A Novel N-terminal Motif is Responsible for the Evolution of Neural Crest-Specific Gene-Regulatory Activity in Vertebrate FoxD3

January 2014

Hiroki ONO

A Novel N-terminal Motif is Responsible for the Evolution of Neural Crest-Specific Gene-Regulatory Activity in Vertebrate FoxD3

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Science

Hiroki ONO

Table of contents

Abstract

The neural crest is unique to vertebrates and has allowed the evolution of their complicated craniofacial structure. During vertebrate evolution, the acquisition of the neural crest must have been accompanied by the emergence of a new gene regulatory network (GRN). Here, to investigate the role of protein evolution in the emergence of the neural crest GRN, I examined the neural crest cell (NCC) differentiation-inducing activity of chordate *FoxD* genes. Amphioxus and vertebrate (*Xenopus*) FoxD proteins both exhibited transcriptional repressor activity in Gal4 transactivation assays and bound to similar DNA sequences in vitro. However, whereas vertebrate FoxD3 genes induced the differentiation of ectopic NCCs when overexpressed in chick neural tube, neither amphioxus FoxD nor any other vertebrate FoxD paralogs exhibited this activity. Experiments using chimeric proteins showed that the N-terminal portion of the vertebrate FoxD3 protein is critical to its NCC differentiation-inducing activity. Further-more, replacement of the N-terminus of amphioxus FoxD with a 39-amino-acid segment from zebrafishFoxD3 conferred neural crest-inducing activity on amphioxus FoxD or zebrafish FoxD1. Therefore, fixation of this N-terminal amino acid sequence may have been crucial in the evolutionary recruitment of FoxD3 to the vertebrate neural crest GRN.

Introduction

Evolutionary developmental biology (evo-devo) has played an important role in elucidating morphological evolution. Especially, through a clarifying the details of gene regulation of the development, evo-devo studies have contributed to elucidate the molecular context that give rise to morphological novelty. The widely-accepted concept of developmental evolution is the "genetic toolkit" (Carroll, 2001) "Toolkit" genes are transcription factor and signaling molecules, that regulators patterning of body and body parts. Notably, "toolkit" genes are broadly conserved among diverse organisms. Therefore, alteration of when and where the "toolkit" genes are expressed during embryogenesis is important for morphological evolution. As shown in a lot of evo-devo studies which have paid attention for *cis*-regulatory element of protein coding genes (Carroll, 2001, 2005; Davidson, 2006).

However, Kawashima et al. (2009) have pointed out that novel genes produced by domain shuffling may also play a critical role in the evolution of novel structures. They showed that genes acquired in the common ancestors of chordates are involved in the development of their characteristic features. In the common ancestors of the vertebrates, for example, the genes encoding Aggrecan, Occludin, and Tectorin alpha were built up by domain shuffling and were perhaps involved in the evolution of cartilage, tight

junctions, and tectorial membranes, respectively (Kawashima et al., 2009).

Novel sequence motifs in transcription factors have also been implicated in the evolution of morphologic features. For example, the glutamine–alanine-rich sequence (QA domain) of insect Ultrabithorax protein is thought to have been important in the evolutionary loss of abdominal appendages (Galant and Carroll, 2002; Ronshaugen et al., 2002). Similarly, the N-terminal motif of the Daphnia Antennapedia protein has also been implicated in the evolution of their specific appendage morphology (Shiga et al., 2002). Lynch et al. (2008) presented evidence that modification to HoxA-11 was essential in the evolution of mammalian pregnancy, as the modified protein has acquired a novel regulatory relationship with the prolactin gene. These studies have revealed that the evolution of morphology is driven not only by the molecular evolution of *cis*-regulatory elements but also by the evolution of protein coding sequences.

Neural crest cells are vertebrate embryonic cell population that originates from ectoderm between neural plate and non-neural ectoderm (neural plate border). Neural crest cells migrate throughout the embryo, and differentiate into numerous cell types. Neural crest cell is vertebrate novelty and first arose in the ancestors of vertebrates and have performed a central role in the evolution of vertebrates, particularly in their complicated craniofacial structures (Gans and Northcutt, 1983). The gene regulatory

network (GRN) underlying neural crest cell differentiation has been intensively studied. The transcription factor genes that are expressed at the neural plate border, including Dlx, Zic, Pax3/7 and Msx, are termed the “neural plate border specifiers”. These neural plate border specifiers define a region between neural plate and non-neural ectoderm, where give rise to neural crest cells (Meulemans and Bronner-Fraser, 2004). The transcription factor genes that are expressed in pre-migratory and migrating neural crest cells, including Slug/Snail, Foxd3, AP-2, Sox9/10 and Twist are termed the “neural crest specifiers”. These neural crest specifiers act downstream of neural plate border specifiers, and regulate the fate of neural crest cell by controlling the expression of neural crest effectors, such as cadherins and collagens (Meulemans and Bronner-Fraser, 2004). Notably, in protochordates (both amphioxus and ascidians), homologs of the neural plate border specifiers are expressed in the border region between the neural and non-neural ectoderm (Holland et al., 1996; Wada et al., 1997; Aniello et al., 1999; Holland et al., 1999; Sharman et al., 1999; Caracciolo et al., 2000; Gostling and Shimeld, 2003; Meulemans and Bronner-Fraser, 2004; Wada and Makabe, 2006; Yu et al., 2008). In contrast, homologs of the neural crest specifiers (with the exception of snail/slug) are not expressed in the corresponding regions of protochordates; thus, the neural crest specifiers are likely to be new recruits to the neural crest GRN (Langeland et al., 1998;

Imai et al., 2002; Meulemans and Bronner-Fraser, 2002, 2004; Yu et al., 2004, 2008; Wada and Makabe, 2006; Meulemans and Bronner-Fraser, 2007; Wada, 2010). It has been proposed that by co-opting neural crest specifier genes into a pre-existing neural plate border specification genetic network during early vertebrate evolution, cells at the neural plate border region acquired new cellular properties, such as migration and the ability to differentiate into diverse cell types, and evolved into neural crest cells (Meulemans and Bronner-Fraser, 2004, 2005; Yu, 2010). This idea is supported by recent experiments in ascidians showing that ectopic expression of homolog of one of the neural crest specifier genes (Twist) can reprogram neural plate border-derived pigment cells into migratory mesenchymal cells (Abitua et al., 2012). During this process of co-option, some transcription factors may have continued to regulate the same downstream genes that they regulated in the ancestral context, only now also in NCCs. In addition, they may have acquired new target genes, possibly by gaining the ability to physically interact with other transcription factors. This process would have activated new target genes in the NCCs that were not activated in the ancestral context. Thus, I reason that neofunctionalization of transcription factors might be accompanied by the evolutionary fixation of new sequence motifs, particularly those involved in intermolecular interactions.

In the present study, I focused on the transcription factor FoxD3 (Forkhead box D3). Because two rounds of genome duplication occurred during the evolution of vertebrates (Putnam et al., 2008), most vertebrate neural crest specifiers have several paralogs in vertebrate species but only a single homolog in protochordate species (reviewed in Wada and Makabe, 2006). For some other neural crest specifiers, including Sox9/10, snail/slug, and AP-2, duplicate paralogs are expressed in vertebrate NCCs (Hilger-Eversheim et al., 2000; Linker et al., 2000; Hong and Saint-Jeannet, 2005), indicating that co-option of these genes occurred before the genome duplications. In contrast, among five known vertebrate paralogs of FoxD, only FoxD3 is expressed in the neural crest; the other paralogs have retained their ancestral chordate roles in the forebrain, somites, and notochord (Kos et al., 2001; Sasai et al., 2001; Yu et al., 2002; Yu, 2010).

Therefore, I decided to focus on FoxD3 in our attempts to detect specific amino acid sequences involved in the neofunctionalization of FoxD underlying neural crest specification. In the present study, I examined the molecular evolution underlying the neofunctionalization of FoxD3 by examining the NCC differentiation-inducing activity of genes of the FoxD family in vertebrates and amphioxus, the most basal group of chordates (Bourlat et al., 2006; Putnam et al., 2008). I found that overexpressed in chick neural tubes, only vertebrate FoxD3 induces the production of ectopic NCCs; neither

amphioxus FoxD nor any other vertebrate FoxD3 paralogs (such as FoxD1 or FoxD5) exhibit this activity. Furthermore, by assaying the activity of chimeric FoxD proteins, I identified the N-terminal region of the FoxD3 protein as the essential region for ectopic induction of NCCs. These results indicate that the involvement of FoxD3 in the GRN of NCC differentiation was accompanied by fixation of the N-terminal sequence motif. Our findings constitute the first evidence linking the evolution of vertebrate NCCs to the molecular evolution of a specific protein sequence.

Materials and Methods

FoxD constructs

FoxD constructs for chick electroporation were made by inserting the complete open reading frames of the amphioxus FoxD (AmphiFoxD), zebrafish FoxD1 (zFoxD1), zebrafish FoxD3 (zFoxD3), zebrafish FoxD5 (zFoxD5), *Xenopus* FoxD1 (xFoxD1), *Xenopus* FoxD2 (xFoxD2), *Xenopus* FoxD3 (xFoxD3), mouse *FoxD4* (mFoxD4) into the expression vector pCAGGS (Momose et al., 1999). The complete open reading frame of zFoxD1, zFoxD3 and zFoxD5 were amplified from total cDNA of adult zebrafish by PRC using each primer (Table 1). The complete open reading frames of xFoxD1, xFoxD2 and AmphiFoxD were amplified from pCS2^{+} vectors inserted xFoxD1, xFoxD2, mFoxD4 and AmphiFoxD respectively by PCR using each primer (Table 1). The amplified FoxD genes were digested by restriction enzymes and inserted into pCAGGS vector by using T4-DNA ligase (Promega).

Chimeric protein constructs were produced by amplifying partial cDNA fragments and inserting them into pCAGGS. Partial lamprey FoxD-A gene was amplified from total ammocoete larva cDNA of *Lethenteron reissneri* collected in GOGYO rever in Tochigi prefecture by PCR using primers (Table 2). Primer sequences and restriction enzyme sites were shown in Table 2. The sequences of the chimeric constructs are shown in

Figures 3-8, respectively. I confirmed that no mutation occurred during plasmid construction by sequencing.

Plasmid preparation and electroporation of plasmid DNA into chick neural tubes

Plasmid DNAs were transfected into *Escherichia coli* cultured in LB (Becton, Dickinson and Company) medium 16 hours. After cultured, Plasmid DNAs were extracted by using QIAGEN Plasmid Midi Kit or Maxi Kit (QIAGEN). Plasmid DNA was electroporated into chick neural tubes essentially as described in Wada et al. (2006). Circular plasmid DNA (3 mg/ml) was injected into the neural tube lumen of chick embryos at Hamburger–Hamilton (HH) stage 09 at the level of the trunk, and five square pulses of 20 mV were applied for 50 msec each. 24 hours after electroporation, the embryos (at HH stage 20–22) were fixed for staining. In order to visualize efficiency of electroporation, GFP expression vector (pCAGGS-GFP; Wada et al., 2006) was co-electroporated.

Immunohistology and in situ hybridization

After electroporation, embryos were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) at 4 ℃ for 40 hours, transferred through a

methanol/PBS gradient, and stored in 100% methanol at –20℃ until use. Specimens were sectioned after frozen in O.C.T. (Optimal Cutting Temperature) compounds (Sakura Finetek Japan) by using CM3050 III (Leica). In situ hybridization was performed on sectioned specimens following Wada et al. (2006). Immunohistochemical analysis was performed with monoclonal antibody of HNK-1 (mouse IgM, Tucker et al., 1988), and polyclonal antibody against GFP (Clonetech).

Results

Overexpression of vertebrate FoxD1, FoxD2, FoxD4, FoxD5, and amphioxus AmphiFoxD do not induce ectopic NCC differentiation in chick embryo

After gene duplication, five vertebrate FoxD paralogs had undergone sub-functionalization and had shared the ancestral function in mesodermal differentiation (Yu et al., 2002; Yu, 2010). In addition, only FoxD3 acquired novel function in neural crest differentiation aside from mesodermal differentiation through neo-functionalization (Yu et al., 2002, Yu 2010). In Hox genes studies, functional redundancy among paralog genes has been shown (Condie et al., 1994; Greer et al., 2000; Tvrdik et al., 2006). On the other hand, Lynch et al. (2008) showed functional difference in mammalian HoxA11 genes. Thus, I questioned whether FoxD family genes potentially have NCC induction activity. Kos et al. (2001) and Dottori et al. (2001) reported that overexpression of chicken FoxD3 in chick neural tubes induces the differentiation of ectopic NCCs, as assessed by the expression of the Sox10 transcription factor gene and the HNK-1 epitope. I first examined whether the overexpression of FoxD3 orthologs from other vertebrate species would exhibit the same activity when overexpressed in chick neural tube at the level of trunk. As shown in Fig. 1A-F, overexpression of *Xenopus* FoxD3 (xFoxD3) or zebrafish FoxD3 (zFoxD3) caused

marked upregulation of the HNK-1 epitope and Sox10 expression. Thus, FoxD3 orthologs from distant species of vertebrates can induce the production of ectopic NCCs when overexpressed in chick neural tube.

I next examined the activities of other vertebrate FoxD paralogs. As shown in Figure 1, neither zebrafish FoxD1 (zFoxD1), *Xenopus* FoxD1 (xFoxD1), *Xenopus* FoxD2 (xFoxD2), mouse FoxD4 or zebrafish FoxD5 (zFoxD5) upregulated HNK-1 or Sox10 expression when overexpressed in chick neural tubes (Fig. 1G–U). Referring to the phylogeny of the FoxD gene family (Yu et al., 2002), I surmised that the sequence motif for ectopic induction of NCCs became fixed only in FoxD3 orthologs after the vertebrate genome duplications. In support of this conclusion, the overexpression of AmphiFoxD also failed to induce any upregulation of HNK-1 or Sox10 expression (Fig. 1V–X).

The N-terminal sequence of FoxD3 is critical for NCC induction

The amino acid sequence of the DNA-binding, winged-helix motif (WHM) of FoxD3 is highly conserved; only one amino acid substitution is specific to the FoxD3 paralogs (Fig. 9). Thus, differences in the sequence outside of WHM are likely to be responsible for specialization of FoxD3 paralog functions. Therefore, to identify the amino acid sequence motif of FoxD3 responsible for NCC induction, I tested the activity of two

chimeric proteins in which the portion of zFoxD3 N-terminal or C-terminal to the WHM was replaced with the corresponding portion of AmphiFoxD. The chimera Z3-Z3-A contains the zFoxD3 (Z) sequence N-terminal to the WHM, the zFoxD3 (Z) WHM, and the AmphiFoxD (A) sequence C-terminal to the WHM (Fig. 2A, Fig. 3). The inverse chimera A-Z3-Z3 contains the AmphiFoxD (A) sequence N-terminal to the WHM, the zFoxD3 (Z) WHM, and the zFoxD3 (Z) sequence C-terminal to the WHM (Fig. 2A, Fig. 4). I found that the overexpression of the Z3-Z3-A FoxD3 chimera in chick neural tube induced the differentiation of ectopic NCCs, as shown by marked upregulation of HNK-1 and Sox10 expression (Fig. 2B-D), the A-Z3-Z3 FoxD3 chimera failed to significant NCC inducing activity (Fig. 2E-G). Although some A-Z3-Z3 embryos did have a small number of ectopic NCCs, the induction activity was rather low relative to that of normal zFoxD3. Thus, I concluded that the portion of the protein N-terminal to the WHM is critical for the NCC differentiation-inducing activity of FoxD3.

An amino acid sequence alignment of the N-terminal portion of FoxD proteins revealed that N-terminus is conserved in FoxD3 but not in other vertebrate paralogs or in amphioxus FoxD (Fig. 2T), suggesting that this conserved region might be important for FoxD3 function. To examine this hypothesis, I produced a chimeric FoxD protein in which the N-terminal 39 amino acids of AmphiFoxD were replaced with the

corresponding amino acids of zFoxD3. This modified AmphiFoxD protein (designated Z3A-A-A, Fig. 2A, Fig. 5) induced differentiation of ectopic NCCs when overexpressed in chick neural tube (Fig. 2H-J), confirming that evolutionary changes in the N-terminal 39 amino acids would have been sufficient to confer NCC differentiation-inducing activity on the ancestral FoxD transcription factor. Similarly, zFoxD1 protein in which the N-terminal 39 amino acids were replaced with those of zFoxD3 (Z3Z1-Z1-Z1, Fig. 2A, Fig. 6) also induced differentiation of ectopic NCCs (Fig. 2K-M). On the other hand, zFoxD3 whose N-terminal 39 amino acids were replaced with those from AmphiFoxD (AZ3-Z3-Z3, Fig. 2A, Fig. 7) scarcely induced ectopic NCCs (Fig. 2N–P). Thus, N-terminal 39 amino acids are necessary for FoxD3 to induce NCC differentiation.

Searches against the NCBI (http://www.ncbi.nlm.nih.gov/guide/proteins/) and PFam protein databases (http://pfam.sanger.ac.uk/) yielded no proteins other than FoxD proteins containing sequences similar to the N-terminal 39-aa sequence of zFoxD3.

I then asked when the conserved N-terminal sequence was fixed in chordate evolution. FoxD from ascidian *Ciona* shows expression in melanocytes and endodermal cells (Imai et al., 2002; Abitua et al., 2012). Ciona FoxD has a highly divergent sequence in N-terminal region, and no conservation observed (Fig. 2T). Thus, the fixation of the N-terminal sequence is likely to have occurred after the divergence of vertebrates from

invertebrate chordates.

Lamprey was reported to possess a FoxD family gene (FoxD-A) that is expressed during neural crest Q4 differentiation (Sauka-Spengler et al., 2007). The N-terminal sequence of lamprey FoxD-A is moderately conserved with those of other vertebrate FoxD paralogues (Fig. 2T). I tested the activity of the N-terminal sequence of the lamprey FoxD-A by a fusion construct with AmphiFoxD (Fig. 2A, Fig. 8), and found that the lamprey N-terminal sequence do not provide HNK-1/Sox10 inducing activity to amphioxus FoxD (Fig. 2Q–S). Therefore, lamprey FoxD-A may not be able to substitute for the role of gnathostome FoxD3 in the context of chick neural tube.

Discussion

Neofunctionalization of transcription factors

The evolution of development is fundamentally attributable to evolving gene regulation changes. It is generally accepted that morphological evolution is driven by co-option of toolkit genes. In other words, acquisitions of novel expression domain of transcription factor through changes in *cis*-regulatory element contribute to altering gene regulation (Carroll et al., 2001; Davidson, 2006). On the other hand, mutations in protein-coding region of transcription factors are barely considered as a driving force of morphological evolution. Because comparing with *cis*-regulatory element change, mutations in protein-coding region have extensive pleiotropic negative effects during development (Lynch et al., 2008, Lynch and Wagner, 2008).

But some evo-devo studies have shown that transcription factors gain a novel function, with conserving ancestral function, through evolving new functional domain (Galant et al., 2002; Lynch et al., 2008). Because transcription factors often regulate gene expression with other transcription factors, new functional domain act as novel interface of protein-protein interaction and contribute to get new target genes. In protein-mediated evolution, this novel interaction with other transcription factors might be essential for acquiring new function.

There are evidences of FoxD genes conserving ancestral function. (1) Amphioxus and vertebrate FoxD cognates get involved in mesoderm differentiation (Yu et al. 2002; Yuasa et al., 1996; Mariani and Harland, 1998; Gomez-Skarmeta et al., 1999; Scheucher et al., 1995; Wu et al., 1998; Chang and Kessler, 2010; Sullivan et al., 2001), (2) amphioxus and vertebrate FoxD cognates act as transcriptional repressors that bind to similar DNA sequences (Ono et al. 2013), (3) FoxD3 is known to work primarily as a transcriptional repressor via a Groucho-like repressor-interaction motif in its C-terminal domain (Sutton et al., 1996; Pohl and Knöchel, 2001; Sasai et al., 2001; Steiner et al., 2006; Yaklichkin et al., 2007; but note that in some context, it was suggested that vertebrate FoxD3 functions as a transcriptional activator; e.g., Liu and Labosky, 2008). This motif is required for FoxD3 to induce the differentiation of dorsal mesoderm in Xenopus embryos (Yaklichkin et al., 2007) and is conserved in AmphiFoxD. In addition, our findings suggest that it is required for FoxD genes to play their ancestral role in mesoderm development and transcription factor FoxD3 underwent "additive manner" of functional evolution via protein changes during the acquisition of its novel ability to induce NCC differentiation.

The NCC differentiation-inducing function of FoxD3 is unique to vertebrates, and has arisen through the fixation of a specific N-terminal amino acid sequence not present in

AmphiFoxD or *Ciona* FoxD. And the result of domain searches suggest N-terminal amino acid sequence of vertebrate FoxD3 has evolved through short linear motifs (SLiM) switches (Neduva and Russell, 2005; Lohr et al., 2001; Galant and Carroll, 2002), neither domain shuffling (Kawashima et al., 2009) nor simple sequence repeats (SSRs) (Sears et al. 2007). Because of its short lengh and discontinuous arrangement of amino acids, in contrast to normal structural domain and SSRs, SLiMs are hard to identify. To find out the advanced neural crest inducing amino acid sequence, more experimental procedures will be needed (e.g., single amino acid replacement experiments). And also, I found that, although lampreys possess migratory neural crest cells, the N-terminal sequence of the lamprey FoxD-A did not provide HNK-1/Sox10-inducing activity when fused with AmphiFoxD. This observation may reflect the variation in the distal part of the lamprey neural crest gene regulatory network compared with that in gnathostomes (Sauka-Spengler et al., 2007; Nikitina and Bronner-Fraser, 2009). In the lamprey embryo, several neural crest specifier genes including c-Myc, Id, AP2 and Snail are deployed earlier than FoxD3 and SoxE family genes, suggesting that the regulatory linkages among lamprey neural crest specifier genes might be slightly different. Alternatively, this lack of activity may simply be due to technical issues; i.e., N-terminal portion of the lamprey FoxD-A may perform the same role during neural crest

differentiation, but just cannot work in the cellular context of the chick neural tube, possibly due to the divergence of the amino acid sequence in the counterpart proteins.

In either case, this N-terminal amino acid sequence must constitute a new interface critical for FoxD3 to function in the GRN of NCC differentiation. Thomas and Erickson (2009) indicated that FoxD3 represses Mitf expression in avian neural crest cells, and thus suppress neural crest cells from differentiation into pigment cells. This effect of FoxD3 on Mitf expression is not dependent on the DNA binding, but on sequestration of Pax3. Abitua et al. (2012) showed that ascidian FoxD also suppresses Mitf expression. Moreover, they indicated that its portion N-terminal to WHM is sufficient for this suppression. These studies may suggest that the N-terminal sequence unique to vertebrate FoxD3 may be involved in the interaction with Pax3 or other transcription factors, and those interactions may confer the new functions of FoxD3 protein in vertebrate neural crest development.

Evolution of the neural crest GRN

For those interested in the evolutionary origin of vertebrates, an understanding of the evolution of the neural crest GRN is critical. That the neural crest regulatory genes can be divided into neural plate border specifiers and NCC specifiers illuminates the

stepwise evolution of the neural crest GRN. Because protochordate neural plate border specifiers, like those of vertebrates, are expressed in the corresponding region between the neural and non-neural ectoderm (Meulemans and Bronner-Fraser, 2004; Yu et al., 2008; Yu, 2010), their eventual involvement in NCC differentiation would not require a change in their expression patterns. Thus, as the first step in the evolution of the neural crest GRN, the border specifiers have to recruit a set of genes (neural crest specifiers) as their downstream targets. These genes may not have been recruited simultaneously. Duplicate paralogs of SoxE, snail/slug, and AP-2 are expressed in NCCs, indicating that recruitment of these genes to the neural crest GRN occurred before the genome duplications (Wada and Makabe, 2006). In contrast, among the five known vertebrate FoxD paralogs, only FoxD3 is expressed in the neural crest (Yu et al., 2002,2004; Wada and Makabe, 2006). Therefore, FoxD3 might have been recruited slightly later than the other neural crest specifiers, after the genome duplications. The second step in the evolution of the neural crest GRN might be the acquisition of target effector genes, such as cadherin and collagen genes, for the neural crest specifiers. Interestingly, these effector genes appear to have been present during the vertebrate genome duplications but, in several cases, only certain paralogs were recruited as neural crest effectors (e.g., cadherin6, cadherin7, col2a1, and rhoB), suggesting that neofunctionalization of some

effectors to NCC development occurred after the genome duplications (Wada and Makabe, 2006). Actually, cadherin7 was suggested as direct FoxD3 target (Dottori et al. 2001). Therefore, the neural crest GRN may have been completed by the recruitment of some novel target genes after the genome duplications. During its evolution, the neural crest GRN must have gained several new regulatory interactions, probably through the acquisition of new cis-regulatory regions by target genes (Yu et al., 2008). In addition, because most of the transcription factor genes in the neural crest GRN function not only in NCCs but also in other cells, interactions between transcription factors may be essential for NCC-specific regulation of target gene expression. Our FoxD fusion construct studies have shown that the N-terminal region of FoxD3 is critical for its role in neural crest development. SoxE, on the other hand, may not have a fixed motif specific to neural crest development, because Drosophila SoxE can substitute functionally for vertebrate SoxE in NCC differentiation (Cossais et al., 2010). Examination of the neural crest GRN from the aspect of interactions between transcription factors may shed new light on neural crest evolution, and will provide more general insight on how novel GRNs emerged during evolution.

Acknowledgement

I thank Jr-Kai Yu for the pCS2^{+} plasmids inserted *Xenopus* FoxD cognates, Yoshio Wakamatsu for pBluescript inserted chicken Sox10, Hiroshi Wada for the leads, helpful discussion of this research, and all colleagues of Wada's lab. for technical advices.

References

Abitua, P.B., Wagner, E., Navarrete, I.A., Levine, M., 2012. Identification of a rudimentary neural crest in a non-vertebrate chordate. Nature 492, 104–107.

Aniello, F., Locascio, A., Villani, M.G., Gregorio, A.D., Fucci, L., Branno, M., 1999. Identification and developmental expression of Ci-msxb: a novel homologue of Drosophila msh gene in Ciona intestinalis. Mech. Dev. 88, 123–126.

Bourlat, S.J., Juliusdottir, T., Lowe C.J., Freeman, R., Aronowicz, J., Kirschner, M., Lander, E.S., Thorndyke, M., Nakano, H., Kohn, A.B., Heyland, A., Moroz, L.L., Copley, R.R., Telford, M.J., 2006. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. Nature 444, 85–88.

Caracciolo, A., Gregorio, A.D., Aniello, F., Lauro, R.D., Branno, M., 2000. Identification and developmental expression of three Distal-less homeobox containing genes in the ascidian Ciona intestinalis. Mech. Dev. 99, 173–176.

Carroll, S.B., Greiner, J.K., Weatherbee, S.D., 2001. From DNA to diversity. Blackwell, Malden.

Carroll, S.B., 2005. Evolution at two levels: on genes and form. PLoS Biol 3,1159–1166

Chang, L.L., Kassler, D.S., 2010. Foxd3 is an essential Nodal-dependent regulator of zebrafish dorsal mesoderm development. Dev. Biol. 342:39-50

Condie, B.G., Capecchi, M.R., 1994. Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions. Nature 370:304–307.

Cossais, F., Sock, E., Horning, J., Schreiner, S., Kellerer, S., Bösl, M.R., Russel, S., Wegner, M., 2010. Replacement of mouse Sox10 by the Drosophila ortholog Sox100B provides evidence for co-option of SoxE proteins into vertebrate-specific gene-regulatory networks through altered expression. Dev. Biol. 341, 267–281.

Davidson, E.H., 2006. The Regulatory Genome. Academic Press, San Diego.

Dottori, M., Gross, M.K., Labosky, P., Goulding, M., 2001. The winged-helix transcription factor Foxd3 supresses interneuron differentiation and promotes neural crest cell fate. Development 128, 4127–4138.

Galant, R., Carroll, S.B., 2002. Evolution of a transcriptional repression domain in an insect Hox protein. Nature 415, 910–913.

Gans, C., Northcutt, R.G., 1983. Neural crest and the origin of vertebrates: a new head. Science 220, 268–274.

Gomez-Skarmeta JL, de la Calle-Mustienes E, Modolell J, Mayor R., 1999. *Xenopus* brain factor-2 controls mesoderm, forebrain and neural crest development. Mech. Dev, 80,15-27

Gostling, N.J., Shimeld, S.M., 2003. Protochordate Zic genes define primitive somite

compartments and highlight molecular changes underlying neural crest evolution. Evol. Dev. 5, 136–144.

Greer, J.M., J. Puetz,. Thomas K.R., Capecchi, M.R., 2000.Maintenance of functional equivalence during paralogous Hox gene evolution. Nature 403, 661–665.

Hilger-Eversheim, K., Moser, M., Schorle, H., Buettner, R., 2000. Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. Gene 260, 1–12.

Holland, L.Z., Schubert, M., Kozmik, Z., Holland, N.D., 1999. AmphiPax3/7, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and the possible evolutionary precursor of definitive vertebrate neural crest. Evo. Dev. 1, 153–165.

Holland, N.D., Panganiban, G., Henyey, E.L., Holland, L.Z., 1996. Sequence and developmental expression of AmphiDll, an amphioxus Distal-less gene transcribed in the ectoderm, epidermis and nervous system: insights into evolution of craniate forebrain and neural crest. Development 122, 2911–2920.

Hong, C.-S., Saint-Jeannet, J.-P., 2005. Sox proteins and neural crest development. Sem. Dev. Biol., 16.

Imai, K.S., Satoh, N., Satou, Y., 2002. An essential role of FoxD gene in notochord

induction in Ciona embryos. Development 129, 3441–3453.

Kawashima, T., Kawashima, S., Tanaka, C., Murai, M., Yoneda, M., Putnam, N.H., Rokhsar, D.S., Kanehisa, M., Satoh, N., Wada, H., 2009. Domain shuffling and the evolution of vertebrates. Genome Res. 19, 1393–1403.

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.

Langeland, J., Tomsa, J.M., Jackman, W.R., Kimmel, C.B., 1998. An amphioxus snail gene: expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. Dev. Genes Evol. 208, 569–577.

Linker, C., Bronner-Fraser, M., Mayor, R., 2000. Relationship between gene expression domains of Xsnail, Xslug, and Xtwist and cell movement in the prospective neural crest of Xenopus. Dev. Biol. 224, 215–225.

Liu, Y., Labosky, P.A., 2008. Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. Stem Cells 26, 2475–2484.

Lohr, U., Yussa, M., Pick, L., 2001. Drosophila fushi tarazu: a gene on the border of homeotic function. Curr. Biol. 11,1403–1412.

Lynch, V.J., Tanzer, A., Wang, Y., Leung, F.C., Gellersen, B., Emera, D., Wagner, G.P.,

2008. Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. Proc. Natl. Acad. Sci. U.S.A. 105, 14928–14933.

Lynch, VJ., Wagner, GP. 2008. Resurrecting the role of transcription factor change in developmental evolution. Evolution 62:2131-2154

Mariani FV, Harland RM. 1998. XBF-2 is a transcriptional repressor that converts ectoderm into neural tissue. Development 125:5019–5031.

Meulemans, D., Bronner-Fraser, M., 2002. Amphioxus and lamprey AP-2 genes: implications for neural crest evolution and migration pattern. Development 129, 4953–4962.

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

Meulemans, D., Bronner-Fraser, M., 2005. Central role of gene cooption in neural crest evolution. J. Exp. Zool. B Mol. Dev. Evol. 304, 298–303.

Meulemans, D., Bronner-Fraser, M., 2007. Insights from amphioxus into the evolution of vertebrate cartilage. PLoS ONE 2, e787.

Momose, T., Tonegawa, A., Takeuchi, J., Ogawa, H., Umesono, K., Yasuda, K., 1999. Efficient targeting of gene expression in chick embryos by microelectroporation. Dev. Growth Differ. 41, 335–344.

Neduva, V., Russell, R.B., 2005. Linear motifs: evolutionary interaction switches. FEBS Lett. 579:3342–3345.

Nikitina, N.V., Bronner-Fraser, M., 2009. Gene regulatory networks that control the specification of neural-crest cells in the lamprey. Biochem. Biophys. Acta 1789, 274–278.

Pohl, B.S., Knöchel, W., 2001. Overexpression of the transcriptional repressor FoxD3 prevents neural crest formation in Xenopus embryos. Mech. Dev. 103, 93–106.

Putnam, N.H., Butts, T., Ferrier, D.E.K., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.K., Benito-Gutirrez, E., Dubchak, I., Garcia-Fernendez, J., Gibson-Brown, J.J., Grigoriev, I.V., Horton, A.C., de Jong, P.J., Jurka, J., Kapitonov, V.V., Kohara, Y., Kuroki, Y., Lindquist, E., Lucas, S., Osoegawa, K., Pennacchio, L.A., Salamov, A.A., Satou, Y., Sauka-Spengler, T., Schmutz, J., Shin-I, T., Toyoda, A., Bronner-Fraser, M., Fujiyama, A., Holland, L.Z., Holland, P.W.H., Satoh, N., Rokhsar, D.S., 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453, 1064–1071.

Ronshaugen, M., McGinnis, N., McGinnis, W., 2002. Hox protein mutation and macroevolution of the insect body plan. Nature 415, 914–917.

Sasai, N., Mizuseki, K., Sasai, Y., 2001. Requirement of FoxD3-class signaling for

neural crest determination in Xenopus. Development 128, 2525–2536.

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

Scheucher M, Dege P, Lef J, Hille S, Kno ¨chel W. 1995. Transcription patterns of four different fork head/HNF-3 related genes (XFD-4, 9 and 10) in Xenopus laevis embryos. Rouxs Arch Dev. Biol. 204,203–211.

Sears, K. E., A. Goswami., Flynn, J.J., Niswander, L.A., 2007. The correlated evolution of Runx2 tandem repeats, transcriptional activity, and facial length in carnivora. Evol. Dev. 9,555–565.

Sharman, A.C., Shimeld, S.M., Holland, P.W.H., 1999. An amphioxus Msx gene expressed predominantly in the dorsal neural tube. Dev. Genes Evol. 209, 260–263.

Shiga, Y., Yasumoto, R., Yamagata, H., Hayashi, S., 2002. Evolving role of antennapedia protein in arthropod limb patterning. Development 129, 3555–3561.

Steiner, A.B., Engleka, M.J., Lu, Q., Piwarzyk, E.C., Yaklinchkin, S., Lefebvre, J.L., Walters, J.W., Pineda-Salgado, L., Labosky, P.A., Kessler, D.S., 2006. FoxD3 regulation of nodal in Spemann organizer is essential xenopus dorsal mesoderm development. Development 133, 4827–4838.

Sullivan SA, Akers L, Moody SA. 2001. foxD5a,a Xenopus winged helix gene, maintains an immature neural ectoderm via transcrip-tional repression that is dependent on the C-terminal domain. Dev Biol 232:439–457.

Sutton, J., Costa, R., Klung, M., Field, L., Xu, D., Largaespada, D.A., Fletcher, C.F., Jenkins, N.A., Copeland, N.G., Klemsz, M., Hromas, R., 1996. Genesis, a winged helix transcriptional repressor with expression restricted to embryonic stem cells. J. Biol. Chem. 271, 23126–23133.

Thomas, A.J., Erickson, C.A., 2009. FOXD3 regulates the lineage switch between neural crest-derived glial cells and pigment cells by repressing MITF through a non-canonical mechanism. Development 136, 1849–1858.

Tvrdik, P. and M. R. Capecchi. 2006. Reversal of Hox1 Gene Subfunctionalization in the Mouse. Develop. Cell 11:239–250.

Wada, H., Makabe, K., 2006. Gnome duplications of early vertebrates as a possible chronicle of the evolutionary history of the neural crest. Int. J Biol. Sci. 2,133-141

Wada, H., 2010. Origin and genetic evolution of the vertebrate skeleton. Zool. Sci. 27, 119–123.

Wada, H., Escriva, H., Zhang, S., Laudet, V., 2006. Conserved RARE localoization in amphioxus Hox clusters and implications for Hox code evolution in the vertebrate

neural crest. Dev. Dyn. 235, 1522–1531.

Wada, H., Holland, P.W.H., Sato, S., Yamamoto, H., Satoh, N., 1997. Neural tube is partially dorsalized by overexpression of HrPax-37: the ascidian homologue of Pax-3 and Pax-7. Dev. Biol. 187, 240–252.

Wada, H., Makabe, K., 2013. Genome duplications of early vertebrates as a possible chronicle of the evolutionary history of the neural crest. Int. J. Bio. Sci(In press).

Wu, S.C., Grindley, J., Winnier, G.E., Hargett, L., Hogan, B.L., 1998. Mouse Mesenchyme forkhead 2 (Mf2): expression, DNA binding and induction by sonic hedgehog during somitogenesis. Mech. Dev. 70, 3–13.

Yaklichkin, S., Steiner, A.B., Lu, Q., Kessler, D.S., 2007. FoxD3 and Grg4 physically interact to repress transcription and induce mesoderm in Xenopus. J. Biol. Chem. 282, 2548–2557.

Yu, J.-K., Holland, N.D., Holland, L.Z., 2002. An amphioxus winged helix/forkhea gene, AmphiFoxD: insights into vertebrate neural crest evolution. Dev. Dyn. 225, 289–297.

Yu, J.-K., Holland, N.D., Holland, L.Z., 2004. Tissue-specific expression of FoxD reporter constructs in amphioxus embryos. Dev. Biol. 274, 452–461.

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

1132.

Yu, J.K.S., 2010. The evolutionary origin of the vertebrate neural crest and its developmental gene regulatory network – insights from amphioxus. Zoology 113, 1–9.

Yuasa J, Hirano S, Yamagata M, Noda M. 1996. Visual projection map specified by topographic expression of transcription factors in the retina. Nature 382:632–635.

Tables and Figures

Table 1

Primer names	Primer sequences (underlines are restriction site)	Restriction enzyme names
Amphioxus FoxD F	GGGAATTCATGCTTCTCGAGGCGGACGC	EcoRI
Amphioxus FoxD R	GGGTCGACTCACGAGTGCACGTGCGGCCA	SalI
zebrafish FoxD1 F	GGGTCGACATGTCGGACAGTTCTGCTCT	SalI
zebrafish FoxD1 R	GGGGATCCCTAGAAATGGCAATTGTTAAG	BamHI
zebrafish FoxD3 F	GGCTCGAGATGACCCTGTCTGGAGGCAC	XhoI
zebrafish FoxD3 R	GGGGATCCTCATTGAGAAGGCCATTTCGA	BamHI
zebrafish FoxD5 F	GGGAATTCATGACCCTCTCCCAGGATTA	EcoRI
zebrafish FoxD5 R	GGCTCGAGTCAACAGTGAGGATAAACCAT	XhoI
Xenopus FoxD1 F	GGGAATTCATGACTCTGAGCTCTGACAT	EcoRI
Xenopus FoxD1 R	GGCTCGAGCTAGTGGTTTGTAAGCACCGT	XhoI
Xenopus FoxD2 F	GGCTCGAGATGACTTTGGGCACAGAGAT	XhoI
Xenopus FoxD2 R	GGGATATCTTAAAACTCACAGCTCTTAAG	EcoRV
Xenopus FoxD3 F	GGGAATTCATGACCCTGTCAAGCAGCGG	EcoRI
Xenopus FoxD3 R	GGCTCGAGTTATTGCGCTGGCCATTTGGC	XhoI
mouse FoxD4 F	GGCTCGAGATGAACTCAGCAAGAGCTGG	XhoI
mouse FoxD4 R	GGGATATCTTAAAATTCGGGCAAGGTCCC	EcoRV

Table 2

Primer names (constructs name)	Primer sequences (underlines are restriction site)	Restriction enzyme names
Z3-Z3-A zebrafish R	GGGTCGACCCTGAGAATGTCCGGCTGATG	SalI
Z3-Z3-A Amphioxus F	GGCTCGAGCCCACGGCCTTCATGGCGGCC	XhoI
A-Z3-Z3 zebrafish F	GGTCTAGAATGACCCTGTCTGGAGGCAC	XbaI
A-Z3-Z3 Amphioxus F	GGTCTAGAATGCTTCTCGAGGCGGACGC	XbaI
A-Z3-Z3 Amphioxus R	GGGAGCTCTCCACGTCTGTATTCTCCGCG	SacI
Z3A-A-A zebrafish F	GGGAATTCATGACCCTGTCTGGAGGCAC	EcoRI
Z3A-A-A zebrafish R	GGTCTAGAGTCCTGCTCCATCCCCTCGTC	XbaI
Z3A-A-A Amphioxus	GGGCTAGCCAGGGGAGCCATCCACAGGGC	NheI
Z3Z1-Z1-Z1 dfoxd3 F	GGCTCGAGATGACCCTGTCTGGAGGCACC	XhoI
Z3Z1-Z1-Z1 dfoxd3 R	GGGCTAGCACTGTCCTGCTCCATCCCCTC	NhaI
Z3Z1-Z1-Z1 dfoxd1 F	GGTCTAGATTGGACAATGACTCCGATGAC	XbaI
Z3Z1-Z1-Z1 dfoxd1 R	GGGGATCCCTAGAAATGGCAATTGTTAAG	BamHI
AZ3-Z3-Z3 dfoxd3 F	GGGATATCGACTGCGAAAGCCAGTGCATG	EcoRV
AZ3-Z3-Z3 dfoxd3 R	GGGGATCCTCATTGAGAAGGCCATTTCGA	BamHI
AZ3-Z3-Z3 afoxd F	GGGTCGACATGCTTCTCGAGGCGGACGCC	SalI
AZ3-Z3-Z3 afoxd R	GGGATATCGCTGGTCATCTCCCGGGGAAG	EcoRV
LA-A-A lfoxda F	GGGAATTCATGACCCCGCTCTCCGGGTCC	EcoRI
LA-A-A lfoxda R	GGGCTAGCAGCGTCGTCGCTGTCACC	NheI
LA-A-A adoxd F	GGGAATTCGCTACGAGCCAGGGGAGCCAT	EcoRI, NheI
LA-A-A adoxd R	GGGATATCTCACGAGTGCACGTGCGG	EcoRV

Figure 1

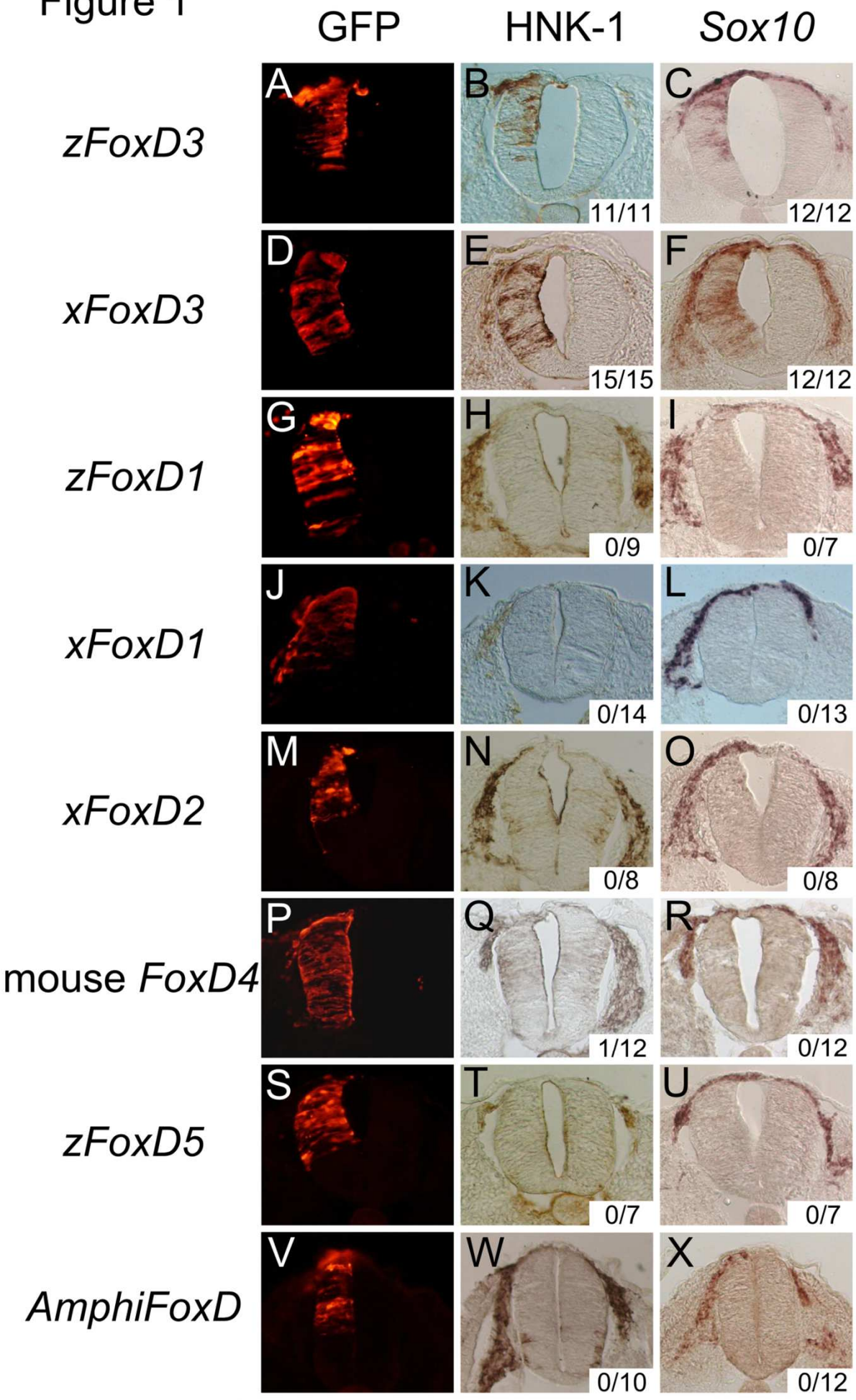
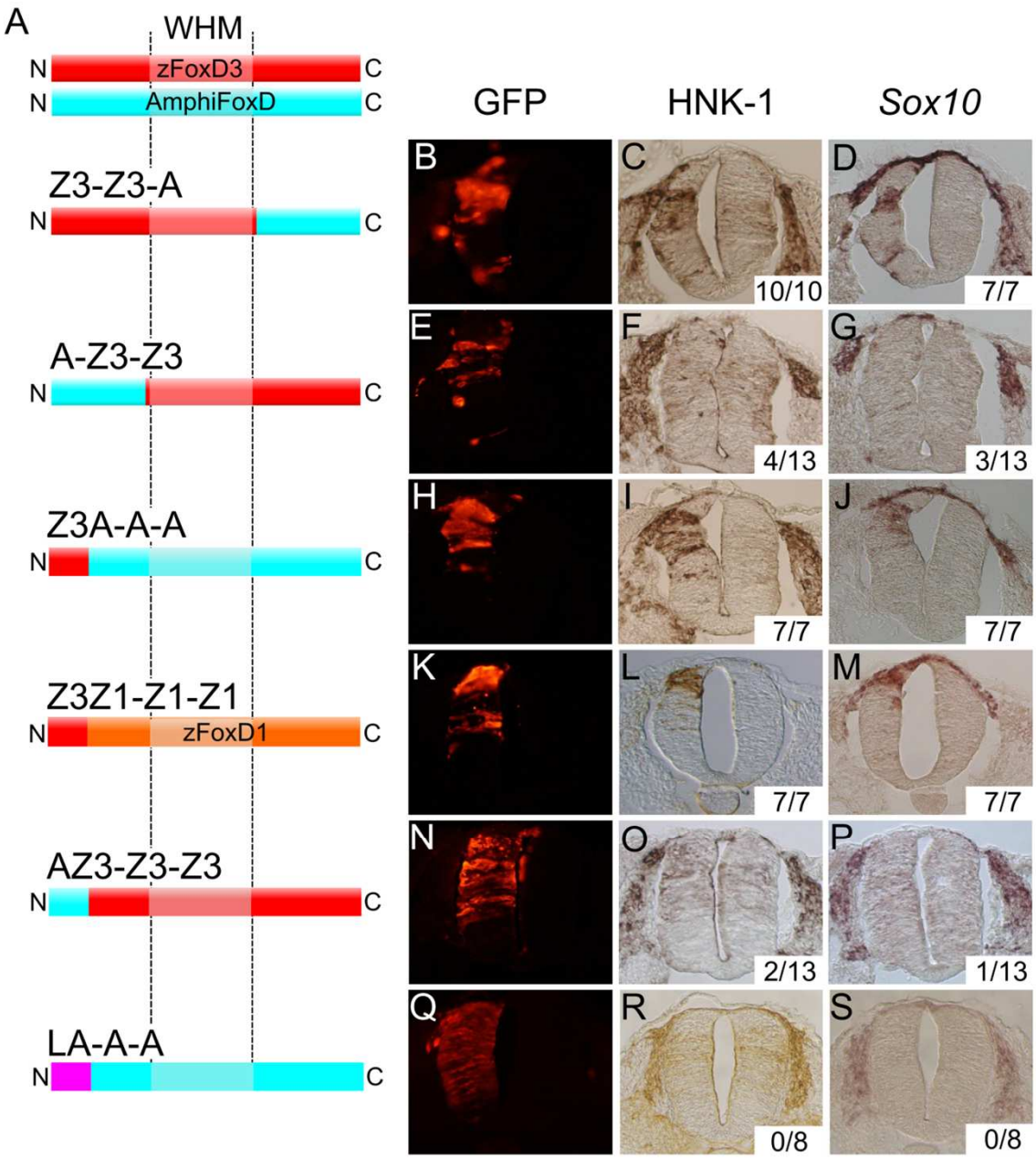

GFP
HNK-1
Sox10
zFoxD3
xFoxD3
zFoxD1
xFoxD1
xFoxD2
mouse FoxD4
zFoxD5
AmphiFoxD
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
11/11
12/12
15/15
12/12
0/9
0/7
0/14
0/13
0/8
0/8
1/12
0/12
0/7
0/7
0/10
0/12

Figure 2


A
WHM
N
zFoxD3
C
N
AmphiFoxD
C
Z3-Z3-A
A-Z3-Z3
Z3A-A-A
Z3Z1-Z1-Z1
zFoxD1
AZ3-Z3-Z3
LA-A-A
GFP
HNK-1
Sox10
B
C
D
10/10
7/7
E
F
G
4/13
3/13
H
I
J
7/7
7/7
K
L
M
7/7
7/7
N
O
P
2/13
1/13
Q
R
S
0/8
0/8

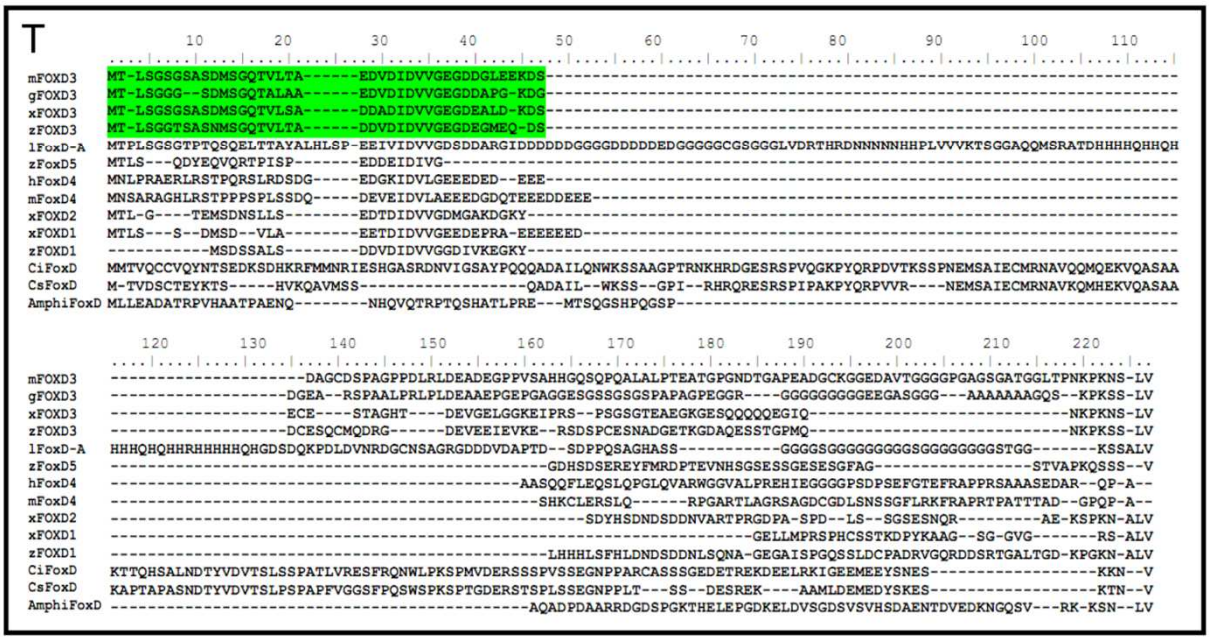

T
mFOXD3
gFOXD3
xFOXD3
zFOXD3
lFoxD-A
zFoxD5
hFoxD4
mFoxD4
xFOXD2
xFOXD1
zFOXD1
CiFoxD
CsFoxD
AmphiFoxD

Figure 3
Z3-Z3-A

Figure 4
A-Z3-Z3

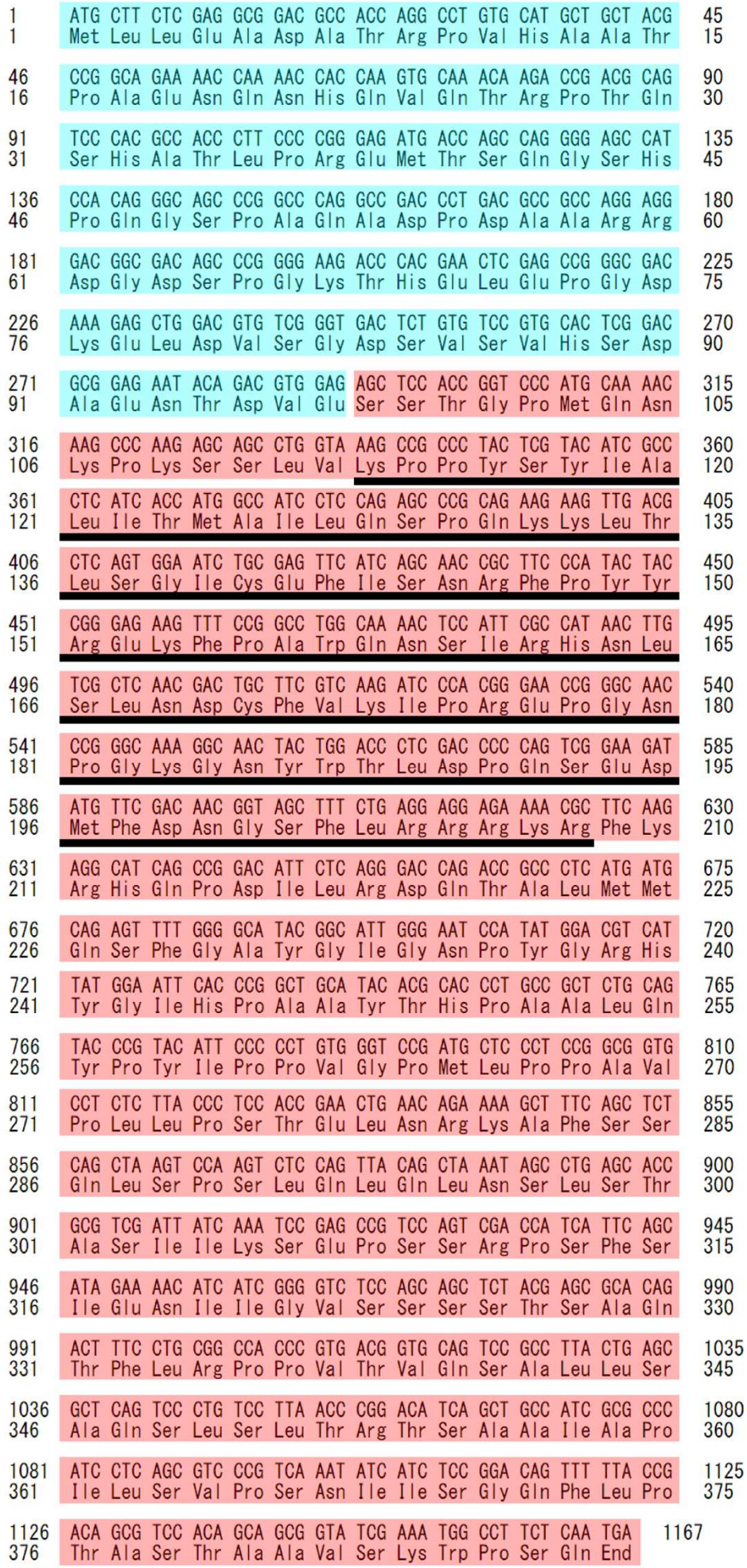


Figure 5

Z3A-A-A

```
1     ATG ACC CTG TCT GGA GGC ACC AGT GCC AGC AAC ATG TCC GGT CAG   45
1     Met Thr Leu Ser Gly Gly Thr Ser Ala Ser Asn Met Ser Gly Gln   15

46    ACC GTG CTC ACA GCT GAC GAT GTG GAT ATC GAC GTG GTC GGG GAG   90
16    Thr Val Leu Thr Ala Asp Asp Val Asp Ile Asp Val Val Gly Glu   30

91    GGT GAC GAG GGG ATG GAG CAG GAC TCT AGC CAG GGG AGC CAT CCA   135
31    Gly Asp Glu Gly Met Glu Gln Asp Ser Ser Gln Gly Ser His Pro   45

136   CAG GGC AGC CCG GCC CAG GCC GAC CCT GAC GCC GCC AGG AGG GAC   180
46    Gln Gly Ser Pro Ala Gln Ala Asp Pro Asp Ala Ala Arg Arg Asp   60

181   GGC GAC AGC CCG GGG AAG ACC CAC GAA CTC GAG CCG GGC GAC AAA   225
61    Gly Asp Ser Pro Gly Lys Thr His Glu Leu Glu Pro Gly Asp Lys   75

226   GAG CTG GAC GTG TCG GGT GAC TCT GTG TCC GTG CAC TCG GAC GCG   270
76    Glu Leu Asp Val Ser Gly Asp Ser Val Ser Val His Ser Asp Ala   90

271   GAG AAT ACA GAC GTG GAA GAC AAG AAT GGA CAG TCT GTA CGG AAG   315
91    Glu Asn Thr Asp Val Glu Asp Lys Asn Gly Gln Ser Val Arg Lys   105

316   AAA TCC AAC CTT GTG AAA CCG CCG TAC TCT TAC ATA GCT CTC ATT   360
106   Lys Ser Asn Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile   120

361   ACC ATG TCA ATT CTG CAG TCT CCG CAG AAG AAA CTG ACT CTC AGC   405
121   Thr Met Ser Ile Leu Gln Ser Pro Gln Lys Lys Leu Thr Leu Ser   135

406   CAG ATC TGT GAG TTT ATC ATG AAC AGG TTT CCT TAC TAC AGG GAG   450
136   Gln Ile Cys Glu Phe Ile Met Asn Arg Phe Pro Tyr Tyr Arg Glu   150

451   CGG TTC CCG GTG TGG CAG AAC TCC ATC CGA CAC AAC CTC TCG CTG   495
151   Arg Phe Pro Val Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu   165

496   AAC GAC TGC TTC GTG AAG ATC CCG CGC GAG CCT GGC AAC CCC GGC   540
166   Asn Asp Cys Phe Val Lys Ile Pro Arg Glu Pro Gly Asn Pro Gly   180

541   AAG GGC AAC TAC TGG ACG CTG GAC CCC GCC AGC GAG GAC ATG TTC   585
181   Lys Gly Asn Tyr Trp Thr Leu Asp Pro Ala Ser Glu Asp Met Phe   195

586   GAC AAC GGC AGT TTC CTG CGG CGC CGA AAA CGG TTC AAG CGG CAG   630
196   Asp Asn Gly Ser Phe Leu Arg Arg Arg Lys Arg Phe Lys Arg Gln   210

631   GCT CCG GAC GTG TTA CGG GAG CCC ACG GCC TTC ATG GCG GCC ACG   675
211   Ala Pro Asp Val Leu Arg Glu Pro Thr Ala Phe Met Ala Ala Thr   225

676   GAT CCG TAC AGA CAC CAC CTG GGT CTG ATC CAC CCG CAC CAC CAC   720
226   Asp Pro Tyr Arg His His Leu Gly Leu Ile His Pro His His His   240

721   CCT CAC CCA GCG GCG CTG CCC TAC CAC TAC ATG TCC CCG CTG CCG   765
241   Pro His Pro Ala Ala Leu Pro Tyr His Tyr Met Ser Pro Leu Pro   255

766   CCG CCC GTC CCC CTG CCC CTC CCC CAC GCG CCG ACC GCC GCA GAC   810
256   Pro Pro Val Pro Leu Pro Leu Pro His Ala Pro Thr Ala Ala Asp   270

811   TTC GCA CGG ACG CAG GCG CTG GCC GCG CAG ATC GCC GGG GGA GTC   855
271   Phe Ala Arg Thr Gln Ala Leu Ala Ala Gln Ile Ala Gly Gly Val   285

856   GGC GCC TTC GCC TCG GTG GGC GGG TTG ACC CTG CCC GTC ACC ACC   900
286   Gly Ala Phe Ala Ser Val Gly Gly Leu Thr Leu Pro Val Thr Thr   300

901   CCC GTC ACG ACG CAC CGG CCG GCG GGG TTC TCC ATA GAA AAC ATC   945
301   Pro Val Thr Thr His Arg Pro Ala Gly Phe Ser Ile Glu Asn Ile   315

946   ATC GGG AGC AGC GCT GCC AGC GAC AAG ACT GTC TCC ACC ACC TTC   990
316   Ile Gly Ser Ser Ala Ala Ser Asp Lys Thr Val Ser Thr Thr Phe   330

991   TCC ATC AGC ACG ACG GGA GCA CCC GCG TTC CGC CCC ACC GTG TCG   1035
331   Ser Ile Ser Thr Thr Gly Ala Pro Ala Phe Arg Pro Thr Val Ser   345

1036  GTC CCC GCC ACC ATC CCG GTC TGC GCC ACC GGA CTC AGA CCC CCG   1080
346   Val Pro Ala Thr Ile Pro Val Cys Ala Thr Gly Leu Arg Pro Pro   360

1081  GAC TCC TTA CCG TTC GGC GGC GGG ACC AGC GCC TTC ACC TCC CCC   1125
361   Asp Ser Leu Pro Phe Gly Gly Gly Thr Ser Ala Phe Thr Ser Pro   375

1126  CTC CAC ATG GAC CTG GAG AAG TAC AGG CAA TGT CTG CAG TGC AAT   1170
376   Leu His Met Asp Leu Glu Lys Tyr Arg Gln Cys Leu Gln Cys Asn   390

1171  GGC AGC GTC CCT TCC TGG CCG CAC GTG CAC TCG TGA   1206
391   Gly Ser Val Pro Ser Trp Pro His Val His Ser End
```

Figure 6
Z3Z1-Z1-Z1

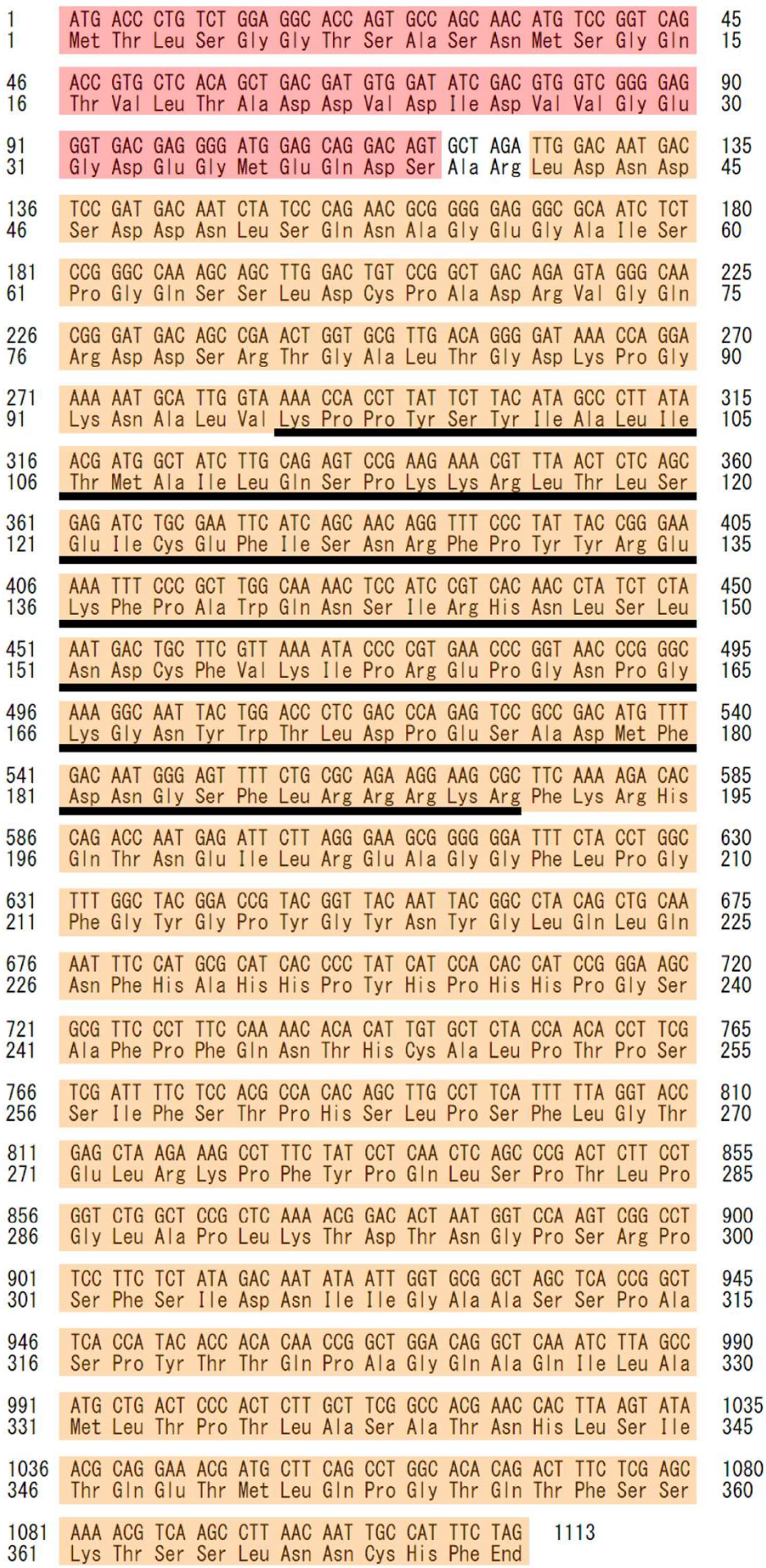


Figure 7
AZ3-Z3-Z3

```
1     ATG CTT CTC GAG GCG GAC GCC ACC AGG CCT GTG CAT GCT GCT ACG  45
1     Met Leu Leu Glu Ala Asp Ala Thr Arg Pro Val His Ala Ala Thr  15

46    CCG GCA GAA AAC CAA AAC CAC CAA GTG CAA ACA AGA CCG ACG CAG  90
16    Pro Ala Glu Asn Gln Asn His Gln Val Gln Thr Arg Pro Thr Gln  30

91    TCC CAC GCC ACC CTT CCC CGG GAG ATG ACC AGC GAT ATC GAC TGC  135
31    Ser His Ala Thr Leu Pro Arg Glu Met Thr Ser Asp Ile Asp Cys  45

136   GAA AGC CAG TGC ATG CAG GAC CGG GGA GAT GAG GTG GAG GAG ATC  180
46    Glu Ser Gln Cys Met Gln Asp Arg Gly Asp Glu Val Glu Glu Ile  60

181   GAG GTG AAG GAG CGC AGC ACC AGT CCC TGC GAG AGC AAC GCT GAC  225
61    Glu Val Lys Glu Arg Ser Thr Ser Pro Cys Glu Ser Asn Ala Asp  75

226   GGA GAG ACC AAG GGG GAT GCT CAG GAG AGC TCC ACC GGT CCC ATG  270
76    Gly Glu Thr Lys Gly Asp Ala Gln Glu Ser Ser Thr Gly Pro Met  90

271   CAA AAC AAG CCC AAG AGC AGC CTG GTA AAG CCG CCC TAC TCG TAC  315
91    Gln Asn Lys Pro Lys Ser Ser Leu Val Lys Pro Pro Tyr Ser Tyr  105

316   ATC GCC CTC ATC ACC ATG GCC ATC CTC CAG AGC CCG CAG AAG AAG  360
106   Ile Ala Leu Ile Thr Met Ala Ile Leu Gln Ser Pro Gln Lys Lys  120

361   TTG ACG CTC AGT GGA ATC TGC GAG TTC ATC AGC AAC CGC TTC CCA  405
121   Leu Thr Leu Ser Gly Ile Cys Glu Phe Ile Ser Asn Arg Phe Pro  135

406   TAC TAC CGG GAG AAG TTT CCG GCC TGG CAA AAC TCC ATT CGC CAT  450
136   Tyr Tyr Arg Glu Lys Phe Pro Ala Trp Gln Asn Ser Ile Arg His  150

451   AAC TTG TCG CTC AAC GAC TGC TTC GTC AAG ATC CCA CGG GAA CCG  495
151   Asn Leu Ser Leu Asn Asp Cys Phe Val Lys Ile Pro Arg Glu Pro  165

496   GGC AAC CCG GGC AAA GGC AAC TAC TGG ACC CTC GAC CCC CAG TCG  540
166   Gly Asn Pro Gly Lys Gly Asn Tyr Trp Thr Leu Asp Pro Gln Ser  180

541   GAA GAT ATG TTC GAC AAC GGT AGC TTT CTG AGG AGG AGA AAA CGC  585
181   Glu Asp Met Phe Asp Asn Gly Ser Phe Leu Arg Arg Arg Lys Arg  195

586   TTC AAG AGG CAT CAG CCG GAC ATT CTC AGG GAC CAG ACC GCC CTC  630
196   Phe Lys Arg His Gln Pro Asp Ile Leu Arg Asp Gln Thr Ala Leu  210

631   ATG ATG CAG AGT TTT GGG GCA TAC GGC ATT GGG AAT CCA TAT GGA  675
211   Met Met Gln Ser Phe Gly Ala Tyr Gly Ile Gly Asn Pro Tyr Gly  225

676   CGT CAT TAT GGA ATT CAC CCG GCT GCA TAC ACG CAC CCT GCC GCT  720
226   Arg His Tyr Gly Ile His Pro Ala Ala Tyr Thr His Pro Ala Ala  240

721   CTG CAG TAC CCG TAC ATT CCC CCT GTG GGT CCG ATG CTC CCT CCG  765
241   Leu Gln Tyr Pro Tyr Ile Pro Pro Val Gly Pro Met Leu Pro Pro  255

766   GCG GTG CCT CTC TTA CCC TCC GCC GAA CTG AAC AGA AAA GCT TTC  810
256   Ala Val Pro Leu Leu Pro Ser Ala Glu Leu Asn Arg Lys Ala Phe  270

811   AGC TCT CAG CTA AGT CCA AGT CTC CAG TTA CAG CTA AAT AGC CTG  855
271   Ser Ser Gln Leu Ser Pro Ser Leu Gln Leu Gln Leu Asn Ser Leu  285

856   AGC ACC GCG TCG ATT ATC AAA TCC GAG CCG TCC AGT CGA CCA TCA  900
286   Ser Thr Ala Ser Ile Ile Lys Ser Glu Pro Ser Ser Arg Pro Ser  300

901   TTC AGC ATA GAA AAC ATC ATC GGG GTC TCC AGC AGT CTA CGA GCG  945
301   Phe Ser Ile Glu Asn Ile Ile Gly Val Ser Ser Ser Leu Arg Ala  315

946   ATA CAG ACT TTC CTG CGG CCA CCC GTG ACG GTG CAG TCC GCC TTA  990
316   Ile Gln Thr Phe Leu Arg Pro Pro Val Thr Val Gln Ser Ala Leu  330

991   CTG AGC GCT CAG TCC CTG TCC TTA ACC CGG ACA TCA GCT GCC ATC  1035
331   Leu Ser Ala Gln Ser Leu Ser Leu Thr Arg Thr Ser Ala Ala Ile  345

1036  GCG CCC ATC CTC AGC GTC CCG TCA AAT ATC ATC TCC GGA CAG TTT  1080
346   Ala Pro Ile Leu Ser Val Pro Ser Asn Ile Ile Ser Gly Gln Phe  360

1081  TTA CCG ACA GCG TCC ACA GCA GCG GTA TCG AAA TGG CCT TCT CAA  1125
361   Leu Pro Thr Ala Ser Thr Ala Ala Val Ser Lys Trp Pro Ser Gln  375

1126  TGA   1128
376   End
```

Figure 8
LA-A-A

```
1    ATG ACC CCG CTC TCC GGG TCC GGG ACT CCG ACC CAG AGC CAG GAG   45
1    Met Thr Pro Leu Ser Gly Ser Gly Thr Pro Thr Gln Ser Gln Glu   15

46   CTG TCG ACC GCT TAC GCG TTG CAC CTC TCG CCC GAG GAG ATC GTC   90
16   Leu Ser Thr Ala Tyr Ala Leu His Leu Ser Pro Glu Glu Ile Val   30

91   ATC GAT GTG GTC GGT GAC AGC GAC GAC GCT GCT AGC AGC CAG GGG   135
31   Ile Asp Val Val Gly Asp Ser Asp Asp Ala Ala Ser Ser Gln Gly   45

136  AGC CAT CCA CAG GGC AGC CCG GCC CAG GCC GAC CCT GAC GCC GCC   180
46   Ser His Pro Gln Gly Ser Pro Ala Gln Ala Asp Pro Asp Ala Ala   60

181  AGG AGG GAC GGC GAC AGC CCG GGG AAG ACC CAC GAA CTC GAG CCG   225
61   Arg Arg Asp Gly Asp Ser Pro Gly Lys Thr His Glu Leu Glu Pro   75

226  GGC GAC AAA GAG CTG GAC GTG TCG GGT GAC TCT GTG TCC GTG CAC   270
76   Gly Asp Lys Glu Leu Asp Val Ser Gly Asp Ser Val Ser Val His   90

271  TCG GAC GCG GAG AAT ACA GAC GTG GAA GAC AAG AAT GGA CAG TCT   315
91   Ser Asp Ala Glu Asn Thr Asp Val Glu Asp Lys Asn Gly Gln Ser   105

316  GTA CGG AAG AAA TCC AAC CTT GTG AAA CCG CCG TAC TCT TAC ATA   360
106  Val Arg Lys Lys Ser Asn Leu Val Lys Pro Pro Tyr Ser Tyr Ile   120

361  GCT CTC ATT ACC ATG TCA ATT CTG CAG TCT CCG CAG AAG AAA CTG   405
121  Ala Leu Ile Thr Met Ser Ile Leu Gln Ser Pro Gln Lys Lys Leu   135

406  ACT CTC AGC CAG ATC TGT GAG TTT ATC ATG AAC AGG TTT CCT TAC   450
136  Thr Leu Ser Gln Ile Cys Glu Phe Ile Met Asn Arg Phe Pro Tyr   150

451  TAC AGG GAG CGG TTC CCG GTG TGG CAG AAC TCC ATC CGA CAC AAC   495
151  Tyr Arg Glu Arg Phe Pro Val Trp Gln Asn Ser Ile Arg His Asn   165

496  CTC TCG CTG AAC GAC TGC TTC GTG AAG ATC CCG CGC GAG CCT GGC   540
166  Leu Ser Leu Asn Asp Cys Phe Val Lys Ile Pro Arg Glu Pro Gly   180

541  AAC CCC GGC AAG GGC AAC TAC TGG ACG CTG GAC CCC GCC AGC GAG   585
181  Asn Pro Gly Lys Gly Asn Tyr Trp Thr Leu Asp Pro Ala Ser Glu   195

586  GAC ATG TTC GAC AAC GGC AGT TTC CTG CGG CGC CGA AAA CGG TTC   630
196  Asp Met Phe Asp Asn Gly Ser Phe Leu Arg Arg Arg Lys Arg Phe   210

631  AAG CGG CAG GCT CCG GAC GTG TTA CGG GAG CCC ACG GCC TTC ATG   675
211  Lys Arg Gln Ala Pro Asp Val Leu Arg Glu Pro Thr Ala Phe Met   225

676  GCG GCC ACG GAT CCG TAC AGA CAC CAC CTG GGT CTG ATC CAC CCG   720
226  Ala Ala Thr Asp Pro Tyr Arg His His Leu Gly Leu Ile His Pro   240

721  CAC CAC CAC CCT CAC CCA GCG GCG CTG CCC TAC CAC TAC ATG TCC   765
241  His His His Pro His Pro Ala Ala Leu Pro Tyr His Tyr Met Ser   255

766  CCG CTG CCG CCG CCC GTC CCC CTG CCC CTC CCC CAC GCG CCG ACC   810
256  Pro Leu Pro Pro Pro Val Pro Leu Pro Leu Pro His Ala Pro Thr   270

811  GCC GCA GAC TTC GCA CGG ACG CAG GCG CTG GCC GCG CAG ATC GCC   855
271  Ala Ala Asp Phe Ala Arg Thr Gln Ala Leu Ala Ala Gln Ile Ala   285

856  GGG GGA GTC GGC GCC TTC GCC TCG GTG GGC GGG TTG ACC CTG CCC   900
286  Gly Gly Val Gly Ala Phe Ala Ser Val Gly Gly Leu Thr Leu Pro   300

901  GTC ACC ACC CCC GTC ACG ACG CAC CGG CCG GCG GGG TTC TCC ATA   945
301  Val Thr Thr Pro Val Thr Thr His Arg Pro Ala Gly Phe Ser Ile   315

946  GAA AAC ATC ATC GGG AGC AGC GCT GCC AGC GAC AAG ACT GTC TCC   990
316  Glu Asn Ile Ile Gly Ser Ser Ala Ala Ser Asp Lys Thr Val Ser   330

991  ACC ACC TTC TCC ATC AGC ACG ACG GGA GCA CCC GCG TTC CGC CCC   1035
331  Thr Thr Phe Ser Ile Ser Thr Thr Gly Ala Pro Ala Phe Arg Pro   345

1036 ACC GTG TCG GTC CCC GCC ACC ATC CCG GTC TGC GCC ACC GGA CTC   1080
346  Thr Val Ser Val Pro Ala Thr Ile Pro Val Cys Ala Thr Gly Leu   360

1081 AGA CCC CCG GAC TCC TTA CCG TTC GGC GGC GGG ACC AGC GCC TTC   1125
361  Arg Pro Pro Asp Ser Leu Pro Phe Gly Gly Gly Thr Ser Ala Phe   375

1126 ACC TCC CCC CTC CAC ATG GAC CTG GAG AAG TAC AGG CAA TGT CTG   1170
376  Thr Ser Pro Leu His Met Asp Leu Glu Lys Tyr Arg Gln Cys Leu   390

1171 CAG TGC AAT GGC AGC GTC CCT TCC TGG CCG CAC GTG CAC TCG TGA   1215
391  Gln Cys Asn Gly Ser Val Pro Ser Trp Pro His Val His Ser End   405

1216 TAG  1218
406  End
```

Figure 9

```
                  10        20        30        40        50        60        70        80        90
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
mFOXD3            KPPYSYIALITMAILQSPMKKLTLSGICDFISNKFPYYKEKFPAWQNSIRHNLSLNDCFIKIPREPGNPGKGNYWSLDPASEDMFDNGSFLRRRKR
gFOXD3            KPPSSYIALITMAILQSPHKRLTLSGICAFISGRFPYYCRKFPAWQNSIRHNLSLNDCFVKIPREPGRPGKGNYWSLDPASQDMFDNGSFLRRRKR
xFOXD3            KPPYSYIALITMAILQSPHKRLTLSGICAFISGRFPYYRRKFPAWQNSIRHNLSLNDCFVKIPREPGHPGKGNYWSLDPASQDMFDNGSFLRRRKR
zFOXD3            KPPYSYIALITMAILQSPQKKLTLSGICEFISNRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPQSEDMFDNGSFLRRRKR
zFOXD1            KPPYSYIALITMAILQSPQKKLTLSGICEFISNRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPQSEDMFDNGSFLRRRKR
zFOXD1            KPPYSYIALITMSILQSPQKKLTLSGICEFISNRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPQSEDMFDNGSFLRRRKR
xFOXD2            KPPYSYIALITMAILQSPQKKLTLSGICEFISNRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPQSEDMFDNGSFLRRRKR
zFOXD5            KPPYSYIALITMSILQSPKKRLTLSEICEFISNRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPESADMFDNGSFLRRRKR
hFOXD4            KPPYSYIALITMSILQSPKKRLTLSEICDFISSRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPESADMFDNGSFLRRRKR
mFOXD4            KPPYSYIALITMAILQSPKKRLTLSEICEFISNRFPYYREKFPAWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPESADMFDNGSFLRRRKR
AmphiFOXD         KPPYSYIALITMSILQSPQKKLTLSQICEFIMNRFPYYRERFPVWQNSIRHNLSLNDCFVKIPREPGNPGKGNYWTLDPASEDMFDNGSFLRRRKR
```

Figure legends

Figure 1
Effect of FoxD overexpression on HNK-1 antigen and Sox10 expressions in chick neural tube. Upregulation of the HNK-1 epitope (middle column: B, E, H, K. N, Q, T,W) and Sox10 (right column: C, F, I, L, O, R, U, X) were induced by zebrafish FoxD3 (zFoxD3) and xFoxD3, but not by zFoxD1, xFoxD1, xFoxD2, mouse FoxD4, zFoxD5, or AmphiFoxD. Transfected cells were visualized by anti-GFP antibody in adjacent sections of embryos in which GFP-pCAGGS were co-electroporated (left column: A, D, G, J, M, P, S, V). Ectopic expression of the FoxD proteins was induced on the left-hand side of the neural tube. Numbers in the panel show the number of embryos in which marker overexpression was observed as a fraction of the number of embryos examined.

Figure 2
Effect of overexpression of chimeric FoxD proteins on Sox10 and HNK-1 epitope expression in chick neural tube. (A) Schematic illustrations of chimeric protein constructs, where amino acid segments from zFoxD3, AmphiFoxD, zFoxD1 and lamprey FoxD-A are shown in red, blue, orange and magenta respectively. Upregulation of the HNK-1 epitope (middle column: C, F, I, L, O, R) and Sox10 (right column: D, G, J, M, P, S) were induced by chimeric constructs: Z3-Z3-A, Z3A-A-A and Z3Z1-Z1-Z1, but only fairly induced by A-Z3-Z3, AZ3-Z3-Z3 or LA-A-A. Transfected cells were visualized by anti-GFP antibody in adjacent sections of embryos in which GFP-pCAGGS were co-electroporated (left column: B, E, H, K. N, Q). Ectopic expression of the FoxD proteins was induced on the left-hand side of the neural tube. Numbers in the panel show the number of embryos in which marker overexpression was observed as a fraction of the number of embryos examined. (T) Amino acid sequence alignment of the N-terminal portions of proteins encoded by genes of the FoxD family. The 39-aa N-terminal segment conserved in FoxD3 genes is shaded green.

Figure 3
The nucleotide sequence and amino acid sequence of chimeric FoxD constract Z3-Z3-A. The segment of zebrafish FoxD3 sequence is shaded magenta, and amphioxus FoxD sequence is shaded cyan. The location of DNA-binding motif, winged-helix motif is underlined.

Figure 4
The nucleotide sequence and amino acid sequence of chimeric FoxD constract A-Z3-Z3. The segment of zebrafish FoxD3 sequence is shaded magenta, and amphioxus FoxD sequence is shaded cyan. The location of DNA-binding motif, winged-helix motif is underlined.

Figure 5
The nucleotide sequence and amino acid sequence of chimeric FoxD constract Z3A-A-A. The segment of zebrafish FoxD3 sequence is shaded magenta, and amphioxus FoxD sequence is shaded cyan. The location of DNA-binding motif, winged-helix motif is underlined.

Figure 6
The nucleotide sequence and amino acid sequence of chimeric FoxD constract Z3Z1-Z1-Z1. The segment of zebrafish FoxD3 sequence is shaded magenta, and zebrafish FoxD1 sequence is shaded orange. The location of DNA-binding motif, winged-helix motif is underlined.

Figure 7
The nucleotide sequence and amino acid sequence of chimeric FoxD constract AZ3-Z3-Z3. The segment of zebrafish FoxD3 sequence is shaded magenta, and amphioxus FoxD sequence is shaded cyan. The location of DNA-binding motif, winged-helix motif is underlined.

Figure 8
The nucleotide sequence and amino acid sequence of chimeric FoxD constract LA-A-A. The segment of lamprey FoxD-A sequence is shaded purple, and amphioxus FoxD sequence is shaded cyan. The location of DNA-binding motif, winged-helix motif is underlined.

Figure 9
The alignment of amino acid sequence of the DNA-binding, winged-helix motif among FoxD cognates. Only one amino acid substitution is specific to the FoxD3 paralogs.