BMC Genomics

RESEARCH Open Access

The identification, adaptive evolutionary analyses and mRNA expression levels of homeobox (*hox*) genes in the Chinese mitten crab *Eriocheir sinensis*

Shasha Chen[1], Xianfeng Jiang[1], Longjie Xia[1], Zhiyi Chen[1], Kaiya Zhou[1], Jie Yan[1*] and Peng Li[1*]

Abstract

Background Arthropods are the largest group in the animal kingdom and are morphologically characterized by heterorhythmic segments. Brachyuran decapod crustaceans undergo brachyurization metamorphosis in the early developmental process, characterized by a reduced abdomen that is folded beneath the cephalothorax and inserted between the pereiopods or in a special cavity. As the main cause of major alterations in the evolution of animal body plans, *Hox* genes encode transcription factors and are involved in bilaterian anterior-posterior axis patterning.

Results We found eight *Hox* genes (*labial*, *proboscipedia*, *Deformed*, *zerknüllt*, *Sex combs reduced*, *Antennapedia*, *Ultrabithorax*, *fushi tarazu*, *abdominal-A* and *Abdominal-B*) in *Eriocheir sinensis*. The phylogenetic topology of 13 arthropod *Hox* genes was closely related to traditional taxonomic groupings. Genome collinearity analysis was performed using genomic data and chromosomal location data of *E. sinensis* and *Portunus trituratus*. We found that their chromosomes were highly collinear, and there was a corresponding collinear relationship between the three *Hox* genes (*lab, ftz* and *Abd-B*). The mRNA expression levels of *Scr* and *Antp* fluctuated significantly in different developmental stages of *E. sinensis*, especially in the brachyurization stages. Evolutionary analysis indicated the presence of positively selected sites in *Ubx*.

Conclusions In this study, we used genome-wide analysis to identify and analyze all members of the *Hox* genes in *E. sinensis*. Our data will contribute to a better understanding of *Hox* genes in *E. sinensis* and provide useful molecular evolutionary information for further investigation on their roles in the brachyurization of crabs.

Highlights

- Eight *Hox* genes (*lab*, *Dfd*, *Scr*, *ftz*, *Antp*, *Ubx*, *abd-A*, and *Abd-B*) were identified in *Eriocheir sinensis*.
- There was a corresponding collinear relationship between the three *Hox* genes (*lab, ftz* and *Abd-B*) in *E. sinensis* and *Portunus trituratus*.
- Positively selected sites in *Ubx* (1 M 0.952*, 3 S 0.998**, 4 Y 0.998**, P = 0.0000) were identified in *E. sinensis*.

*Correspondence:
Jie Yan
yanjie@njnu.edu.cn
Peng Li
lipeng@njnu.edu.cn

Full list of author information is available at the end of the article

- The full-length cDNA of *Scr* and *Antp* in *E. sinensis* were cloned and characterized. *Hox* genes (*Scr* and *Antp*) were differentially expressed during the larval development of *E. sinensis*.

Keywords *Eriocheir sinensis*, Homeobox genes, Developmental regulation, Adaptive evolution

Introduction

In the lifecycle of an organism, the formation of the body and its organs during ontogeny or regeneration is known as morphogenesis. Homeobox genes are essential for controlling embryonic patterning and regulating these anatomical changes. These genes are mainly characterized by an approximately 180 bp (bp) homeodomain conserved region [1]. Homeobox genes were first identified in *Drosophila melanogaster* [2], followed by sequential discovery in all multicellular organisms, from parasites to vertebrates, plants and fungi, and they are highly evolutionarily conserved [3, 4].

Traditionally, homeobox genes are divided into two subfamilies. One contains genes that are arranged in clusters on chromosomes and expressed along the main anterior-posterior (A-P) axis of the animal body, which are called *Hox* genes or type-I homeobox genes. The other are non-A-P homeobox genes, which are not arranged in clusters but rather are scattered across different chromosomes, and they are named type-II homeobox genes based on sequence similarity [5]. In particular, *Hox* genes are specifically related to the regulation of body plan development during the embryonic stage of the life cycle [6]. *Hox* genes are mostly involved in the morphogenesis and differentiation of tissues or organs such as limbs [7], brain [8], muscle [9], blood [10], and bone [11]. Each *Hox* gene contains a conserved homeobox domain (homeodomain, HD) involved in the regulation of downstream genes [12]. The homeobox domain is a 60-amino-acid protein sequence with an alpha-helical secondary structure [13]. The activity of these regulatory proteins encoded by homeobox genes is most pronounced at the early embryonic stage, when the embryo is forming a body axis [14]. The expression patterns of *Hox* genes show a linear pattern: genes at the 3'-end are expressed first and regulate the anterior development of the embryo, while genes at the 5'-end are expressed later and regulate the posterior development of the embryo [15]. This phenomenon is called spatial collinearity and temporal collinearity.

Arthropods have maximum species richness in the animal world, and they are also the most morphologically diverse group, accounting for approximately 80% of all animal species [16]. To date, ten ancestral homologs have been confirmed in arthropods: *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *zerknüllt* (*zen*), *Sex combs reduced* (*Scr*), *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *fushi tarazu* (*ftz*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*) [17]. *Drosophila melanogaster* has eight *Hox genes* and was found to be divided into two homologous heterogeneous gene clusters on chromosome 3: the antennal foot complex Antp-C (antennapedia complex) gene cluster and the double-thorax complex BX-C (bithorax complex) gene cluster [18]. Antp-C is mainly involved in the developmental regulation of the head and chest, and BX-C is mainly involved in the regulation of posterior thoracic and abdominal development [19].

In the early developmental process, brachyuran decapod crustaceans undergo brachyurization metamorphosis, which is characterized by a reduced abdomen that is folded beneath the cephalothorax and inserted between the pereiopods or in a special cavity [20, 21]. The Chinese mitten crab (*Eriocheir sinensis*) is a limnic and intertidal crab mainly distributed in eastern and northern China. The juvenile development period of *E. sinensis* normally consists of five zoeae stages and a megalopa stage, and it subsequently enters the first juvenile crab stage and develops into an adult crab after multiple molts [22]. The most conspicuous feature of morphological change from megalopa to adult crab is brachyurization, including abdominal degeneration. A previous study showed that *Hox* genes are involved in embryonic development and appendage development in crustaceans [23], but the molecular regulatory mechanisms of brachyurization are poorly studied [24].

In this study, we used genome-wide analysis to identify all members of the *Hox* gene family between *E. sinensis* and 12 other arthropod species. These sequences were aligned and used to construct a phylogenetic tree. Moreover, the full-length cDNA of the *Scr* and *Antp* genes were cloned and characterized from the Chinese mitten crab (*E. sinensis)* using expressed sequence tag (EST) and rapid amplification of cDNA end (RACE) techniques. The mRNA expression profiles of *Scr* and *Antp* transcripts in various samples (different developmental stages of juvenile crabs) were measured by fluorescent real-time quantitative PCR (RT–qPCR). The selective pressure of *Scr* and *Antp* was determined to reveal the relationship between brachyurization and adaptive evolution of *Hox* genes in Brachyura crabs. Our data will contribute to a better understanding of *Hox* genes in *E. sinensis* and provide useful molecular evolutionary information for further investigation on their roles in brachyurization of crabs.

Result

Identification and analyses of *Hox genes* in the Chinese mitten crab

Eight *Hox* genes (*lab*, *Dfd*, *Scr*, *ftz*, *Antp*, *Ubx*, *abd-A* and *Abd-B*) were identified through BLAST analysis of the *E. sinensis* genome database. A sequence alignment of the eight Hox proteins in *E. sinensis* with those of other species, such as *D. melanogaster* and *L. vannamei*, revealed the presence of a residue sequence of YPWM and a conserved domain of homeodomain except *Abd-B* (Supplementary file S1). We analyzed the physical properties and amino acid composition of the Hox proteins in Chinese mitten crab and found that it was rich in Ala, Gly, Pro, Gln, Ser and Thr (Supplementary file S2 and Supplementary file S3).

The amino acid sequences of eight Hox proteins from *E. sinensis* were clustered. The results showed that the eight *Hox* genes were classified into two major clades. *Scr*, *Dfd*, *ftz*, *Ubx* and *abd-A* clustered together, and the other group included *Antp*, *Abd-B* and *lab*. The genomes of *E. sinensis* and *P. trituratus* were analyzed for collinearity, and there was a high degree of collinearity among multiple chromosomes (Fig. 1). This reflects the sister group relationship of *E. sinensis* and *P. trituratus*. Among them, the *Hox* genes *(lab, ftz* and *Abd-B*) present on chromosome 21 in *E. sinensis* are colinear with those in the *P. trituratus* genome, retaining more ancestral traits.

Phylogenetic analysis of the *Hox* gene family

We found a total of 112 *Hox* genes in 13 selected species and mapped their transcription orientation from genomic data with annotation files (Fig. 2). Aligning Hox protein sequences, a total of 112 Hox protein sequences from 13 selected species were used to build a phylogenetic tree using the maximum likelihood approach (Fig. 3, Supplementary file S5). *E. sinensis* and *P. trituberculatus*, belonging to Eubrachyura, are sister groups to shrimps and grouped with hermit crabs into one class. Each *Hox* gene of *Eriocheir sinensis* was first clustered with the *Hox* gene of crabs and shrimps. This clustering phenomenon

Fig. 1 Collinearity patterns between genomic regions from *Eriocheir sinensis* and *Portunus trituratus*

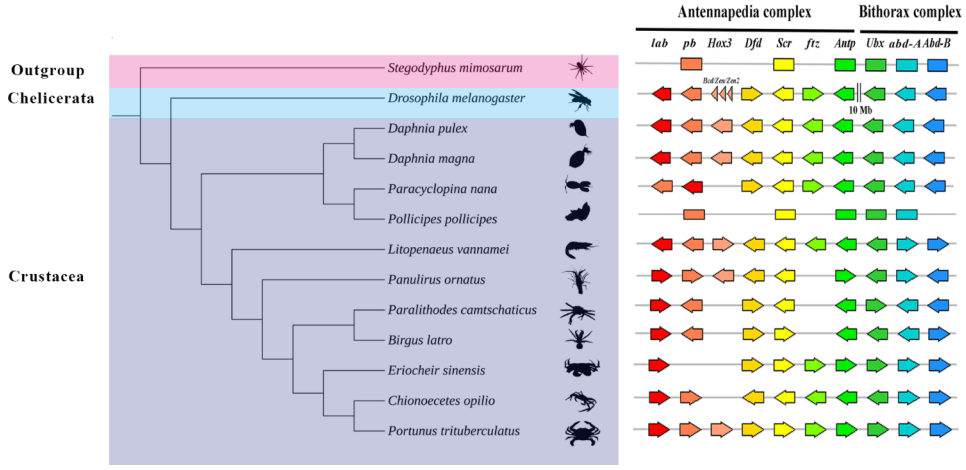

Fig. 2 Comparison of Hox genes among arthropods

shows that crustaceans are relatively closely related to insects and slightly distantly related to Chelicerata species. The results support the traditional taxonomy.

Adaptive evolution analyses of *Hox* genes

Different ω ratio models were used to determine if *Hox* genes in the arthropod species have experienced positive selection. In the site model, the positively selected site with a posterior probability greater than 0.95 was identified only in *Ubx.* In the branching model, *E. sinensis* was set as the foreground branch, and the other branches were set as background branches. Out of the seven *Hox* genes (*lab, Scr, Antp, Ubx, ftz, abd-A* and *Abd-B*), the results showed that there was a significant difference in the evolution rate of *Ubx* and *abd-A* between *E. sinensis* and other animals, so *Ubx* and *abd-A* were considered rapidly evolving genes. The results of branch-site model A revealed evidence of positive selection on *Ubx* (1 M 0.952*, 3 S 0.998**, 4 Y 0.998**) among the arthropod lineages (Tables 1, 2 and 3). Setting Brachyura as the foreground branch for selection pressure analysis, we found that in the branch model, the evolution rate of *Ubx* was different in Brachyura from the background branch. In the branch site model, *Antp* (57 K 1.000**) had a positive selection site, and its mutation site was the same as the positive selection site of the Chinese mitten crab. This indicates that the *Antp* gene is possibly involved in the brachyurization development of crabs (Supplementary file S6 and Supplementary file S7).

Full-length cDNA structure of *Scr* and *Antp* and mRNA expression of the *Antp* and *Scr* genes in *E. sinensis* during different developmental periods

The full-length cDNA sequences of *Scr* (1770 bp, **KP822930**) and *Antp* (2413 bp, **KP822927**) were obtained by DNAStar Lasergene 7.1 after the sequence splicing of the *Scr* and *Antp* genes. Through analysis, it was found that the full-length cDNA sequence of *Scr* has a 5'-end untranslated region (UTR) of 424 bp and a 3'-end untranslated region of 350 bp, and its open reading frame (ORF) contains 996 bp, encoding 332 amino acids. The full-length cDNA sequence of *Antp* has a 5'-end untranslated region (UTR) of 324 bp and a 3'-end untranslated region of 1123 bp, and its ORF contains 966 bp, encoding 322 amino acids. Their open reading frames both contain a conserved domain (Homeobox) (Supplementary files S4).

The mRNA transcripts of *Scr* and *Antp* were detected in all tested developmental stages of the Chinese mitten crab. The mRNA expression of *Scr* in *E. sinensis* was relatively higher at stages O and Z1 to Z4 and lower at stages Z5, M and J1 to J3 (Fig. 4A). The mRNA expression level of *Antp* in *E. sinensis* was the highest at stage O, and the expression of *Antp* gradually decreased with the development and growth of crabs (Fig. 4B).

Discussion

In this study, we identified a single *Hox* gene cluster including seven typical homeotic genes and one additional homeotic gene from the genome of *E. sinensis*. However, orthologs of *pb* and *Hox3* were not found in *E. sinensis* based on comparison with the sequences of arthropods, including *D. melanogaster* and *L. vannamei*. However, because 10 *Hox* genes were extracted from the

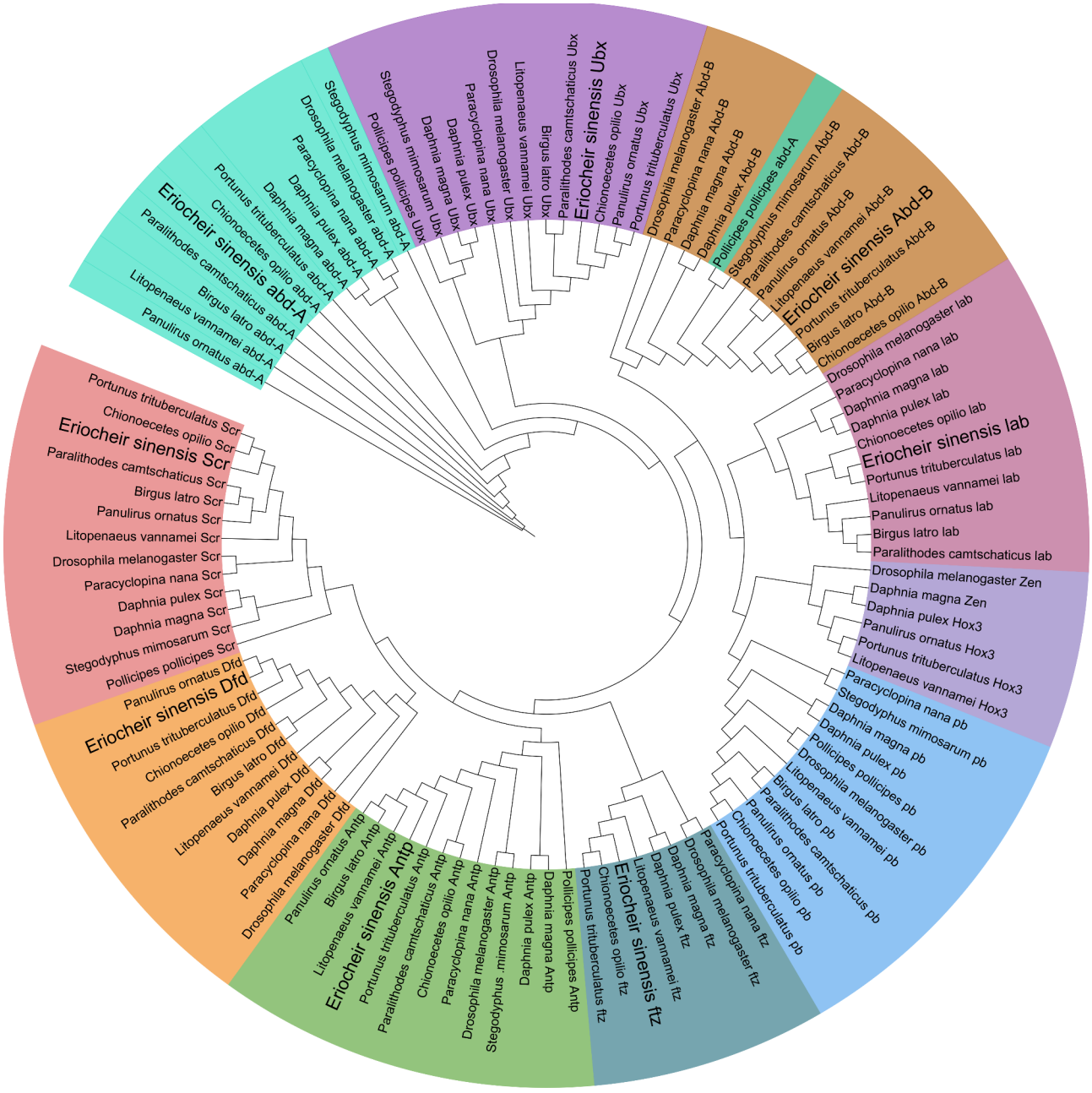

Fig. 3 Phylogenetic tree of Hox among arthropods

transcriptome of the Chinese mitten crab in a previous study [25], the *pb* and *Hox3* genes were not found in this study, perhaps due to incomplete genomic assembly for *E. sinensis*.

Previous studies have found that the repertoire of *Hox* genes differs among different species. *D. melanogaster*, *L. vannamei*, *Daphnia pulex* and *Daphnia magna* are represented by complete sequences for all ten *Hox* genes [26, 27]. However, only nine *Hox* genes from *Paracyclopina nana* were identified, all except *Hox3* [28]. A *Hox3* gene ortholog was also not present in the *Hox* gene cluster of *Macrobrachium olfersii* [29]. Moreover, *pb*, *Hox3* and *ftz* were missing in two species of copepods, *Tigriopus japonicus* and *T. kingsejongensis* [30]. Because *Hox* gene clusters are expressed along the main anterior-posterior (A-P) axis, dynamic mutation in the *Hox* gene causes the ectopic development of a given organ. Normal development of *D. melanogaster* results in a pair of balancing rods in the posterior thorax, but a mutation of *Ubx* will cause flies to grow a pair of wings in the posterior thorax [31]. The data of *Hox* gene sequences in arthropods, especially crustaceans, will contribute to explaining the

Table 1 Selective pressure analysis of *Hox* based on site model

Gene	Models	np	-lnL	LRT P- values	positively selected sites (*PP* > 0.95)
lab	M7	23	1453.233	0.2312	Not allowed
	M8	25	1451.769		————
Scr	M7	26	3338.035	0.0000	Not allowed
	M8	28	3318.876		98 Q 0.802
Antp	M7	24	1394.409	0.9993	Not allowed
	M8	26	1394.409		————
Ubx	M7	18	2034.683	0.0173	Not allowed
	M8	20	2030.628		120 T 0.999*
ftz	M7	16	4599.532	0.9998	Not allowed
	M8	18	4599.532		57 P 0.533
abd-A	M7	18	2365.844	0.9990	Not allowed
	M8	20	2365.845		————
Abd-B	M7	21	2656.995	0.9997	Not allowed
	M8	23	2656.995		————

Table 2 Selective pressure analysis of *Hox* based on branch model

Gene	Models (branch *E. sinensis*)	np	-lnL	LRT P- values	Omega values
lab	one-ratio (M0)	22	1455.098	0.2931	$\omega_0 = 0.0850$
	two-ratio	23	1454.546		$\omega_1 = 0.0001$
Scr	one-ratio (M0)	25	3344.156	0.2087	$\omega_0 = 0.1283$
	two-ratio	26	3343.366		$\omega_1 = 0.6928$
Antp	one-ratio (M0)	23	1464.302	0.4098	$\omega_0 = 0.0509$
	two-ratio	24	1463.961		$\omega_1 = 0.4207$
Ubx	one-ratio (M0)	17	2067.615	0.0015	$\omega_0 = 0.0336$
	two-ratio	18	2062.555		$\omega_1 = 0.2125$
ftz	one-ratio (M0)	15	4713.532	0.6572	$\omega_0 = 0.1134$
	two-ratio	16	4713.434		$\omega_1 = 0.1601$
abd-A	one-ratio (M0)	17	2359.625	0.5271	$\omega_0 = 0.0047$
	two-ratio	18	2359.425		$\omega_1 = 0.0001$
Abd-B	one-ratio (M0)	20	2680.479	0.1622	$\omega_0 = 0.0232$
	two-ratio	21	2681.455		$\omega_1 = 20.9758$

Table 3 Selective pressure analysis of *Hox* based on branch site model

Gene	Models (branch *E. sinensis*)	np	-lnL	LRT P- values	positively selected sites (*PP* > 0.95)
lab	Model A	25	1452.449	1.0000	————
	Model A null	24	1452.449		Not allowed
Scr	Model A	28	3316.316	1.0000	21 G 0.943,45 S 0.937
	Model A null	27	3316.316		
Antp	Model A	26	1392.478	0.0456	56 M 0.805,57 K 0.961*
	Model A null	25	1394.476		
Ubx	Model A	20	2037.891	0.0001	1 M 0.960*, 3 S 0.998**, 4 Y 0.998**, 5 E 0.552, 6 Q 0.684
	Model A null	19	2045.864		
ftz	Model A	18	4640.752	0.9989	————
	Model A null	17	4640.752		Not allowed
abd-A	Model A	20	2359.625	0.9989	————
	Model A null	19	2359.625		Not allowed
Abd-B	Model A	23	2678.448	0.9975	————
	Model A null	22	2678.448		Not allowed

common morphological changes and evolutionary process in crustaceans during metamorphosis.

Compared with higher vertebrates, studies on *Hox* genes in arthropods lag far behind. *Hox* genes have been reported to play important roles in cell division, paired appendage development, axial morphogenesis regulation and skeleton maintenance [32, 33]. The basal expression levels of *Hox* genes demonstrated different patterns of expression over different developmental stages. We selected *Scr* and *Antp* for further investigation because we only cloned these two *Hox* genes (failed to acquire the full-length cDNA of other *Hox* genes in *E. sinensis*) in the early stage of our experiment (before the genome of *E. sinensis* was determined) and examined the mRNA expression patterns of these two genes. In this experiment, the mRNA expression patterns of the *Scr* and *Antp* genes were detected at all 10 stages of juvenile development in *E. sinensis*, suggesting that *Scr* and *Antp* may play a regulatory role in the juvenile development of *E. sinensis*. Previous studies indicated that *Scr* controls the development of the insect prothorax [34], and *Antp* is involved in regulating morphological changes in the abdominal appendages in crustaceans [35, 36]. The main morphological changes in brachyurization development mainly occurred in stages Z5, M and J1 of *E. sinensis* (Fig. 4C), and *Scr* and *Antp* mRNA were expressed at the lowest levels in stages Z5 ~ J3, indicating the possibility that these *Hox* genes are partially involved in brachyurization and further developmental regulation. The thoracic function of *E. sinensis* is increased with shortened abdomen, whereas the abdominal function is weakened, certain abdominal segments heal and fold under the cephalothorax, and the motor function is lost. The mRNA expression of *Scr* and *Antp* changed dramatically during their respective larval developmental phases, mirroring changes in the morphology of *E. sinensis* larvae. During the egg stage, when the Chinese mitten crab individuals were fully formed, the relative expression of the *Scr* and *Antp* gene mRNAs peaked. In contrast, during the Z5 stage, their expression reached its lowest point, which caused abdominal degeneration in *E. sinensis*. This significant change ($p < 0.01$) implied that these two genes might be involved in regulating the short tailing process of *E. sinensis*. Young et al. indicated that *Hox* genes are highly expressed in the vertebrate posterior embryonic region and play a role in posterior tissue generation [37]. This finding fits with the so-called dose effect that quantitative variation in the levels of *Hox* gene products can affect segment morphology in subtle ways [38].

The rate of gene evolution depends on the substitution rate of the nucleotides in the nucleic acid molecule over a certain period of time, and the selection pressure reacts to the rate of gene evolution. Based on selective pressure analyses of *Hox* genes, *E. sinensis* was set as

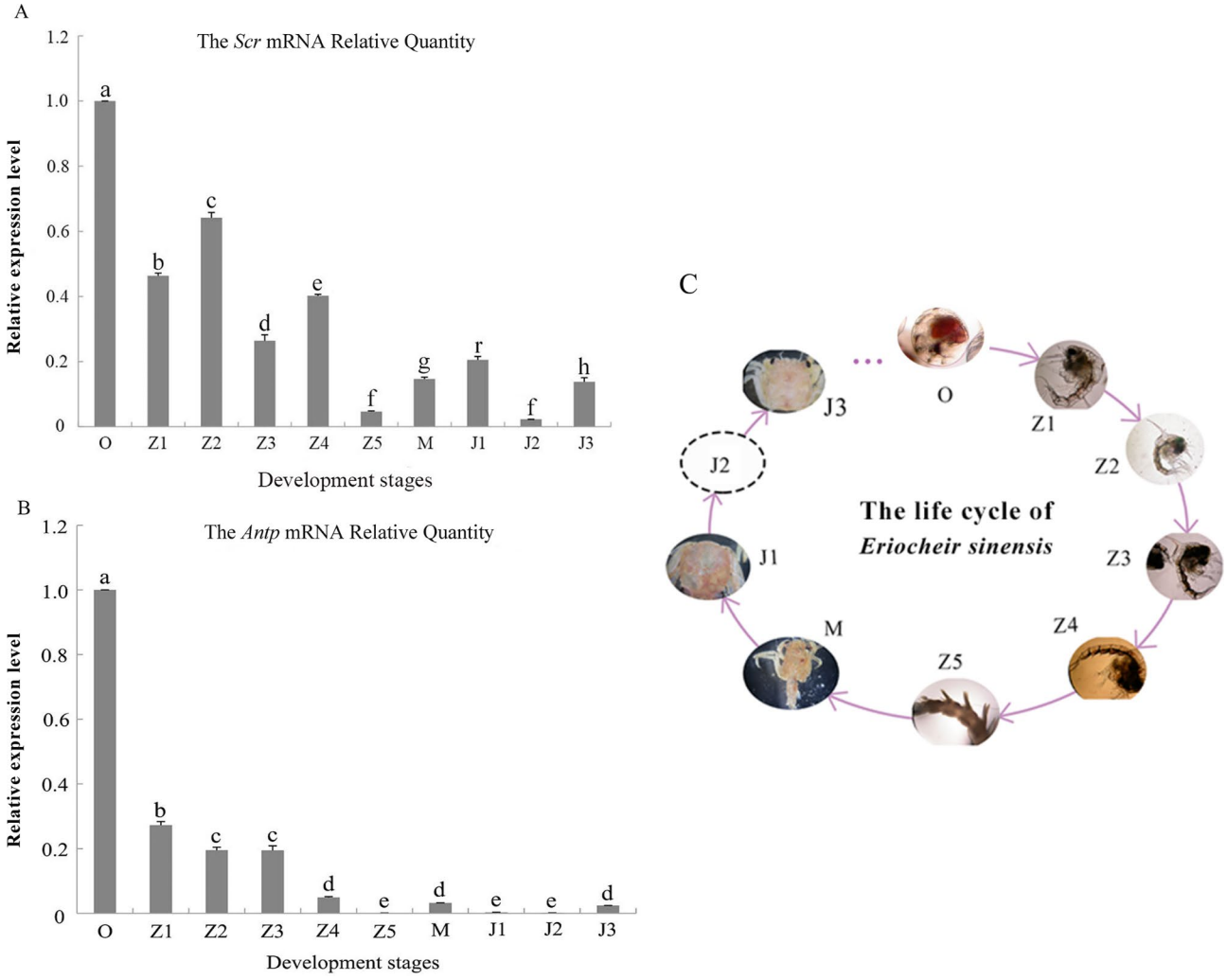

Fig. 4 The developmental periods of *Eriocheir sinensis*. A, B : the relative mRNA expression of *Antp* and *Scr* of *Eriocheir sinensis* in different development stages (n = 9, mean ± SD). C: the pictures of life cycle of the Chinese mitten crab photoed by Peng Li

the foreground branching clade in the branching model and branch site model. The results revealed that six *Hox* genes had no positive selection signal, but *Antp and Ubx* did, indicating that the evolutionary rates of *Antp and Ubx* were perhaps different among different species. A previous study reported that in the early course of gene evolution, the evolutionary rate of the *Hox* gene in actinopterygian and sarcopterygian fishes is faster than that in terrestrial vertebrates [39]. The different rates of *Hox* gene evolution imply that *Hox* genes may regulate animal morphology to adapt to the environment. Overall, there have been a relatively detailed investigations of the evolution and expression pattern of *Hox* genes in Brachyura crabs, but further work is needed to focus on the molecular evolutionary mechanism and functions in developmental regulation of brachyurization.

Conclusion

Here, we explored eight *Hox* genes from the chromosomal level genome of the Chinese mitten crab *E. sinensis.* A total of 112 putative *Hox* genes were identified in 13 arthropod species. According to the collinear analysis, there are a large number of collinear fragments between *E. sinensis* and *P. trituratus*, which implies that they accumulate less variation with shorter differentiation times and retain more ancestral traits. Evolutionary analyses indicated the presence of positively selected sites in *Ubx*. In the branch model, the *Ubx* and *abd-A* genes in *E. sinensis* were identified as rapidly evolving genes. The full-length cDNA sequences of *Scr* and *Antp* were cloned, and the mRNA expression patterns of the two *Hox* genes in *E. sinensis* were determined at different developmental stages. The expression levels of *Scr* and *Antp* in *E. sinensis* were the highest at stage O and fluctuated significantly in different developmental periods. Our data will contribute

to a better understanding of the *Hox* gene family in *E. sinensis* and provide useful molecular evolutionary information for further investigation into their roles in the brachyurization of crabs.

Materials and methods

Sample collection

Tissue samples of *E. sinensis* were collected from the healthy Chinese mitten crabs that were bred in an aquatic nursery of Rudong County, Nantong City, Jiangsu Province, China. The samples were collected at the fertilized egg stage (stage O), all zoea stages (stage Z1 ~ Z5), megalopa stage (stage M) and juvenile crab stages (stage J1 ~ J3, with approximately 30 mg for each sample). These fresh tissues were stabilized immediately in RNA*later* RNA Stabilization Reagent (QIAGEN, Germany), kept overnight at +4 °C, and finally stored at −20 °C before RNA isolation.

Preparation of total RNA

Following the manufacturer's instructions, we used an RNeasy Mini Kit (Qiagen, Germany) to extract total RNA from different samples from the Chinese mitten crab. The RNA concentration and purity were assessed spectrophotometrically by measuring the absorbance of the solution at 260 and 280 nm in a biophotometer (Eppendorf, Germany). The RNA fragmentation state was evaluated by 1% agarose denatured gel electrophoresis.

Table 4 The GenBank accession used in this study

subphylum	phylum	scientific name	accession
Crustacea	Malacostraca	*Eriocheir sinensis*	GCA_003336515.1
	Malacostraca	*Chionoecetes opilio*	GCA_016584305.1
	Malacostraca	*Portunus trituberculatus*	GCA_008373055.1
	Malacostraca	*Birgus latro*	GCA_018397915.1
	Malacostraca	*Paralithodes camtschaticus*	GCA_018397895.1
	Malacostraca	*Litopenaeus vannamei*	GCA_003789085.1
	Malacostraca	*Panulirus ornatus*	GCA_018397875.1
	Branchiopoda	*Daphnia magna*	GCA_020631705.2
	Branchiopoda	*Daphnia pulex*	GCA_021134715.1
	Copepoda	*Paracyclopina nana*	GCA_019096065.1
	Maxillopoda	*Pollicipes pollicipes*	GCA_011947565.3
Chelicerata	Arachnida	*Stegodyphus mimosarum*	GCA_000611955.2
Hexapoda	Insecta	*Drosophila melanogaster*	GCA_003401745.1

Phylogenetic analysis of *Hox* genes

For comparative genomics analyses, we used all the annotated genes in the Chinese mitten crab and thirteen other arthropod species. Genome assemblies of 13 species were obtained from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) and GigaDB database (http://gigadb.org/) (Table 4). The Hox amino acid sequences of *Litopenaeus vannamei* and *Drosophila melanogaster* were used as queries. BLASTP in BLAST v2.2.23 was used for similarity searches in the genomes of the selected species. An e-value of 10^{-5} was used as an initial cutoff to identify significant matches. Hits < 100 codons and overlapping sequences were excluded.

The ProtParam software tool at the ExPASy portal (https://web.expasy.org/protparam/) was used to calculate the amino acid percentages of Hox gene-encoded proteins. CLUSTALW 2.0.10 [40] was used to compare the homology of Hox protein sequences in *E. sinensis*, *D. melanogaster* and *L. vannamei*. The gene structure of the *Hox* genes and the distribution of motifs and domains at each sequence were visualized using TBtools version 1.074 [41]. Mapping of gene family members in chromosome locations was performed using the R 4.0.5 package ("ggplot2" and "gggenes") [42]. Amino acid sequences were aligned with MAFFT, and phylogenetic analyses were performed with IQtree on the Phylosuit platform. The phylogenetic analysis of Hox protein sequences used the JTT model for maximum likelihood analysis with 1000 bootstrap replicates. Using genomic data and chromosomal location data of *E. sinensis* and *Portunus trituratus*, cross-species genome collinearity analysis was performed using MCScanX software and then visualized in Circos 0.67.

Analyses of evolutionary pressure

We used comparisons of nonsynonymous/synonymous substitution ratios (dN/dS) to quantify natural selection. Different types of selection were represented by different values: $\omega < 1$ means purifying selection, $\omega = 1$ indicates neutral selection, and $\omega > 1$ means positive selection. Due to the incomplete sequence of *Dfd*, we only performed selective pressure analysis on the other seven *Hox* genes (*lab, Scr, Antp, Ubx, ftz, abd-A* and *Abd-B*) in *E. sinensis*. Analyses of selective pressure were conducted using the Codeml program implemented in the Phylogenetic Analysis by Maximum Likelihood (PAML4.9) package. Different *Hox* genes in crustacean species have experienced positive selection. The selected site model (Site model), branch model (Branch model), and branch site model (Branch-site model) were used to test for the presence of positive branches on the divergent clades of the crustaceans for *Hox*.

Table 5 Primers used for *Scr* and *Antp* genes amplification and real-time quantitative PCR detection

primer name	primer sequences (5′→3′)	primer length (bp)
Scr-F0	5′-CAGATCTACCCGTGGATGAAGAG-3′	23
Scr-R0	5′-CGTTCATGCTCGCCATCTTGTGC-3′	23
Scr-F1	5′-GAACTCCAACGGCGAGACCAAG-3′	22
Scr-R1	5′-CCATCTTGTGCTCCTTCTTCCA-3′	22
Scr-F2	5′-CCTATTGGTTGCTGTCGGTCAC-3′	22
Scr-R2	5′-ATATTTTGATCTGTCGCTCGGT-3′	22
Scr-R3	5′-GCGTAGTCCTGAAGGGTAGTCCAT-3′	24
Antp-F0	5′-CAGCAACAGCAGGCACAGCAACA-3′	23
Antp-R0	5′-CTGCGAGGGGGACGTTGGGGTC-3′	22
Antp-F1	5′-CCAGGATACCTCCCTCCACATG-3′	22
Antp-R1	5′-GGGTCGGCGTATCTGAAAGGCT-3′	22
Antp-R2	5′-TGCTCTTGTTTTCTTTCTTCCA-3′	22
Antp-R3	5′-AGGGCCGATGAGGTGTTTTGTT-3′	22
Antp-F2	5′-TCATAGTGTAAAGTGTATTCTGT-3′	23
Antp-F3	5′-CGGGAAACAGCCTACCAGTGCC-3′	22
Antp-F4	5′-CACGCCGACTCCTGTCATCCCT-3′	22
Antp-F5	5′-CAAGAGCAAAGTGGAGAACGGGAACA-3′	26
Antp-R4	5′-CAAACAATACCAGAGTAGGAGAAGC-3′	25
Antp-R5	5′-GTCTCGTCCGCTTCTCGTTCCTG-3′	23
primer name	primer sequences (5′→3′)	primer length (bp)
QEs-ScrF	5′-GAGTCTGGAAGGAGCCTCTG-3′	20
QEs-ScrR	5′-CGGGTAGATCTGTGGTTGTG-3′	20
Qβ-actinF	5′-CTCCTGCTTGCTGATCCACATC-3′	22
Qβ-actinR	5′-GCATCCACGAGACCACTTACA-3′	21
QEs-AntpF	5′-ATTCCACTTCAACCGCTACC-3′	20
QEs-AntpR	5′-CTGTTCCCGTTCTCCACTTT-3′	20
Qβ-actinF	5′-CTCCTGCTTGCTGATCCACATC-3′	22
Qβ-actinR	5′-GCATCCACGAGACCACTTACA-3′	21

Full-length cDNA amplification and the mRNA expression patterns of *Antp* and *Scr* during different periods

Gene-specific primers for *Scr* and *Antp* gene amplification were designed based on the high-throughput sequencing of the Chinese mitten crab transcriptome sequence obtained earlier in our laboratory (data unpublished). The primers are shown in Table 5. The cDNA was synthesized using a 3′ Full RACE Core Set with PrimeScript™ RTase (Qiagen, Germany). PCR amplification was followed by using TaKaRa LA *Taq* with GC Buffer (TaKaRa, Dalian, China). The *β-actin* gene (GenBank accession no. **HM053699.1**) of *E. sinensis*, as an effective internal control [26, 43], was selected to calibrate the cDNA template. Gene-specific primers for RT−qPCR were designed according to the sequences of the two genes (*Scr* and *Antp*), and these primers are shown in Table 5. Real-time quantitative PCR (RT−qPCR) was conducted using the SYBR® Premix *Ex*-Taq™ II kit (TaKaRa, Dalian, China). The RT−qPCR cycling conditions were as follows: initial denaturation at 95 °C/30 sec; 40 cycles of 95 °C/5 sec and 60 °C/30 sec; 95 °C/15 sec; 60 °C/1 min, 95 °C/15 sec. The RT−qPCR assay was carried out in triplicate on 96-well plates. Employing the formula $RATE = 2^{-\Delta\Delta Ct}$, the comparative Ct method was used to analyze the relative expression levels of *Scr* and *Antp*. Statistical analysis of the *Scr* and *Antp* gene expression data was performed using STATISTICA 10.0. One-way analysis of variance (ANOVA) with Tukey's post hoc tests were performed to assess statistical significance, in which statistical significance was accepted at $p \leq 0.05$ and marked with asterisks or different letters.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-023-09489-w.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (Grant No. 31000954 to PL), Natural Science Foundation of the Jiangsu Higher Education Institutions (19KJA330001) awarded to PL, and Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22_1614) awarded to ZYC, and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

.

Authors' contributions

PL and KYZ designed and supervised the experiments. SSC and XFJ performed the experiments. SSC, LJX, ZYC, XFJ and PL analyzed data. SSC, YJ, and PL prepared the manuscript. All authors discussed the results, and implications and commented on the manuscript.

Funding information

This work was supported by grants from the National Natural Science Foundation of China (Grant No. 31000954 to PL), Natural Science Foundation of the Jiangsu Higher Education Institutions (19KJA330001) awarded to PL, and Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22_1614) awarded to ZYC, and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Data Availability

All data generated or analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

In the present work, the experimental procedures complied with current laws on animal welfare and research in China and the Guide for the Care and Use of Laboratory Animals (8th edition), and were approved by the Institutional Animal Care and Use Committee of Nanjing Normal University and were conducted in accordance with related guidelines [SYXK (Jiangsu) 2020-0047 and IACUC-20220258]. The authors confirmed that animals did not suffer unnecessarily at any stage of experiments in this study.

Consent for publication
Not applicable.

Competing interests
The authors declared that there are no conflicts interests.

Author details
[1]Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210023, China

Received: 18 October 2022 / Accepted: 28 June 2023
Published online: 03 August 2023

References

1. Kaufman TC, Lewis R, Wakimoto B. Cytogenetic analysis of chromosome 3 in Drosophila melanogaster: the homoeotic gene complex in polytene chromosome interval 84a-B[J]. Genetics. 1980;94(1):115–33.
2. Gehring WJ, Affolter M, Bürglin T. Homeodomain Proteins[J]. Annual Rev Biochem Annual Reviews. 1994;63(1):487–526.
3. Krumlauf R. *Hox* genes, clusters and collinearity[J]. Int J Dev Biol. 2018;62(11–12):659–63.
4. Holland PWH. Evolution of homeobox genes[J]. Wires Dev Biol. 2013;2(1):31–45.
5. Mark M, Rijli FM, Chambon P. Homeobox genes in embryogenesis and pathogenesis[J]. Pediatr Res. 1997;42(4):421–9.
6. Grenier JK, Garber TL, Warren R, et al. Evolution of the entire arthropod *hox* gene set predated the origin and radiation of the onychophoran/arthropod clade[J]. Curr Biol. 1997;7(8):547–53.
7. Izpisúa-Belmonte J-C, Duboule D. Homeobox genes and pattern formation in the vertebrate limb[J]. Dev Biol. 1992;152(1):26–36.
8. Vollmer J-Y, Clerc RG. Homeobox genes in the developing mouse Brain[J]. J Neurochem. 1998;71(1):1–19.
9. Tsumagari K, Baribault C, Terragni J et al. DNA methylation and differentiation: *Hox* genes in muscle cells[J]. Epigenetics & Chromatin, BioMed Central, 2013, 6(1): 25.
10. Abuhantash M, Collins EM, Thompson A. Role of the *HOXA* cluster in HSC emergence and blood cancer[J]. Biochem Soc Trans. 2021;49(4):1817–27.
11. Song JY, Pineault KM, Dones JM, et al. *Hox* genes maintain critical roles in the adult skeleton[J]. Volume 117. National Academy of Sciences; 2020. pp. 7296–304. 13.
12. Treisman J, Harris E, Wilson D, et al. The homeodomain: a new face for the helix-turn-helix?[J]. BioEssays. 1992;14(3):145–50.
13. Bürglin TR, Affolter M. Homeodomain proteins: an update[J]. Volume 125. Springer Berlin Heidelberg: Chromosoma; 2016. pp. 497–521. 3.
14. McGinnis W, Krumlauf R. Homeobox genes and axial patterning[J]. Cell. 1992;68(2):283–302.
15. Dressler GR, Gruss P. Anterior boundaries of *hox* gene expression in mesoderm-derived structures correlate with the linear gene order along the chromosome[J]. Differentiation. 1989;41(3):193–201.
16. Pascual-Anaya J, D'Aniello S, Kuratani S, et al. Evolution of *hox* gene clusters in deuterostomes[J]. BMC Dev Biol. 2013;13:26.
17. Akam M. *Hox* genes in arthropod development and evolution[J]. Biol Bull. 1998;195(3):373–4.
18. Lohman FP, Medema JK, Gibbs S, et al. Expression of the SPRR cornification genes is differentially affected by carcinogenic transformation[J]. Exp Cell Res. 1997;231(1):141–8.
19. Lewis RA, Wakimoto BT, Denell RE, et al. Genetic analysis of the Antennapedia gene complex (Ant-C) and adjacent chromosomal regions of *Drosophila melanogaster*. II. Polytene chromosome segments 84A-84B1,2[J]. Genetics. 1980;95(2):383–97.
20. Števčić Z, Systematic, Zoology. [Oxford University Press, Society of Systematic Biologists, Taylor & Francis, Ltd.], 1971, 20(3): 331–40.
21. Guinot D, Bouchard J-M. Evolution of the abdominal holding systems of brachyuran crabs (Crustacea, Decapoda, Brachyura)[J]. Zoosystema. 1998;20:613–94.
22. Song C, Cui Z, Hui M, et al. Comparative transcriptomic analysis provides insights into the molecular basis of brachyurization and adaptation to benthic lifestyle in *Eriocheir sinensis*[J]. Gene. 2015;558(1):88–98.
23. Li P, Zha J, Sun H, et al. Identification of differentially expressed genes during the brachyurization of the chinese Mitten crab Eriocheir japonica sinensis[J]. Biochem Genet. 2011;49(9):645–55.
24. Averof M, Akam M. *Hox* genes and the diversification of insect and crustacean body plans[J]. Nature. 1995;376(6539):420–3.
25. Li P. Cloning and analysis on morphology and sex regulation related *hox* gene family and fruitless gene in *Eriocheir sinensis*[J]. Qingdao: Institute of Oceanology, Chinese Academy of Sciences; 2017.
26. Sun X. The primary structure and expression study of *hox* genes in pacific white shrimp, Litopeneaus vannamei[D]. Qingdao: Institute of Oceanology,Chinese Academy of Science; 2015.
27. Colbourne J, Pfrender ME, Gilbert D, et al. The ecoresponsive genome of *Daphnia pulex*[J]. Volume 331. Science (New York, N.Y.),; 2011. pp. 555–61. 6017.
28. Kim H-S, Kim B-M, Lee B-Y, et al. Identification of *hox* genes and rearrangements within the single homeobox (*hox*) cluster (192.8 kb) of the cyclopoid copepod (*Paracyclopina nana*)[J]. J Experimental Zool Part B: Mol Dev Evol. 2016;326(2):105–9.
29. Jaramillo ML, Ammar D, Quispe RL, et al. Identification of hox genes and their expression profiles during embryonic development of the emerging model organism, Macrobrachium olfersii. J Experimental Zool Part B: Mol Dev Evol. 2022;2022(Jan 1):1–9.
30. Kim D-H, Jeong H, Kim M-S et al. Identification and characterization of homeobox gene clusters in harpacticoid and calanoid copepods[J]. J Experimental Zool Part B: Mol Dev Evol, 2021, 2021 Dec 2(n/a).
31. Li H. Hua. *Hox* genes and their evolutionary mechanisms[J]. Chin J Zool. 2011;46(1):136–42.
32. Mallo M. Reassessing the role of *hox* genes during vertebrate development and evolution[J]. Trends Genet. 2018;34(3):209–17.
33. Mallo M, Alonso CR. The regulation of *hox* gene expression during animal development[J]. Development. 2013;140(19):3951–63.
34. Elias-Neto M, Belles X. Tergal and pleural structures contribute to the formation of ectopic prothoracic wings in cockroaches[J]. Volume 3. Royal Society Open Science, The Royal Society; 2016. p. 160347. 8.
35. Khadjeh S, Turetzek N, Pechmann M, et al. Divergent role of the *hox* gene *Antennapedia* in spiders is responsible for the convergent evolution of abdominal limb repression[J]. Proc Natl Acad Sci United States Am Natl Acad Sci. 2012;109(13):4921–6.
36. Shiga Y, Yasumoto R, Yamagata H, et al. Evolving role of Antennapedia protein in arthropod limb patterning[J]. Development. 2002;129(15):3555–61.
37. Young T, Rowland JE, Van de Ven C, et al. *Cdx* and *hox* genes differentially regulate posterior axial growth in mammalian embryos[J]. Dev Cell. 2009;17(4):516–26.
38. Lewis EB. A gene complex controlling segmentation in *Drosophila melanogaster*[J]. Nature. 1978;276(5688):565–70.
39. Málaga-Trillo E, Meyer A. Genome duplications and accelerated evolution of *hox* genes and cluster architecture in teleost fishes1[J]. Am Zool. 2001;41(3):676–86.
40. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX[J]. Current Protocols in Bioinformatics, 2003, 00(1): 2.3.1–2.3.22.
41. Chen C, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. Mol Plant. 2020;13(8):1194–202.
42. R Core Team. R: A language and environment for statistical computing.[J]. R Foundation for Statistical Computing, Vienna, Austria., 2018.
43. Jin X-K, Li W-W, Cheng L, et al. Two novel short C-type lectin from chinese mitten crab, Eriocheir sinensis, are induced in response to LPS challenged.[J]. Fish Shellfish Immunol. 2012;33(5):1149–58.

Publisher's Note