CCGTCCAGCACTTGACTCACG-3'*

PmSoxE1 (MN410933): *F: 5'-CGGGCTGAGTCATTACTCGCATCG-3'; R: 5'-CTCTCGTCGCTGTCGGAAGC-3'*

PmEts1a (SIMRbase: PMZ-0040201): *F: 5'-GGACCTTCAAGGAGTACATGAGC-3'; R: 5'-GAGAGCGGTACTCGTGGAAAGTC-3'*

PmDmbx1 (Ensembl: ENSPMAG00000008114): *F: 5'-GCGCATGAATACCGGCCGTCG-3'; R: 5'-TTGCTTTGATGCTGTTACAAGG-3'*

PmLhx5 (MN410936): *F: 5'-CGTGCGTTCGTGACCCCATC-3'; R: 5'-GAGGCCAGGTAGTCCTCCTTG-3'*

PmBrn3 (SIMRbase: PMZ-0005302): *F: 5'-CGAGTCTCCTTAACGCGTTAGCTC-3'; R: 5'-GCTCTGGTGGGAGACAATATCCACG-3'*

LeTfap2b (MN410937): *F: 5'-TCCCACTTCCACAGAAGAAT-3'; R: 5'-TCCTTGTCTCCAGTTTTGGTG-3'*

LeSox10 (MN410938): *F: 5'-ACCCCCGTTCTGTGTGTCT-3'; R: 5'-GGCAGGTACTGGTCGAACTC-3'*

LeSox9 (MN410939): *F: 5'-CCCAGCCACTACAATGAGCAG-3'; R: 5'-CCGTACGGCATCAGCAAATG-3'*

LeSox8 (MN410944): *F: 5'-CAACTCCGCCCACCACTCC-3'; R: 5'-TGGCCTAGTCAGGGTTGTGTAG-3'*

LeEts1 (MN410940): *F: 5'-TTCAGCCTGAAGAACGTGGAC-3'; R: 5'-GCAAGACTTGTCCGTCAGGAG-3'*

LeDmbx1 (MN410941): *F: 5'-CAATCAACACGACAGGGACA-3'; R: 5'-GTAAGCTGTCAAGCCCCAGA-3'*

LeLhx5 (MN410942): *F: 5'-TCATCGACGAAAACAAATTTGTGTG-3'; R: 5'-TGAATAACCCGCATGTTGAGGC-3'*

LeBrn3c (MN410943): *F: 5'-CTTCAAGCCGGACATCACCTAC-3'; R: 5'-TAGATCCCTGCTTGTTCCTGC-3'*

Drtfap2a (NM_176859): *F: 5'-GTCACGGCATTGATACTGGACTC-3'; R: 5'-TCATTGGCACACTGCTTTACTGAT-3'*

Drsox10 (NM_131875): *F: 5'-GTGAAACACACTTCCCTGGGGATAC-3'; R: 5'-GTGGAGACATGTGTGTATGGCGTC-3'*

Drsox8b (NM_001025465): *F: 5'-ATGAGCGAGGAGCGGGAAAAGTG-3'; R: 5'-GGGTCTGGACAGAGTGGTGTAGAC-3'*

Drets1 (NM_001017558): *F: 5'-CAGACAGCGGATCTTGTTGAGGGA-3'; R: 5'-CAGTCCAGCTGATGAAGGACTGG-3'*

Drdmbx1a (NM_152977): *F: 5'-CGTGCCAGTCCTACTATCAGTCTC-3'; R: 5'-CTGCTGTGTAGTGCATGCAACC-3'*

Drlhx5 (NM_131218): *F: 5'-CACGGACATGATATCCCATGCAGAC-3'; R: 5'-CTAGCTCACTTCTGACCATCAGATGC-3'*

Drbrn3c (NM_131278): *F: 5'-ATGATGACCATGAACGGCAAGC-3'; R: 5'-GTGCACTGCTGAATACTTCATCC-3'*

Phylogenetic analysis of Dmbx proteins

Candidate Dmbx sequences were assembled as an ungapped Fasta file and imported into the TCoffee server (http://tcoffee.crg.cat) and processed using default parameters in an Expresso (http://tcoffee.crg.cat/apps/tcoffee/do:expresso) into a protein alignment [26,27,28,29,30]. The TCoffee fasta alignment was imported into MegAlign Pro (DNAstar ver 15.0.0) and ambiguous regions with poor alignment scores were removed, leaving only large, contiguous regions of well-aligned sequence. This alignment of 218 amino acid resides was exported as nexus format. The start of the *P. marinus* sequence is missing, and so residues were recoded from gaps to indicate missing sequence. The file was modified to include a MrBayes block, with aamodelpr=mixed, stopval=0.01, ngen=200000, and burninfrac=0.25. The file was executed within MrBayes3.2.1, and resulting consensus tree visualized in FigTree v1.4.2 to show posterior probabilities (as %) at corresponding branch labels. Image output files from Megalign Pro and FigTree v1.4.2 were combined in Adobe Illustrator 2019 (Adobe Creative Suite 2019) (Supplemental Figure 6). NCBI accession numbers or Ensembl identifiers for Dmbx sequences used in phylogenetic analyses are as follows:

XP_003725762.1 (*S. purpuratus)*

NP_001161526.1 (*S. kowalevskii)*

AAT66431.1 (*B. floridae)*

Ensembl ENSPMAG00000008114 (*P. marinus)*

XP_020369662.1 (*R. typus)*

AAI34895.1 (Dmbx1a, *D. rerio)*

NP_001017625.1 (Dmbx1b, *D. rerio)*

XP_017949066.1 (*X. tropicalis)*

XP_001234036.2 (*G. gallus)*

NP_671725.1 (*H. sapiens)*

In situ hybridization of lamprey, skate, and zebrafish embryos, Sectioning, Imaging, and Biotapestry modeling

Whole mount *in situ* hybridization was was performed using previously published protocols[8, 16, 31,32]. Cryosections of lamprey, skate, or zebrafish embryo *in situs* were sectioned at 18 μm with a *Microm* HM550 *cryostat. In situ* analysis of S25 skate embryos sections was performed using paraffin sections as follows: After fixation, embryos were embedded in paraffin and sections were prepared at 5 (skate) or 10μm (lamprey) thickness on a Zeiss microtome. After paraffin removal with histosol, sections were hybridized with 1ng/μl anti-sense digoxygenin-labelled probes overnight at 70°C in a humidifying chamber. After hybridization, sections were washed with 50% formamide/50% 1X SSCT buffer followed by washes with MABT and a blocking step in 1% Roche blocking reagent. Sections were then incubated overnight at room temperature with a 1:2000 dilution of anti-DIG-Alkaline Phosphatase antibody (Roche). After several washes with MABT, chromogenic color was developed using NBT/BCIP precipitation (Roche). Imaging was performed on a Zeiss AxioImager.M2 equipped with an Apotome.2. Gene network models were assembled using Biotapestry[31].

Chicken embryo electroporation, dissociated, and cell sorting

Cranial and trunk neural crest cells were labeled using previously published neural crest enhancers FoxD3-NC1.1 and FoxD3-NC2, respectively[32]. To isolate cranial neural crest cells, stage Hamilton-Hamburger (HH) 4 embryos were bilaterally electroporated with FoxD3-NC1.1>eGFP and cultured *ex ovo* until stage HH9+ [33]. For each biological replicate, at least 15 embryo heads were dissected in Ringers and washed thrice in chilled 1x PBS. For trunk neural crest cells, stage HH10 embryos were bilaterally electroporated with FoxD3-NC2>eGFP and cultured *in ovo* until stage HH18. Based on the expression of the reporter, five embryo trunks spanning the length of five somites were dissected in Ringers and washed thrice in chilled 1x PBS. The tissues were dissociated in Accumax (Innovative Cell Technologies, Inc.) for 15 minutes at 37°C and GFP+ cells were collected using Fluorescence Activated Cell Sorting.

Library preparation and sequencing

Chicken libraries were prepared using the SMART-Seq Ultra Low Input RNA Kit (Takara) according to the manufacturer's protocol. For lamprey embryos, tissue was dissected from the cranial dorsal neural tubes of n=100 T21 and trunk neural tube of n=100 T23.5 embryos. Total RNA was extracted using the RNAqueous kit (Ambion). RNA-Seq was performed at the Millard and Muriel Jacobs Genetics and Genomics Laboratory (California Institute of Technology, Pasadena, CA) at 50 million reads on 2 biological replicates for both the T21 cranial and T23.5 trunk neural tube samples. Sequencing libraries were built according to Illumina Standard Protocols. SR50 sequencing was performed in a HiSeq Illumina machine. Databases have been deposited to NCBI (BioProject # PRJNA497902)

Statistical analysis of lamprey and chicken axial population RNAseq

To identify orthologous genes between lamprey and chicken, the lamprey proteome obtained from SIMRbase[34] was aligned to the chicken proteome using the BLAT alignment software

available on the UCSC genome browser [35, 36]. Briefly, every lamprey protein sequence was locally queried against the chicken proteome, following which regions with the longest alignment were matched to the respective chicken gene. Using this alignment-based approach, proteins with an alignment percentage score between 52-100 (see Supplementary table for exact scores for each orthologue) were identified as orthologues, and their respective cDNA sequences were obtained from the chicken and lamprey databases. Chicken cranial and trunk libraries were aligned to the chicken sequences, while the lamprey cranial and trunk libraries were aligned to the lamprey sequences using Bowtie2 [37]. Transcript counts were calculated using HTSeq-Count and differential gene expression analysis was performed using DESeq2 [38,39]. Using chicken gene annotations as a reference, we added the transcript counts for duplicated orthologues found in the lamprey genome to calculate an "aggregated" transcript count for each gene. These aggregated transcript counts were then normalized using the formula:

$$Z_i = \frac{T_i - \min(T)}{\max(T) - \min(T)}$$

where Z_i– Normalized transcript count

T_i– Absolute transcript count

A subset of genes previously identified as being part of the neural crest gene regulatory network [15] was then isolated from the count matrix and plotted as a heatmap to obtain the gene expression matrix.

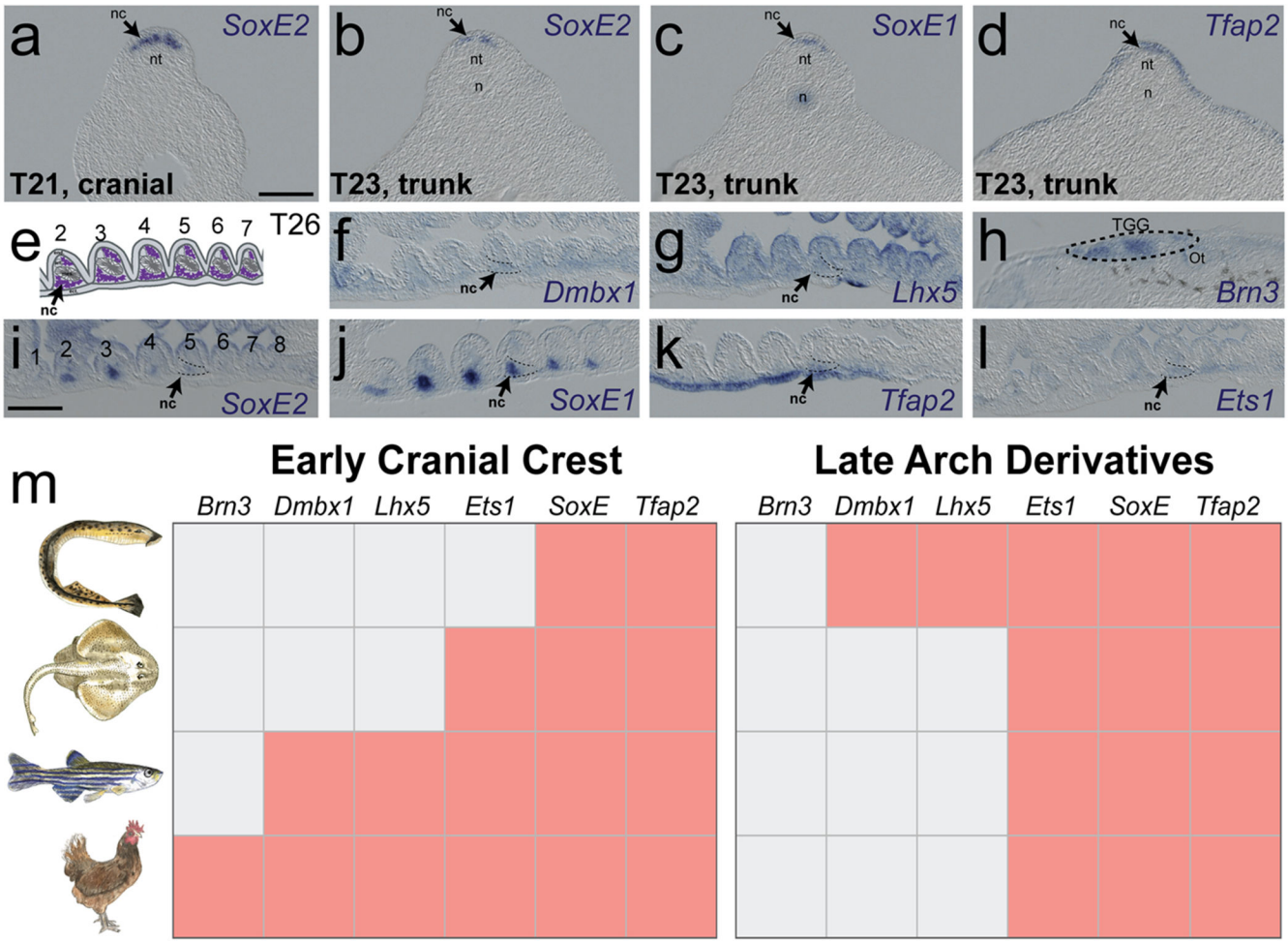
Data availability

All raw sequencing data for all RNAseq libraries (Figure 3) and merged reference transcriptomes are available online (NCBI BioProject# PRJNA497902). Sequences of *in situ* probe templates for Figures 1B, 1C, 2A, and 2C are available through GenBank accession codes found in the methods.

Code Availability

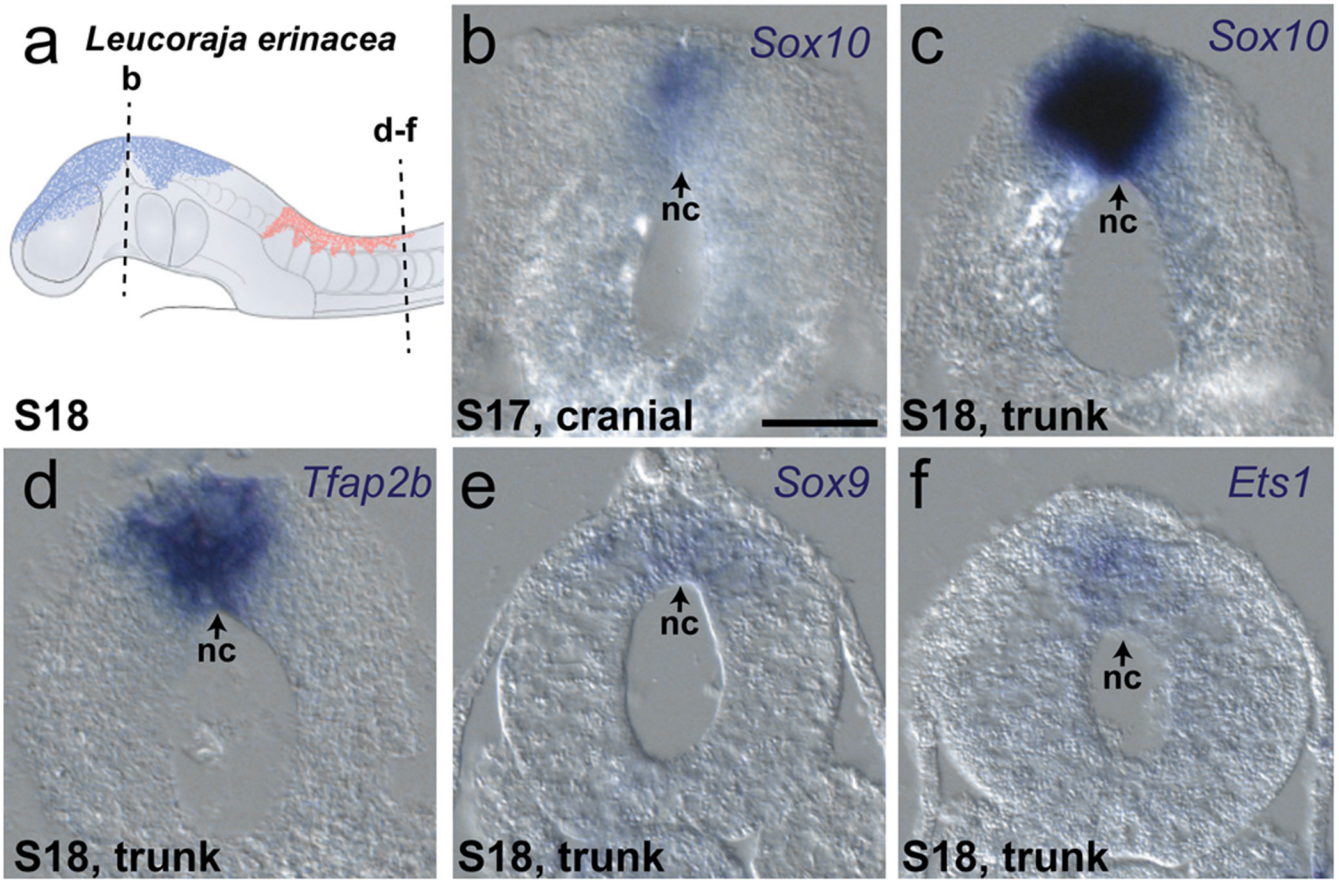
Code used to analyze sequencing datasets are available from the corresponding author upon request.

Extended Data



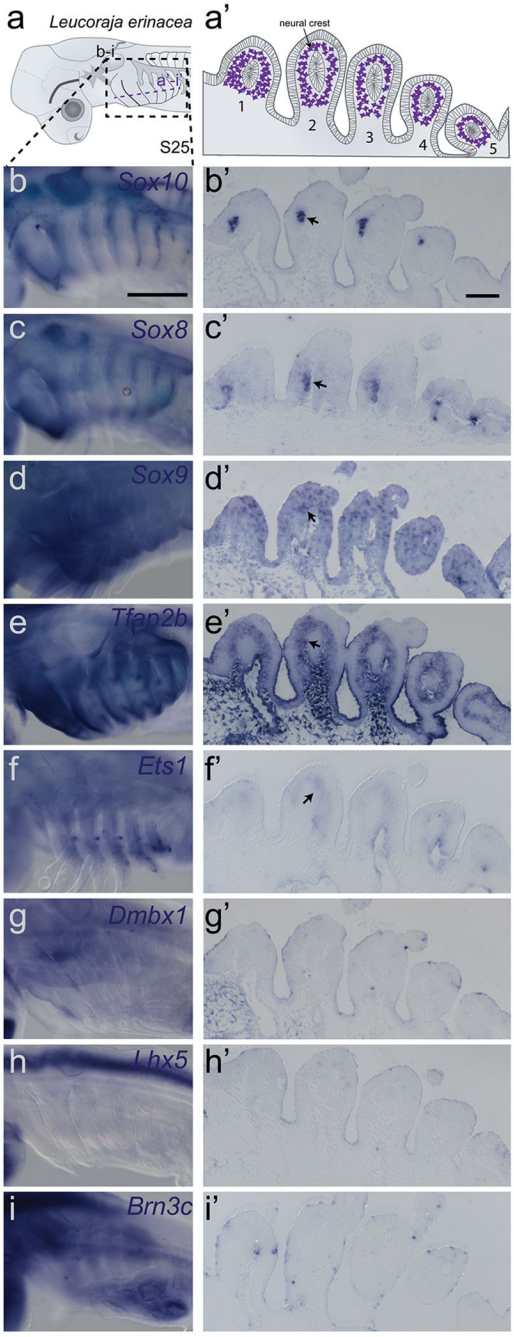
Extended Data 1. Heterochronic shifts of cranial specific gene regulatory nodes from later neural crest derivatives to an early specification program happened gradually throughout gnathostome evolution.

a-d.) Expression of lamprey orthologues of amniote cranial specific genes at T21 (cranial) and T23 (trunk) in cross-section. e.) Pharyngeal neural crest derivative expression in Tahara stage 26 *Petromyzon marinus* frontal section (illustration based on Damas, et al (1944)[40]). f-l.) Cranial circuit orthologues are expressed in pharyngeal arch derivatives, with the exception of *Brn3* which is present in the neural crest-derived cranial sensory ganglia[41] in lamprey frontal sections. m.) Gene expression matrix summarizing the heterochronic shift of cranial crest specific circuit nodes. nc=neural crest, nt= neural tube, n=notochord, end.=endoderm, ect.=ectoderm, mes.=mesoderm. Scale bars= 100μm. Cryosections of *in situs* were reproducible on n≥5 embryos per time point for n≥2 experiments.



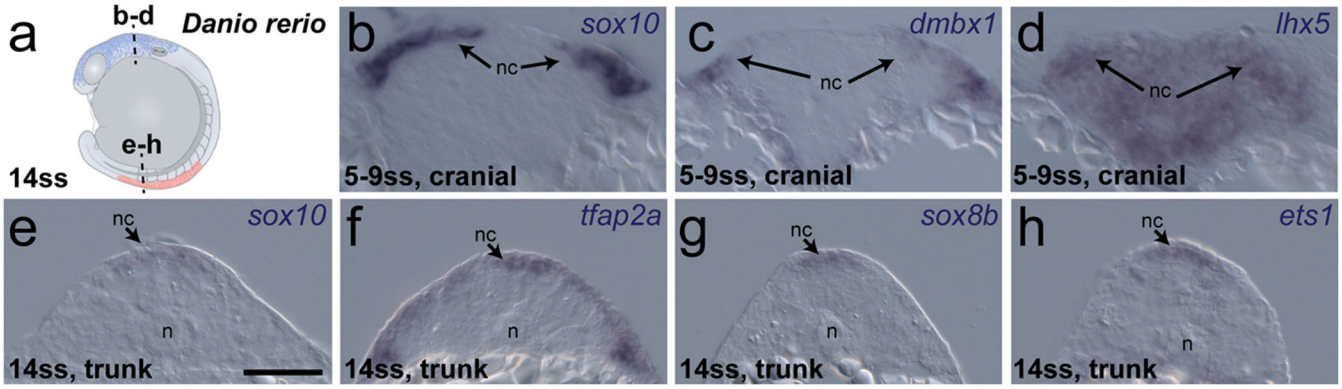
Extended Data 2. Expression of cranial circuit genes in the neural crest of the little skate.

a.) Schematic of a stage 18 *Leucoraja erinacea* embryo with the neural crest illustrated as blue (cranial) and red (trunk). Placement of cross-sections depicted on the illustration for figures b-f. nc=neural crest, scale bar= 50μm. Cryosections of *in situs* were reproducible on n≥2 embryos for n≥2 experiments.



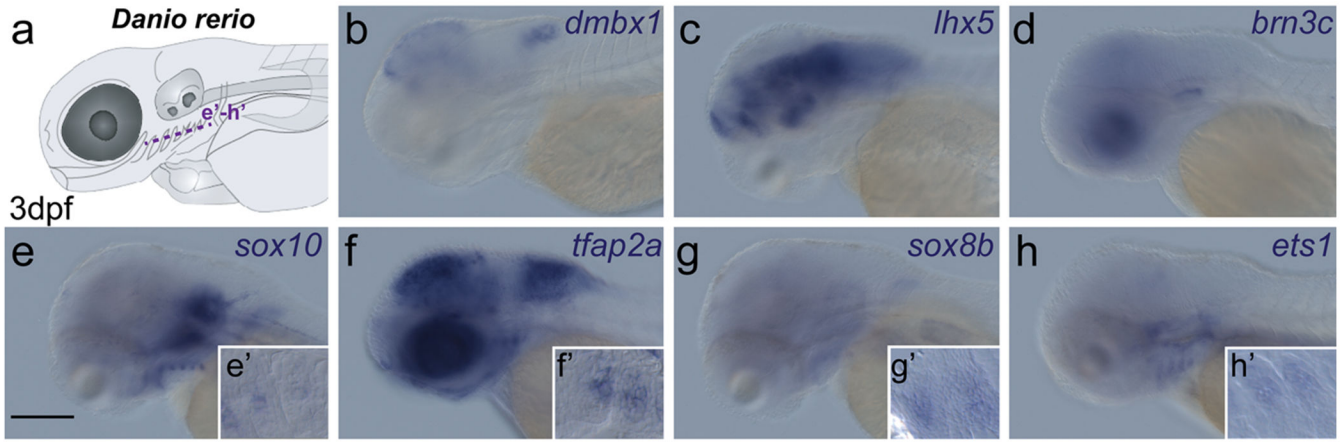
Extended Data 3. Pharyngeal neural crest derivative expression of cranial circuit orthologues in stage 25 *Leucoraja erinacea* embryos.

a.) Dashed box on the illustration represents the region of the head for each embryo imaged in figures b-i, and the purple dashed line depicts the location of the frontal section for figures a'-i'. Pharyngeal neural crest derivative expression of cranial circuit orthologues is seen in panels b-f'. *Dmbx1, Lhx5,* and *Brn3c* are absent in pharyngeal arch derivatives at stage 25 (g-i'). b-i, scale bar=500μm. b'-i', scale bar= 100μm. *in situs* were reproducible on n≥2 embryos.



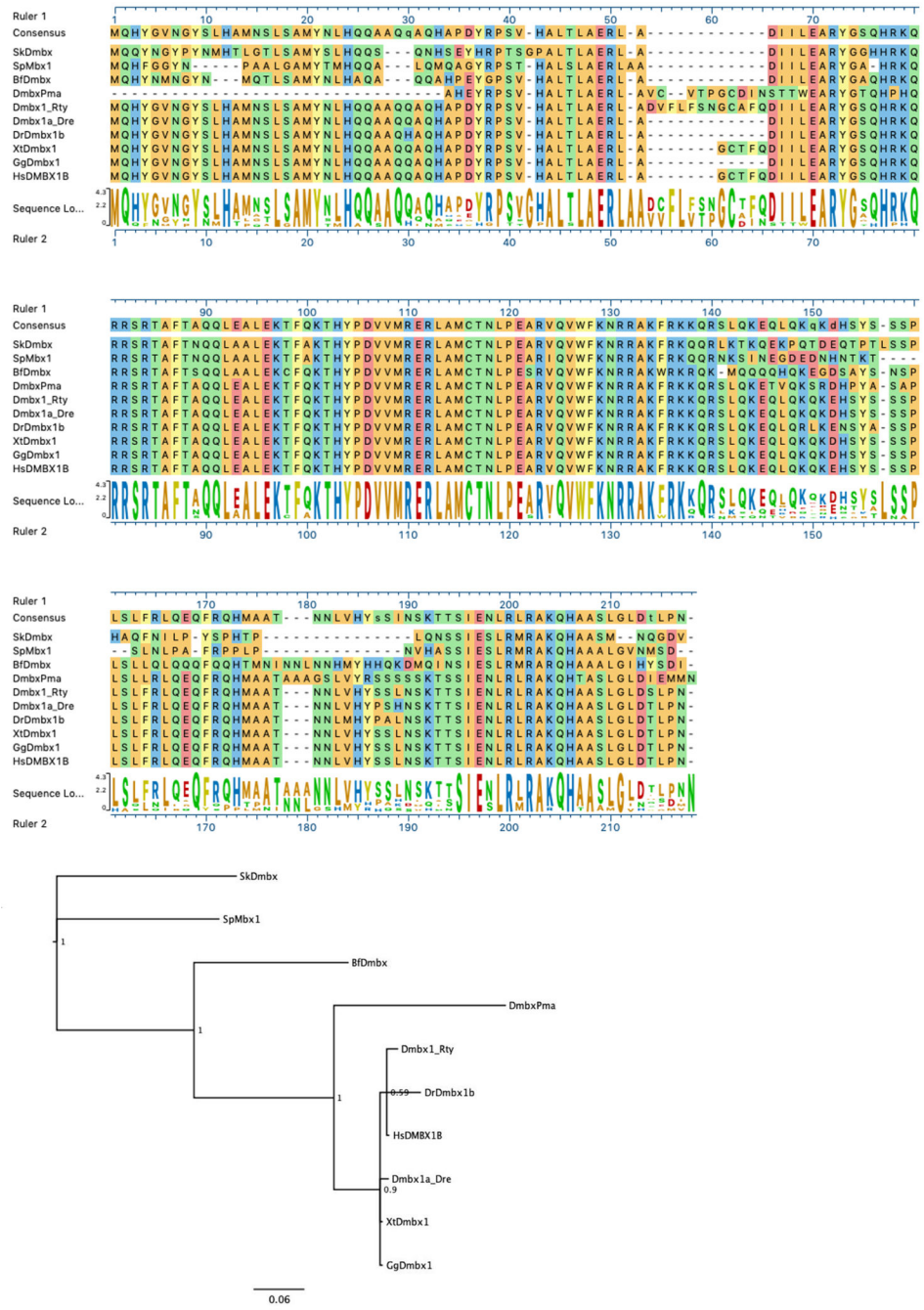
Extended Data 4. Expression of cranial circuit genes in the neural crest of the zebrafish.

a.) Schematic of a 14ss *Danio rerio* embryo with the neural crest illustrated as blue (cranial) and red (trunk). Placement of cross-sections depicted on the illustration for figures b-h. nc=neural crest, n=notochord, scale bars= 50μm. *in situs* were reproducible on n≥10 embryos.



Extended Data 5. Expression of cranial circuit orthologues in pharyngeal neural crest derivatives of 3dpf *Danio rerio* embryos.

a.) Purple dashed line depicts the location of the frontal section for figures e'-h'. Expression of cranial circuit orthologues in pharyngeal arches is seen in panels e-h'. *Dmbx1, Lhx5,* and *Brn3c* are absent from pharyngeal arch derivatives at 3dpf (b-d). Scale bar= 150μm. *in situs* were reproducible on n≥10 whole mount and cryosectioned embryos.



Extended Data 6. P. marinus *Dmbx* is homologous to gnathostome *Dmbx* genes.

a.) Truncated alignment of Dmbx protein sequences. An alignment of full length Dmbx protein sequences was assembled using TCoffee and contiguous regions tagged by the program as poorly or moderately well-aligned were removed, leaving 218 well-aligned residues. b.) Bayesian consensus phylogenetic tree, with posterior probabilities are shown at corresponding nodes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank Joanne Tan-Cabugao and Eli Grossman for technical assistance. We would also like to thank David Mayorga and Ryan Fraser for help with fish husbandry. We thank Brennah Martik for illustrating the adult animals for our expression matrices and acknowledge the Caltech Millard and Muriel Jacobs Genetics and Genomics Laboratory, in particular, Igor Antoshechkin for sequencing of our RNAseq libraries. We thank Rochelle Diamond, Jamie Tijerina, and Patrick Cannon of the The Caltech Flow Cytometry Cell Sorting Facility for cell sorting assistance.

Funding:

This work is supported by NIH grant R01NS086907, R01DE024157, and R35NS111564 to MEB. MLM is supported by a Helen Hay Whitney Foundation postdoctoral fellowship. SG is supported by a graduate fellowship from the American Heart Association (18PRE34050063).

References:

1. Simoes-Costa M & Bronner ME Reprogramming of avian neural crest axial identity and cell fate. Science 352, 1570–1573 (2016). [PubMed: 27339986]
2. Gans C & Northcutt RG Neural crest and the origin of vertebrates: a new head. Science 220, 268–273 (1983). [PubMed: 17732898]
3. Glenn Northcutt R The new head hypothesis revisited. J. Exp. Zool. B Mol. Dev. Evol 304B, 274–297 (2005).
4. Le Douarin NM Development Of The Peripheral Nervous System From The Neural Crest. Annu. Rev. Cell Dev. Biol 4, 375–404 (1988).
5. Le Douarin N The Neural Crest. (Cambridge University Press, 1982).
6. Sauka-Spengler T, Meulemans D, Jones M & Bronner-Fraser M Ancient evolutionary origin of the neural crest gene regulatory network. Dev. Cell 13, 405–420 (2007). [PubMed: 17765683]
7. Nikitina N, Sauka-Spengler T & Bronner-Fraser M Dissecting early regulatory relationships in the lamprey neural crest gene network. Proc. Natl. Acad. Sci. U.S.A 105, 20083–20088 (2008). [PubMed: 19104059]
8. Green SA, Uy BR & Bronner ME Ancient evolutionary origin of vertebrate enteric neurons from trunk-derived neural crest. Nature 544, 88–91 (2017). [PubMed: 28321127]
9. Häming D et al. Expression of Sympathetic Nervous System Genes in Lamprey Suggests Their Recruitment for Specification of a New Vertebrate Feature. PLoS ONE 6, e26543 (2011). [PubMed: 22046306]
10. Betancur P, Bronner-Fraser M & Sauka-Spengler T Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. Proc. Natl. Acad. Sci. U.S.A 107, 3570–3575 (2010). [PubMed: 20139305]
11. Haldin CE & LaBonne C SoxE factors as multifunctional neural crest regulatory factors. The International Journal of Biochemistry & Cell Biology 42, 441–444 (2010). [PubMed: 19931641]
12. McCauley DW & Bronner-Fraser M Importance of SoxE in neural crest development and the evolution of the pharynx. Nature 441, 750–752 (2006). [PubMed: 16760978]
13. Lee EM et al. Functional constraints on SoxE proteins in neural crest development: The importance of differential expression for evolution of protein activity. Dev. Biol 418, 166–178 (2016). [PubMed: 27502435]
14. Green SA, Simoes-Costa M & Bronner ME Evolution of vertebrates as viewed from the crest. Nature 520, 474–482 (2015). [PubMed: 25903629]
15. Martik ML & Bronner ME Regulatory Logic Underlying Diversification of the Neural Crest. Trends Genet. 33, 715–727 (2017). [PubMed: 28851604]

16. Gillis JA, Alsema EC & Criswell KE Trunk neural crest origin of dermal denticles in a cartilaginous fish. Proc. Natl. Acad. Sci. U.S.A 114, 13200–13205 (2017). [PubMed: 29158384]
17. Sauka-Spengler T & Bronner-Fraser M Evolution of the neural crest viewed from a gene regulatory perspective. Genesis 46, 673–682 (2008). [PubMed: 19003930]
18. Wada H, Kobayashi M & Zhang S Ets identified as a trans-regulatory factor of amphioxus Hox2 by transgenic analysis using ascidian embryos. Dev. Biol 285, 524–532 (2005). [PubMed: 16051213]
19. Meulemans D & Bronner-Fraser M Amphioxus and lamprey AP-2 genes: implications for neural crest evolution and migration patterns. Development 129, 4953–4962 (2002). [PubMed: 12397104]
20. Takahashi T & Holland PWH Amphioxus and ascidian Dmbx homeobox genes give clues to the vertebrate origins of midbrain development. Development 131, 3285–3294 (2004). [PubMed: 15201221]
21. Yu J-K, Meulemans D, McKeown SJ & Bronner-Fraser M Insights from the amphioxus genome on the origin of vertebrate neural crest. Genome Res. 18, 1127–1132 (2008). [PubMed: 18562679]
22. Davidson EH & Erwin DH Gene regulatory networks and the evolution of animal body plans. Science 311, 796–800 (2006). [PubMed: 16469913]
23. Parker HJ, Bronner ME & Krumlauf R A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. Nature 514, 490–493 (2014). [PubMed: 25219855]
24. Parker HJ, Bronner ME & Krumlauf R The vertebrate Hox gene regulatory network for hindbrain segmentation: Evolution and diversification: Coupling of a Hox gene regulatory network to hindbrain segmentation is an ancient trait originating at the base of vertebrates. Bioessays 38, 526–538 (2016). [PubMed: 27027928]

Methods References:

25. Nikitina N, Bronner-Fraser M & Sauka-Spengler T Culturing lamprey embryos. Cold Spring Harb Protoc 2009, pdb.prot5122–pdb.prot5122 (2009).
26. Di Tommaso P et al. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res. 39, W13–7 (2011). [PubMed: 21558174]
27. Armougom F et al. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res. 34, W604–8 (2006). [PubMed: 16845081]
28. O'Sullivan O, Suhre K, Abergel C, Higgins DG & Notredame C 3DCoffee: combining protein sequences and structures within multiple sequence alignments. J. Mol. Biol 340, 385–395 (2004). [PubMed: 15201059]
29. Poirot O, Suhre K, Abergel C, O'Toole E & Notredame C 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. Nucleic Acids Res. 32, W37–40 (2004). [PubMed: 15215345]
30. Notredame C, Higgins DG & Heringa J T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol 302, 205–217 (2000). [PubMed: 10964570]
31. Longabaugh WJR, Davidson EH & Bolouri H Computational representation of developmental genetic regulatory networks. Dev. Biol 283, 1–16 (2005). [PubMed: 15907831]
32. Simões-Costa MS, McKeown SJ, Tan-Cabugao J, Sauka-Spengler T & Bronner ME Dynamic and differential regulation of stem cell factor FoxD3 in the neural crest is Encrypted in the genome. PLoS Genet. 8, e1003142 (2012). [PubMed: 23284303]
33. HAMBURGER V & HAMILTON HL A series of normal stages in the development of the chick embryo. J. Morphol 88, 49–92 (1951). [PubMed: 24539719]
34. Smith JJ et al. The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. Nat. Genet 1 (2018). doi:10.1038/s41588-017-0036-1 [PubMed: 29273803]
35. Kent WJ BLAT--the BLAST-like alignment tool. Genome Res. 12, 656–664 (2002). [PubMed: 11932250]

36. Karolchik D et al. The UCSC Genome Browser Database. Nucleic Acids Res. 31, 51–54 (2003). [PubMed: 12519945]

37. Langmead B & Salzberg SL Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012). [PubMed: 22388286]

38. Anders S, Pyl PT & Huber W HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015). [PubMed: 25260700]

39. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014). [PubMed: 25516281]

40. Damas H Recherches sur le développement de *Lampetra fluviatilis* L. Contribution à l'étude de la céphalogenèse des vertébrés. Arch. Biol 1–284 (1944).

41. Modrell MS et al. A fate-map for cranial sensory ganglia in the sea lamprey. Dev. Biol 385, 405–416 (2014). [PubMed: 24513489]

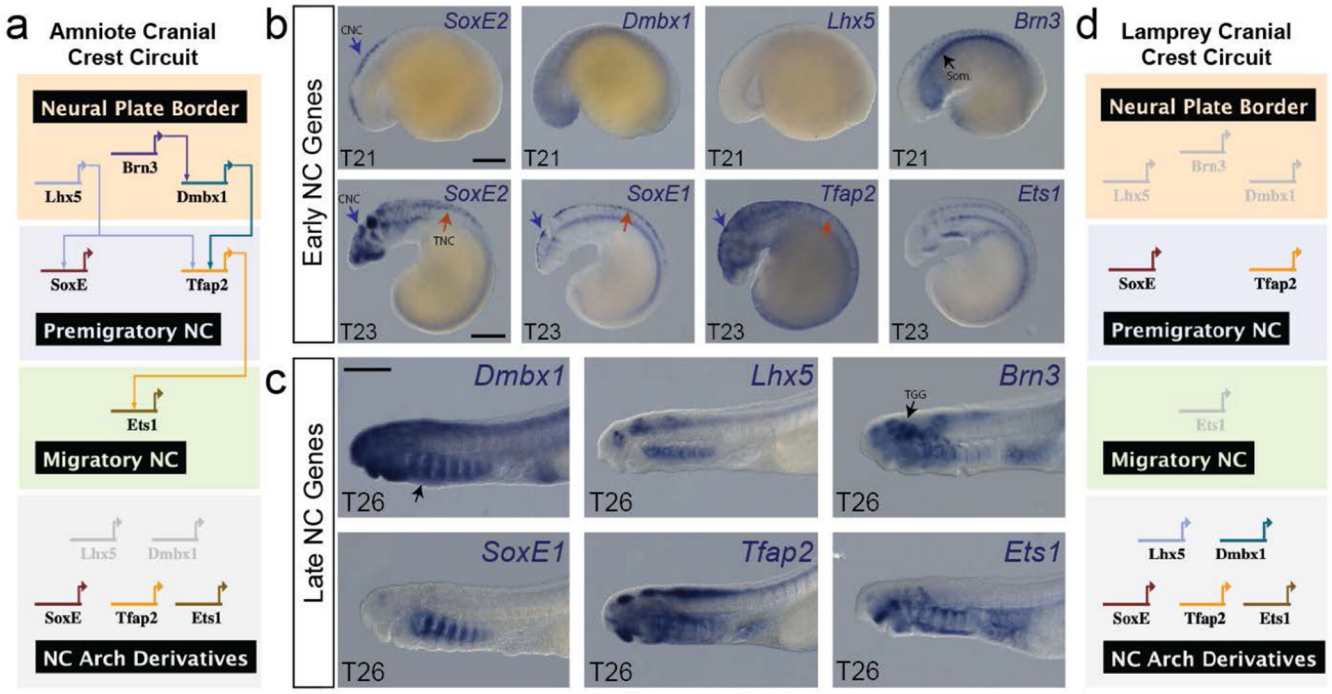


Figure 1. Lamprey cranial neural crest lacks most components of a chick "cranial crest circuit".

a.) Biotapestry model of cranial specific gene regulatory circuit driving skeletal differentiation in amniotes. b.) Expression of lamprey orthologues of amniote cranial specific genes at T21 and T23. Blue arrows represent expression in the cranial neural crest (CNC), and red arrows represent expression in the trunk neural crest (TNC). c.) Late expression of cranial specific orthologues in the pharyngeal arch neural crest derivatives (black arrow). d.) Biotapestry model of the lamprey circuit with the addition of late module expression of markers in the pharyngeal arch neural crest derivatives. TGG, trigeminal ganglia. Scale bars, 250μm. Reproducible on n≥5 embryos per time point for n≥10 experiments.

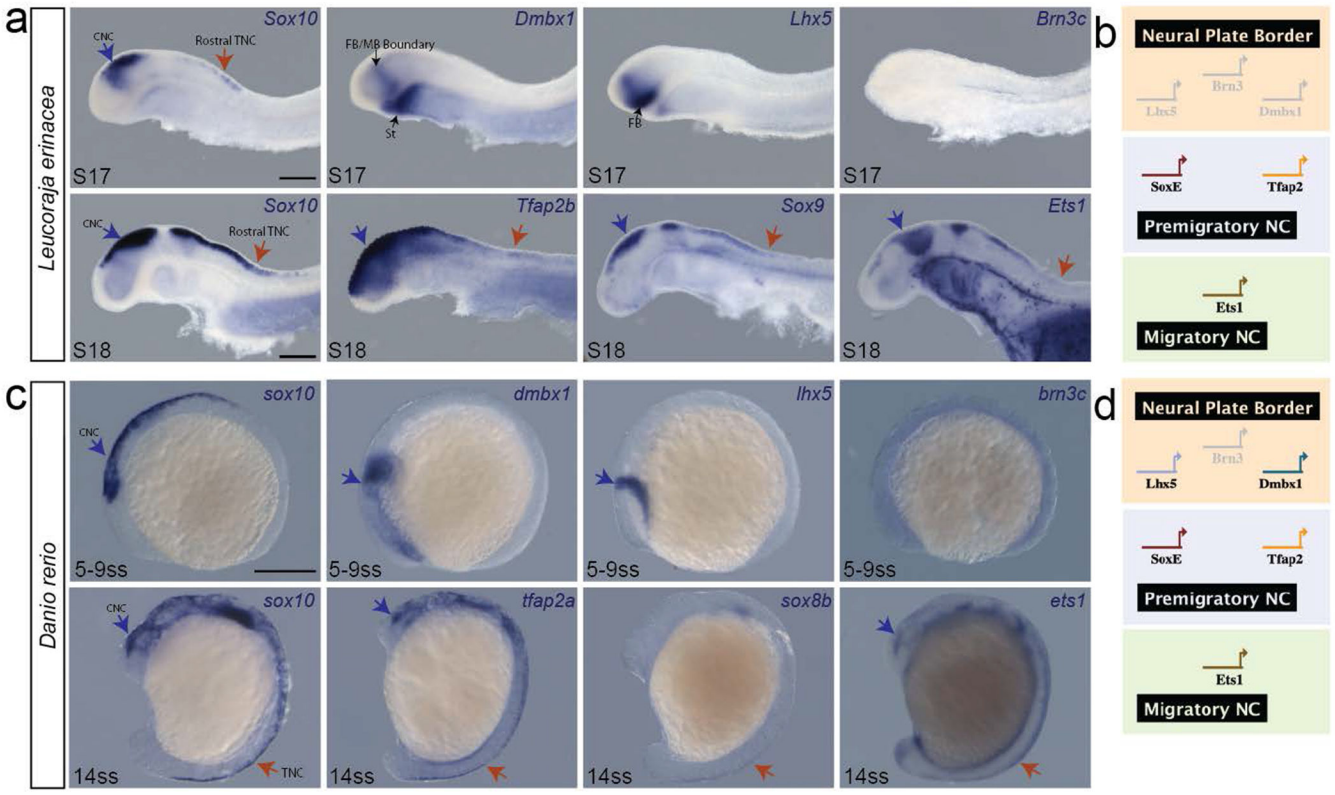


Figure 2. Nodes of an early cranial-specific circuit were acquired in and restricted to the cranial neural crest progressively throughout gnathostome evolution.

a.) Expression of cranial-specific orthologues in the Little Skate, *Leucoraja erinacea* at stage (S) 17 and 18. Expression of orthologues in the cranial neural crest (CNC) are depicted with a blue arrow and rostral trunk neural crest (TNC) with red arrows. b.) Biotapestry model of the skate circuit with the addition of a novel node, *Ets1.* c.) Expression of cranial-specific orthologues in the zebrafish, *Danio rerio* at 5-9 somite stage (ss) and 14ss. d.) Biotapestry model of the zebrafish circuit with the addition of novel early nodes, *lhx5* and *dmbx1.* FB/MB, forebrain/midbrain. St, stomodeum. Scale bars, 250μm. For skates, *in situs* were reproducible on n≥2 embryos for n≥2 experiments. For zebrafish, *in situs* were reproducible on n≥5 embryos per time point for n≥10 experiments.

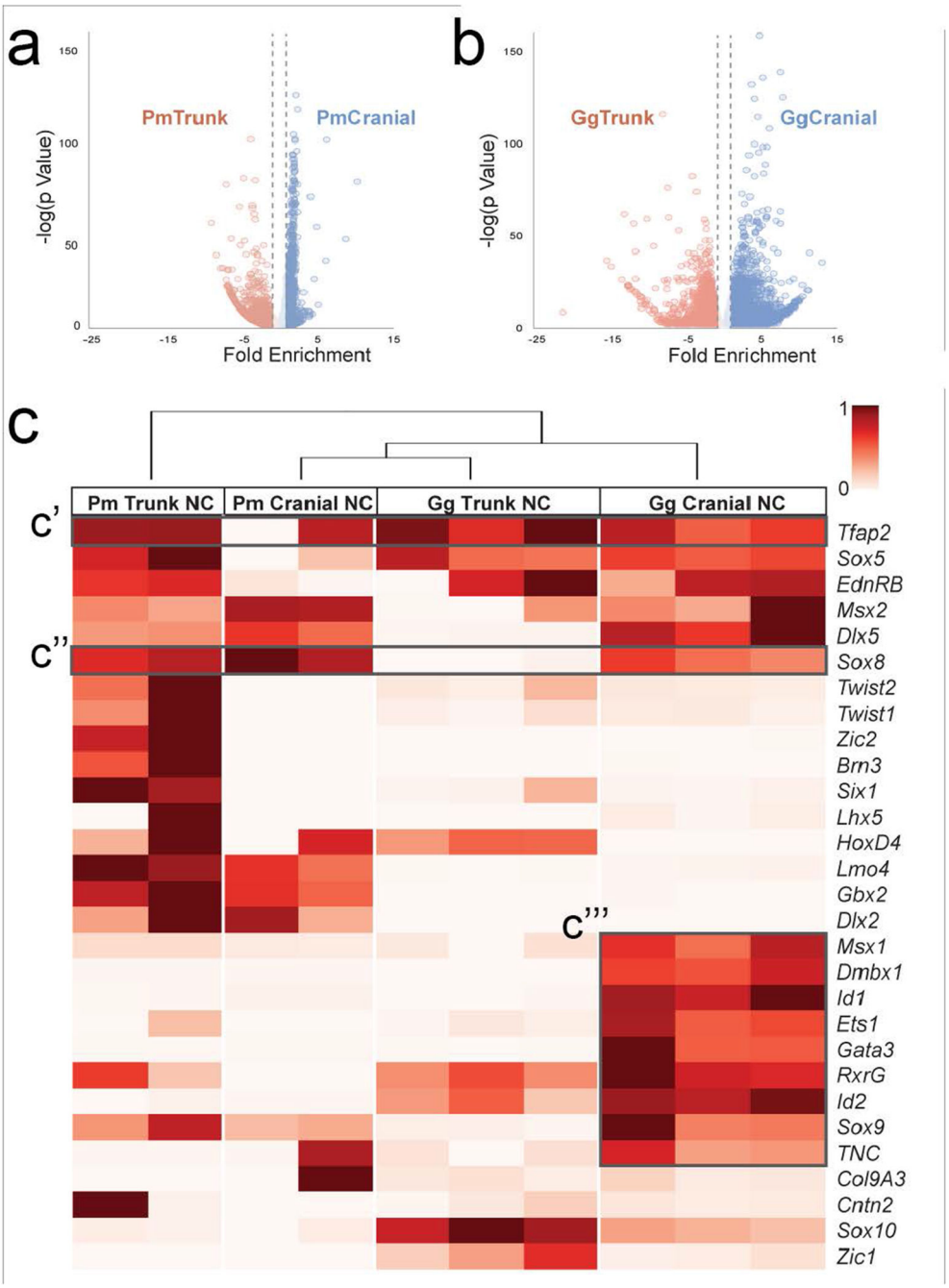


Figure 3. Tissue-specific RNA-seq comparisons between lamprey and chicken reveal ancestral neural crest had a more trunk-like identity.

a.) Volcano plot showing lamprey differential enrichments of cranial (blue) and trunk (red) genes by population RNAseq (100 embryos were dissected for each of n=2 biological replicates) (adjusted p-value<0.05). b.) Volcano plot showing enrichment of genes in the cranial (blue) versus the trunk (red) neural crest in chicken, *Gallus gallus* (≥15 heads and 5 trunks were dissected and prepared for FAC-sorting for each of n=3 biological replicates) (adjusted p-value<0.05). c.) Hierarchical clustering analysis of all RNAseq libraries focused

on the neural crest GRN reveals similarities and differences between axial levels among species.

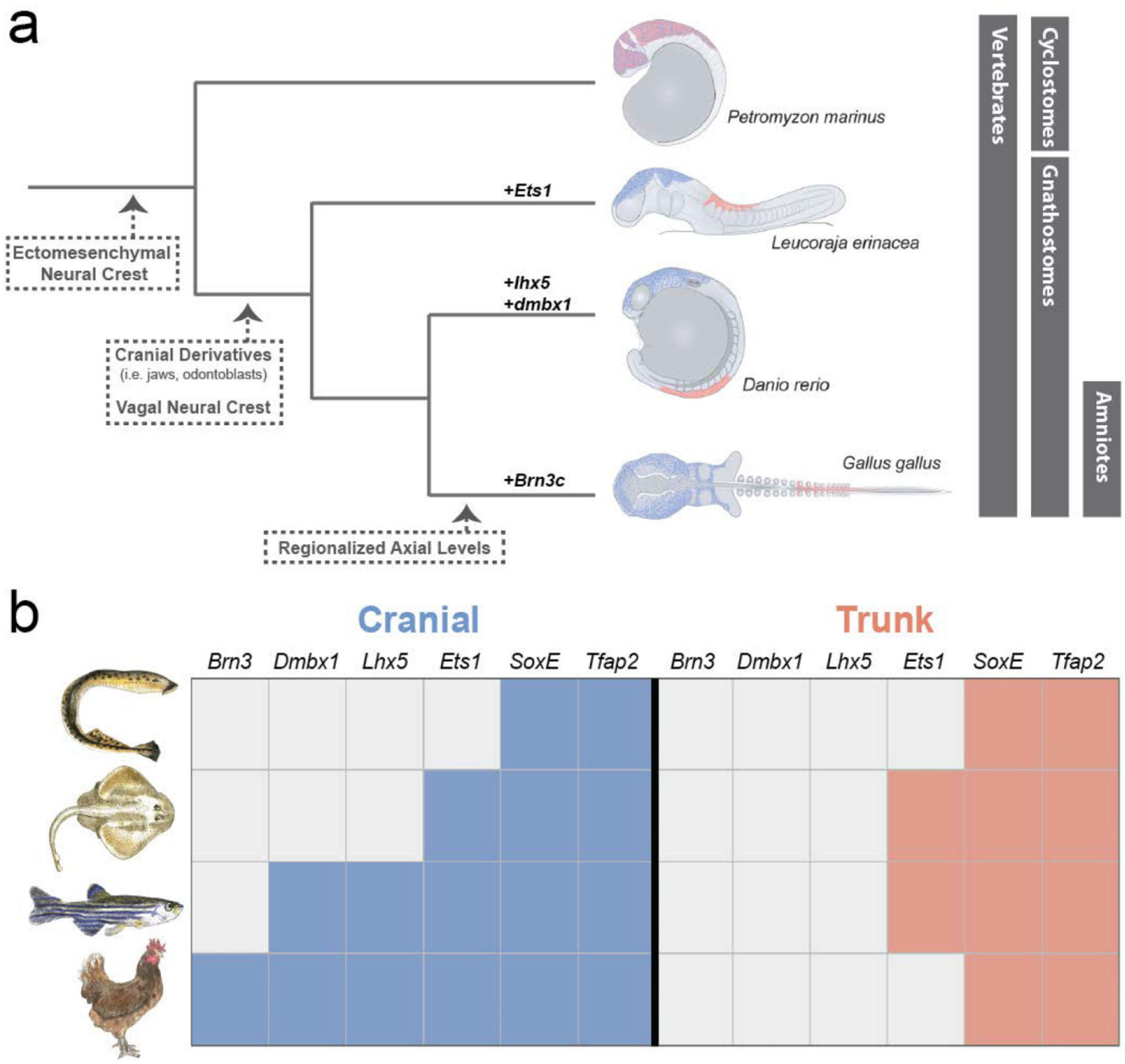


Figure 4. Model of the evolution of neural crest axial levels during vertebrate evolution.

a.) Our data suggest that the ancestral neural crest was a more uniform population of cells along the body axis that underwent gradual regulatory modifications during gnathostome evolution. b.) Progressive restriction of the “cranial circuit genes” to only the cranial axial level led to axial specialization of the neural crest regulatory program.