Genetics and Molecular Biology, 30, 2, 494-503 (2007)
Copyright by the Brazilian Society of Genetics. Printed in Brazil
www.sbg.org.br

Research Article

Characterization of partial *Hox* gene sequences in annual fish of the subfamily Cynolebiatinae (Cyprinodontiformes, Rivulidae)

Verónica Gutiérrez[1], María José Arezo[2] and Graciela García[1]

[1]*Sección Genética Evolutiva, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay.*
[2]*Sección Biología Celular, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay.*

Abstract

Hox genes encode a family of transcription factors implicated in conferring regional identity along the anteroposterior axis in developing animal embryos. These genes are organized in genomic clusters, expressed collinearly and highly conserved in vertebrates. Among teleost, South American annual killifishes of the Cynolebiatinae subfamily represent an excellent model in development studies because their embryos are capable of undergoing reversible developmental arrest (diapause) at three well-defined morphological stages. They are also an excellent model for evolutionary studies due to the high rates of mutation of their mitochondrial genome, their karyotypic divergence and their morphological variability. In this study, three partial homeobox sequences were isolated from different species of the Cynolebiatinae subfamily. Phylogenetic analyses and sequence comparisons revealed that they belong to the anterior Hox complex group, specifically to paralogue groups 1 and 3. This is the first time that partial *Hox* genes have been described in species of the Cynolebiatinae subfamily.

Key words: subfamily Cynolebiatinae, annual killifishes, *Hox* genes.

Received: April 17, 2006; Accepted: November 6, 2006.

Introduction

The homeobox genes play important roles in the developmental processes of many multicellular organisms. The homeobox was originally described as a conserved protein-coding sequence of about 180 base pairs, in many developmental control genes of *Drosophila*. The encoded proteins are evolutionarily conserved transcription factors that share a 60 amino acids DNA-binding domain (the homeodomain, or HD) and regulate axial patterning segment or cell identity and proliferation (Lewis, 1978; McGinnis *et al.*, 1984a, b). Homeodomain proteins regulate diverse developmental programs by modulating expression patterns of target genes in a temporal, spatial and tissue-specific manner (Gehring *et al.*, 1994a). X-ray crystallographic and NMR spectroscopic analyses on several members of this family revealed that the structure consists of a tri α-helical core and an N-terminal arm that becomes partially ordered in the presence of DNA (Qian *et al.*, 1989; Kissinger *et al.*, 1990; Gehring *et al.*, 1994a; Wolberger, 1996; Passner *et al.*, 1999). The N-terminal region, just upstream of helix 1, contacts the minor groove of DNA, whereas helix 3, known as the "recognition helix", binds to the major groove. The mode of HD-DNA interaction appears to be highly conserved within HD-containing proteins (Gehring *et al.*, 1994b; Chauvet *et al.*, 2000).

In vertebrates, the *Hox* genes (a class of homeobox genes) are organized in linked chromosomal clusters and show a striking colinearity in their 5'-3' chromosomal position. They also show remarkably similar expression patterns along the anteroposterior (AP) axis, specifying the identities of different body regions during embryogenesis (McGinnis and Krumlauf, 1992). While invertebrates studied to date have a single Hox cluster, tetrapod vertebrates have four separated Hox clusters (termed clusters A to D) lying on four different chromosomes. Within each cluster, the genes have been classified into 13 paralogue groups according to sequence homology and location in the genome (Krumlauf, 1994). The expression pattern of *Hox* genes during development and homeodomain sequence comparisons indicate relatedness between anterior (*Hox* 1-3), medial (*Hox* 4-8) and posterior (*Hox* 9-13) paralogue groups (Ruddle *et al.*, 1994). Some paralogue groups can be recognized by the homeodomain sequence alone, others only by using characteristic residues outside the homeodomain (Sharkey *et al.*, 1997).

Although it is widely assumed that vertebrates have four Hox clusters, studies on ray-finned fish have shown that they have more Hox clusters than tetrapods. These

Send correspondence to Graciela García. Facultad de Ciencias, Iguá 4225, 11400 Montevideo, Uruguay. E-mail: ggarcia@fcien.edu.uy.

extranumeral Hox clusters result from a genome duplication event that is specific for the fish (actinopterygian) lineage. Sets of seven Hox clusters have been described in *Danio rerio* (Amores *et al.*, 1998; Prince *et al.*, 1998), in three pufferfish species, *Takifugu rubripes*, *Tetraodon nigroviridis* and *Sphenoides nephelus* (Aparicio *et al.*, 1997, 2002; Jaillon *et al.*, 2004; Amores *et al.*, 2004) and in *Oryzias latipes* (Naruse *et al.*, 2000). However, preliminary data indicates four *Hox* clusters in the non-annual killifish *Fundulus heteroclitus* (Misof and Wagner, 1996). The four-cluster situation is also retained in the sarcopterygian lineage (Koh *et al.*, 2003), in basal ray-finned fishes, such as the bichir *Polypterus senegalus* (Ledje *et al.*, 2002), and in the horn shark *Heterodontus francisci* (Kim *et al.*, 2000). The study of Hox clusters, therefore, provides the opportunity to understand the relationship between molecular evolution, genome organization and gene expression.

The Cynolebiatinae subfamily (Cyprinodontiformes: Rivulidae) is a specious group of annual fish distributed from northeastern Brazil to northeastern and southern Argentina, Uruguay and Paraguay (Costa, 1995). They live in temporary ponds and each generation completes a full life cycle within one year. The population survives dry seasons in the form of eggs buried in the mud. During the subsequent rainy season, the ponds refill, the eggs hatch and the larvae rapidly grow to sexual maturity and reproduce (Wourms, 1967, 1972a). The developmental pattern of annual killifish embryos is unique compared to other teleosts, and is characterized by dispersion and subsequent reaggregation of blastomeres and the occurrence of embryonic diapause. Diapause is a state of developmental arrest that precedes the onset of unfavorable environmental conditions, is promoted by genetic and environmental factors and typically occurs as part of the natural developmental program. Thus, embryos may enter diapause even under conditions considered optimal for development. Annual killifish embryos can undergo reversible developmental diapause at one or all of three well-defined morphological stages (diapause I, II and III) (Wourms, 1967, 1972a, b; Hand and Podrabsky, 2000). Studies in diapausing embryos of the annual killifish *Austrofundulus limnaeus* reveal that during diapause II protein synthesis is substantially diminished due to depressed metabolism. At this stage embryos appear to be more resistant to environmental stresses such as anoxia and dehydration (Podrabsky and Hand, 1999, 2000; Podrabsky *et al.*, 2001).

To date, early development has been studied in some *Austrolebias* species (subfamily Cynolebiatinae, *sensu* Costa 1995; 1998) and available data show that members of this genus display a dispersion-reaggregation process and that they undergo facultative arrest at diapause I and II and obligate arrest at diapause III (Wourms, 1972a; Carter and Wourms, 1991; Arezo *et al.*, 2005).

The development of annual fish embryos exhibits several characteristics that are of interest and there is a lack of useful information about their developmental genetic control. The study presented here is the first description concerning *Hox* genes in different teleost fish species belonging to the Cynolebiatinae. We established their cluster affiliation and investigated if amino acid changes in homeodomain primary sequence affect its ternary structure and the homeodomain-DNA interaction. Our results suggest that the sequences isolated by us are highly homologues to the anterior genes of the Hox complex of other vertebrates.

Materials and Methods

Specimens

We studied four *Austrolebias* species (*A. bellottii, A. charrua, A. cheradophilus and A. viarius*) from temporary ponds in Uruguay and *A. adloffi* from the southernmost Brazilian state of Rio Grande do Sul (which borders Uruguay) kindly supplied by L. Malabarba (1991). A specimen of *Brevoortia aurea* (Clupeiformes: Alosinae) was also examined (Table 1). Tissues and voucher specimens are deposited in the Evolutionary Genetics Section, Science Faculty, University of the Republic (Sección Genética Evolutiva, Facultad de Ciencias, Universidad de la República), Montevideo, Uruguay.

Genomic DNA extraction, amplification, cloning and sequencing

Genomic DNA was isolated from ethanol-fixed liver tissue using sodium chloride protein precipitation followed

Table 1 - Capture sites of the *Austrolebias* (*A*) and *Brevoortia* (*B*) specimens analyzed by us. The *A. adloffi* specimens came from Brazil, all other specimens came from Uruguay.

Species: accession number	Capture site
B. aurea (Spinx and Agassiz, 1829): 217	Puerto del Buceo, Montevideo City
A. adloffi (Ahl, 1922): GP303	Ponte do Gravatai, Rio Grande do Sul
A. bellottii (Steindachner, 1881): GP363; GP400	Carmelo town, Colonia Department; Bañado Verocay, Salto Department
A. charrua (Costa and Cheffe, 2001): GP1498; GP334	Bañados del Este (pond 32), Rocha Department; Route 14, km 489, Rocha Department
A. cheradophilus (Vaz-Ferreira *et al.*, 1964)*: GP950	Castillos City, Rocha Department
A. viarius (Vaz-Ferreira *et al.*, 1964)*; GP329; GP1242	Route 10, km 267, Rocha Department; Route 10 (pond 3), Rocha Department

*Full citation: Vaz-Ferreira, Sierra de Soriano and Scaglia de Paulete, 1964.

by ethanol precipitation of total DNA (modified from Medrano *et al.*, 1990). To ascertain the quality of DNA, an aliquot of the extract DNA was run in ethidium bromide stained 1% (w/v) agarose gel using 1XTAE (Tris-Acetate-EDTA).

Amplification of a homeodomain region was performed by the polymerase chain reaction (PCR) with the specific primers Sog1 (5'-TGAGCTGGAAAAGGAGTT C-3') and Sog2 (5'-CTACGGTTTGGAACCAG-3'). These primers were designed from a partial *Hox* sequence previously isolated in our laboratory from *Austrolebias gymnoventris* (Pereiro and García, unpublished data) using the degenerate primers SO1 and SO2 (Bayascas *et al.*, 1997). The PCR was carried out in a 10 μL total volume using 6.4 μL of H_2O, 1 μL of 10X buffer, 0.3 μL of $MgCl_2$ (50 mM), 0.2 μL of dNTPs (10 mM), 0.5 μL of each primer (10 μM), 0.1 μL of *Taq* DNA polymerase (5 U/μL) (Invitrogen) and 1 μL of DNA, under the following conditions: one denaturation step at 94 °C for 5 min, followed by 30 cycles of 94 °C for 45 s, 48 °C for 45 s and 72 °C for 1 min, and a final elongation step at 72 °C for 7 min. The PCR products were separated by electrophoresis in a 6% (w/v) non-denaturing polyacrylamide gel and visualized by silver staining (Sanguinetti *et al.*, 1994).

Every PCR fragment was subsequently cloned into the pGEM®-T Easy vector using the reagents and protocols supplied by the manufacturer (Promega). The recombinant DNA was obtained by DNA minipreparation of individual clones by alkaline lysis (Sambrook *et al.*, 1989). The clones were screened for correctly sized inserts by electrophoresis in ethidium bromide stained 0.8% (w/v) agarose gel using 1XTAE, followed by a PCR with the specific Sog1 and Sog2 primers. Sequencing reactions were performed on each template using the Sog2 primer and were analyzed in an automated ABI PRISM 377 DNA Sequencer. Nucleotide sequences from *A. bellottii*, *A. cheradophilus* and *A. viarius* were conceptually translated using the commercially available computer software Gene Runner 3.05 for Windows (Hastings Software). The amino acid sequences were compared against the National Center for Biotechnology Information (NCBI) protein database by using the BLASTp program on the Basic Local Alignment Search Tool (BLAST) network service (Altschul *et al.*, 1990).

Southern blot analysis

To verify the existence of *Hox* sequences in the Cynolebiatinae subfamily we first performed a Southern blot analysis of genomic DNA from two *Austrolebias* sister taxa (*A. viarius* and *A. charrua*) using the *A. cheradophilus* homeobox amplified fragment as a probe to screen for homologous sequences. We also carried out a Southern blot analysis of genomic DNA from three other *Austrolebias* species (*A. adloffi, A. bellottii and A. charrua*) and *B. aurea* using the *A. bellottii* amplified fragment as a probe.

For Southern blotting analysis genomic DNA was isolated as describe above from different species of *Austrolebias* and the distantly related taxon *B. aurea* (Table 1). Approximately 10 μg of DNA was digested with Hinf-I (10000 U/mL) (Promega) at 37 °C for 24 h. After digestion, the DNA was separated using 0.5X TBE (Tris-Borate-EDTA) and 1.5% (w/v) agarose gel stained with ethidium bromide and transferred onto Hybond-N+ nylon membranes (Amersham Life Science) and the blot hybridized consecutively with two different probes using the ECL direct nucleic acid labeling and detection systems (catalogue no. RPN3001; Amersham Biosciences). The probes used were the vectors containing the amplified *A. cheradophilus* and *A. bellottii* fragment, both digested with 20000 U mL^{-1} of the Nde-I restriction enzyme (Biolabs). Probe hybridization was carried out overnight in 0.5 M NaCl at 42 °C and the hybridized bands visualized after 2 h exposure according to the instructions of the kit.

Phylogenetic analyses

Homeodomain amino acid sequences from *A. bellottii*, *A. viarius*, *A. cheradophilus* and partial homeodomain sequences for anterior (1-3), medial (4-8) and posterior (9-13) Hox paralogue groups from other metazoans obtained from GenBank, were aligned by using the CLUSTALX 1.8 program (Thompson *et al.*, 1997) (Table 2). Partial homeodomain sequences from *A. gymnoventris* and *A. luteoflammulatus*, previously isolated in our laboratory (Pereiro and Garcia, unpublished data) were also included in our alignment. The MEGA 3 program (Kumar *et al.*, 2004) was used to construct a neighbor-joining (NJ) tree (Saitou and Nei, 1987) using a p-distance model and complete deletion. Bootstrap values for the nodes were determined by analyzing 1000 bootstrap replicates to estimate the strength of the groupings.

Nucleotide substitution, structural and functional analyses

For nucleotide substitution analysis we used the DnaSP 4.00 computer program (Rozas *et al.*, 2003) to compare synonymous (Ks) and non-synonymous (Ka) sites, to calculate ratio rates (ω = Ka/Ks) and to evaluate the codon usage bias between Cynolebiatinae species and their respective paralogue groups, in order to detect positive selection in highly conserved *Hox* gene fragments.

For structural and functional analysis the homeodomain three-dimensional structure was predicted by comparative protein modeling methods using the automated Swiss Model program (Schwede *et al.*, 2003). The structure was modeled on the basis of its structural similarity with the human HoxB1 homeodomain (Protein data Bank entry 1972). The degree of identity between the template and our isolated partial Hox sequences enabled a preliminary model to be generated using the SWISS-MODEL program. The Swiss-PdbViewer 3.7 program, service pack 5, (Guex

Table 2 - Abbreviations and GenBank accession numbers of species used in Figures 2 and 3.

Species (abbreviation)	Gene (Italic type) GenBank accession number (Roman type)
Austrolebias bellottii (A. bell)	*Hox3* DQ242532
Austrolebias cheradophilus (A. che)	*Hox1* DQ242533
Austrolebias gymnoventris (A. gymno)	*Hox1* DQ242530
Austrolebias luteoflammulatus (A. luteo)	*Hox1* DQ242531
Austrolebias viarius (A. via)	*Hox1* DQ242534
Danio rerio (Dr)	*HoxA1a* CAD52137; *HoxB1b* NP_571217; *HoxA2b* NP_571181; *HoxB2a* AAH65967; *HoxA3a* NP_571609; *HoxB3a* AAH95559; *HoxB4a* NP_571193; *HoxA5a* NP_571615; *HoxC5a* NP_571219; *HoxC6b* NP_571605; *HoxB8a* AAH53287; *HoxB8b* XP_691976.
Heterodontus francisci (Hf)	*HoxA1* AAF44639; *HoxA2* AAF44640; *HoxA3* AAF44641; *HoxD8* Q9IA12; *HoxA9* AAF44646; *HoxA10* AAF44647; *HoxD13* Q9IA17.
Homo sapiens (Hs)	*HoxA1* AAB35423; *HoxB1* NP_002135; *Hox D1* AAH14477; *HoxA2* EAL24227; *HoxA3* EAL24226; *HoxD3* AAH05124; *HoxB4* NP_076920; *HoxC5* NP_061826; *HoxB6* P17509; *HoxB8* NP_076921; *HoxA9* AAP35636; *HoxD9* NP_055028; *HoxC10* NP_059105; *HoxA11* NP_005514; *HoxC11* NP_055027; *HoxD11* AAH14477; *HoxD12* NP_067016; *HoxB13* NP_006352; *HoxC13* NP_059106.
Latimeria menadoensis (Lm)	*HoxC1* AAO43034; *HoxD1* AAO43041; *HoxB2* AAO43027; *HoxB3* AAO43028; *HoxA6* AAO43019; *HoxC8* AAO43037; *HoxD9* AAO43045; *HoxA10* AAO43022; *HoxC10* AAO43039; *HoxA11* AAO43023; *HoxC12* AAO38042; *HoxD12* AAO43047; *HoxA13* AAO43024; *HoxC13* AAO43040.
Morone saxatilis (Ms)	*HoxB2a* AAN52289; *HoxB3a* AAN52288; *HoxA9* AAD46396; *HoxA10* AAD46395.
Oncorhynchus mykiss (Om)	*HoxA2bi* AAX63749; *HoxD9ai* AAX63774; *HoxD9aii* AAX63776; *HoxD10ai* AAX63771.
Oryzias latipes (Olat)	*HoxA1A* BAA86231; *HoxD3A* BAA86235; *HoxB4A* BAA86239; *HoxD4A* BAA86241; *HoxA5A* BAA86244; *HoxB5B* BAA86245; *HoxB6B* BAA86247; *HoxC8A* BAA86248; *HoxA9A* BAA86255; *HoxA9B* BAA86254.
Petromyzon marinus (Pm)	*Hox1w* AAL61641; *HoxE2* AAM19466; *HoxN6* AAM19474; *HoxR8* AAC04330; *HoxX10* AAM19480; *HoxY11* AAM19481.
Salmon salar (Ss)	*HoxB2* P09638.
Sphoeroides nephelus (Sn)	*HoxB6b* AAQ72844.
Takifugu rubripes (Tr)	*HoxB1a* DAA05215; *HoxB2a* DAA05216; *HoxB3a* DAA05217; *HoxB4* O13074; *HoxC6* AAB68682; *HoxC8* AAB68681; *HoxA9* O42506; *HoxC9* O42502; *HoxA10* AAB68683; *HoxD10* Q9IA14; *HoxD11* Q9IA15 *HoxD12* Q9IA16.

and Peitsch, 1997) was then used to analyze and visualize the structures.

Results

PCR products and Southern blot *Hox* genes analyses

The PCR amplifications using the Sog1 and Sog2 specific primers detected fragments of 105 bp in *A. bellottii,* 115 bp in *A. cheradophilus* and 115 bp in *A. viarius*. The deduced amino acid sequences analyses (BLASTp and sequence comparison) identified them as *Hox* genes. The *A. cheradophilus* and *A. viarius* amino acid sequence alignments were identical, so we will not differentiate between these *Hox* genes in the text. The nucleotide and deduced amino acid sequences of these subfamily species have been deposited in the GenBank database (Table 2).

The Southern blot analysis detected a single band in both *A. viarius* and *A. charrua* using the *A. cheradophilus* homeobox amplified fragment as a probe to screen for homologous *Hox* sequences (Figure 1a). A single band was also detected in *B. aurea*, *A. adloffi*, *A. bellottii* and *A. charrua* using the *A. bellottii* amplified fragment as a probe (Figure 1b).

Phylogenetic analyses

To determine the cluster affiliation of the isolated partial *Hox* genes, we performed a combined NJ-phylogenetic tree for anterior (*Hox* 1-3), medial (*Hox* 4-8) and posterior (*Hox* 9-13) paralogue groups, using partial homeodomain (HD) sequences of known vertebrates *Hox* genes (Table 2, Figure 2). As shown in Figure 2, *A. cheradophilus*, *A. viarius*, *A. gymnoventris* and *A. luteoflammulatus* grouped with members of the paralogue group 1 (PG1) with 69% bootstrap support. It is interesting to note that *A. cheradophilus*, *A. viarius* and *A. gymnoventris Hox* genes appear closely related to *Danio rerio* HoxB1b (55%) and that of *A. luteoflammulatus* to *Takifugu rubripes* HoxB1a (82%). On the other hand, the *A. bellottii Hox* gene grouped within the paralogue group 3 (PG3) with 78% bootstrap support, specifically with members from clusters A and D but without robust statistical support.

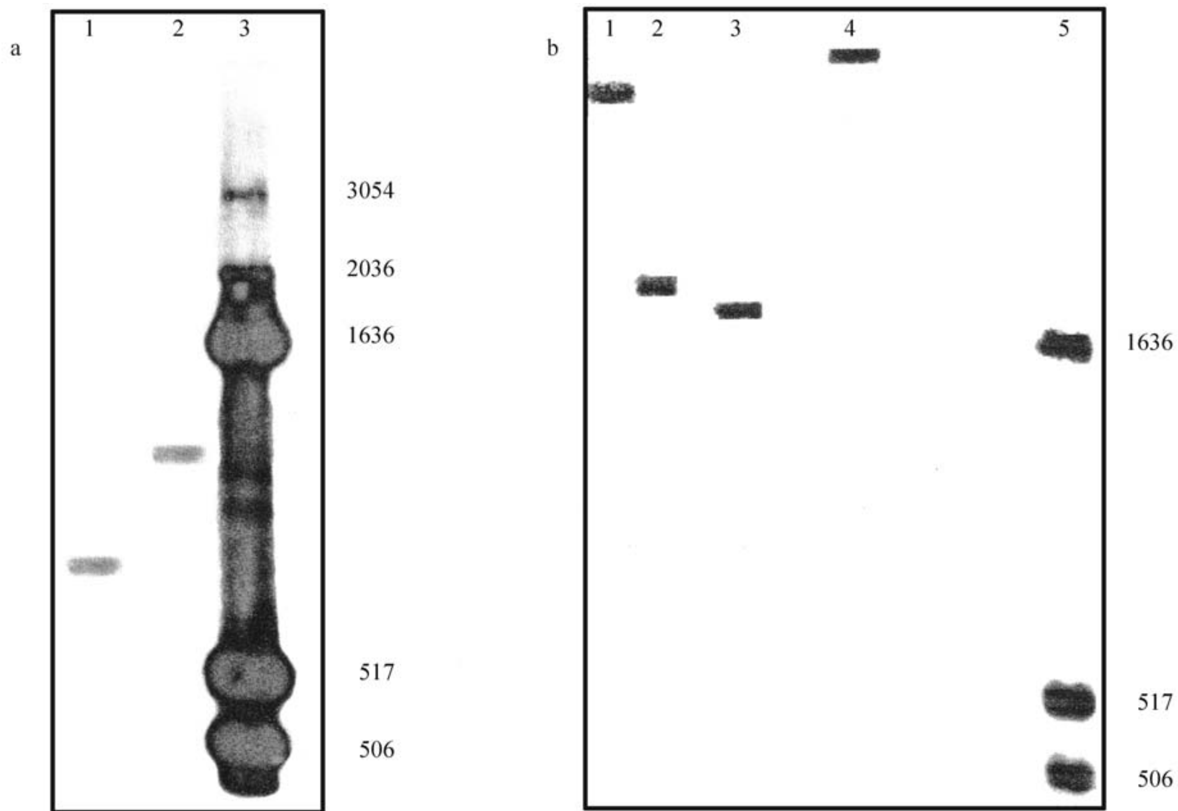

Figure 1 - a) Southern blot analysis of *Austrolebias viarius* (lane 1) and *Austrolebias charrua* (lane 2) genomic DNA digested with Hinf-I. The probe used was the amplified *Hox* fragment in *Austrolebias cheradophilus*. Lane 3: 1 kb DNA ladder molecular marker (GIBCO). b) Southern blot analysis of *Brevoortia aurea* (lane 1), *Austrolebias adloffi* (lane 2), *Austrolebias charrua* (lane 3) and *Austrolebias bellottii* (lane 4) genomic DNA digested with Hinf-I. The probe used was the amplified *Hox* fragment in *Austrolebias bellottii*. Lane 5: 1 kb DNA ladder molecular marker (GIBCO).

Synonymous *vs.* Non-synonymous nucleotide substitutions

Although amino acid substitution rates increase, they never exceed those of synonymous substitutions because relaxed purifying selection is acting (Hartl, 2000). However, the non-synonymous (Ka) to synonymous (Ks) substitution ratio (ω) was significantly larger than 1, if positive selection is acting. Our analysis of synonymous nucleotide substitutions reveals that the species analyzed had a Ks value of 26, whereas the other PG1 species had a Ks value of 23. For non-synonymous substitutions, we obtained a Ka value of 89 for the Cynolebiatinae subfamily species and 81 for the remaining PG1 species (Table 2). In both cases, Cynolebiatinae subfamily species presented higher values with respect to other species of the paralogue group 1. Furthermore, when we compared the ratio rates (= Ka/Ks) between the Cynolebiatinae subfamily and other PG1 species we found that they varied from 0.14 to 1.88, indicating that a weak footprint of positive selection was acting (Hartl, 2000).

We also calculated the codon bias index (CBI) to estimate the codon usage bias. When the Cynolebiatinae subfamily sequences were excluded, the CBI was 0.670 for the PG1 species, 0.736 for PG3 species and 0.846 for Cynolebiatinae subfamily species. Additionally, it is worth noting that the G+C3s value for the *Austrolebias* species was 0.913, whereas 0.703 was obtained for the other PG1 species.

Structural and functional analyses

To ascertain which homeodomain (HD) amino acids are conserved within species of this annual fish subfamily, we compared Hox protein sequences from species that grouped with them (Figure 3). The alignment of PG1 partial HD sequences revealed that, as compared with the human HoxB1 sequence, the sequences for the fish studied had the following substitutions: *A. cheradophilus* and *A. viarius* had Thr-27, Val-34 and Ser-37 substitutions; *A. gymnoventris* had Thr-27, Val-34, Val-36 and Ser-37 substitutions; and *A. luteoflammulatus* had Phe-16, Ser-23 and Thr-27 substitutions (Figure 3a). Furthermore, the comparison of the primary structures of *Austrolebias Hox1* sequences revealed some differences between them. While *A. cheradophilus* shared high sequence identity with *A. gymnoventris*, sequence similarity was considerably lower when compared with *A. luteoflammulatus*. Interestingly, these sequences shared some unexpected residues with other PG1 species. For example, the Thr-27 present in *A.*

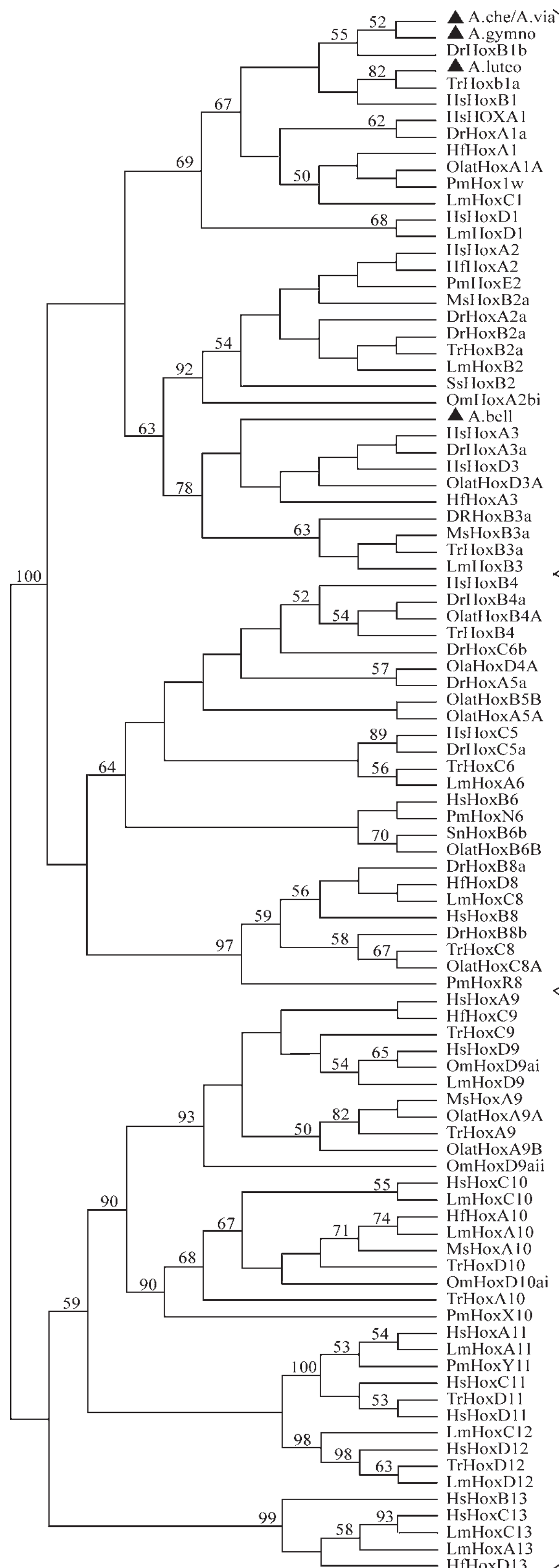

Figure 2 - Neighbor-joining tree using partial homeodomain amino acid sequences. The anterior (1-3), medial (4-8) and posterior (9-13) Hox paralogue groups are indicated. The *Austrolebias* species are marked with a black triangle (5). The number at each node represents bootstrap values above 50% recovered in 1000 replicates. GenBank accession numbers and abbreviations are given in Table 2.

cheradophilus, *A. gymnoventris* and *A. luteoflammulatus* was present in almost all sequences analyzed, except in HsHoxB1 and LmHoxD1, whereas the Val-34 present in *A. cheradophilus* and *A. gymnoventris* was only present in DrHoxB1b sequence while Ser-37 was present in both HsHoxA1 and DrHoxA1a. On the other hand, Ser-23 present in the *A. luteoflammulatus* sequence only was also present in TrHoxB1a. However, it is worth noting that conserved and paralogue-characteristic amino acids were present in all the sequences we analyzed.

The analysis of the PG3 HD sequences revealed that *A. bellottii* had Arg-24, Cys-27, Pro-29, Met-34, Asn-36, Leu-37, Asn-39, Thr-41 and Arg-43 substitutions compared to the human HoxB1 sequence. However, when compared to human HoxA3 we founded that the only difference was a Val-45 substitution (Figure 3b). In this sequence, conserved and paralogue-characteristic amino acids were also present. The HD 3D structure modeling of *A. bellottii*, *A. cheradophilus*, *A. gymnoventris* and *A. luteoflammulatus* Hox proteins showed that the ternary structure of the HD and the HD-DNA interaction had not been altered (Figure 4).

Discussion

Although *Hox* genes have been characterized in many species of teleost fish, this study is the first report about partial *Hox* gene isolation and characterization in annual killifishes belonging to the subfamily Cynolebiatinae.

We employed PCR amplification with specific primers to isolate partial homeodomain sequences from three *Austrolebias* species, *A. bellottii, A. cheradophilus* and *A. viarius*. These sequences and those previously isolated from *A. gymnoventris* and *A. luteoflammulatus* (Pereiro and García, unpublished data) represent the first homeobox fragments reported for Cynolebiatinae subfamily species. The isolated sequences show a high percentage of similarity with those of other vertebrates. Accordingly with the homeobox sequence conservation in virtually all metazoans (McGinnis *et al.*, 1984a), our Southern blot analyses revealed that homeobox sequences are conserved not only among species of the Cynolebiatinae subfamily but also in the unrelated clupeid fish *Brevoortia aurea* (Figure 1).

Because of homeodomain amino acid conservation the phylogenetic tree topology (Figure 2) followed by comparative analyses (Figure 3) allowed us to predict the distribution of *Austrolebias* species partial Hox sequences within the anterior group of *Hox* genes. So far, we have identified partial homeodomain sequences of three *Hox1* sequences and a *Hox3* sequence in annual fish. Interestingly, even though homeodomain sequences are extraordinarily conserved throughout evolution, comparative analyses have shown unexpected amino acid shifts among Cynolebiatinae subfamily species. The alignment of the HD sequences from the PG1 group reveals that two of these residues reside in two different *Austrolebias* species,

Val-36A in *A. gymnoventris* and Phe-16L in *A. luteoflammulatus* (Figure 3a). Note that Leu-16 is a conserved and paralogue-characteristic amino acid present in members of the PG1 (Sharkey *et al.*, 1997). Similarly, analysis of the HD sequences from the PG3 group revealed that *A. bellottii* HD shared high sequence identity with members of clusters A and D, although this species presents a Val-45I residue (Figure 3b).

Although some residues distinguish Hox paralogue groups from one another, it has been shown that the HD structures are remarkably similar and that the mode of HD-DNA interaction has been extraordinarily well conserved (Sharkey *et al.*, 1997). Because of this and the fact that the homeodomain of the human HoxB1 protein and those of Cynolebiatinae subfamily species share a greater percentage of amino acid identity (Figure 3), we used the three-dimensional structure of HoxB1 as a template for homology modeling (Figure 4). We found no ternary structure changes even if primary sequences diverge. Nevertheless, *in vivo* assays in which the function of homologous

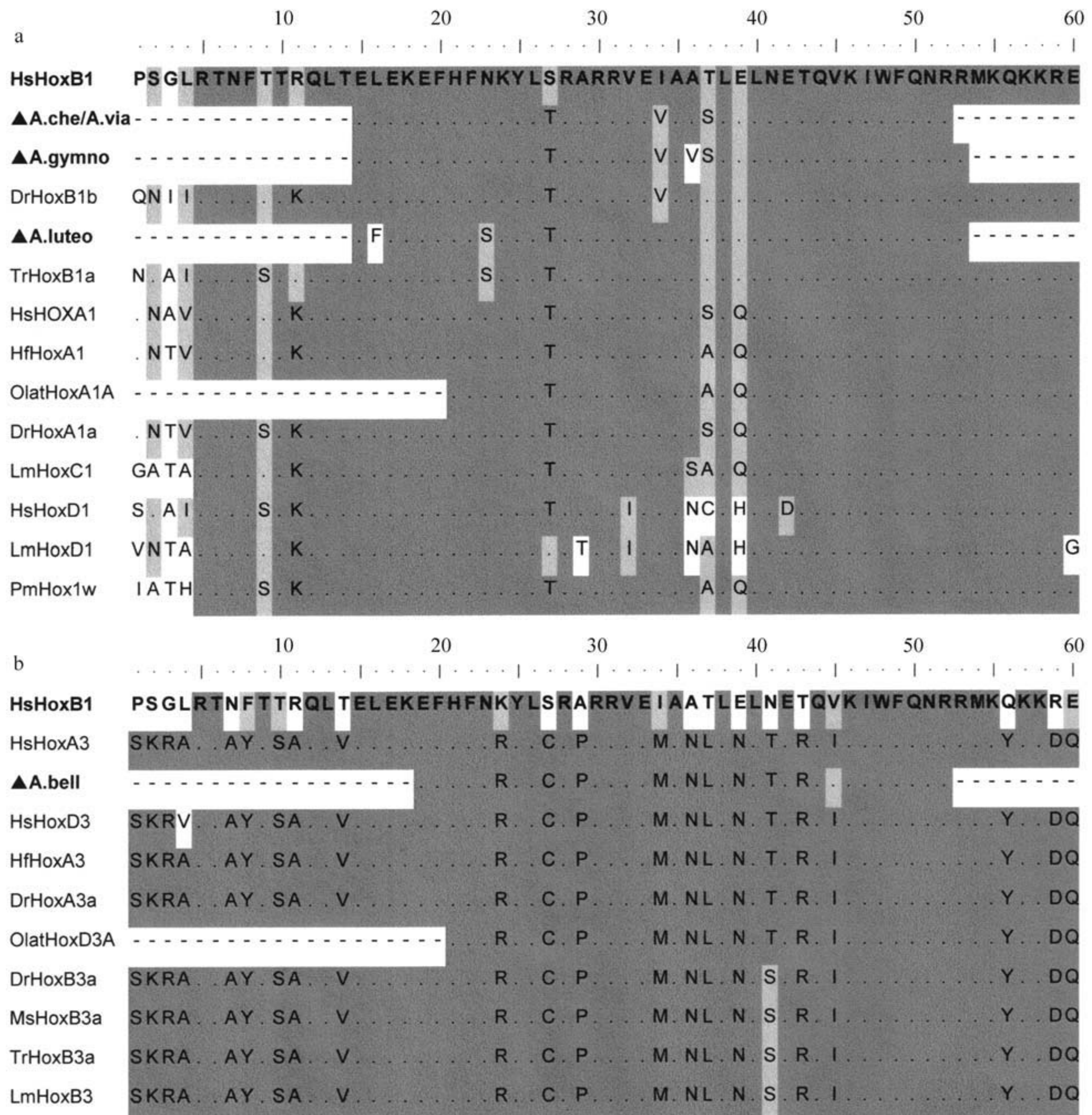

Figure 3 - a) Alignment of partial homeodomain amino acid sequences from species that clustered with the *Austrolebias* species (*A. cheradophilus, A. gymnoventris, A. luteoflammulatus and A. viarius*) b) Alignment of partial homeodomain amino acid sequences from species that clustered with *Austrolebias bellottii*. In both alignments, the homeodomain sequence of human HoxB1 (bold type) is shown at the top of the figure for reference. Identical amino acids are shaded and dots indicate amino-acid identity to human HoxB1 homeodomain.

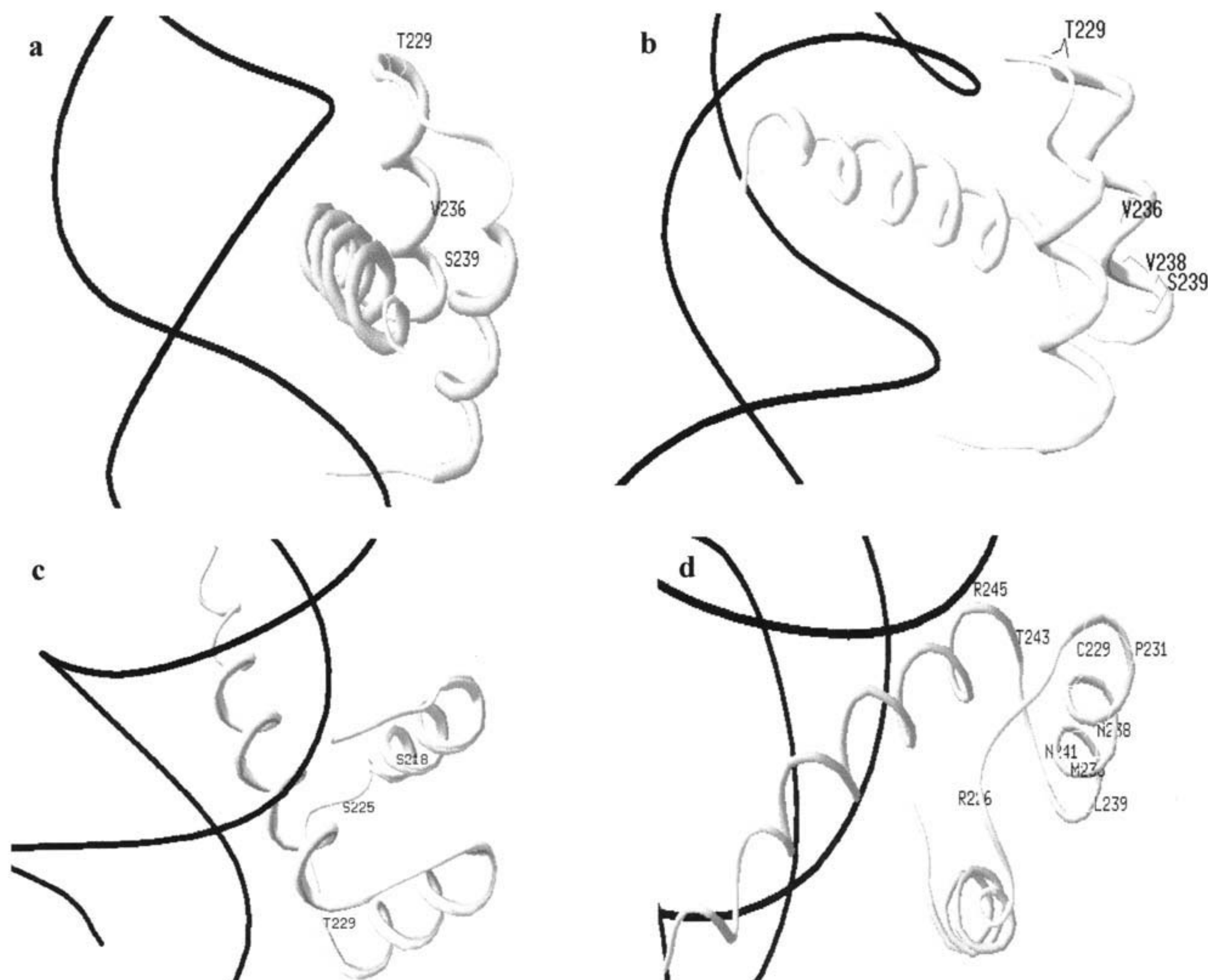

Figure 4 - Swiss-PdbViewer (ribbon diagram) representation of the homology-modeled homeodomain-DNA (HD-DNA) complex for (a) *Austrolebias cheradophilus*, (b) *Austrolebias gymnoventris*, (c) *Austrolebias luteoflammulatus* and (d) *Austrolebias bellottii*. Selected residues within the homeodomain that were identified in our alignments are indicated in each model. In all four species none of the altered residues was predicted to interact directly with the DNA.

proteins is tested in the context of site-directed mutagenesis (Golding and Dean, 1998) may resolve the issue.

As we expected, *Hox* genes from Cynolebiatinae subfamily species showed the signature of adaptive amino acid replacements since non-synonymous substitutions greatly exceeded the synonymous substitution sites among distantly related taxa of the paralogue group 1. Although they were not statistically significant, the rates of both synonymous and non-synonymous nucleotide substitutions were higher in the Cynolebiatinae subfamily species than in the other PG1 species. Despite the fact that positive selection was acting, the substitution pattern detected among Cynolebiatinae subfamily and PG1 species could be indicating that for more distantly related species two or more independent mutations (multiple hits) might have occurred at the same site. Thus, in a molecular phylogenetic context, amino acid sites were identical not because of identity by descent from a common ancestor but because different types of mutation (*e.g.* parallel, convergent or reverse mutations) could arise at the same site.

Our analyses based on a restricted highly conserved region of the HD showed that both synonymous and non-synonymous rates increased in many comparisons among Cynolebiatinae subfamily species and other PG1 Hox sequences. However, the highest codon usage bias value among them (CBI = 0.846) could be indicating that some synonymous substitutions were also constrained for these *Hox* genes. Further substitution pattern analysis among Cynolebiatinae *Hox* genes could be important in detecting the characteristic evolutionary signature of duplicated divergent evolution among the *Hox* genes, as has previously been found in the ray-finned fish HoxA cluster (Wagner *et al.*, 2005).

Many molecular evolutionary questions remain open about the selective constraints acting on *Hox* genes of annual fish of the Cynolebiatinae subfamily since high levels

of mitochondrial sequence divergence and high rates of amino acid replacements among coding regions have been previously detected (García *et al.*, 2000, 2002). In this case, *Hox* genes were assigned to a series of physiological and life-history variables such as generation time, life span, age of first reproduction, rate of population increase and metabolic rate that are certainly not exclusively confined to mitochondria.

While the limited homeobox sequences described in the present paper allowed us to assign them to specific paralogue groups, the *Austrolebias* partial Hox sequences could not be assigned to specific Hox clusters.. More sequence data, *e.g.*, from the entire homeobox region and flanking regions, are needed to assign these genes to their correct clusters. Furthermore, it would be interesting to determine the number of Hox clusters in even one species of the Cynolebiatinae subfamily, since at least seven Hox clusters have been described for many teleosts (Amores *et al.*, 1998, 2004; Aparicio *et al.*, 2002; Naruse *et al.*, 2004) and only four in the killifish *Fundulus heteroclitus* (Misof and Wagner 1996). Given that the development of annual killifishes shows reversible arrest at three well-defined stages, it would also be interesting to compare *Hox* gene expression between species of the subfamily Cynolebiatinae.

Acknowledgments

We wish to thank the following people: L. Malabarba for kindly donate individuals of *A. adloffi* from RS, Brazil (1991), E. Castillo for kindly provided the degenerate primers SO1 and SO2, G. Bedó and Y. Panzera for their technical assistance, F. Panzera for his helpful corrections and critical reading and an anonymous reviewer for useful comments on an earlier draft of this article. The authors are also grateful to the Japanese government for the donation of equipment.

References

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410.

Amores A, Suzuki T, Yan Y-L, Pomeroy J, Singer A, Amemiya C and Postlethwait JH (2004) Developmental roles of pufferfish Hox clusters and genome evolution in ray-finned fish. Genome Res 14:1-10.

Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, *et al.* (1998) Zebrafish hox clusters and vertebrate genome evolution. Science 282:1711-1714.

Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, *et al.* (2002) Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. Science 297:1301-1310.

Aparicio S, Hawker K, Cottage A, Mikawa Y, Zuo L, Venkatesh B, Chen E, *et al.* (1997) Organization of the *Fugu rubripes* Hox clusters: Evidence for continuing evolution of vertebrate Hox complexes. Nat Genet 16:79-83.

Arezo MJ, Pereiro L and Berois N (2005) Early development in the annual fish *Cynolebias viarius*. J Fish Biol 66:1357-1370.

Bayascas JR, Castillo E, Muñoz-Mármol AM and Saló E (1997) Planarian Hox genes: Novel patterns of expression during regeneration. Development 124:141-148.

Carter CA and Wourms JP (1991) Cell behaviour during early development in the South American annual fishes of the genus *Cynolebias*. J Morphol 210:247-266.

Chauvet S, Merabet S, Bolder D, Scott MP, Pradel J and Graba Y (2000) Distinct Hox protein sequences determine specificity in different tissues. Proc Natl Acad Sci USA 97:4064-4069.

Costa WJE (1998) Phylogeny and classification of Rivulidae revisited: Origin and evolution of annualism and miniaturization in rivulid fishes (Cyprinodontiformes, Aplocheiloidei) J Comp Biol 3:33-92.

Costa WJEM (1995) Pearl Killifishes: The Cynolebiatinae: Systematics and Biogeography of a Neotropical Annual Fish Subfamily (Ciprinodontiformes, Rivulidae). T.F.H. Publications, Neptune City, 128 pp.

García G, Alvarez-Valin F and Gómez N (2002). Mitochondrial genes: Signals and noise in the phylogenetics reconstruction of the annual killifish genus *Cynolebias* (Cyprinodontiformes, Rivulidae). Biol J Linn Soc Lond 76:49-59.

García G, Wlasiuk G and Lessa EP (2000). High levels of mitochondrial cytochrome b divergence in annual killifishes of de genus *Cynolebias* (Cyprinodontiformes, Rivulidae). Zool J Linn Soc 129:93-110.

Gehring WJ, Affolter M and Burglin T (1994a) Homeodomain proteins. Annu Rev Biochem 63:487-526.

Gehring WJ, Qian YQ, Billeter M, Furokubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, *et al.* (1994b) Homeodomain-DNA recognition. Cell 78:211-223.

Golding GB and Dean AM (1998) The structural basis of molecular adaptation. Mol Biol Evol. 15:355-369.

Guex N and Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18:2714-2723.

Hand SC and Podrabsky JE (2000) Bioenergetics of diapause and quiescence in aquatic animals. Thermochim Acta 349:31-42.

Hartl DL (2000) A Primer of Population Genetics. 3rd edition. Sinauer Associates, Sunderland, 221 pp.

Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, Bouneau L, *et al.* (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature 431:946-957.

Kim CB, Amemiya C, Bailey W, Kawasaki K, Mezey J, Miller W, Minoshima S, *et al.* (2000) Hox cluster genomics in the horn shark, *Heterodontus francisci.* Proc Natl Acad Sci USA 97:1655-1660.

Kissinger CR, Liu BS, Martin-Blanco E, Kornberg TB and Pabo CO (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: A framework for understanding homeodomain-DNA interactions. Cell 63:579-590.

Koh EG, Lam K, Christoffels A, Erdmann MV, Brenner S and Venkatesh B (2003) *Hox* gene clusters in the Indonesian coelacanth, *Latimeria menadoensis*. Proc Natl Acad Sci USA 100:1084-1088.

Krumlauf R (1994) *Hox* genes in vertebrate development. Cell 78:191-201.

Kumar S, Tamura K and Nei M (2004) MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. Brief Bioinform 5:150-163.

Ledje C, Kim CB and Ruddle FH (2002) Characterization of *Hox* genes in the bichir, *Polypterus palmas*. J Exp Zool 294:107-111.

Lewis EB (1978) A gene complex controlling segmentation in *Drosophila*. Nature 276:565-570.

McGinnis W and Krumlauf R (1992) Homeobox genes and axial patterning. Cell 68:283-302.

McGinnis W, Levine MS, Hafen E, Kuroiwa A and Gehring WJ (1984a) A conserved DNA sequence in homeotic genes of the *Drosophila* Antennapedia and bithorax complexes. Nature 308:428-433.

McGinnis W, Garber RL, Wirz J, Kuroiwa A and Gehring WJ (1984b) A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. Cell 37:403-408.

Medrano JF, Aasen E and Sharrow L (1990) DNA extraction from nucleated red blood cells. Biotechniques 8:43.

Misof BY and Wagner GP (1996) Evidence for four Hox clusters in the killifish, *Fundulus heteroclitus* (Teleostei). Mol Phylogenet Evol 5:309-322.

Naruse K, Tanaka M, Mita K, Shima A, Postleithwait J and Mitani H (2004) A medaka gene map: The trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping. Genome Res 14:820-828.

Naruse K, Fukamachi S, Mitani H, Kondo M, Matsuoka T, Kondo S, Hanamura N, *et al.* (2000) A detailed linkage map of Medaka, *Oryzias latipes*: Comparative genomics and genome evolution. Genetics 154:1773-1784.

Passner JM, Ryoo HD, Shen L, Mann RS and Aggarwal AK (1999) Structure of a DNA-bound ultrabithorax-extradenticle homeodomain complex. Nature 397:714-719.

Podrabsky JE, Carpenter JF and Hand SC (2001) Survival of water stress in annual fish embryos: Dehydration avoidance and egg envelope amyloid fibers. Am J Physiol Regul Integr Comp Physiol 280:R123-R131.

Podrabsky JE and Hand SC (2000) Depression of protein synthesis during diapause in embryos of the annual killifish *Austrofundulus limnaeus*. Physiol Biochem Zool 73:799-808.

Podrabsky JE and Hand SC (1999) The bioenergetics of embryonic diapause in an annual killifish, *Austrofundulus limnaeus*. J Exp Biol 202:2567-2580.

Prince VE, Joly L, Ekker M and Ho RK (1998) Zebrafish *hox* genes: Genomic organization and modified colinear expression patterns in the trunk. Development 125:407-420.

Qian YQ, Billeter M, Otting G, Müller M, Gehring WJ and Wüthrich K (1989) The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressors. Cell 59:573-580.

Rozas J, Sánchez-Delbarrio JC, Messeguer X and Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496-2497.

Ruddle FH, Bentley KL, Murtha MT and Risch N (1994) Gene loss and gain in the evolution of the vertebrates. Development Suppl:155-161.

Saitou N and Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425.

Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning. A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory Press, New York.

Sanguinetti CJ, Dias Neto E and Simpson AJ (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques 17:914-921.

Schwede T, Kopp J, Guex N and Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res 31:3381-3385.

Sharkey M, Graba Y and Scott MP (1997) *Hox* genes in evolution: Protein surfaces and paralog group. Trends Genet 13:145-151.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876-4882.

Wagner GP, Takahashi K, Lynch V, Prohaska SJ, Fried C, Stadler PF and Amemiya C (2005) Molecular evolution of duplicated ray finned fish HoxA clusters: Increased synonymous substitution rate and asymmetrical co-divergence of coding and non-coding sequences. J Mol Evol 60:665-676.

Wolberger C (1996) Homeodomain interactions. Curr Opin Struct Biol 6:62-68.

Wourms JP (1972a) The developmental biology of annual fishes. I. Stages in the normal development of *Austrofundulus myersi* Dahl. J Exp Zool 182:143-168.

Wourms JP (1972b) The developmental biology of fishes. III. Pre-embryonic diapause of variable duration in the eggs of annual fishes. J Exp Zool 182:389-414.

Wourms JP (1967) Annual fishes. In: Wilt FW and Wessels NK (eds) Methods in Development Biology. Thomas Y. Crowell, New York, pp 123-137.

Internet Resources

Gene Runner software, http://www.generunner.com.

Basic Local Alignment Search Tool (BLAST), http://www.ncbi.nlm.nih.gov/BLAST (August 1, 2005).

Swiss Model software, http://swissmodel.expasy.org.

Associate Editor: Ricardo Guelerman P. Ramos