



**UNIVERSIDADE FEDERAL DA PARAÍBA**  
**CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS**  
**ÁREA DE CONCENTRAÇÃO EM ZOOLOGIA**

**Dissertação de Mestrado**

**Evolução dos genes Hox anteriores e a diversidade dos arquétipos  
metazoários**

**Victor Montenegro Marcelino**  
**João Pessoa 2019**



**Evolução dos genes Hox anteriores e a diversidade dos arquétipos  
metazoários**

**Victor Montenegro Marcelino**

**Dissertação apresentada ao Programa de  
Pós-graduação em Ciências Biológicas,  
área de concentração Zoologia, da  
Universidade Federal da Paraíba, como  
parte dos requisitos para obtenção do Grau  
de Mestre em Ciências Biológicas (Zoologia).**

**Orientador: Dr.Sávio Torres de Farias**

**Coorientador: Dr.Gustavo Henrique  
Calazans Vieira**

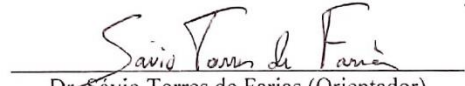
**Victor Montenegro Marcelino**

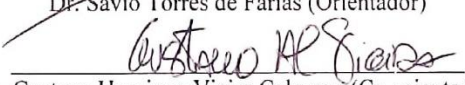
**João Pessoa 2019**

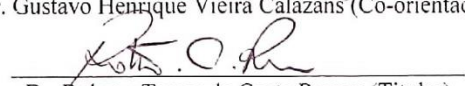
Ata da 317ª Apresentação e Banca de Defesa  
de Mestrado de Victor Montenegro  
Marcelino

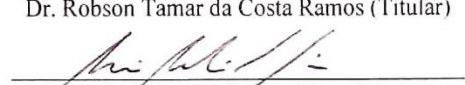
Ao(s) Sete dias do mês de junho de dois mil e dezenove, às 14:00 horas, no(a) Sala do PPGCB, da Universidade Federal da Paraíba, reuniram-se, em caráter de solenidade pública, membros da banca examinadora para avaliar a dissertação de mestrado de **Victor Montenegro Marcelino**, candidato(a) ao grau de Mestre em Ciências Biológicas. A banca foi composta pelos seguintes professores/pesquisadores: **Dr. Sávio Torres de Farias (Orientador)**, **Dr. Gustavo Henrique Vieira Calazans (Co-orientador)**, **Dr. Robson Tamar da Costa Ramos (Titular)** e **Dr. Márcio Bernardino da Silva (Titular)**. Compareceram à solenidade, além do(a) candidato(a) e membros da banca examinadora, alunos e professores do PPGCB. Dando início à sessão, a coordenação fez a abertura dos trabalhos, apresentando o(a) discente e os membros da banca. Foi passada a palavra para o(a) orientador(a), para que assumisse a posição de presidente da sessão. A partir de então, o(a) presidente, após declarar o objeto da solenidade, concedeu a palavra a **Victor Montenegro Marcelino**, para que dissertasse, oral e sucintamente, a respeito de seu trabalho intitulado "**Evolução dos genes Hox anteriores e a diversidade dos arquétipos metazoários**". Passando então a discorrer sobre o aludido tema, dentro do prazo legal, o(a) candidato(a) foi a seguir arguido(a) pelos examinadores na forma regimental. Em seguida, passou a Comissão, em caráter secreto, a proceder à avaliação e julgamento do trabalho, concluindo por atribuir-lhe o conceito APROVADO. Perante a aprovação, declarou-se o(a) candidato(a) legalmente habilitado(a) a receber o grau de **Mestre em Ciências Biológicas**, área de concentração **Zoologia**. Nada mais havendo a tratar eu, **Dr. Sávio Torres de Farias**, como presidente, lavrei a presente ata que, lida e aprovada, assino juntamente com os demais membros da banca examinadora.

João Pessoa, 07/06/2019.

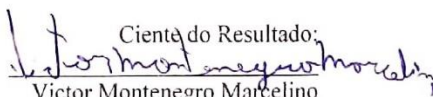
  
Dr. Sávio Torres de Farias (Orientador)

  
Dr. Gustavo Henrique Vieira Calazans (Co-orientador)

  
Dr. Robson Tamar da Costa Ramos (Titular)

  
Dr. Márcio Bernardino da Silva (Titular)

Ciente do Resultado:

  
Victor Montenegro Marcelino



**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

M314e Marcelino, Victor Montenegro.

Evolução dos genes Hox anteriores e a diversidade dos  
arquétipos metazoários / Victor Montenegro Marcelino. -  
João Pessoa, 2020.

51 f. : il.

Orientação: Sávio Torres de Farias.

Coorientação: Gustavo Henrique Calazans Vieira.

Dissertação (Mestrado) - UFPB/CCEN.

1. Bauplan; Família multigênicas; Modelo birth-and-dea.  
I. de Farias, Sávio Torres. II. Vieira, Gustavo  
Henrique Calazans. III. Título.

UFPB/BC

## **Agradecimentos**

Agradeço ao professor Sávio Torres de Farias por ter sugerido o tema desta dissertação e me orientado nesses dois anos de mestrado.

Ao professor Gustavo Henrique Calazans Vieira por ter-me co-orientado e dado apoio intelectual nesse trabalho.

À Pós-doutoranda Janaína Lima de Oliveira do Laboratório de Genética Evolutiva Paulo Leminski por ter auxiliado na produção do artigo e tirado “dúvidas moleculares”

Ao meu amigo Paulo Eduardo Toscano Soares por sua contribuição em scripts de programação para facilitar o manuseio dos dados.

A minha família – mãe, Ana Lucia Montenegro; pai, Francisco Tomaz Marcelino Neto; irmã, Inngryd Montenegro Marcelino e a minha Namorada Maria Vanessa Pontes da Costa Espinola por terem dado suporte emocional e me incentivado durante todo o período do mestrado.

Aos meus amigos Vitor Polanski, Victor Augusto e Matheus Estrela por terem apoiado e ajudarem a focar na escrita desta dissertação.

A minha psicóloga Anna Cecília por ter me ajudado psicologicamente com o mestrado.

## Resumo

Filogenias de Metazoa baseadas em genes envolvidos com a especificação de planos corporais fornecem uma poderosa alternativa às baseadas em características morfológicas e/ou marcadores moleculares clássicos. Eles fornecem pistas não apenas sobre a história evolutiva dos animais, mas também sobre as mudanças associadas em sua arquitetura molecular. Aqui, usamos sequências publicamente disponíveis de conjuntos de genes Hox envolvidos com padrões de região anterior (*Hox1-Hox4*), para reconstruir a história evolutiva de metazoários. Mostramos que a informação contida nesses genes reflete a evolução e a diversificação da maioria dos arquétipos de animais. Além disso, nossos achados revelam que a evolução dos genes Hox (e clusters) como uma família multigênica é consistente com um modelo de birth-and-death restrito pelo desenvolvimento, onde há um trade-off entre uma modificação de genes relativamente rápido e seus papéis centrais de desenvolvimento. Nossas descobertas fornecem tanto uma ampla filogenia de metazoários quanto um modelo evolutivo fortemente apoiado para explicar as mudanças dos genes Hox - uma família multigênica que é fundamental para o desenvolvimento animal e que provavelmente influenciou a diversificação de arquétipos de animais.

**Palavras Chaves** : Bauplan; Família multigênicas; Modelo birth-and-death; Archelosauria

## ABSTRACT

Metazoan phylogenies based on genes involved with specification of body plans provides a powerful alternative to those based on morphological traits and/or classical molecular markers. They give clues not only on animal evolutionary history, but also on associated changes in their molecular architecture. Here, we use publicly available sequences of *Hox* gene clusters involved with anterior region patterning (*Hox1-Hox4*), to reconstruct the evolutionary history of metazoans. We show that information harbored by these genes reflects the evolution and diversification of most animal archetypes. Moreover, our findings reveal that evolution of *Hox* genes (and clusters) as a multigene family is consistent to a birth-and-death model constrained by development, where there is a trade-off between a relatively fast gene turnover and their central developmental roles. Our findings provide both a strongly supported broad phylogeny of metazoans and an evolutionary model to explain the evolution of *Hox* genes – a multigene family that is fundamental to animal development and which is likely to have influenced the diversification of animal archetypes.

**Key words** :Body plan;Multigenefamily;Birth-and-deathmodel;Archelosauria

## **Lista de figuras**

|                                                                   |    |
|-------------------------------------------------------------------|----|
| Figura 1. Filogenia Hox1                                          | 23 |
| Figura 2. Filogenia Hox2                                          | 24 |
| Figura 3. Filogenia Hox3                                          | 25 |
| Figura 4. Filogenia Hox 4                                         | 28 |
| Figura 5. relações inferidas entre clusters de Hox de vertebrados | 30 |

## Sumário

|                                          |    |
|------------------------------------------|----|
| <b>Resumo</b>                            | 4  |
| ABSTRACT                                 | 5  |
| 1. Introdução                            | 9  |
| 1.4 Novidades gênicas durante a evolução | 14 |
| 2. Objetivos                             | 15 |
| 2.1 Objetivo geral                       | 15 |
| 2.2 Objetivos específicos                | 15 |
| 3. ARTIGO CIENTÍFICO                     | 16 |
| 4. Considerações finais                  | 34 |
| REFERÊNCIAS                              | 35 |

# PREFÁCIO

Esta dissertação está estruturada em formato de artigo científico, com uma introdução geral, objetivos em português e em seguida um draft (versão preliminar do artigo) que pretendemos publicar em inglês. O artigo está dividido em Abstract, introdução (main), resultados e discussão (results and discussion), conclusão (conclusion) e métodos (methodes). Por fim temos as considerações finais em português e as referências da dissertação e artigo e anexos referentes ao artigo.

# 1. Introdução

## 1.1 Diversidade

A classificação dos seres vivos teve início na Grécia antiga, com o considerado primeiro biólogo Aristóteles (384-322.a.c). Aristóteles classificou aproximadamente 400 animais, dividindo-os em dois grandes grupos: os de sangue vermelho, vertebrados e os que não têm sangue, os invertebrados. Aristóteles também foi o primeiro a utilizar homologia estrutural nos caracteres morfológicos se opondo à analogia funcional<sup>1</sup>. Muitas outras formas de classificar os seres vivos foram propostas até o surgimento da escola Lineana criada por Carl Linnaeus (1707-1778), que reunia as espécies em táxons buscando agrupar aqueles que tinham uma essência ontológica em comum, em que essas essências podem ser compartilhadas. Este método é utilizado até hoje, no qual os pesquisadores procuram classificar os seres vivos de acordo com as semelhanças compartilhadas e muitas vezes sem um entendimento claro dos aspectos ontológicos envolvidos<sup>2</sup>.

Um dos grupos de organismos mais conhecido são os chamados Metazoários ou popularmente conhecidos como animais, são um grupo taxonômico de organismos eucariotos multicelulares, e eles têm como caracteres sinapomórficos a presença de colágeno, ovogénese, espermatogénese (corpos polares), estrutura espermática especial e redução do genoma mitocondrial<sup>3</sup>.

Atualmente são reconhecidos 34 filos animais viventes, cada um com um bauplan único e características distintas que permitem cada filo ser diferente um do outro. Esses filos tiveram sua origem aproximadamente 600 milhões de anos durante a explosão do cambriano. Cada filo se diferencia do outro no grau de organização corporal, na simetria, no número de folhetos embrionários ou germinativos e no número de cavidades corpóreas<sup>4</sup>. Todos os animais vivos estão distribuídos em cinco clados: Porifera, Ctenophora, Cnidaria, Placozoa e Bilateria. Estes clados possuem monofilia fortemente sustentada. O clado Bilateria subdivide-se em Protostomia, Deuterostomia e Xenacoelomorpha. O clado Deuterostomia (clivagem indeterminada e blastóporo como neoformação) tem relações internas bem suportadas. Dentro de Deuterostomia

tem os seguintes clados: Ambulacraria (formado por Hemichordata e Echinodermata), Urochordata, Cephalochordata e Craniata.

Em Protostomia (clivagem determinada e blastopore dando origem à boca) têm três clados suportados: Chaetognatha, Ecdysozoa e Spiralia. Ecdysozoa (cutícula em muda, epicutícula trilaminada) subdivide-se em três clados: Panarthropoda, Nematoda e Scalidophora<sup>3</sup>. O monofiletismo de Scalidophora é questionável porque eles só têm uma característica que os une em um clado e é a presença de escalídeos inervados<sup>5</sup>. O clado Spiralia (clivagem em espiral) subdivide-se em: Rouphozoa, Gnathifera e Lophotrochozoa. Outros pequenos clados como orthonectida, Cyclophora e Rhombozoa não são claros quanto ao seu posicionamento em Spiralia. Lophotrochozoa é um clado formado por trochozoa (Annelida e Mollusca) e Lophophorata<sup>3</sup>.

## 1.2 Problemas Taxonômicos

Dentro de Craniata temos uma superclasse chamada tetrapoda cujas características são a presença de apêndices pares distintos e a porção proximal dos apêndices com um único osso e a porção distal com dois ossos. Dentro de Tetrapoda temos as seguintes classes: Lissamphibia e Amniota (Possuem ovo amniótico). Amniota formado por Mammalia, Aves, Crocodylia, Lepidosauria e Testudines<sup>6</sup>.

A posição de Testudines (tartarugas) dentro de Amniota é tema de discussão entre os biólogos, existindo três hipóteses de posicionamento dos Testudines. A primeira são os testudines sendo Anaspidas (sem fenestra temporal), colocando-os depois de Mammalia (Sinápsida=uma única fenestra temporal) e na base dos Diapsida (Duas fenestras temporais). A segunda hipótese agrupa Testudines como grupo irmão de Lepidosauria. As duas primeiras hipóteses são utilizando dados morfológicos e a terceira hipótese que coloca Testudines como grupo irmão de Archosauria (Aves + Crocodylia) é uma hipótese de dados moleculares<sup>7,8</sup>.

### 1.3 Dados moleculares

Com o advento da biologia celular e molecular, a variação nos níveis cromossômicos (número e estrutura) e moleculares (mais precisamente sequências de nucleotídeos e aminoácidos) forneceram ricas fontes de informação para estudos filogenéticos<sup>4</sup>. Em eucariotos, isso pode ser exemplificado pelo amplo uso de marcadores moleculares como, a subunidade I da Citocromo C oxidase (COI) para bioidentificar (identificação baseada em marcador molecular) espécies (DNA barcoding)<sup>9</sup> e a pequena subunidade ribossomal 18S (18S rRNA) para entender sua história evolutiva<sup>10</sup>.

Apesar dos esforços para integrar métodos tradicionais à taxonomia molecular e filogenética<sup>11</sup>, preocupações foram levantadas em relação a ambos os lados. Por um lado, os planos do corpo animal são muito úteis para identificar filos, mas os caracteres para unificá-los em grupos maiores são escassos e frequentemente de homologia incerta<sup>10</sup>. Por outro lado, filogenias baseadas em marcadores moleculares têm ajudado a entender as relações evolutivas amplas e específicas de linhagens de metazoários<sup>10,12,13</sup>, mas o grande foco em dados moleculares tem frequentemente chamado a atenção para características morfológicas, cuja evolução pretendemos entender em primeiro lugar<sup>14</sup>. Essas preocupações de integração de métodos tradicionais à taxonomia molecular foram, pelo menos, parcialmente melhoradas pelo uso de genomas completos em vez de poucos genes para reconstruir a história evolutiva (frequentemente denominada filogenômica), seguida pelo mapeamento de caracteres morfológicos para a árvore resultante<sup>14</sup>. Embora essa abordagem se aproxime de dois métodos concorrentes de abordagem à sistemática filogenética, ela ainda não fornece um vínculo mecanicista entre a evolução e a diversificação de formas corporais e as mudanças moleculares subjacentes.

Uma terceira via depende da investigação de traços potencialmente homólogos sob uma perspectiva de desenvolvimento. Dados embriológicos têm sido sugeridos há muito tempo como uma poderosa fonte para identificar relações entre táxons, desde o século XIX, quando Karl Ernst von Baer observou que características comuns em táxons superiores como Craniata, emergem mais frequentemente no início do desenvolvimento, do que caracteres específicos em táxons mais exclusivos como ordem ou famílias, chamado de lei de von Baer. Depois de Von Baer, o biólogo Ernst Haeckel reconduziu a lei de Baer e disse - "ontogenia recapitula a filogenia", mas no final do século XIX descobriu que estava errado, porque os traços (caracteres) têm taxas

evolutivas diferentes uma em relação a outra, em descendentes, e nos seus ancestrais, e cada estagio ontogenético possui adaptações específicas<sup>15</sup>.

No entanto, os traços de desenvolvimento foram deixados de lado em estudos filogenéticos, devido à limitada compreensão de seus mecanismos subjacentes<sup>10</sup>, até alguns anos atrás. Com os avanços na compreensão das bases genéticas do desenvolvimento, o campo evolutivo tem agora a oportunidade de conectar a diversidade observada de formas de vida animal existentes e extintas, a genes e vias de desenvolvimento envolvidos com inovação morfológica, e avaliar os passos da história evolutiva escrita em sequências de genes subjacentes. A integração dessas fontes complementares de informação resultou no surgimento do novo campo da biologia evolutiva do desenvolvimento, ou evo-devo, que pode ser estendido ainda mais aos estudos filogenéticos ("filo-evo-devo")<sup>16</sup>.

Os genes Hox são um exemplo clássico para entender as bases genéticas do desenvolvimento animal, e que podem iluminar a compreensão dos caracteres morfológicos. Os genes *Hox* são um conjunto de genes reguladores estruturais que codificam fatores de transcrição com um domínio ligante ao DNA chamado de domínio homeobox. Este tem cerca de 60 aminoácidos definindo o padrão e a identidade do segmento, determinando o plano corporal dos metazoários<sup>17-20</sup>, impressos por meio de um perfil complexo de expressão no eixo ântero-posterior de Bilateria - o 'código de Hox'<sup>21</sup>. E funcionam como um exemplo clássico de como a compreensão da base genética do desenvolvimento animal pode esclarecer aspectos da evolução morfológica.

Lewis (1978), ao estudar mutações homeóticas, aquelas que causam uma transformação de uma região do corpo em outra região do corpo, por exemplo, no caso do gene *Antennapedia* - um gene *Hox* 'posterior' - encontrado em *Drosophila melanogaster*, foi comprovado que o *Antennapedia* controla o desenvolvimento de pernas - mutações na região reguladora deste gene resulta no desenvolvimento de ectópico de pernas no lugar de antenas (mutações de ganho de função)<sup>22,23</sup>, ou antenas ectópicas em local das pernas (mutações de perda de função)<sup>24</sup>. Os genes Hox ficam situados nos aglomerados de cromossomos, e seus *clusters* parálogos estão em geral organizados de maneira colinear em relação aos seus distintos domínios de expressão<sup>18</sup>.

O grupo mais basal a apresentar genes *Hox* são os cnidários. Os *Hox* de cnidários são expressos no polo oral aboral diferente dos de Bilateria<sup>24</sup> e assim, o ancestral de todos os genes *Hox* deve ter aparecido no ancestral dos Eumetazoa. Ao longo da morfogênese dos metazoários, os genes *Hox* apresentam-se expressos ao longo do eixo ântero-posterior nos triploblásticos e polo oral e aboral em diblásticos do corpo do embrião metazoário<sup>25</sup>. O ancestral hipotético do

*Hox*, é chamado *protoHox*, e deve ser composto por um agrupamento de dois genes e poderia ter surgido em um ancestral antes da divisão entre cnidários e bilatérios<sup>26</sup>. Outros eventos de duplicação gênica originaram um grupo *Hox* de artrópodes ancestrais, consistindo de 10 genes<sup>27,28</sup> seguidos de dois eventos: tetraploidização- um nó ancestral em Craniata e um segundo em Teleostei<sup>27-30</sup>.

A proteína Ultrabitorax (*Ubx*), codificada por genes *Hox*, determina a segmentação em animais<sup>31</sup>. Uma mutação nessa proteína, que confere uma cauda poli-alanina, aconteceu em Hexapoda, conferindo a eles a segmentação do tórax em protórax, mesotórax e metatórax. Essa mesma proteína no táxon irmão Crustacea não codifica o mesmo padrão de segmentação e ela não apresenta a cauda polia-alanina<sup>31</sup>.

Nos vertebrados existe uma diferenciação na quantidade de vértebras em uma das regiões cervical, torácica, lombar, sacral e caudal. Na maioria dos mamíferos existem 7 vértebras cervicais, enquanto em aves existem entre 13 e 15. Tal variação está ligada à mudança na expressão espacial do gene *Hoxc6*<sup>32</sup>.

As mutações que ocorrem nos genes *Hox* contribuem para a evolução do plano corporal dos animais, visto que, há evidências de que os padrões corporais (*Bauplans*) dos metazoários foram alterados por mutações que ocorrem em muitos níveis de hierarquias reguladoras dos genes *Hox*. Além disso, existem evidências de que mutações em sequências de proteínas transcritas pelos genes *Hox* contribuíram para a diversidade morfológica ao longo da evolução<sup>18</sup>.

Os *Hox* da região anterior de vertebrados estão envolvidos na segmentação do Rombencéfalo, que é uma região evolutivamente conservada do encéfalo de vertebrados e tem um papel importante no controle sensorial e motor da cabeça e controle de atividades automáticas como respiração<sup>33</sup>. O Rombencéfalo é caracterizado por sua subdivisão em segmentos repetitivos chamados rombômeros<sup>34</sup>. Já em protostomados, os genes *Hox*, da região anterior definem a posição das estruturas da região cefálica tais como, as maxilas, probóscide, glândulas salivares e palpos maxilares<sup>35</sup>.

## 1.4 Novidades gênicas durante a evolução

Novos genes surgem por dois mecanismos: duplicação e divergência de genes já existentes e por evolução *de novo*, que pressupõe que os genes podem surgir diretamente do DNA não codificante<sup>36,37</sup>. Na duplicação de genes, um gene surge de um evento que ocorre durante a meiose, entre cromossomos homólogos desalinhados, fazendo com que a cópia do gene fique pareada em outro cromossomo homólogo durante um erro no *crossing over*, podendo levar a diversos rearranjos cromossômicos, causando um mau pareamento que pode acumular mutações nesse gene duplicado e manter uma função semelhante a original, ou sofrer processo de divergência e passar a ter uma função completamente nova. A duplicação pode acontecer também quando um elemento de transposição insere uma cópia de uma parte do gene em outro local, surgindo um novo segmento de gene duplicado, que estará propenso a sofrer mutações e processo de divergência. Genes podem surgir também por “exon shuffling” onde no exon ou “domain shuffling”, podem ser recombinados para formar muitos genes codificadores de proteínas quiméricas<sup>38</sup>.

A evolução das famílias gênicas ocorrem através de três modelos, sendo eles: o modelo *birth-and-death evolution* no qual novos genes surgem por várias duplicações consecutivas e alguns desses genes podem permanecer no genoma por longo tempo enquanto que outros genes são retirados ou inativados por mutações deletérias, o modelo de evolução em concerto onde os genes da família multigênica (que é um grupo de genes que descendem de um gene ancestral comum tendo funções semelhantes e sequência de DNA semelhante) evoluem simultaneamente de maneira que uma mutação se espalha pela família através de permutação; ou pode evoluir pelo modelo de conversão gênica e o modelo de evolução divergente que é a diferenciação gradual de genes filogeneticamente relacionados à medida que os genes se duplicam e adquirem novas funções<sup>39</sup>. Tem sido hipotetizado, que a evolução dos clusters Hox é consistente com um modelo de *birth-and-death evolution* de famílias multigênicas. Traçar esses eventos de duplicação e perda de genes, juntamente com a análise da informação abrigada pelos

genes Hox, provou ser útil para entender vários aspectos da filogenia animal, tais como o apoio de linhagens de Ecdysozoa e Lophotrochozoa em protostômios<sup>40</sup>, relações entre os quatro subfilos existentes de artrópodes. (Chelicerata, Crustacea, Hexapoda, e Myriapoda)<sup>41</sup> e grupos de vertebrados<sup>38</sup>.

## 2. Objetivos

### 2.1 Objetivo geral

Elucidar a história evolutiva dos genes hox anteriores buscando relacionar com a evolução dos metazoários, propondo um modelo para explicar a evolução de Hox como uma família multigênica.

### 2.2 Objetivos específicos

- *Reconstruir filogenias específicas a partir dos genes Hox e/ou suas sequências ancestrais.*
- *Propor um modelo de evolução da família multigênica dos genes Hox;*

### 3. ARTIGO CIENTÍFICO

#### **Anterior *Hox* genes evolution and the diversity of metazoan archetypes**

Victor Montenegro Marcelino<sup>1\*</sup>, Janaina Lima de Oliveira<sup>1,2</sup>, Paulo Eduardo Toscano Soares<sup>1,3</sup>, Gustavo Henrique Calazans Vieira<sup>2</sup>, Sávio Torres de Farias<sup>1\*</sup>

**1.** Laboratório de Genética Evolutiva Paulo Leminsk, Departamento de Biologia Molecular, Universidade Federal da Paraíba, João Pessoa, Brazil

**2.** Departamento de Sistemática e Ecologia, Universidade Federal da Paraíba, João Pessoa, Brazil.

**3.** Laboratório de Biologia Molecular Aplicada, Universidade Federal do Rio Grande do Norte, Natal, Brazil

\*Correspondence to: victor-montenegro@hotmail.com & stfarias@yahoo.com.br

**Keywords:** Bodyplan; Multigene family; Birth-and-death model; Archelosauria

## ABSTRACT

Metazoan phylogenies based on genes involved with specification of body plans provides a powerful alternative to those based on morphological traits and/or classical molecular markers. They give clues not only on animal evolutionary history, but also on associated changes in their molecular architecture. Here, we use publicly available sequences of *Hox* gene clusters involved with anterior region patterning (*Hox1-Hox4*), to reconstruct the evolutionary history of metazoans. We show that information harbored by these genes reflects the evolution and diversification of most animal archetypes. Moreover, our findings reveal that evolution of *Hox* genes (and clusters) as a multigene family is consistent to a birth-and-death model constrained by development, where there is a trade-off between a relatively fast gene turnover and their central developmental roles. Our findings provide both a strongly supported broad phylogeny of metazoans and an evolutionary model to explain the evolution of *Hox* genes – a multigene family that is fundamental to animal development and which is likely to have influenced the diversification of animal archetypes.

## MAIN

From ancient systems proposed by Aristotle to modern Linnaean taxonomy, classification of the great diversity of animal life forms into groups of related species has been grounded on morphological traits and their putative homology<sup>1,2</sup>. With the advent of cellular and molecular biology, variation harbored at the chromosomal (number and structure) and molecular (most obviously nucleotide and amino acid sequences) levels have provided rich sources of information to phylogenetic studies<sup>4</sup>. In eukaryotes, this can be exemplified by the broad use of molecular markers such as the Cytochrome C oxidase subunit I (COI) to bioidentify species (DNA barcoding)<sup>9</sup> and the small ribosomal subunit 18S (18S rRNA) to understand their evolutionary history<sup>10</sup>.

Despite efforts to integrate traditional methods to molecular taxonomy and phylogenetics<sup>11</sup>, concerns have been raised regarding both sides. On the one hand, animal body plans are very useful to identify phyla, but characters to unify them into larger (inclusive) groups are both scarce and frequently of uncertain homology<sup>10</sup>. On the other, phylogenies based on molecular markers have helped to understand broad and specific evolutionary relationships of metazoan lineages<sup>10,12,13</sup>, but the large focus on molecular data have often shadowed the attention to morphological traits, whose evolution we aim to understand in the first place<sup>14</sup>. These concerns have been at least partially ameliorated by usage of whole genomes instead of a handful of genes to reconstruct evolutionary history (often named phylogenomics), followed by mapping of morphological characters to the resulting tree<sup>14</sup>. Although this approach appraises two competing methods of addressing phylogenetics, it still does not provide a mechanistic link between the evolution and diversification of body forms and the underlying molecular changes.

A third avenue relies on investigating potentially homologous traits under a developmental perspective. Embryological data has long been suggested as a powerful source to identify relationships among taxa since nineteenth century when Karl Ernst von Baer realized that

features common in higher taxa like Craniata emerge in development more frequently than specific traits in low taxa like order or families. And this was called the von Baer's law. After Von Baer, the biologist Ernst Haeckel reappointed the Baer's law, stating - "ontogeny recapitulates phylogeny". Nevertheless, since the final of nineteenth century that paradigm has been challenged because biologists realized traits that have different rates from each other in descendants than in their ancestors and specific adaptation for different stages of embryo<sup>15</sup>. However, developmental traits were relatively put aside in phylogenetic studies due to the limited understanding of its underlying mechanisms<sup>10</sup> – until a few years ago. With the astonishing advances on comprehending the genetic bases of development, evolutionary biology has now the opportunity to connect (I) the observed diversity of extant and extinct animal life forms to (II) developmental genes and pathways involved with morphological innovation and (III) assess the steps of evolutionary history written in underlying gene sequences. Integration of these complementary sources of information has resulted in the emergence of the new field of evolutionary developmental biology, or evo-devo, which can be further extended to phylogenetic studies ('phylo-evo-devo')<sup>16</sup>.

The *Hox* genes are a stereotypical example of how the understanding of genetic basis of animal development can illuminate aspects of morphological evolution. These genes play fundamental roles on defining metazoan segmental patterning and identity<sup>6,20</sup>, imprinted through a complex profile of expression across the anteroposterior axis in Bilateria – the '*Hox* code'<sup>21</sup>. Changes on the correct local expression of these genes result in homeotic transformations of the body regions involved (*i.e.* misplaced development of morphological structures)<sup>21,22</sup>. For instance, in *Drosophila melanogaster*, mutations in the regulatory region of *Antennapedia* – a 'posterior' *Hox* gene – result in the development of ectopic of legs in place of antennae (gain-of-function mutations)<sup>24,26</sup>, or ectopic antennae in place of legs (loss-of-function mutations)<sup>24</sup>.

Because *Hox* genes are already present in Cnidaria<sup>21</sup>, the minimal *ProtoHox* cluster must have emerged before the Cnidaria-Bilateria split, composed by only two anterior genes<sup>35</sup>. Further

events of gene duplication have originated an ancestral arthropod *Hox* cluster consisting of 10 genes<sup>27,28</sup>, followed by two events of tetraploidization: one in the Craniata and a second one in teleosts<sup>27–29</sup>. In fact, it has been hypothesized that evolution of *Hox* clusters is consistent to a birth-and-death model of evolution of multigene families, where a few duplicated gene copies remain in the genome for a long time, whereas others are inactivated, deleted or go through a process of pseudogenisation<sup>39</sup>. Tracing these events of gene duplication and loss, together with analysis of the information borne by *Hox* genes has been proved useful to understand several aspects of animal phylogeny, such as the support of Ecdisozoan and Lophotrocozan lineages within Protostomes<sup>40</sup>, relationships among the four extant arthropod subphyla (Chelicerates, Crustaceans, Hexapods, and Myriapods)<sup>41</sup> and Vertebrate groups<sup>42</sup>.

Here, we elucidate the evolutionary history of metazoans from information held by anterior *Hox* genes, highlighting their contribution to evolution of vertebrate bauplan, and propose a model to explain *Hox* evolution as a multigene family.

## RESULTS & DISCUSSION

We used publicly available protein sequences from the four anterior *Hox* genes (*Hox*-1 to *Hox*-4) to reconstruct the evolutionary history of *Hox*gene and their implications in metazoan phylogeny – each gene analyzed separately but including all copies of the same gene in a single alignment (see Methods). The four resulting trees show clear shared features. First, they have high support values overall, with slight (non-significant) differences among them: *Hox*-1 with the best distribution of support values, followed by *Hox*-4, *Hox*-2, and *Hox*-3, respectively (Figure S1). Second, support values are higher for basal nodes, supporting previous findings that *Hox* gene sequences may be a powerful source of information to resolve deep evolutionary relationships, - and this is why recent branches were collapsed to major metazoan groups in our phylogenies (Figures 1-4). Finally, *Hox* genes from Protostomes and non-Chordata Deuterostomes form

separate groups from Vertebrate clusters. Our phylogenies are consistent with the Meyer and Zardoya<sup>6</sup> hypothesis of evolution of vertebrates and the position of Testudines is consistent with the phylogenetic hypothesis of Chiari et al.<sup>43</sup> in *Hox-B1*, *Hox-A2*, *Hox-D3*, *Hox-A4*, *Hox-B4*, and *Hox-D4* phylogenies and provides additional support for the hypothesis of Testudines as a sister group of Archosauria.

However, the four resulting trees also have particularities, so they are discussed separately below.

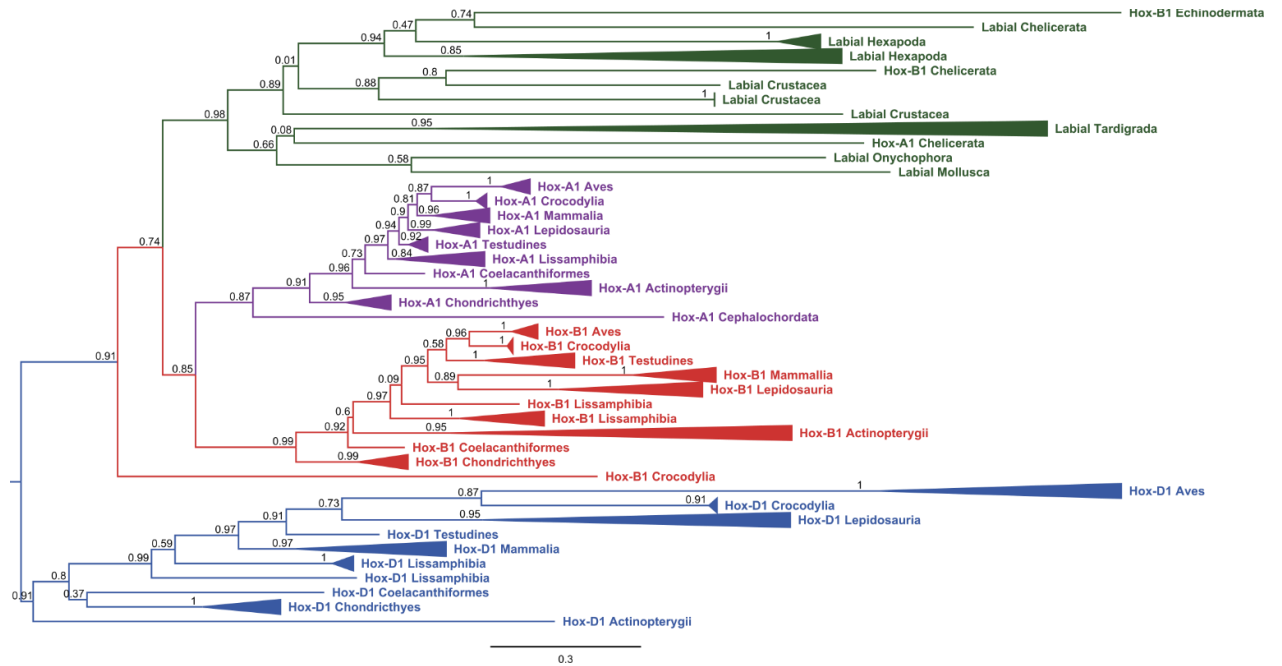
### ***Hox1***

Support values from the *Hox-1* phylogenetic tree range from 0.7 to 0.9, but this is not uniform across all of the topology: derived nodes, in general, show weaker supports (Figure 1). *Hox-D1* is expressed in posterior neuroectoderm<sup>44</sup> and is phylogenetically located at the base of the tree, showing evolutionary relationships within the superphylum Gnathostomata, characterized by the presence of jaws. This group is split in two main branches: a more basal one formed by bony (Actinopterygii) and other group formed by cartilaginous (Chondrichthyes) and lob-finned fishes (Coelacanthiformes), and Tetrapoda (Lissamphibia, Amniota with Synapsida and Diapsida (Testudines is a basal group)).

*Hox-A1* and *B1* are sister groups located more exclusively in the tree. *Hox-A1* is formed by Chordata, a group characterized by the presence of notochord, dorsal tubular nerve cord, pharyngeal slit, endostyle and post-anal tail<sup>45</sup>, and is represented in our data by two lineages, Cephalochordata and Vertebrata (Figure 1). Cephalochordata is placed at the base, followed by a group composed of Chondrichthyes, Actinopterygii, Coelacanthiformes, and Amniota. *Hox-B1* produced a similar topology, but Mammalia formed a group with lepidosauria different from *Hox-A1* that Mammalia is a sister group of Crocodylia and Aves. Also, one crocodilian *Hox-B1* copy is placed outside of the *Hox-B1* branch, which is likely due to a poor protein prediction and/or

annotation. In both *Hox-A1* and *B1*, Mammalia is placed within Diapsida, despite being the sister taxon of the latter – within (or representing) the Synapsid. Interestingly, these two genes act in synergy to form the branchial arch artery, which are transitional embryo blood vessels<sup>46</sup>, and this synergy may explain their close relationship between *Hox-A1* and *Hox-B1* in *Hox-1* phylogeny. This hypothesis is further supported by observation that *Hox-1* copies have functional redundancy in modelling the hindbrain in *Xenopus laevis* embryos<sup>47</sup> because when *Hox-A1* is knockout the fifth rhombomer turn absent or reduced, and the fourth rhombomer become reduced. Additionally, when the *Hox-B1* is knockout the identity of the fourth rhombomer are altered and segmentation is not affected. However, when both *Hox-A1* and *Hox-B1* are deleted and the phenotype is fifth rhombomere and fourth rhombomer being mis-specified probably deleted and the loss of second pharyngeal arch<sup>46</sup>.

Finally, our *Hox-1* phylogeny shows a split between derived Deuterostomes and a group formed by Protostomes and non-Chordata Deuterostomes. The latter is represented by *labial* (*Hox-1* denomination for *Drosophila melanogaster*) in Onychophora, Tardigrada, Insect, Crustacea, Mollusca, Chelicerata, and Echinodermata, which is involved with development of the salivary gland and cephalo-pharyngeal skeleton<sup>35</sup>. Echinodermata is placed within the Protostome branch, as a sister group of Chelicerata. This might be a bias resulting from the fact that basal Deuterostomes are poorly sampled in the data we had used herein<sup>6</sup>. In general, phylogenetic relationships inferred from *labial* in Protostomes do not form monophyletic groups. This can be potentially explained by stronger selective pressures in *labial*, since it is present in single copy (resulting in higher sequence similarity between groups), in contrast to the presence of up to three in *Hox-1* copies in Deuterostomes (diluting selective pressures in each copy).

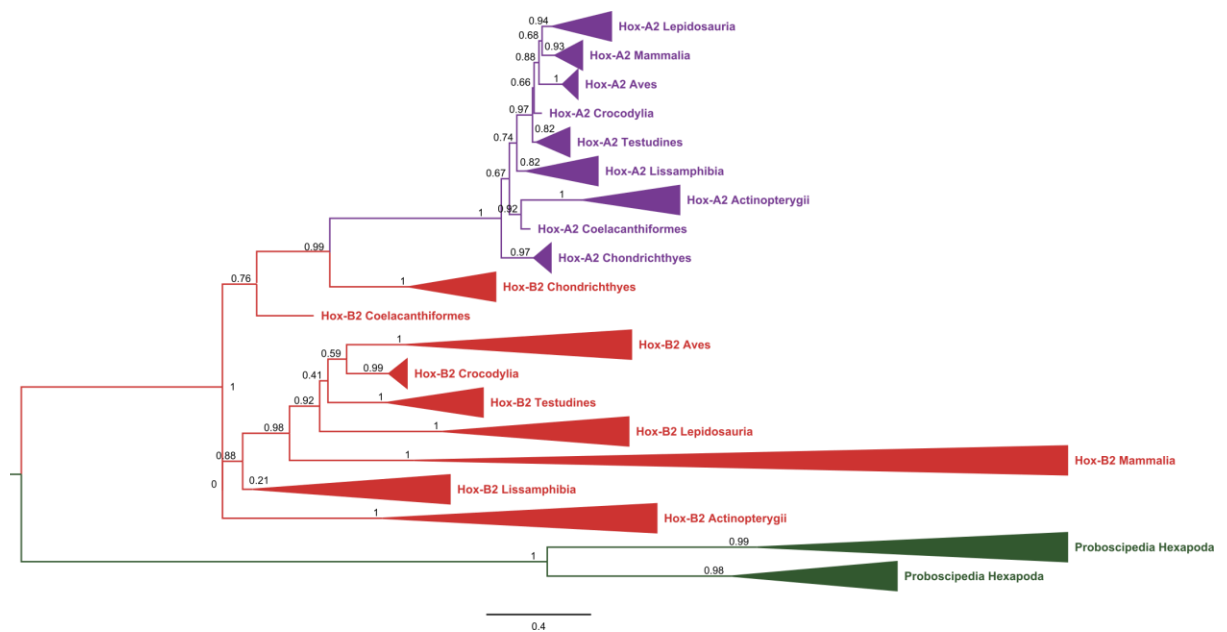


**Figure 1.** Hox1 derived phylogeny. The green branches show the Labial gene, while purple ones represent the Hox-A1, the red as Hox-B1, and in blue Hox-D1. The support values are represented by Shimodaira-Hasegawa-like.

## Hox2

*Hox-A2* and *Hox-B2* control the basal plate and alar plate of the Hindbrain rhombomers<sup>47</sup>. The overlapping function between the *Hox-A2* and *Hox-B2* genes may then predict *Hox-A2* and *Hox-B2* as closer groups. The probocipedia gene is the equivalent of *Hox2* in insects and participates in the development of labial palps and maxillary and buccal parts<sup>35</sup>. In the phylogeny *Hox2* (Figure 3) there was a clear separation between protostomy and deuterostomy, the Deuterostomates being the Chordata represented by the *Hox-A2* and *Hox-B2* proteins and the Protostomia by Hexapoda. In *Hox-A2* sequences, vertebrates are organized according to the most

accepted hypotheses: cartilaginous fish appearing first<sup>6</sup> and in this group is a sister group of Teleostomi. In the base of Teleostomi we have a group formed by Coelacanthiformes and Actinopterygii. The position of the mammalian sequences as a sister group of Lepidosauria, may be a convergence of the sequences of these two groups or an incongruity since this group has a low support (0.68)<sup>43</sup>. The *Hox-B2* presented an incongruity with respect to the positioning of the ray-finned fish sequences and affinities to *Hox-A2* may be because of the lack of support of this group (0).



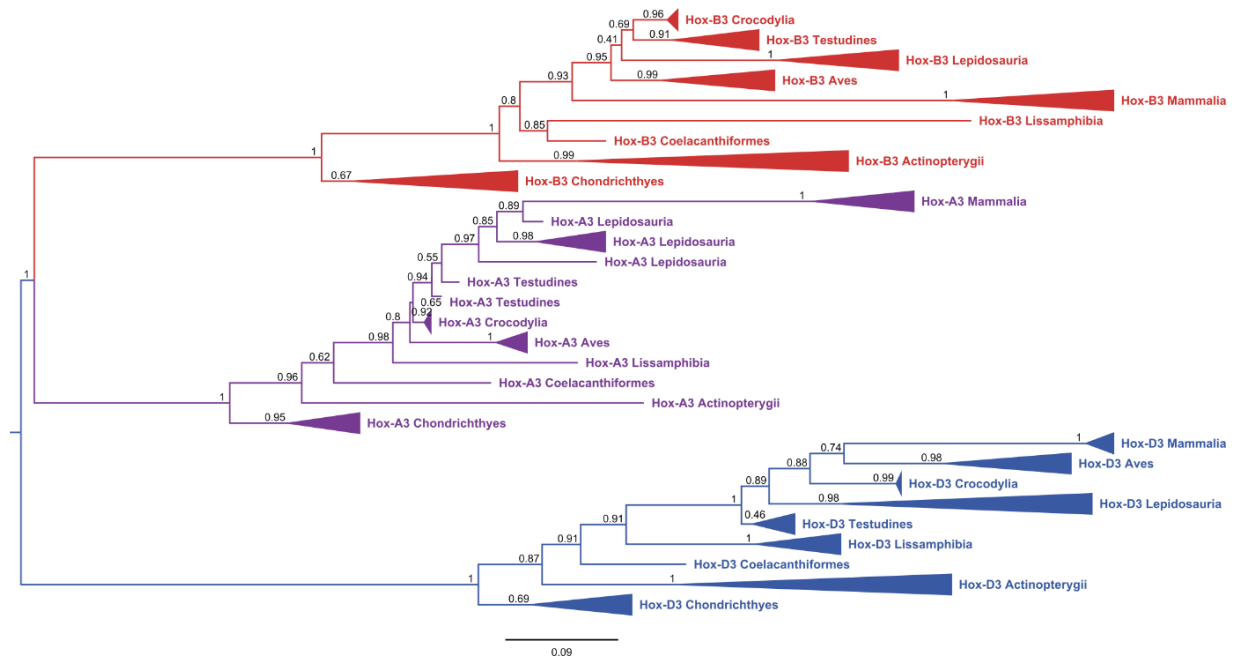
**Figura2.** Hox2 phylogeny in green we have the Proboscipedia, in purple have the Hox-A2 and in red Hox-B2. The support values are represented by Shimodaira-Hasegawa-like.

## Hox3

In the phylogeny of *Hox3* (Figure 3) most of the basal nodes showed a value greater than 0.7. *Hox-A3* and *Hox-B3* appeared as the sibling gene groups. The group of the *Hox-D3* was represented by the Gnathostomata and its main synapomorphy is the appearance of the mandible which originates from the skeleton of the anterior pharyngeal arches<sup>6,43</sup>. The *Hox-D3* and *Hox-*

A3 proteins are associated with differentiation of the first cervical vertebra Atlas<sup>48</sup>. *Hox-A3* affects the intrinsic ability of the neural crest cell population to differentiate and/ or induce adequate differentiation of the surrounding pharyngeal arch and tissue sacs. *Hox-A3* mutants lose the thymus<sup>49</sup>. In *Hox-A3* the relations within Amniota shows Aves, instead of Mammalia, as the first diversification within Amniota. *HoxB-3* is associated with migration of motor neurons in the hindbrain<sup>50</sup>. In *Hox-B3* phylogeny, Testudines are located as sister group of Crocodylia and Archosauria and Lepidosauria are polyphyletic.

No invertebrate Hox 3 (*zen*, *zen2* and *bcd*) has homeotic roles and their duplications are not well resolved<sup>51</sup>.



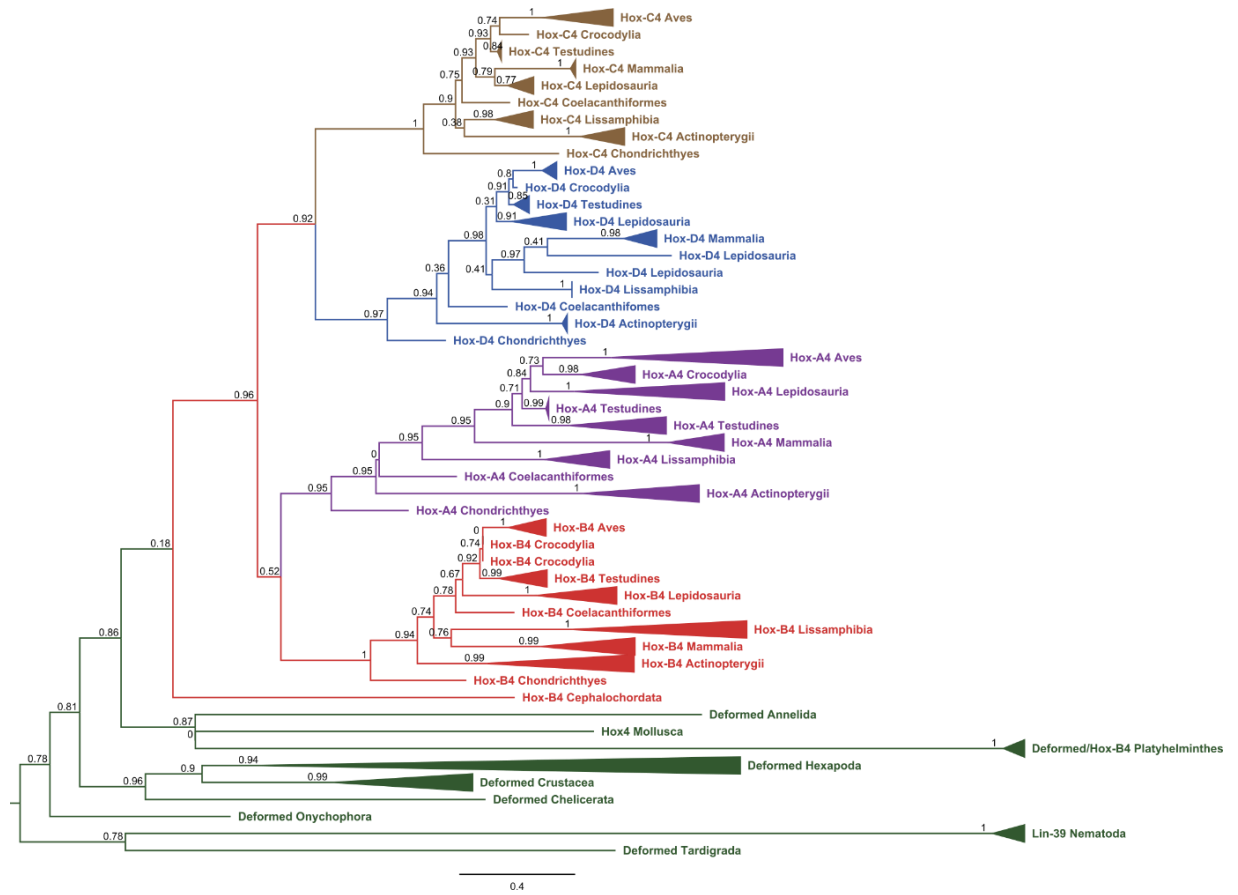
**Figura 3.** Hox3 phylogeny , in purple have the Hox-A3 and in red Hox-B3 and in blue Hox-D3. The support values are represented by Shimodaira-Hasegawa-like.

## Hox4

In the phylogeny of *Hox4* (Figure 4) most of the basal nodes exhibited support values greater than 0.7 (with the exception of two clades) and the more derived nodes had a fluctuation between low and high support values. The vertebrate-containing groups were divided into two *Hox4* sister groups, with *Hox-B4* and *Hox-A4* showing evolutionary affinities and *Hox-D4* and *Hox-C4* comprising another group. *Lin-39* controls cell fate in the central region of the body of *Caenorhabditis elegans*, prevents its cell fusion of the vulva, and is also associated with the fate of Nematoda epidermal cells<sup>52–54</sup>. The group formed by *Lin-39* was distinct from the other sequences of *Hox* of protostomates because they are the most basal sequences of *Hox* with respect to the *Deformed*, except for the sequence of Tardigrada, being a phylogenetic approximation with the clade of the Nematoda, since the relation within Ecdysozoa has not yet been fully elucidated. The *Deformed* protein maintains the boundary between the maxilla and the mandible, activating apoptosis located in the lobe of the head<sup>55</sup>. *Deformed* succeeded in dividing the sequences belonging to Ecdysozoa groups (a group in which all taxa undergo ecdysis) and Platytrchozoa (Platyhelminthes and Lophotrochozoa) and our results corroborate with the phylogenetic hypothesis of De Rosa *et al.*<sup>40</sup>, where posterior *Hox* were used for this phylogenetic reconstruction. Also, *Hox-D4* defines the boundaries between rhombencephalon and somites<sup>56</sup>. Within the branch formed by *Hox-D4*, Testudines showed evolutionary affinities with Archosauria as the sister group Aves + Crocodylia. *Hox-C4* protein is associated with the activation and proliferation of lymphocytes<sup>57</sup> and the differentiation of keratinocytes<sup>58</sup>. The *Hox-C4* gene recovered a topology in which Coelacanthiformes is the sister taxon of Amniota and such an incongruous result may be due to bad annotation of this sequence and/or by the low value of support and Lissamphibia sequences closer to Actinopterygii but also not well supported. Furthermore, *Hox-C4* also recovered affinities of Testudines with Archosauria. The group of *Hox-B4* the sequence of Coelacanthiformes, which is related to the development of the epidermis<sup>59</sup> and/or formation of cervical vertebrae<sup>60</sup> as a Diapsida sibling group. As *Hox-C4* and *Hox-D4*, *Hox-B4* also placed Testudines as a sister taxon of Archosauria. The *Hox-B4*, *Hox-C4*, and *Hox-*

*D4* genes have similar expressions within the Archosauria clade, where it determines the thoracic cervical limit<sup>61</sup> and this may be the cause of the grouping patterns recovered in this study. The *Hox* sequence of Cephalochordata became the sister of all *Hox* deuterostome, possibly because it was a more basal deuterostome to the others and because there was no duplication of clusters in that group<sup>30</sup>.

The *Hox* phylogenies that are present in Gnathostomata do not always equate the phylogenetic hypothesis of the Craniata proposed by Meyer and Zardoya<sup>6</sup>, and this is because our phylogeny reconstructs the relation of *Hox* genes present in Gnathostomata and differences can be explained with future experimental tests like in the *Hox-D3* phylogeny that sequences of Aves and Mammalia are sibling and the function of *Hox-D3* associated with differentiation of the first cervical vertebra Atlas may explain the possible convergence of the sequences between Mammalia and Aves because the atlas vertebra within Amniota provides vertical movement of the head that are made by the articulation of the skull with the Atlas. The invertebrate *Hox* proteins seem to reflect the phylogeny of the Ecdysozoa<sup>62</sup> and Platytrichochozoa<sup>40</sup>. Our work corroborates the hypothesis that developmental genes are good genes for phylogenies of larger (or more inclusive) groups, such as phyla and most classes<sup>63</sup>.



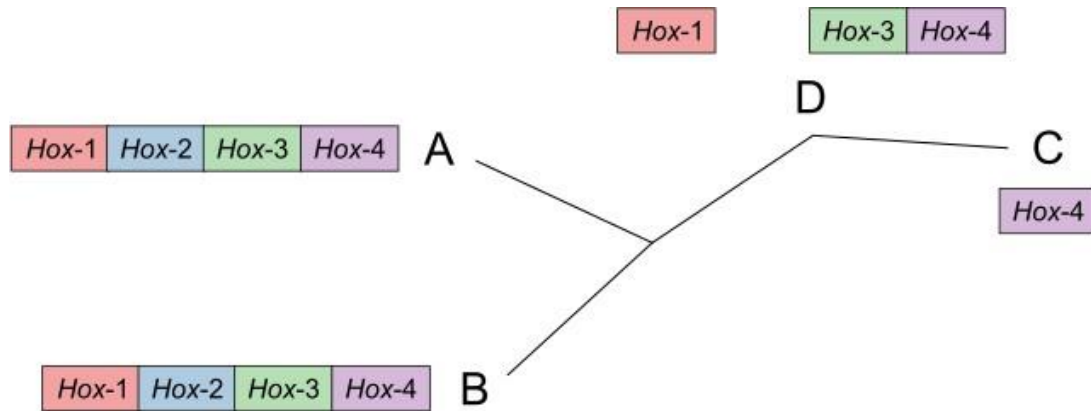
**Figura4.** Hox4 phylogeny in green we have the Deformed , in purple have the Hox-A4 , in red Hox-B2 , in blue Hox-D1 and in brown Hox-C4. The support values are represented by Shimodaira-Hasegawa-like.

### Anterior *Hox* genes evolution as part of a multigene family

*Hox* genes are part of a multigene family, and as such, their inclusion in phylogenetic studies should also regard potential relationships between gene copies within the same genome. By analyzing paralogs for each gene aligned altogether (*A*, *B*, *C* and *D*; see Methods), it is possible to infer the relative time of gene duplication and the evolutionary dynamics shaping the multigene family. For example, observation that gene clusters are shared by many clades suggests that the duplication event took place in ancient times, before diversification of more recent groups. Moreover, the resulting tree topologies and relationship among gene copies can provide information to distinguish between three alternative scenarios for evolution of multigene

families<sup>39</sup>: i) divergent evolution, where genes are phylogenetically related and have diverged gradually as the duplicated genes have gone through neofunctionalization; ii) concerted evolution, where all members of a gene family are assumed to evolve in a concerted manner (not independently), such that sequences from all copies are roughly homogeneous (*i.e.* a mutation occurring in a gene copy spreads to all gene copies in the family) by unequal crossover or gene conversion; iii) birth-and-death evolution, where new genes are created by gene duplication, and some duplicated genes are maintained in the genome for a long time, whereas others are lost either by selection (for a compact genome or to regulate gene dosage) or pseudogenisation (accumulation of deleterious mutations under relaxed selection)<sup>39</sup>.

Orthologue search of *Hox* genes show that most vertebrate groups carry more than one cluster, whereas protostomes and lower deuterostomes (Echinodermata and Cephalochordata) carry only one (Table S1). Namely, vertebrates harbor three copies of *Hox*-1 (A, B and D; Figure 1), two copies of *Hox*-2 (A and B; Figure 2), three copies of *Hox*-3 (A, B and D; Figure 3), and four copies of *Hox*-4 (A, B, C and D; Figure 4). Because all vertebrate species share the same gene content in each cluster (all species carry all four anterior genes in clusters A and B, only *Hox*-4 in cluster C and all genes but *Hox*-2 in cluster D), the establishment of gene content in the four *Hox* clusters seem to have emerged after Urochordata and before cyclostomes (hagfish and lampreys)<sup>64</sup>. For all four *Hox*-based phylogenies, each gene copy is closer to orthologues in other species than to paralogs within the same genome (Figures 1-4). For instance, Avian *Hox*-A1 protein sequences share more similarities to Crocodilian *Hox*-A1 than to Avian *Hox*-B1 or D1 (Figure 1). Moreover, in cases where more than two clusters are present in vertebrate species, sequences from clusters A and B are more closely related to each other than sequences from clusters C and/or D (Figures 1,3 and 4). These new and interesting findings suggest a tree topology in which clusters A and B are more closely related to each other than to C and D, with C more distantly related to all others (Figure 5).



**Figura5.** Inferred relationships between vertebrate Hox clusters. Unrooted tree topology was inferred from gene content in each of the four vertebrate clusters, but also corroborated by phylogenetic analysis of protein sequences (see below Figures 2-5).

Among the three models for multigene family evolution, our results are more consistent to the birth-and-death model, characterized by a fast gene turnover across groups, and consistent with previous work of Holland (2012)<sup>51</sup>. However, the pattern of gene gain/loss in *Hox* gene clusters appears to be restricted by their developmental roles and relationships with other genes associated with the establishment of each bauplan<sup>65</sup>. Thus, because evolutionary rates of *Hox* genes seem to be dependent and constrained by evolutionary rates of other developmental genes, we name this process a 'birth-and-death model constrained by development'. This model provides fundamental support to the hourglass model, which predicts a basic embryonic bauplan common to large taxonomic groups and has been associated to the collinear expression of *Hox* genes<sup>66–68</sup>.

## Conclusion

This work contributes to the understanding of how *Hox* genes evolve. *Hox* genes may evolve in a model constrained by bauplan and pleiotropic genes regulation following the phylotype hypothesis of the hourglass model. These genes showed their importance as molecular markers of large groups like Protostomia and Deuterostomia, Vertebrata, and Ecdysozoa and Platytrchozoa.

## METHODS

Homologous *Hox* protein sequences involved with body patterning of the anterior region (*Hox1-Hox4*) were retrieved from NCBI (<http://ncbi.nlm.nih.gov/>) using the BLASTp algorithm<sup>69</sup>. With the exception of *Hox3* (whose sequences included in our dataset comprise only vertebrates, since evolution of this gene is characterized by gene duplication and neofunctionalization to *bicoid*, *zerknüllt*, and *z2* across some protostome taxa<sup>35,70</sup>, turning assertion of orthologs in basal metazoan a challenging task), all anterior *Hox* sequences were analyzed individually, but grouping together genes from different clusters (*A*, *B*, *C* and *D*, where applicable). This approach resulted in four independent sets of data, which were aligned using MAFFT v7 webserver<sup>71</sup> and a combination of available parameters to refine the alignment: iterative refinement methods (E-INS-i, L-INS-i and G-INS-i) and scoring matrix (BLOSUM 45 and BLOSUM 62) for amino acid sequences. The best evolutionary model for each *Hox* protein/alignment was JTT+I+G+F and was inferred using ProtTest 3.2<sup>72</sup> considering the Akaike information criterion (AIC). Unrooted maximum likelihood trees of each *Hox* was inferred using PhyML 3.0 webserver<sup>73</sup>. FigTree v1.4.3<sup>74</sup> was used to visualise and edit the trees.

## **Data Availability**

All data used in the current study are publicly available. Species names, gene names, and respective accession numbers in NCBI are available in Supplementary Table SX.

## **ACKNOWLEDGEMENTS**

This work was funded by a studentship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to V.M.M. and a fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (153407/2018-2) to J.L.O. We would like to thank *Núcleo de Processamento de Alto Desempenho* of the *Universidade Federal do Rio Grande do Norte - NPAD / UFRN* for the access to its supercomputer.

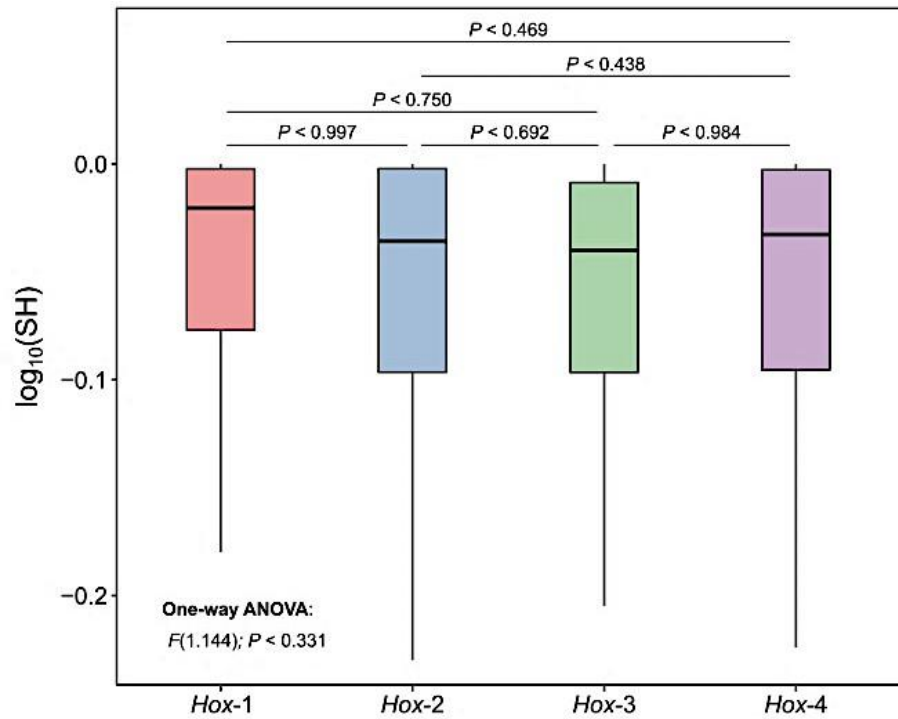
## **AUTHOR CONTRIBUTIONS**

V.M.M. run all of the evolutionary analyses, V.M.M., G.H.C.V. and S.T.F. conceived the project and analyses, P.E.T.S. contributed to data processing & bioinformatics analyses, V.M.M., J.L.O., and S.T.F. analyzed results and wrote the paper.

## **COMPETING INTERESTS**

The authors declare no competing financial interests.

## SUPPLEMENTARY FIGURES



**Figura S1.** Shimodaira-Hasegawa (SH) support values variation within and among the four Hoxgenes. The middle line, bottom and top of show the median, 25th and 75th percentiles respectively; whiskers present the 95% confidence interval of the distributions. Statist

## 4. Considerações finais

Este trabalho contribui para a compreensão de como os genes Hox anteriores evoluem. Os genes *Hox* podem evoluir em um modelo de restrição por bauplan e regulação de genes pleiotrópicos de desenvolvimento, onde os organismos são restringidos no desenvolvimento, seguindo a hipótese do filotipo da ampulheta, mostrando que os estágios de embriogênese média em vertebrados têm uma maior similaridade morfológica. Entretanto, precisa-se de mais estudos experimentais para expandir o entendimento para outros grupos filogenéticos. Esses genes mostraram sua importância como marcadores moleculares de grandes grupos como Protostomia e Deuterostomia, Vertebrata e Ecdysozoa e Platytrochozoa. Para o clado Platytrochozoa, seria interessante procurar caracteres morfológicos que servissem como características sipomórficas do grupo.

As filogenias de Hox que estão presentes em Gnatosthomata, nem sempre igualam a hipótese filogenética dos Craniata proposta por Meyer e Zardoya<sup>6</sup>, e isso porque nossa filogenia reconstrói a relação dos genes Hox presentes em Gnatosthomata, e diferenças podem ser explicadas com futuros testes experimentais como na filogenia Hox-D3, em que as sequências de Aves e Mammalia são irmãos e a função de Hox-D3 associada à diferenciação da primeira vértebra cervical Atlas pode explicar a possível convergência das sequências entre Mammalia e Aves, porque a vértebra atlas dentro de Amniota fornece movimento vertical da cabeça que são feitos pela articulação do crânio com o Atlas. Nosso trabalho também corrobora com o trabalho de Chiari et al (2012)<sup>8,43</sup> através das filogenias Hox-B1, Hox-A2, Hox-D3, Hox-A4, Hox-B4 e Hox-D4 e também fornece suporte adicional para a hipótese de Testudines como um grupo irmão de Archosauria.

# REFERÊNCIAS

1. Mare, R. D. *A concepção da teoria evolutiva desde os gregos*. (EDIPUCRS).
2. Amorim, D. de S. *Fundamentos de sistemática filogenética*. (Holos, 2002).
3. Dunn, C. W., Giribet, G., Edgecombe, G. D. & Hejnol, A. Animal Phylogeny and Its Evolutionary Implications. *Annu. Rev. Ecol. Evol. Syst.***45**, 371–395 (2014).
4. Autores, V. *Princípios integrados de zoologia*. (Guanabara Koogan, 2013).
5. Laumer, C. E. *et al.* Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. R. Soc. B Biol. Sci.***286**, 20190831 (2019).
6. Meyer, A. & Zardoya, R. Recent Advances in the (Molecular) Phylogeny of Vertebrates. *Annu. Rev. Ecol. Evol. Syst.***34**, 311–338 (2003).
7. Hedges, S. B. Amniote phylogeny and the position of turtles. *BMC Biol.***10**, 64 (2012).
8. Crawford, N. G. *et al.* A phylogenomic analysis of turtles. *Mol. Phylogenet. Evol.***83**, 250–257 (2015).
9. Hebert Paul D. N., Cywinska Alina, Ball Shelley L. & deWaard Jeremy R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B Biol. Sci.***270**, 313–321 (2003).
10. Field, K. G. *et al.* Molecular phylogeny of the animal kingdom. *Science***239**, 748–753 (1988).
11. Pires, A. C. & Marinoni, L. DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica***10**, 339–346 (2010).
12. Halanych, K. M. *et al.* Evidence from 18S ribosomal DNA that the lophophorates are protostome animals. *Science***267**, 1641–1643 (1995).
13. Aguinaldo, A. M. A. *et al.* Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature***387**, 489 (1997).

14. Telford, M. J. The Animal Tree of Life. *Science***339**, 764–766 (2013).
15. Futuyma, D. J. & Kirkpatrick, M. *Evolution*. (Oxford University Press, 2017).
16. Minelli, A. Phylo-evo-devo: combining phylogenetics with evolutionary developmental biology. *BMC Biol.***7**, 36 (2009).
17. Gehring, W. J. & Hiromi, Y. Homeotic Genes and the Homeobox. *Annu. Rev. Genet.***20**, 147–173 (1986).
18. Pearson, J. C., Lemons, D. & McGinnis, W. Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.***6**, 893–904 (2005).
19. Kardong, K. V. & Hoenen, S. M. M. *Vertebrados: anatomia comparada, função e evolução*. (Roca, 2014).
20. Wolpert, L. & Tickle, C. *Principles of Development*. (OUP Oxford, 2011).
21. Lewis, E. B. A Gene Complex Controlling Segmentation in *Drosophila*. in *Genes, Development and Cancer: The Life and Work of Edward B. Lewis* (ed. Lipshitz, H. D.) 205–217 (Springer US, 2004). doi:10.1007/978-1-4419-8981-9\_13
22. Schneuwly, S., Klemenz, R. & Gehring, W. J. Redesigning the body plan of *Drosophila* by ectopic expression of the homoeotic gene *Antennapedia*. *Nature***325**, 816–818 (1987).
23. Struhl, G. A homoeotic mutation transforming leg to antenna in *Drosophila*. *Nature***292**, 635 (1981).
24. Reddy, P. C., Unni, M. K., Gungi, A., Agarwal, P. & Galande, S. Evolution of Hox-like genes in Cnidaria: Study of *Hydra* Hox repertoire reveals tailor-made Hox-code for Cnidarians. *Mech. Dev.***138**, 87–96 (2015).
25. Ryan, J. F. *et al.* Pre-Bilaterian Origins of the Hox Cluster and the Hox Code: Evidence from the Sea Anemone, *Nematostella vectensis*. *PLOS ONE***2**, e153 (2007).

26. Chourrout, D. *et al.* Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature***442**, 684–687 (2006).
27. Amores, A. *et al.* Developmental Roles of Pufferfish Hox Clusters and Genome Evolution in Ray-Fin Fish. *Genome Res.***14**, 1–10 (2004).
28. Sundström, G., Larsson, T. A. & Larhammar, D. Phylogenetic and chromosomal analyses of multiple gene families syntenic with vertebrate Hox clusters. *BMC Evol. Biol.***8**, 254 (2008).
29. Dehal, P. & Boore, J. L. Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLOS Biol.***3**, e314 (2005).
30. Pascual-Anaya, J. *et al.* Broken colinearity of the amphioxus Hox cluster. *EvoDevo***3**, 28 (2012).
31. Galant, R. & Carroll, S. B. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature***415**, 910–913 (2002).
32. Burke, A. C., Nelson, C. E., Morgan, B. A. & Tabin, C. Hox genes and the evolution of vertebrate axial morphology. *Development***121**, 333–346 (1995).
33. Parker, H. J., Bronner, M. E. & Krumlauf, R. The vertebrate Hox gene regulatory network for hindbrain segmentation: Evolution and diversification. *BioEssays***38**, 526–538 (2016).
34. Parker, H. J., Bronner, M. E. & Krumlauf, R. A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. *Nature***514**, 490–493 (2014).
35. Hughes, C. L. & Kaufman, T. C. Hox genes and the evolution of the arthropod body plan1. *Evol. Dev.***4**, 459–499 (2002).
36. Neme, R. & Tautz, D. Phylogenetic patterns of emergence of new genes support a model of frequent de novo evolution. *BMC Genomics***14**, 117 (2013).

37. Kaessmann, H. Origins, evolution, and phenotypic impact of new genes. *Genome Res.***20**, 1313–1326 (2010).
38. Chen, S., Krinsky, B. H. & Long, M. New genes as drivers of phenotypic evolution. *Nat. Rev. Genet.***14**, 645–660 (2013).
39. Nei, M. & Rooney, A. P. Concerted and Birth-and-Death Evolution of Multigene Families. *Annu. Rev. Genet.***39**, 121–152 (2005).
40. Struck, T. H. *et al.* Platyzoan Paraphyly Based on Phylogenomic Data Supports a Noncoelomate Ancestry of Spiralia. *Mol. Biol. Evol.***31**, 1833–1849 (2014).
41. Cook, C. E., Smith, M. L., Telford, M. J., Bastianello, A. & Akam, M. Hox genes and the phylogeny of the arthropods. *Curr. Biol.***11**, 759–763 (2001).
42. Málaga-Trillo, E. & Meyer, A. Genome Duplications and Accelerated Evolution of Hox Genes and Cluster Architecture in Teleost Fishes. *Integr. Comp. Biol.***41**, 676–686 (2001).
43. Chiari, Y., Cahais, V., Galtier, N. & Delsuc, F. Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). *BMC Biol.***10**, 65 (2012).
44. Gamse, J. T. & Sive, H. Early anteroposterior division of the presumptive neurectoderm in *Xenopus*. *Mech. Dev.***104**, 21–36 (2001).
45. Lowe, C. J., Clarke, D. N., Medeiros, D. M., Rokhsar, D. S. & Gerhart, J. The deuterostome context of chordate origins. *Nature***520**, 456–465 (2015).
46. Roux, M. *et al.* Hoxa1 and Hoxb1 are required for pharyngeal arch artery development. *Mech. Dev.***143**, 1–8 (2017).
47. McNulty, C. L., Peres, J. N., Bardine, N., Akker, W. M. R. van den & Durston, A. J. Knockdown of the complete Hox paralogous group 1 leads to dramatic hindbrain and neural crest defects. *Development***132**, 2861–2871 (2005).

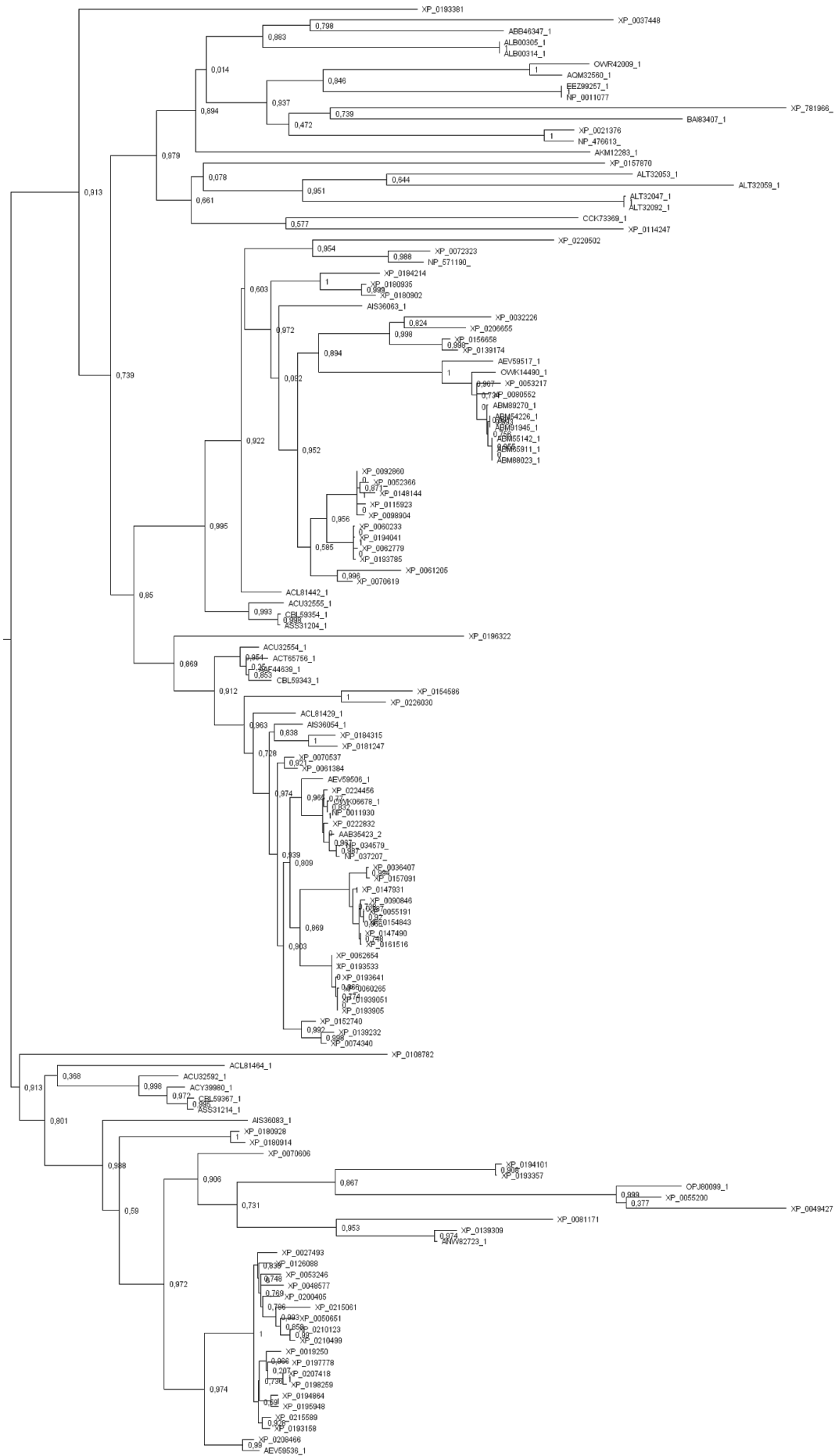
48. Condie, B. G. & Capecchi, M. R. Mice with targeted disruptions in the paralogous genes *hoxa*-3 and *hoxd*-3 reveal synergistic interactions. *Nature***370**, 304 (1994).
49. Manley, N. R. & Capecchi, M. R. The role of *Hoxa*-3 in mouse thymus and thyroid development. *Development***121**, 1989–2003 (1995).
50. Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A. & Krumlauf, R. Altered segmental identity and abnormal migration of motor neurons in mice lacking *Hoxb*-1. *Nature***384**, 630–634 (1996).
51. Holland, P. W. H. Evolution of homeobox genes. *Wiley Interdiscip. Rev. Dev. Biol.***2**, 31–45 (2013).
52. Clark, S. G., Chisholm, A. D. & Horvitz, H. R. Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39*. *Cell***74**, 43–55 (1993).
53. Eizinger, A. & Sommer, R. J. The Homeotic Gene *lin-39* and the Evolution of Nematode Epidermal Cell Fates. *Science***278**, 452–455 (1997).
54. Koh, K. *et al.* Cell fates and fusion in the *C. elegans* vulval primordium are regulated by the EGL-18 and ELT-6 GATA factors — apparent direct targets of the LIN-39 Hox protein. *Development***129**, 5171–5180 (2002).
55. Lohmann, I., McGinnis, N., Bodmer, M. & McGinnis, W. The *Drosophila* Hox Gene *Deformed* Sculpts Head Morphology via Direct Regulation of the Apoptosis Activator *reaper*. *Cell***110**, 457–466 (2002).
56. Morrison, A., Ariza-McNaughton, L., Gould, A., Featherstone, M. & Krumlauf, R. *HOXD4* and regulation of the group 4 paralog genes. *Development***124**, 3135–3146 (1997).
57. Meazza, R. *et al.* Expression of *HOXC4* homeoprotein in the nucleus of activated human lymphocytes. *Blood***85**, 2084–2090 (1995).

58. Rieger, E. *et al.* Expression of the Homeobox Gene HOXC4 in Keratinocytes of Normal Skin and Epithelial Skin Tumors Is Correlated with Differentiation. *J. Invest. Dermatol.***103**, 341–346 (1994).
59. Kömüves, L. G. *et al.* HOXB4 homeodomain protein is expressed in developing epidermis and skin disorders and modulates keratinocyte proliferation. *Dev. Dyn.***224**, 58–68 (2002).
60. Ramfrez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R. & Bradley, A. Hoxb-4 (Hox-2.6) mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. *Cell***73**, 279–294 (1993).
61. Mansfield, J. H. & Abzhanov, A. Hox expression in the American alligator and evolution of archosaurian axial patterning. *J. Exp. Zoolog. B Mol. Dev. Evol.***314B**, 629–644 (2010).
62. Giribet, G. & Edgecombe, G. D. Current Understanding of Ecdysozoa and its Internal Phylogenetic Relationships. *Integr. Comp. Biol.***57**, 455–466 (2017).
63. Tschopp Patrick & Tabin Clifford J. Deep homology in the age of next-generation sequencing. *Philos. Trans. R. Soc. B Biol. Sci.***372**, 20150475 (2017).
64. Pascual-Anaya, J. *et al.* Hagfish and lamprey Hox genes reveal conservation of temporal colinearity in vertebrates. *Nat. Ecol. Evol.***2**, 859–866 (2018).
65. Smith, J. M. *et al.* Developmental Constraints and Evolution: A Perspective from the Mountain Lake Conference on Development and Evolution. *Q. Rev. Biol.***60**, 265–287 (1985).
66. Hu, H. *et al.* Constrained vertebrate evolution by pleiotropic genes. *Nat. Ecol. Evol.***1**, 1722 (2017).
67. Duboule, D. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development***1994**, 135–142 (1994).

68. Irie, N. & Kuratani, S. The developmental hourglass model: a predictor of the basic body plan? *Development***141**, 4649–4655 (2014).
69. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.***215**, 403–410 (1990).
70. Stauber, M., Jäckle, H. & Schmidt-Ott, U. The anterior determinant bicoid of *Drosophila* is a derived Hox class 3 gene. *Proc. Natl. Acad. Sci.***96**, 3786–3789 (1999).
71. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.***20**, 1160–1166 (2019).
72. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics***27**, 1164–1165 (2011).
73. Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.***59**, 307–321 (2010).
74. Rambaut, A. *Automatically exported from code.google.com/p/figtree: rambaut/figtree.* (2019).

# Anexos

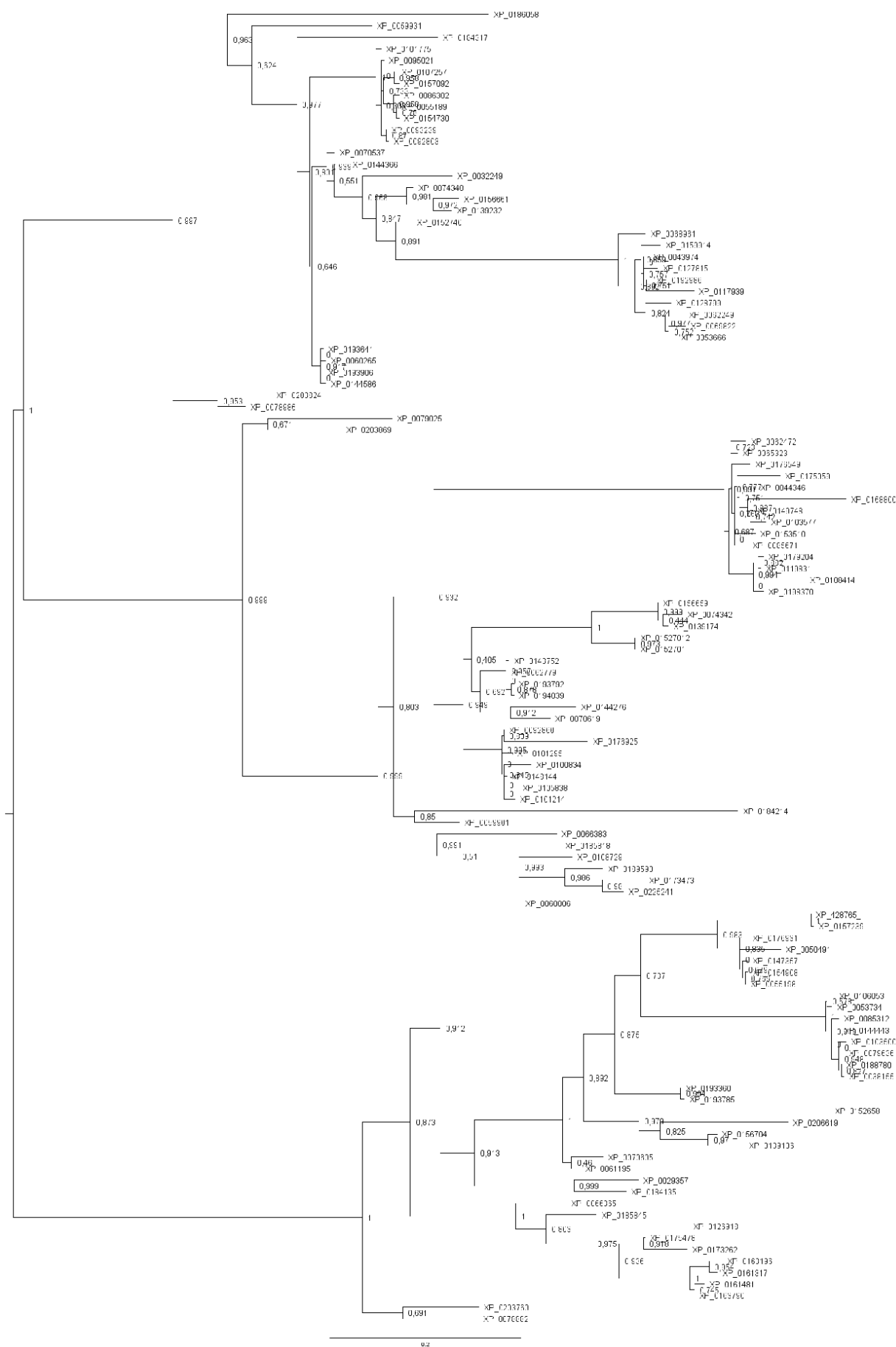
*Hox1*



**Hox2**



***Hox3***



***Hox4***

