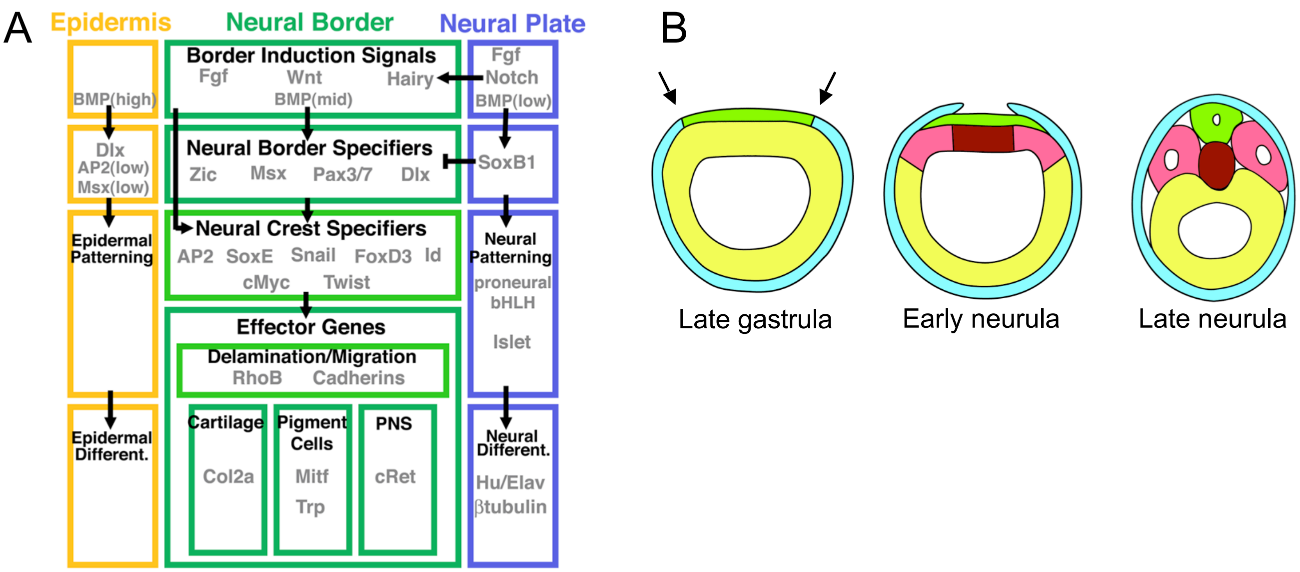

Figure S1 | A. The neural crest gene network in vertebrates. Black arrows indicate empirically verified regulatory interactions. In vertebrates, signals from ventral ectoderm and underlying mesendoderm pattern the dorsal ectoderm, inducing expression of neural border specifiers and neural plate markers. These inductive signals then work with neural border specifiers to activate expression of neural crest specifiers. The neural crest specifiers cross-regulate and activate various effector genes, each of which mediates a different aspect of the neural crest phenotype. These include cassettes controlling neural crest delamination and migration as well as differentiation programs for neural crest derivatives like cartilage, pigment cells and peripheral neurons. **B. Cross sections showing neurulation process in amphioxus embryos**. In late gastrula, the embryo is ovoid with presumptive neural plate (green) becoming flattened. The boundaries of presumptive neural plate and epidermis (blue) at this stage are indicated by arrows, and the mesendoderm is colored in yellow. In early neurula, the epidermis (blue) at the edges of the neural plate detaches from the neural plate and begins to overgrow the neural plate, and the somites (pink) and notochord (red) start to form. Eventually the overgrowing dorsal epidermis fuse together at the dorsal midline, and the underneath neural plate curls up to form a dorsal, hollow neural tube in late neurula stage. Now, neurulation is complete, and the rudiment of notochord, anterior somites, as well as the endodermal tube rudiment that will form the pharyngeal tube and gut (yellow) become distinguishable.

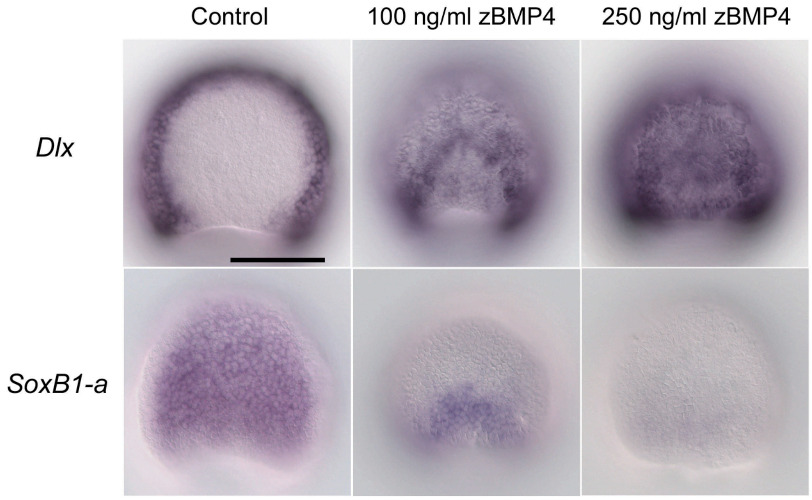

Figure S2 | Dose-dependent effect of zBMP4 treatment on amphioxus epidermal and neural gene expression. Whole-mount in situ hybridization for *Dlx* and *SoxB1-a* at late gastrula stage. All embryos are oriented in dorsal view with anterior at the top. Scale bar, 50 μm. Expression of epidermal marker *Dlx* was expanded by zBMP4 treatment in a dose-dependant manner; conversely, expression of neural marker *SoxB1-a* was reduced by low concentration of zBMP4 treatment and completely disappeared by high concentration of zBMP4.

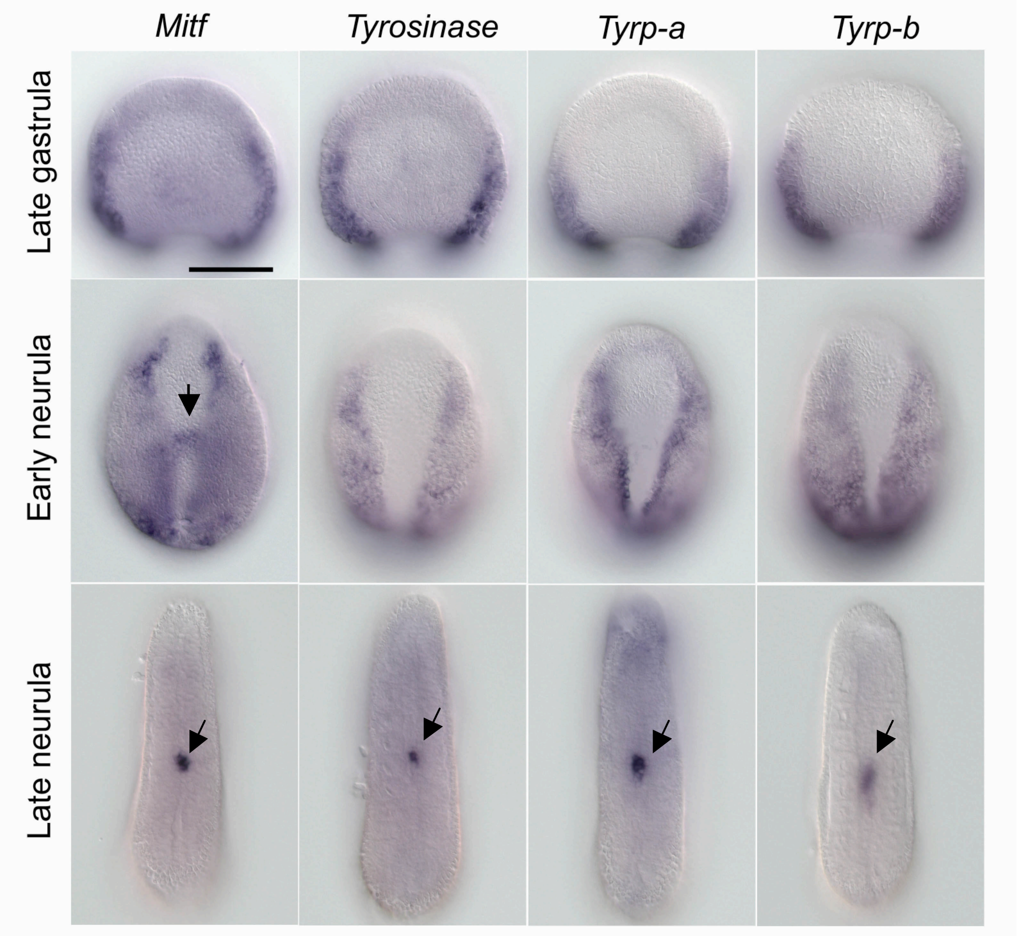

Figure S3 | Expression of *Mitf*, *Tyrosinase*, and two *Tyrosinase-related proteins* in amphioxus embryos. All embryos are oriented in dorsal view with anterior at the top. Scale bar, 50 μm. All these four genes are co-expressed throughout the epidermal ectoderm in gastrula stage and early neurula stage. During neurula stage, a patch of cells inside the developing neural tube (indicating by arrows) start to expressed these four genes, and the onset of Mitf transcription factor gene expression is earlier than the other melanogenic genes.

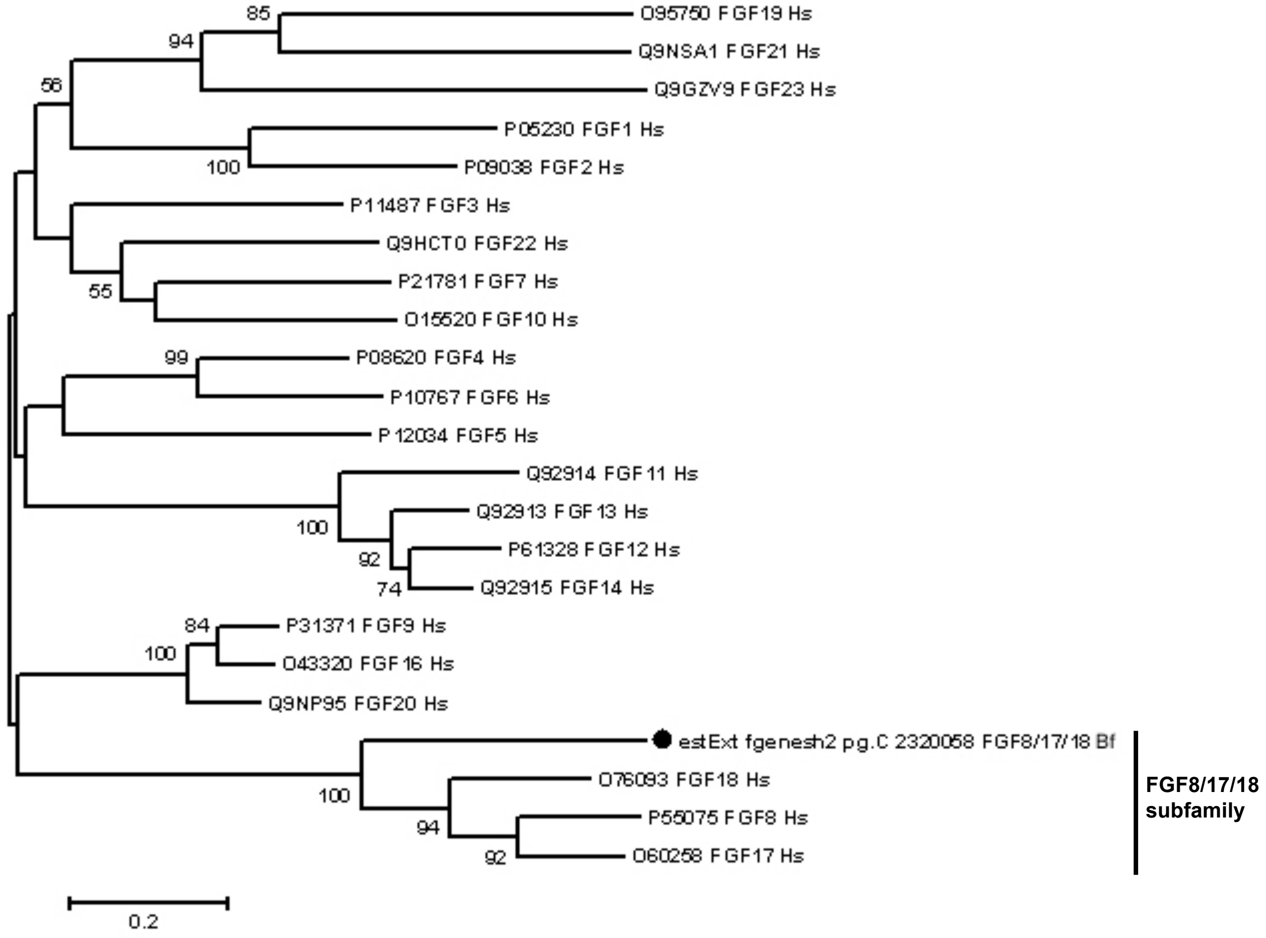

Figure S4 | Phylogenetic tree of FGF proteins constructed by MEGA program using the neighbor-joining method with default settings. Amphioxus protein is shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.2 amino acid substitutions per position. FGF8 has been implicated in playing roles in neural induction and neural crest development. Our survey identified an amphioxus candidate belonging to the FGF8/17/18 subfamily with high bootstrap value of 100%.

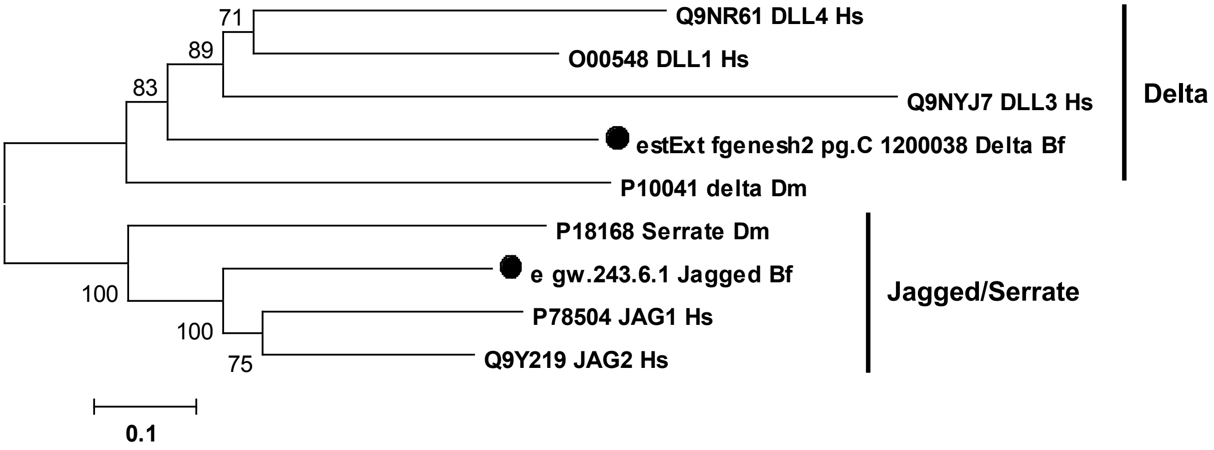

Figure S5 | Phylogenetic tree of Delta/Jagged/Serrate family proteins generated by MEGA program using the neighbor-joining method with default settings. Amphioxus proteins are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. The unrooted tree is shown as a rooted tree for simplicity. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. Delta and Jagged/Serrate proteins are ligands for Notch receptors. Our survey in amphioxus genome identified one candidate gene for Delta and one for Jagged/Serrate. The robust grouping of the tree supports their orthology assignment.

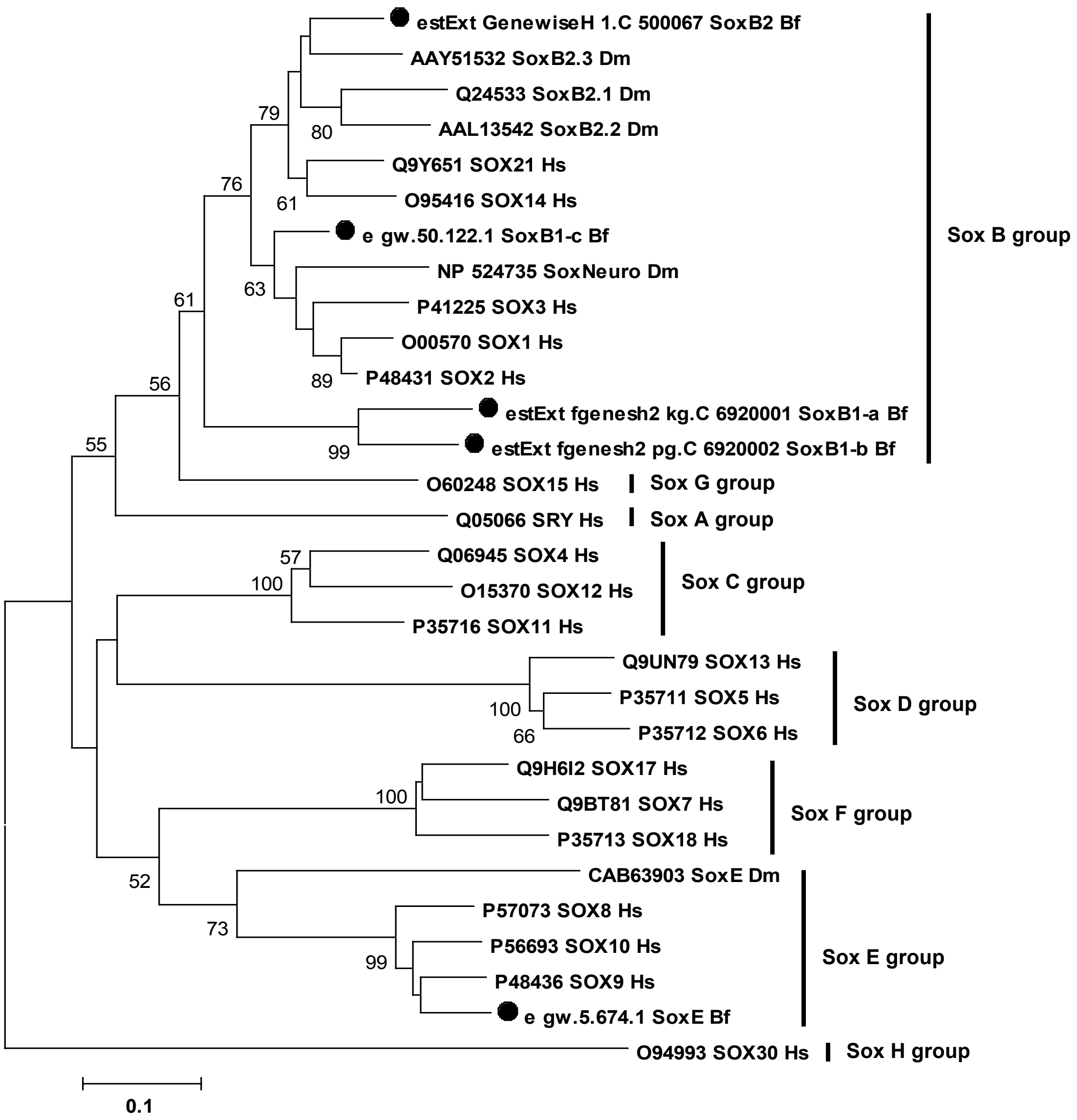

Figure S6 | Phylogenetic tree of Sox proteins constructed by the neighbor-joining method with default settings. Amphioxus protein are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. Our survey of amphioxus Sox family genes was focused on the SoxB and SoxE groups. We identified an amphioxus SoxE group candidate which groups with vertebrate SoxE proteins with high bootstrap value(99%). We identified four candidates of amphioxus SoxB members. Although the bootstrap value of SoxB grouping was not very high, we designated the protein names based on the best hit analysis by blast search against a human proteome.

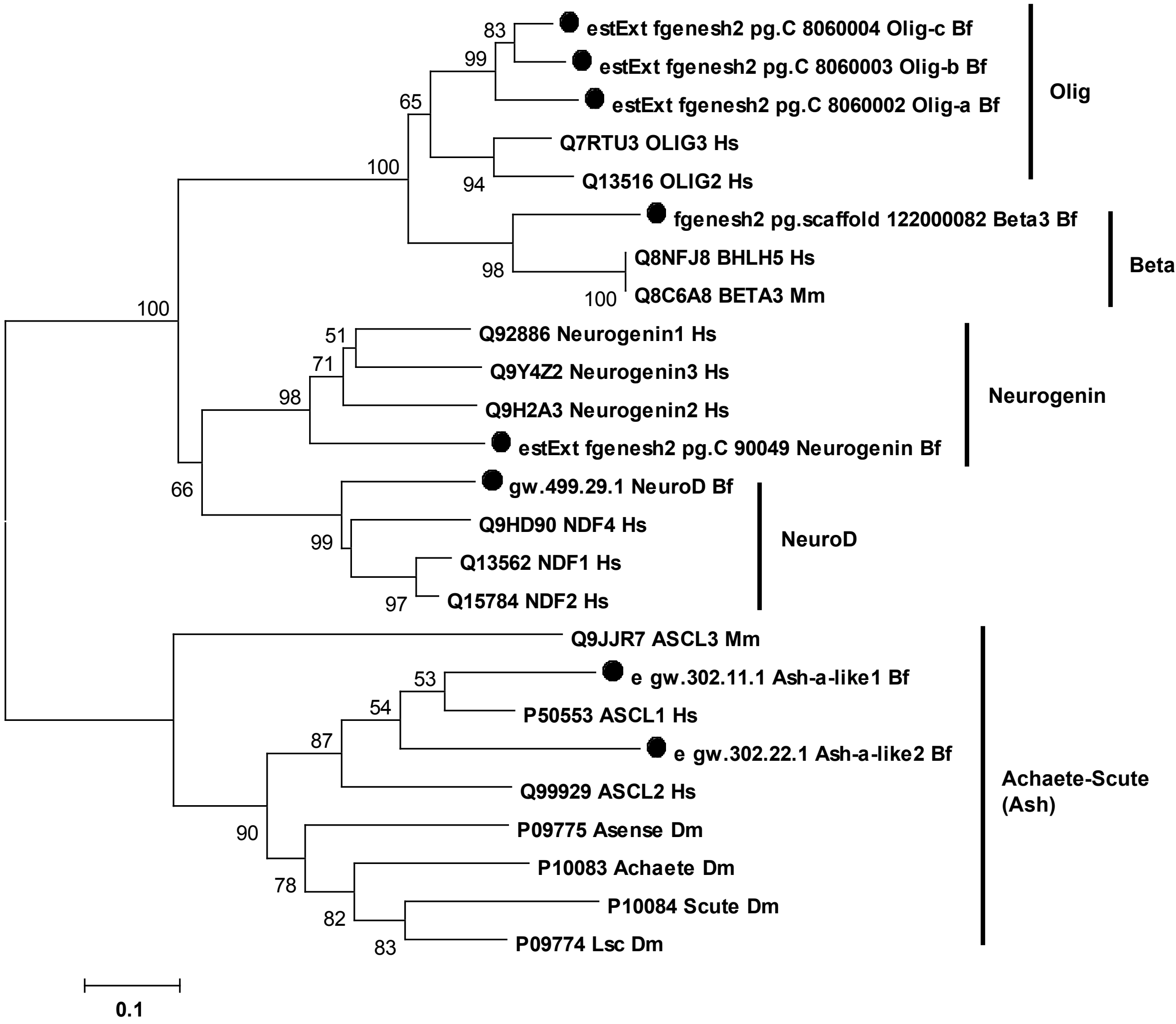

Figure S7 | Phylogenetic tree of Proneural bHLH factors generated by the neighbor-joining method with default settings. Amphioxus proteins are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudo-replications. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. We identified three possible Olig family orthologues in the amphioxus genome and one candidate for the closely related Beta3 family. The three amphioxus Olig genes are tandemly aligned in the genome within a 62-kb region, suggesting that these genes occurred by amphioxus-specific gene duplications. We also identified amphioxus orthologues of Neurogenin and NeuroD family with strong bootstrap values. There are two possible Ash family members in amphioxus genome, however, the phylogenetic resolution was low within the family. Therefore, we designated these two factors as Ash-a-like1 and Ash-a-like2 based on the best hit analysis.

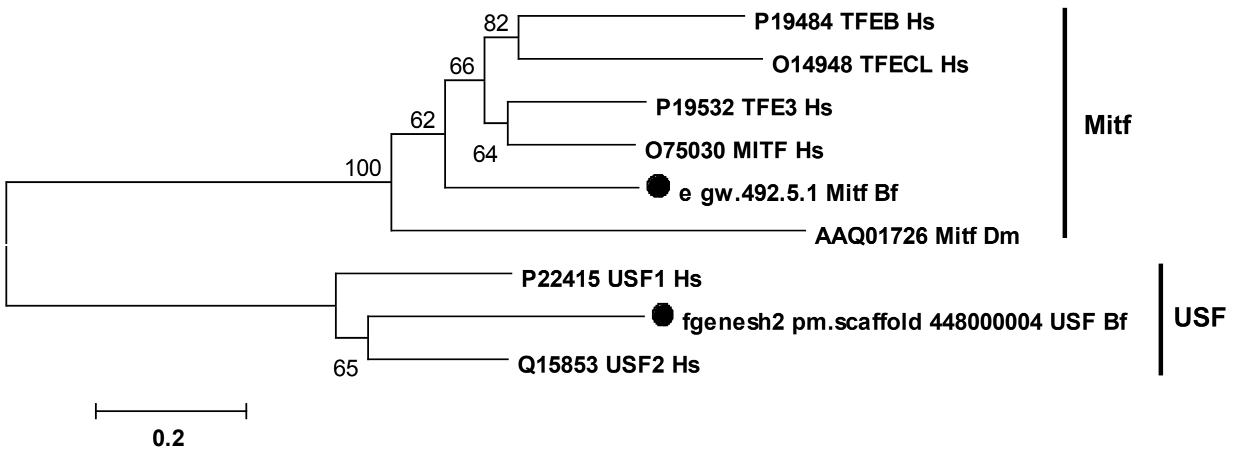

Figure S8 | Phylogenetic tree of Mitf family proteins generated by MEGA program using the neighbor-joining method with default settings. Amphioxus proteins are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. The unrooted tree is shown as a rooted tree for simplicity. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.2 amino acid substitutions per position. Our survey in amphioxus genome identified one orthologue of Mitf, and it was included with other Mitf family factors as a clade with high bootstrap support (100%). We also identified a candidate for the USF family member, which is a closely related family to Mitf.

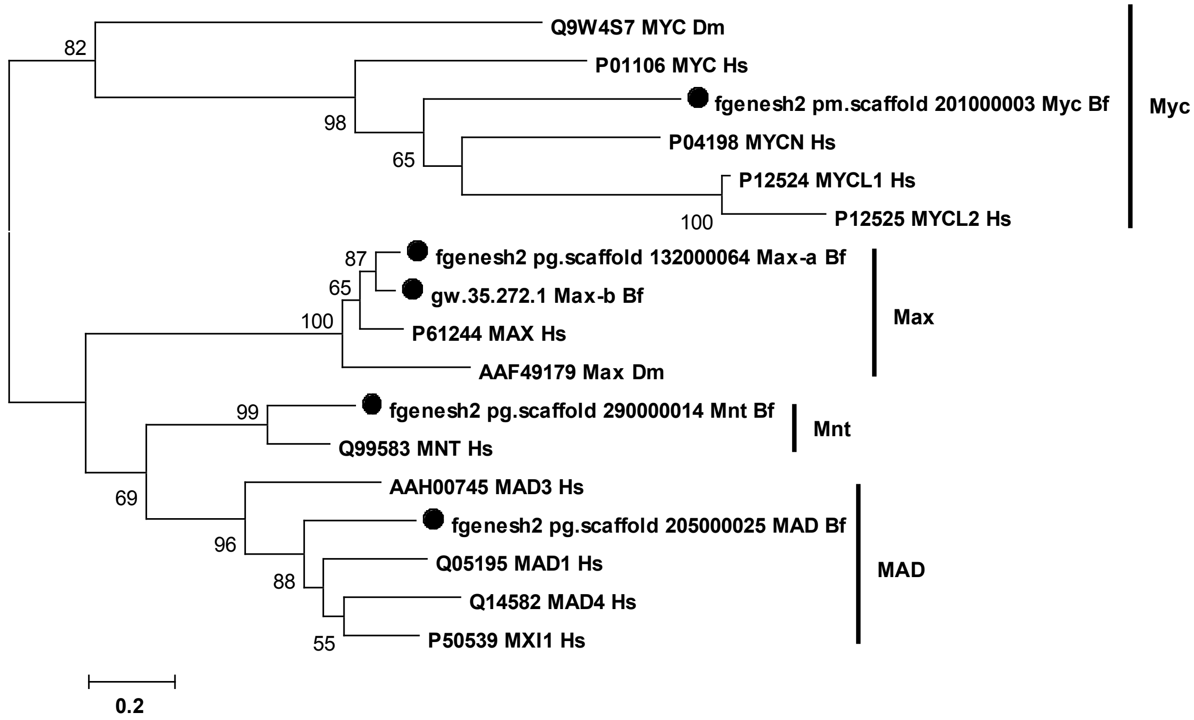

Figure S9 | Phylogenetic tree of Myc, Max, Mnt, and MAD family proteins generated by the neighbor-joining method with default settings. Amphioxus proteins are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. The unrooted tree is shown as a rooted tree for simplicity. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.2 amino acid substitutions per position. Our survey in amphioxus genome identified one candidate orthologue of Myc family factor with good bootstrap support (100%). We also identified orthologues of the closely related Max, Mnt, and MAD family members from amphioxus genome. There are two amphioxus Max members and the phylogenetic analysis suggests they occurred by amphioxus-specific duplication.

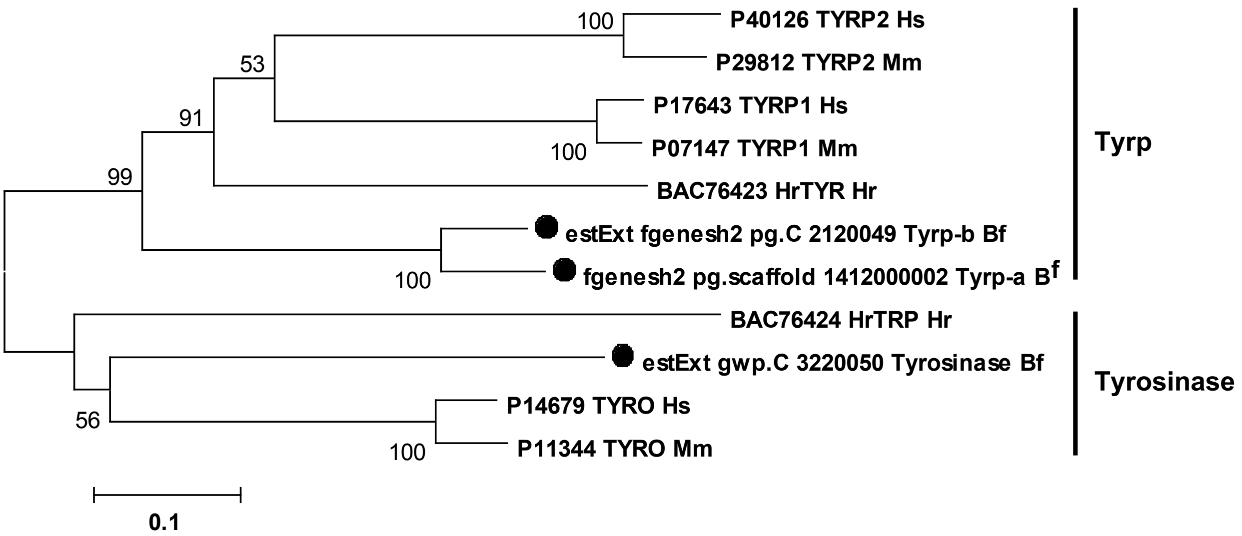

Figure S10 | Phylogenetic tree of Tyrosinase and Tyrosinase-related proteins (Tyrp) generated by the neighbor-joining method with default settings. Amphioxus proteins are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. The unrooted tree is shown as a rooted tree for simplicity. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. Tyrosinase and Tyrosinase-related proteins are separated clearly as two groups in our phylogenetic tree. Our phylogenetic analysis suggests that the tow Tyrosinase-related proteins we identified are amphioxus specific duplications, therefore we designated them as Tyrp-a and Tyrp-b.

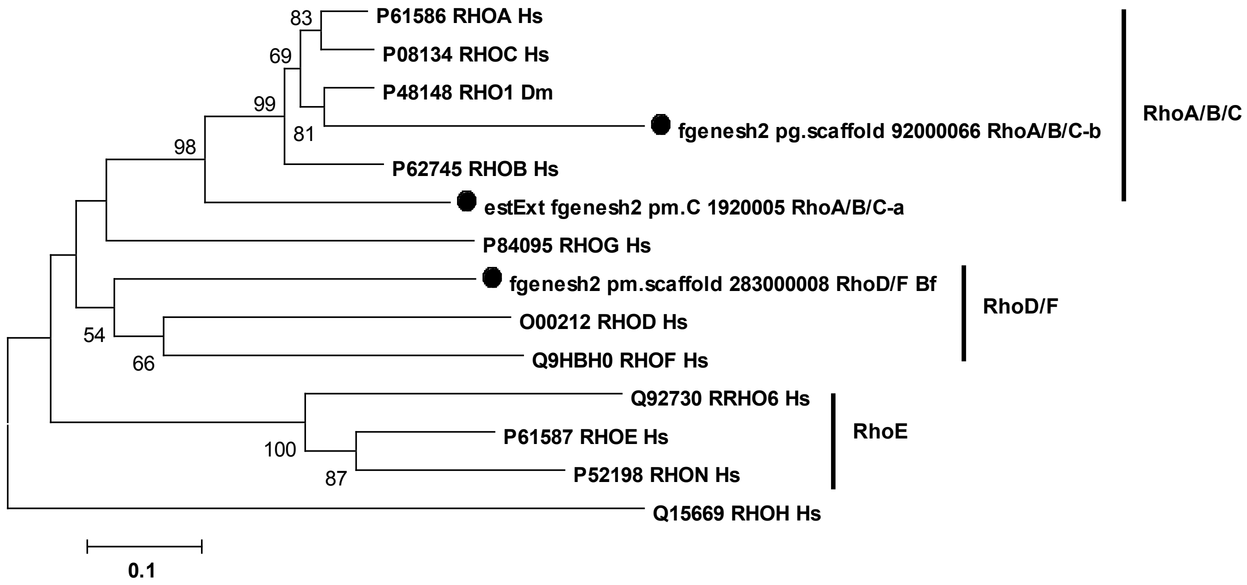

Figure S11 | Phylogenetic tree of Rho small G-proteins generated by MEGA program using the neighbor-joining method with default settings. Amphioxus proteins are shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. The unrooted tree is shown as a rooted tree for simplicity. Protein names are as explained in the Methods. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. We identified two candidate gene models encoding Rho-like G-proteins belonging to the RhoA/B/C group (bootstrap value of 98%). There is also a possible RhoD/F orthologue in amphioxus, however, the bootstrap supporting value is not very high.

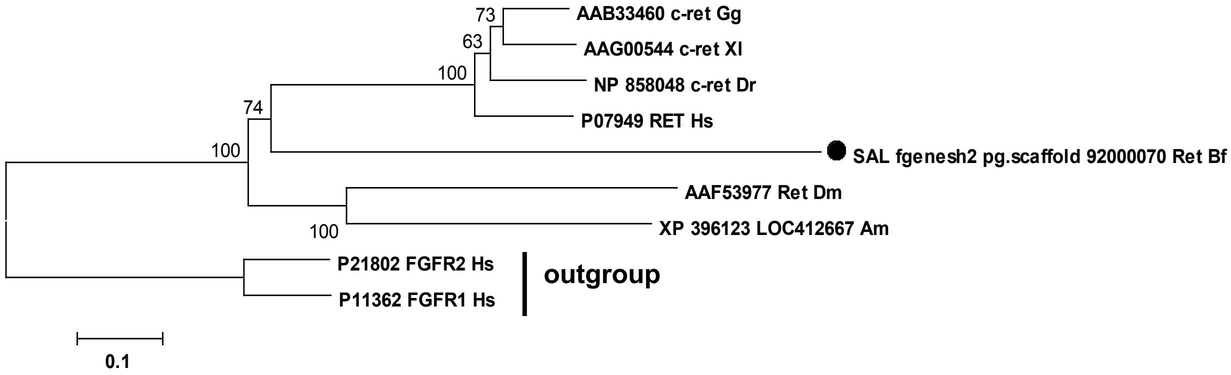

Figure S12 | Phylogenetic tree of Ret receptor tyrosine kinase proteins generated by the neighbor-joining method based on the alignment of a tyrosine kinase domain. Amphioxus protein is shown by the large black dot. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. Protein names are as explained in the Methods. Human FGF receptors were used as outgroup proteins. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. Our survey in amphioxus genome identified one candidate belonging to Ret family.

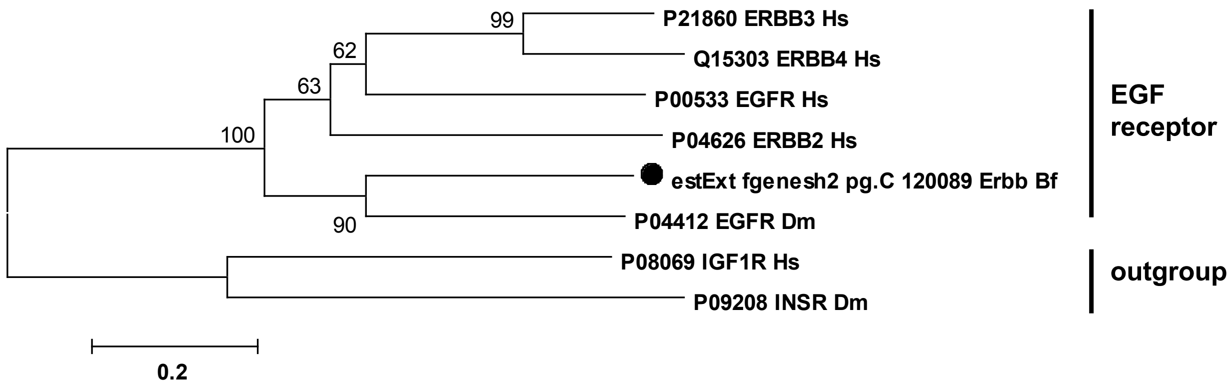

Figure S13 | Phylogenetic tree of EGF receptors generated by the neighbor-joining method based on the alignment of a tyrosine kinase domain. Amphioxus protein is shown by large black dots. The number matching each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. Protein names are as explained in the Methods. Human and fly Insulin receptors were used as outgroup proteins. The scale bar indicates an evolutionary distance of 0.2 amino acid substitutions per position. In vertebrates, there are four members of epidermal growth factor (EGF) receptor family of receptor tyrosine kinase including ERBB1 (EGFR), ERBB2, ERBB3 and ERBB4. We identified only one amphioxus Erbb candidate in the genome, and the orthology among this amphioxus protein and EGFRs from other organisms is supported by high bootstrap value of 100%.

Table S1. Homologues of vertebrate neural crest network genes in the amphioxus *Brachiosoma floridae* genome

Gene name	The best hit model in the version 1.0 assembly	The best hit protein in the human proteome	cDNA clone ID
Neural plate border induction signals			
BMP2/4	estExt_fgenesh2_pg.C_3470056	P12643 BMP-2	bfga034p10
Wnt1	estExt_fgenesh2_pm.C_120023	P04628 WNT-1	bfne123k05
Wnt3	estExt_fgenesh2_pg.C_120044	P56704 WNT-3A	bflv058m15
Wnt6	fgenesh2_pm.scaffold_12000022	Q9Y6F9 WNT-6	bflv009a17
Wnt7	fgenesh2_kg.scaffold_10000002	P56706 WNT-7B	bfne126m18
Wnt8	fgenesh2_kg.scaffold_413000001	Q93098 WNT-8B	bfne137b06
FGF8/17/18	estExt_fgenesh2_pg.C_2320058	O60258 Fibroblast growth factor 17	bfne018g21
Notch	estExt_gwp.C_110607	Q5SXM3 NOTCH1	bfne093p03
Delta	estExt_fgenesh2_pg.C_1200038	O00548 Delta-like 1	bfga003d05
Neural plate border specifiers			
Dlx	estExt_GenewiseH_1.C_530082	P56177 DLX-1	bflv011k11
Msx	estExt_gwp.C_540071	P28360 MSX-1	bflv048l06
Zic (AmphiZic)	estExt_gwp.C_400285	Q2M3N1 ZIC1	bfne006l12
Pax3/7	fgenesh2_kg.scaffold_211000002	Q2PJS5 PAX7	bfne062d03
Neural patterning/differentiation genes			
SoxB1-a *	estExt_fgenesh2_kg.C_6920001	P48431 SOX-2	bfne037g13
SoxB1-b	estExt_fgenesh2_pg.C_6920002	P48431 SOX-2	bfne072e07
SoxB1-c	e_gw.50.122.1	P48431 SOX-2	bfne126o15
SoxB2	estExt_GenewiseH_1.C_500067	O95416 SOX-14	bfad048c06
NeuroD	gw.499.29.1	Q13562 Neurogenic differentiation factor 1	N/A
Neurogenin	estExt_fgenesh2_pg.C_90049	Q92886 Neurogenin 1	N/A
Ash-a-like1	e_gw.302.11.1	P50553 Achaete-scute homolog 1	N/A
Ash-a-like2	e_gw.302.22.1	P50553 Achaete-scute homolog 1	N/A
Beta3	fgenesh2_pg.scaffold_122000082	Q8NFJ8 Basic helix-loop-helix protein 5	N/A
Olig-a	estExt_fgenesh2_pg.C_8060002	Q7RTU3 Oligodendrocyte transcription factor 3	bfne042a07
Olig-b	estExt_fgenesh2_pg.C_8060003	Q7RTU3 Oligodendrocyte transcription factor 3	bfne044b12
Olig-c	estExt_fgenesh2_pg.C_8060004	Q7RTU3 Oligodendrocyte transcription factor 3	bfne004n21
Islet	estExt_fgenesh2_kg.C_1460001	P61371 Islet-1	bfne013h18
Hu/elav	fgenesh2_pg.scaffold_219000066	Q12926 ELAV-like protein 2	bflv027g23

Neural crest specifiers			
Snail	estExt_fgenesh2_kg.C_190001	O43623 Zinc finger protein SLUG	bfne095n03
FoxD	estExt_fgenesh2_kg.C_2440002	Q9UJU5 FOXD3	bfne002k18
Twist	fgenesh2_pm.scaffold_217000008	Q8WVJ9 Twist-related protein 2	bfne115j15
Myc	fgenesh2_pm.scaffold_201000003	P01106 Myc proto-oncogene protein	bfne011k03
Id	fgenesh2_pg.scaffold_166000081	Q02363 ID-2	bfne018j05
AP-2	e_gw.327.57.1	P05549 AP-2 alpha	bfne111f20
SoxE	e_gw.5.674.1	P48436 SOX-9	bfne111n21
Neural crest effector genes			
Mitf	e_gw.492.5.1	O75030 MITF	bflv016p23
Tyrosinase	estExt_gwp.C_3220050	P14679 Tyrosinase	bfga039h23
Tyrp1	fgenesh2_pg.scaffold_1412000002	P17643 TRP-1	bfga030e07
Tyrp2	estExt_fgenesh2_pg.C_2120049	P40126 L-dopachrome tautomerase precursor	bfga006j06
RhoA/B/C-a	estExt_fgenesh2_pm.C_1920005	P61586 RHOA	bfga002n08
RhoA/B/C-b	fgenesh2_pg.scaffold_92000066	P61586 RHOA	bfne019m11
Ret	SAL_fgenesh2_pg.scaffold_92000070†	P07949 RET	N/A
Erbb	estExt_fgenesh2_pg.C_120089	P21860 ErbB-3	N/A
Col2a	fgenesh2_pg.scaffold_53000100	P02458 alpha 1 type II collagen	bflv014e16
cKit	Not found		
Myelin protein P0	Not found		

N/A: not available

* This gene was named AmphiSox1/2/3 in previous literatures

† This gene model was annotated manually by combining two predicted gene models fgenesh2_pg.scaffold_92000069 and fgenesh2_pg.scaffold_92000070.

Table S2. Expression of EGFP in chick embryos electroporated at stage 3+ to 4.

Stage examined	Somites/paraxial mesoderm	Notochord	Neural tube (hindbrain)	Head mesenchyme
Stage 8 (n=4)	4/4	4/4	0/4	4/4
Stage 10 (n=5)	5/5	3/5*	5/5	4/5**
Stage 11 (n=4)	4/4	4/4	4/4	4/4

* The two embryos with no EGFP in the notochord did not show any expression of RFP in the notochord.

** The embryo with no EGFP expression in the head mesenchyme did not show any expression of RFP in the head mesenchyme.