Evolution of *hes* gene family in vertebrates: *hes5* cluster genes were specifically increased in *Xenopus*

Aya Kuretani¹, Takayoshi Yamamoto², Masanori Taira¹,³ and Tatsuo Michiue¹,²*

¹ Graduate School of Science, The University of Tokyo 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. ² Graduate School of Arts and Sciences, The University of Tokyo 3-8-1, Komaba, Meguro-ku, Tokyo 153-8902, Japan. ³ Department of Biological Sciences, Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan

*To whom correspondence should be addressed (tmichiue@bio.c.u-tokyo.ac.jp).
Abstract

Background

*hes* genes are chordate homologs of *Drosophila* genes, *hairy* and *enhancer of split*, which encode a basic helix-loop-helix (bHLH) transcriptional repressor with a WRPW motif. Various developmental functions of *hes* genes, including early embryogenesis and neurogenesis, have been elucidated in vertebrates. However, their orthologous relationships remain unclear partly because of less conservation of relatively short amino acid sequences, less conserved synteny, and species-specific gene duplication. This results in complicated gene names in vertebrates, which are not consistent in orthologs. In a previous study, we revealed that *Xenopus* frogs have two clusters of *hes5*, named “the hes5.1 cluster” and “the hes5.3 cluster.” The origin has not yet been revealed.

Results

Here, we elucidated the orthologous and paralogous relationships of all *hes* genes of human, mouse, chicken, gecko, zebrafish, medaka, coelacanth, spotted gar, elephant shark, and *Xenopus* frogs (*X. tropicalis* and *X. laevis*) by phylogenic and synteny analysis. Any clusters of *hes5* were not found in amniotes, whereas duplicated *hes5* clusters in teleost were found although not as many genes as *Xenopus*. In addition, *hes5* cluster-like structure was found in the elephant shark genome, but not found in cyclostomata.

Conclusion

These data suggest that the *hes5* cluster existed in the gnathostome ancestor, but was lost in amniotes.
**Background**

*hes genes* are chordate homologs of *Drosophila hairy* and *enhancer of split* genes and encode a basic helix-loop-helix (bHLH) transcriptional repressor with a WRPW motif at the C terminus [1]. These genes are known to have various developmental functions, including Notch signaling target and neurogenesis [2], somitogenesis, and early development of the presumptive midbrain–hindbrain boundary (pre-MHB) [3, 4].

Mammals including human and mouse have seven *hes* genes which form subfamilies [5, 6]. Most of the *hes* homologs in zebrafish are called *her* [7]. As in zebrafish, their orthologs in vertebrates remain unclear partly because the two domains, bHLH and Orange domains, and the WRPW motif at the C terminus are not well conserved, and the sequences are relatively small to compare (their total sizes are about 200 aa). Another possible cause was that the genomes of the model organisms including *Xenopus laevis* (*X. laevis*) and *X. tropicalis* had yet to be sequenced.

Recently, many animal genome analyses, including frogs, *X. laevis* and *X. tropicalis*, have been reported. *Xenopus* includes diploid to dodecaploid species, although polyploidy is considered to be rare in amniotes. *X. tropicalis* has a diploid genome, and *X. laevis* has an allottetraploid genome [8]. The genomic analysis showed that the allottetraploidization was caused by interspecific crosses between two species that have a diploid genome. Thus, *X. laevis* has two subgenomes, called L and S [9,10].

In a previous study, we annotated all *hes* genes of *X. tropicalis* and *X. laevis* by phylogenetic analysis and synteny analysis [11]. In brief, for *X. tropicalis*, we revealed the phylogenetic and synteny relationships of the 18 *hes* genes and renamed them properly. *X. laevis* has 37 *hes* genes including 18 homeologs, one *laevis*-specific gene, *hes5.7*, and a pseudogene, *hes7.4*. Although the number of genes doubled after allottetraploidization, *hes* genes, except for *hes2*, have been conserved in *X. laevis*. In addition, *Xenopus* has more than two paralogs of *hes5, hes6*, and *hes7* subfamily genes,
in contrast to human \textit{hes} genes. In particular, the number of \textit{hes5} genes in \textit{Xenopus} is quite high. Interestingly, they form two clusters, which we call “the \textit{hes5.1} cluster” and “the \textit{hes5.3} cluster”. The \textit{hes5.3} cluster is formed with eight genes (\textit{hes5.3} to \textit{hes5.10}).

Clustered genes, such as the Hox gene cluster, the human $\beta$-globin gene cluster, and four clustered human growth hormone (hGH)/chorionic somatomammotropin genes, have various functions with unique regulatory mechanisms. The cluster is considered to be formed as a result of gene duplication and divergence [12, 13].

Some of \textit{hes} genes are already known to be indispensable in neurogenesis [4] and the genes are well conserved despite having many genes, forming two clusters at least in \textit{Xenopus}. This implies that the \textit{hes5} cluster may also play an important role during embryogenesis.

To understand the evolution and role of \textit{hes} genes in vertebrates, it is important to reveal the orthologous relationship. In this study, we first elucidated orthologous and paralogous relationships of the \textit{hes} gene family using phylogenetic and syntenic analyses of human, mouse, chicken, zebrafish, medaka, frogs (\textit{X. tropicalis} and \textit{X. laevis}), \textit{Gekko japonicus}, coelacanth, spotted gar, elephant shark, lamprey, and amphioxus. With this analysis, we revealed that \textit{hes} genes are specifically increased in \textit{Xenopus} and also discussed the evolution of the two \textit{hes5} clusters.

Results

Classification of \textit{hes} genes in sarcopterygian

Previous studies on the annotation of \textit{hes} genes have shown that there are ten \textit{hes5} paralogs in \textit{X. laevis}, which we refer to as "the \textit{hes5.1} cluster" or "the \textit{hes5.3} cluster" [11]. To determine when \textit{hes5} clusters emerged, we first performed a phylogenetic analysis of sarcopterygian \textit{hes5} genes (Fig. 1, Fig. S2; complete tree is shown in Fig. S4). In this
analysis, other hes paralogues were included to clarify the outgroup. Maximum likelihood (ML) phylogenetic tree construction revealed that all of the hes5 genes we observed were assigned in a single clade with high bootstrap percentage (93%; Fig. 1A).

In the hes5 clade, the genes, Hosahes5, Mumuhes5, Gagahes5chr21-3, Gejahes5sc135-1, which include human (Homo sapiens), mouse (Mus musculus), chick (Gallus gallus), and gecko (Gekko japonicus) genes, formed a single clade. Gagahes5chr21-1 and Gajahes5Sc135-2 also formed a single clade. These two clades were located close to each other in the tree (Fig. 1B, blue letters). Interestingly, Xenopus hes5.3-5.9 showed a monophyletic group, suggesting that gene duplication of hes5.3-5.9 occurred independently. In Coelacanth (Latimeria chalumnae (Lach)), four putative hes5 genes were found (red letters). Although one of them formed a clade (Lachhes5sc00059), three of these genes (Lachhes5sc00001 and Lachhes5sc00199) formed a clade with Xenopus hes5.3-5.9. This result indicates that three coelacanth genes may be related to the hes5.3 cluster.

Next, to examine the presence of hes5 clusters, we analyzed synteny of hes5 locus in chicken, geckos, and coelacanth genomes (synteny of other hes genes are shown in Fig. S1A, C). In the chicken genome, hes5 genes were located on a single chromosome, chromosome 21 (Fig. 2B). In gecko, synteny around hes5 was observed in scaffolds 135 and 31595 (Fig. 2C). In coelacanth, we found four hes5 genes in scaffold00199, 00001, 00319, and 00059 (Fig. 2D). Hes5chr21-1-3 genes in chick, hes5sc135-1-2, and hes5-like genes in gecko, and hes5sc00319 gene in coelacanth were all located next to pank4, suggesting that these genes correspond to the hes5.1 cluster (orange background). In chicken and gecko, however, there were no hes5 genes between nol9 and zbtb48, defined as hes5.3 cluster genes in Xenopus (blue background). In contrast, the coelacanth hes5sc00199 gene was located near nol9. This result suggested that coelacanth hes5sc00199 may be homologous to the hes5.3 cluster gene. In coelacanth Sc00001, a
hes5 gene was found near chd5 (Fig. 2D). In Xenopus, the chd5 gene (chd5-like) was located next to rnf207 near the hes5 clusters, suggesting the relevance of the coelacanth gene to the hes5 clusters. Lachhes5c00059 was found near ppi2, which is located on the 1st chromosome in Xenopus, indicating that the syntenic property was different from other hes5 genes. Phylogenetic analysis also indicated that Lachhes5sc00059 was first divided in the hes5 gene family, suggesting a distinct evolution of this gene. Together with these results, it is suggested that the hes5 genes of both chicken and gecko correlate more with the hes5.1 cluster, whereas coelacanth hes5 genes are located in both hes5.1 and hes5.3 clusters.

**Comparison of hes genes between teleosts and Xenopus**

It is known that whole genome duplication (WGD) occurred 500 million year ago in the common ancestor of vertebrates. In addition to this, in teleost, another WGD occurred 3.7 million years ago after divergence from the common ancestor of gnathostomata [14, 15]. Thus, in teleost genome, the two loci having similar gene order to each other, which are called doubly conserved synteny, are often found. [16]. In zebrafish (Danio renio) and medaka/Japanese ricefish (Oryzias latipes), hes genes have not been well characterized as mammalian orthologue. Indeed, many genes that seem to be hes were described as "her" genes. Therefore, we attempted to identify the orthologous relationship of teleost hes genes based on their amino acid sequences. By our phylogenetic analysis, we found that many zebrafish and medaka “her” genes formed a clade with the Xenopus hes subfamily genes (Fig. 3A; complete tree is shown in Fig. S5; detailed gene annotations are shown in Table S1). Zebrafish her6 and medaka her6/her6.2 formed a single clade with Xenopus hes1. Medaka her4 and zebrafish her9 were located near Xenopus hes4 clade. Orlaher8.2 and Dareher8.2/8a showed high similarity with Xenopus hes6.2. Dareher13, Darehes6, Orlahes62orf2, and orlahes6 belonged to a clade with
Xenopus hes6.1/human hes6 genes. The genes annotated as hes3 and hes2 in zebrafish, medaka, and Xenopus belonged to each homologous clade. her5/7 of both medaka and zebrafish belonged to Xenopus hes7.1 clade, whereas zebrafish her1, her11, and medaka her7 belonged to the Xenopus hes7.3 clade.

In hes5, many homologous genes, Dareher4.1·her4.4, two Dareher4.2, Dareher2/12/15.1·15.2, and Orlaher4.2/12 were found in both zebrafish and medaka (Fig. 3B). We examined if the hes5 cluster first appeared in the common ancestor of Teleostei and Sarcopterygii. However, Dareher4.1·4.4, Orlaher4.2, and Dareher15/12/2 and Orlaher12 formed teleost-specific monophyly among the large hes5 clade. These results indicated that it is not clear which Xenopus subclade zebrafish/medaka hes5 belongs to. Thus, we next performed synteny analysis around her4.1·4.4/12 and her2/15.1·15.2 to clarify whether these genes formed clusters similar to Xenopus hes5.1 and hes5.3 clusters (the synteny of other teleost hes gene were shown in Fig. S1B, D).

In the zebrafish genome, the her4.1·4.4/12 cluster and her2/15.1·15.2 cluster were present on chromosomes 23 and 11, respectively. dnajc11 and rnf207 genes were found in the genomic region around the clusters. In addition, icmt, kcnab2, nol9, and chd5 genes, which are located in Xenopus hes5 locus, were also found on either chromosome 23 (DRE23) or chromosome 11 (DRE11). These results suggested that DCSs were found in the hes5 region of the zebrafish genome. Near the her2/15 cluster on DRE11, dnajc11, which is located near Xenopus hes5.3 cluster, was found (Fig. 4A). However, other typical features of the hes5.3 cluster were not observed in the locus. For instance, nol9 or zbtb48 was not located near the hes2/15 locus.

On DRE23, the her4.1·4.4/12 cluster was located between emc1 and icmt (Fig. 4A). The icmt gene was located near the hes3 gene in Xenopus (Fig. 4B). No her/hes gene was detected between zbtb48 and nol9, as in the chicken genomes (Fig. 2B). From these results, it was difficult to determine whether the her4.1·4.4/12 cluster corresponds to
Xenopus hes5.1 cluster or hes5.3 cluster. It should be pointed out that the sequence homology of the zebrafish genes with Xenopus hes5 genes appeared to be higher for the hes5.1 cluster genes than for the hes5.3 cluster genes (Table 1), suggesting that her2/15 and her 4.1-4.4/12 genes of zebrafish might correspond to the hes5.1 cluster genes in Xenopus.

In medaka, the her7 gene was found to be located near grik5, which was located near the hes5.1 cluster in Xenopus (Fig. 4C). However, phylogenetic analysis showed that OLA her7 was in the Xenopus hes7.1 subclade (Fig. 3A). Conversely, OLA her4.4 and her12 were located around these genes, espn, acot7, and hes2.2, which are near the hes5.3 cluster locus in Xenopus, even though no genes were located between nol9 and zbtb48 (Fig. 4D). A gene order similar to the Xenopus hes5 region was also observed in chromosome 1 in medaka, but no hes-related genes were found in the locus (Fig. 4D).

**Classification of hes genes in gnathostomata**

To determine the origin of the hes5 cluster, we carried out phylogenetic analysis with spotted gar (Lepisosteus oculatus), elephant shark (Callorhinchus milii), lamprey (Petromyzon marinus), and amphioxus (Branchiostoma floridae) (Fig. 5A, B, S3: the complete tree is shown in Fig. S6). As a result, genes of hes7 and hes5 were clearly separated from the other genes with high bootstrap values. First, we counted the number of hes genes in these species with the exception of hes7- and hes5-classified genes, although the bootstrap values were low. Spotted gar was considered to have two hes3, two hes7, and three hes6 (Fig. 5A, shown in red letter). Elephant shark had one hes1, hes2, hes4, and hes6 (Fig. 5A, shown in blue letter). In lamprey, there were one hes4, three hes2, and one hes3 (Fig. 5A, shown in purple letter). In amphioxus, the hes A·G gene was found, but not in any of the hes subfamily clades (Fig. 5A, shown in green letter).
In the *hes5* clade, both spotted gar and elephant shark possessed four genes (Fig. 5B), but all the genes were separately classified from the *Xenopus* genes (Fig. 5B, red and blue letters). This feature was different from other *hes*-related genes (Fig. 5A). In lamprey and amphioxus, any *hes5* gene was not found (Fig. 5B). These results indicated that the phylogenetic analysis could not identify the homologous relationship of *hes5* genes between *Xenopus* and gar/elephant shark.

Next, we compared the gene order around *Xenopus hes5* cluster region in spotted gar and elephant shark. In the spotted gar linkage group (LG) 25, four clustered *hes5-like* genes were located next to *pank4*, but no *hes5* genes were found near *nol9* (Fig. 6B). This suggests that gar had a *hes5* cluster, and the cluster was closer to the *hes5.1* cluster than to the *hes5.3* cluster in *Xenopus*. In contrast, three of the four genes in elephant shark were clustered near *nol9* on KI635912.1 (Fig. 6C). This suggested that the clustered genes might be related to the *hes5.3* cluster in *Xenopus*. In addition, the gene named *her3* was located near *pank4*, which is located near the *hes5.1* cluster in *Xenopus*, on HMISc93. Although the gene may have been given a wrong name because the sequence lacking WPRW domain, the synteny analysis suggested that the gene might be the homolog of *hes5*, and thus, the *hes5.1* cluster might be conserved in elephant shark. Another *hes5* gene in elephant shark was located next to *ppil2*. The order was conserved in coelacanth (Fig. 2D), but not in the *Xenopus* *hes5* cluster. This suggests the possibility that the common ancestor of teleost and cartilaginous fishes had another *hes5* next to *ppil2*, but later lost the gene.

**Evolutionary phylogenetic relationships of *hes* gene in gnathostomes**

To further confirm the classification, we performed a phylogenetic analysis with human *hey* genes, which are reported to be close to *hes* genes, as an outgroup (Fig. 7A: complete tree is shown in Fig. S7) [17]. As it can be seen from the low bootstrap values
in the *hes* clades, the phylogenetic tree was not solved well. However, as we already discussed above, eight zebrafish genes (*her4.1*-4.4, *her12, her15.1*-15.2), two medaka genes (*Orlaher4.2/12*), and three gar genes formed a single clade with *Xenopus hes5.1/5.2* (Fig. 7B, B’). However, monophyletic groups including *Xenopus hes5.3*-5.10, three coelacanth *hes5* (LachSc00001/00119/00319), Zebrafish *her2*, and one gar *hes5* (LeocLG25-1) were formed (Fig. 7B, B’), even though LachSc00319 and LeocLG25-1 showed syntenic similarity with the *hes5.1* cluster (Fig. 2D, Fig. 6B).

**Discussion**

Phylogenetic analysis showed that the *hes5* genes were absent in lamprey and amphioxus (Fig. 5). However, eight hairy genes have been reported in amphioxus, four of which have conserved gene expression patterns in vertebrates (in the central nervous system, presomitic mesoderm, somites, notochord, and gut) [18]. Some instances have been reported that the gene names of *hes* and the function were mismatched [19]. It remains unclear why *hes5* was specifically absent in lamprey and amphioxus, but other *hes* genes might substitute for *hes5* function in these species.

We found that elephant shark possessed *hes5* (Figs. 5 and 6). Interestingly, synteny analysis indicated that three *hes5* genes might be the orthologue of the *hes5.3* cluster in *Xenopus*. Together with the result that the putative *hes5* gene (*her3*) existed near *pank4* in the shark, it is thought that a common ancestor of gnathostomata acquired both *hes5.1* and *hes5.3* genes. In spotted gar, the *hes5.3* cluster was not found (Fig. 6B). One of the possibilities is that, after divergence into cartilaginous fishes and neopterygii, *hes5* near *nol9* was translocated to the locus next to *pank4*. Another possibility is that *her3* (*=hes5.1* cluster gene) was duplicated, and three *hes5-like* genes (*=hes5.3* cluster gene) in elephant shark were lost in the spotted gar. Unfortunately, we have not yet obtained
direct evidence for these possibilities from phylogenetic analysis (Fig. 5).

The synteny of both hes5.1 and hes5.3 cluster seemed to be maintained in both teleost and neopterygian, even though the gene orders of these species in these loci were highly divergent (Figs. 4 and 6). In addition, although no cluster was formed probably due to the insufficient scaffold connection, many hes5 genes were found in coelacanths (Fig. 2), suggesting that the prototype of the hes5.1/5.3 clusters would be retained in the common ancestor of amphibians and sarcopterygians. However, all amniote hes5 genes seemed to be the hes5.1 cluster genes, and not hes5.3 cluster genes (Figs. 1 and 2), suggesting that the hes5.3 clusters was lost after branching into amniotes. We further examined the number of exons in the coding regions of each hes5 gene. In both X. tropicalis and X. laevis, almost all hes5 consisted of three exons, except for hes5.8. On the other hand, hes5 genes of many actinopterygian including zebrafish, medaka, and spotted gar genes possessed two exons in coding region (Table 2). This might reflect that hes5 genes in actinopteryzoa and osteichthyes were increased in an independent manner.

In this study, we showed that the number of hes5 genes is specifically high in Xenopus, especially the number of hes5.3 cluster genes. To estimate the duplication process, comparison of the transcriptional direction among hes5 genes may be considered important [11]. As we previously reported, the directions of hes5.5, 5.6, 5.7, and 5.9 are the same. Phylogenetic analysis also indicated that these genes were closely mapped in the tree (Fig. 1A), suggesting that these genes may share a common origin and may be tandemly duplicated in Xenopus. Phylogenetic analysis also indicated that hes5.1, hes5.2, and hes5.10 showed high similarity (Fig. 1B, 3B, 5B, and 7B). This result suggests another possibility that hes5.10 duplicated from hes5.1/5.2 (or vice versa).

It is known that hes5 functions downstream of Notch signaling and inhibits neuronal differentiation [20, 21]. RNA-seq analysis revealed that the expression of
almost all \textit{hes5} genes is high during the gastrula and neurula stages in \textit{Xenopus} [11]. These results suggests that the function of \textit{hes5} is conserved between mouse and \textit{Xenopus}. How \textit{hes5} works in neurogenesis should be investigated, and this may elucidate the significance of the high number of \textit{hes5} genes in \textit{Xenopus}.

\textbf{Conclusion}

In this study, to understand the evolutionary process of \textit{hes} genes, we estimated the evolutionary origins of two \textit{hes5} clusters. Although the \textit{hes5} gene was found in other jawed vertebrates, the number of \textit{hes5} genes was highest in \textit{Xenopus} (Fig. 8). The rudiment of the two clusters was found in elephant shark, suggesting that ancestral species of chondrichthyans might have these clusters. In addition, we reorganized the orthologous relationship of \textit{hes} genes in vertebrates using phylogenic and synteny analyses. These findings go a step further in the research on the function of all \textit{hes} genes in vertebrates as well as the understanding of the evolutionary process of large gene clusters.

\textbf{Methods}

\textbf{Protein sequencing comparison}

A multiple alignment of protein sequence of \textit{hes} genes were visualized with MUSCLE [22].

\textbf{Phylogenetic analysis}

Phylogenetic analysis was performed using RAxML (v8.2.0) [23]. Multiple alignments of protein sequence were carried out using MAFFT (v7.221) [24] with the \texttt{–auto} strategy.
Unaligned regions were trimmed with TrimAl (v1.2rev59) [25] using the --gappyout option and phylogeny trees were constructed by the maximum likelihood method with PROTGAMMAAUTO.

**Annotation of Genes: GenBank accessions of hes genes**

Genomic synteny of hes genes was analyzed using genome assemblies of *X. laevis* (v9.1), *X. tropicalis* (v9) from Xenopus genome project (http://viewer.shigen.info/Xenopus). Other species Gene ID and accession number for these analysis are from NCBI (https://www.ncbi.nlm.nih.gov/) and Ensambl (https://asia.ensembl.org/index.html).

*hes5* genes are as follows: *hes5*chr21-1 ID: 419390. *hes5*chr21-2 ID: 107057363. *hes5*chr21-3 ID: 419392. *hes5*sc135-1 ID: 107122264. *hes5*sc135-2 ID: 107122267. Lach *hes5*Sc00059 ID: 102346872. *hes5*Sc00199 ID: 102346203. Lepisosteus oculatus *hes5* LG25-1 ID: 102684766. *hes5*LG25-2 ID: 102684967. *hes5* LG25-3 ID: 102683774. *hes5*LG25-4 ID: 102685165. Callorhinchus milii *hes5*Sc221 ID: 103188596. *hes5*Sc58-2 ID: 103181452. *hes5*Sc58-3 ID: 103181453. *hes5*Sc58-1 ID: 103181414.

**Author’s contribution**

Experiments were planned by A.K., T.Y., M.T. and T.M., and conducted by A.K., T.M. The manuscript was prepared by A.K., T.Y., M.T. and T.M.

**Ethics approval**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.
References

1. Fisher AL, Ohsako S, Caudy M. The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. Mol Cell Biol. 1996; 16(6):2670-2677.

2. Pourquie O. Vertebrate somitogenesis: a novel paradigm for animal segmentation? Int. J. Dev. Biol. 2003; 47(7-8):597-603.

3. Davis, R., Turner, D. Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning. Oncogene 2001; 20:8342–8357

4. Takada H, Hattori D, Kitayama A, Ueno N, Taira M. Identification of target genes for the *Xenopus* Hes-related protein XHR1, a prepattern factor specifying the midbrain–hindbrain boundary. Dev. Biol.2005; 283: 253-267.

5. Shinga J, Itoh M, Shiokawa K, Taira S, Taira M. Early patterning of the prospective midbrain-hindbrain boundary by the HES-related gene XHR1 in Xenopus embryos. Mech Dev. 2001; 109(2):225-239.

6. Katoh M, Katoh M. Identification and characterization of human HES2, HES3, and HES5 genes in silico. Int J Oncol. 2004; 25(2):529-534.

7. Kageyama R, Ohtsuka T, Kobayashi T. The Hes gene family: repressors and oscillators that orchestrate embryogenesis. Development. 2007; 134(7):1243-1251.
8. Session, A., Uno, Y., Kwon, T. et al. Genome evolution in the allotetraploid frog *Xenopus* laevis. Nature 2016; 538:336–343.

9. Bisbee CA, Baker MA, Wilson AC, Haji-Azimi I, Fischberg M. Albumin phylogeny for clawed frogs (*Xenopus*). Science. 1977; 195:785–787.

10. Harland RM, Grainger RM. *Xenopus* research: metamorphosed by genetics and genomics. Trends Genet. 2011; 27(12):507–515.

11. Watanabe M, Yasuoka Y, Mawaribuchi S, Kuretani A, Ito M, Kondo M, Ochi H, Ogino H, Fukui A, Taira M, Kinoshita T. Conservatism and variability of gene expression profiles among homeologous transcription factors in *Xenopus* laevis. Dev Biol. 2017; 426(2):301–324.

12. Lawrence J. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. Curr Opin Genet Dev. 1999; 9(6):642–648.

13. Lawrence JG, Roth JR. Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics 1996; 143(4):1843–1860.

14. Ohno S. Evolution by gene duplication. 1970; Berlin: Springer-Verlag.

15. Van de Peer, Y., Maere, S. & Meyer, A. The evolutionary significance of ancient genome duplications. Nat Rev Genet. 2009; 10:725–732
16. Kellis, M., Birren, B. & Lander, E. Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature 2004; 428:617–624.

17. Zhou M, Yan J, Ma Z, Zhou Y, Abbood NN, et al. Comparative and Evolutionary Analysis of the HES/HEY Gene Family Reveal Exon/Intron Loss and Teleost Specific Duplication Events. PLOS ONE 2012; 7(7):e40649.

18. Minguillón C, Jiménez-Delgado S, Panopoulou G, Garcia-Fernández J. The amphioxus Hairy family: differential fate after duplication. Development 2003; 130: 5903-5914.

19. Gajewski, M. and Voolstra, C. Comparative analysis of somitogenesis related genes of the hairy/Enhancer of split class in Fugu and zebrafish. BMC Genomics 2002; 3:21.

20. Tanigaki K, Nogaki F, Takahashi J, Tashiro K, Kurooka H, and Honjo T. Notch1 and Notch3 instructively restrict bFGF-responsive multipotent neural progenitor cells to an astroglial fate. Neuron 2001; 29:45-55.

21. Hatakeyama J, Bessho Y, Katoh K, Ookawara S, Fujioka M, Guillemot F, Kageyama R. Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. Development 2004; 131:5539-5550.

22. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004: 32(5):1792-1797.
23. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 2014; 30(9):1312–1313.

24. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30(4):772-780.

25. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009; 25(15):1972-1973.
**Figure legends**

**Fig. 1. Phylogenetic analysis of hes genes in sarcopterygian**

A. The outline of phylogenetic tree in hes gene family. Complete tree was shown in Fig. S1.  
B. The phylogenetic tree was constructed based on the amino acid sequences using the ML method. Bootstrap support values for nodes are indicated (n=100). Evolutionary analysis was conducted in RAxML. Blue and red letters indicate amniote and coelacanth genes, respectively. The abbreviation of animals are as follows: Hosa: human, Mumu: Mouse, Gaga: Chicken, Geja: Japanese gecko, Xetr: *X.tropicalis*, Xela: *X.laevis*, Lach: Coelacanth.

**Fig. 2. Syntenic analysis of hes gene locus in sarcopterygian**

The syntenies in frog (A), chicken (B), gecko (C) and coelacanth (D) are shown. Chromosome number is described as “XLA7L” in each panel. Pentagon arrows show genes with the 5’ to 3’ direction. A magenta one shows hes gene and a broken lined one indicates a pseudogene. Orange and blue square shows hes5-1 and hes5-3 cluster region, respectively. The broken lined circle shows “partial” hes gene.

**Fig. 3. Phylogenetic analysis of teleost hes genes**

The phylogenetic tree was constructed by ML method. Hes genes except hes5 (A) and only hes5 genes (B) were presented. Complete tree was shown in Fig. S2. Red and blue letters indicate zebrafish/medaka and human hes genes, respectively. Hosa, human; Xela, *Xenopus* laevis; Xetr, *Xenopus* tropicalis; Dare, zebrafish; Orla, medaka.

**Fig.4. Comparison with hes gene locus among Xenopus, zebrafish and medaka**

Chromosome number is described as “DRE23”. Pentagon arrows show genes a gene with
5’ to 3’ direction. The magenta one shows hes gene. Broken arrow means same region on DRE23. DRE, zebrafish; OLA, medaka; XLA, Xenopus laevis.

**Fig. 5. Phylogenetic analysis of hes genes of several jawed vertebrates**

The phylogenetic tree was constructed by ML method. Hes genes except hes5 (A) and only hes5 genes (B) of spotted gar, elephant shark, lamprey and amphioxus were presented. Blue, red, purple and green letters indicate spotted gar, elephant shark, lamprey and amphioxus, respectively. Hosa, human; Xetr, Xenopus tropicalis; Leoc, spotted gar; Cami, elephant shark; Pema, lamprey; Brfl, amphioxus.

**Fig.6. Comparison with hes gene locus among Xenopus(A), spotted gar(B) and elephant shark (C)**

Pentagon arrows show genes with the 5’ to 3’ direction. The magenta one shows hes gene and broken lined one is a pseud gene. Orange square shows hes5-1 cluster region. Orange and blue square shows hes5.1 and hes5-3 cluster region, respectively.

**Fig. 7. Comprehensive phylogenetic analysis of hes genes except hes5 (A) and hes5 genes (B) in jawed vertebrate**

Evolutionary analysis was conducted in RAxML. Human HEYL, X.tropicalis hey2 and sponge (Amphimedon queenslandica) HEY1-like gene sequences were used as an outgroup. For easy comprehension, the outline of B is described in B’. Hosa, human; Mumu, Mouse; Gaga, Chicken; Geja, Japanese gecko; Xetr, Xenopus tropicalis; Lach, Coelacanth; Dare, zebrafish; Orla, medaka; Leoc, spotted gar; Cami, elephant shark; Amqu, sponge.

**Fig.8. Evolutionary acquisition of hes5 genes and hes5 cluster**
The tree shows the phylogenetic relationship of jawed vertebrate. The list shows the number of paralogous hes5 genes which synteny conserved with hes5.1cluster and hes5.3 cluster genes and the number of paralogous hes5 genes derived from other hes5 gene.

**Table 1. Amino acid sequence identities of Xenopus hes5 genes with hes5 genes of zebrafish**

Amino acid sequence identities of zebrafish (Dare) hes5 genes are in the left column and that of Xenopus tropicalis hes5 genes are in the right column.

**Table 2. The list of number of exon, protein size and conserved Xenopus hes5 synteny of hes5 genes**

Ortologus hes5 genes are in the left column and the number of exon, protein size and Xenopus conserved synteny in the right column. X1, X2 means isoform.