

---

---

## INFERINDO RELAÇÕES FILOGENÉTICAS A PARTIR DA CONSERVAÇÃO DA PROTEÍNA HOMEÓTICA HOXA-1 DE DIFERENTES VERTEBRADOS POR ANÁLISES GENÉTICAS EVOLUTIVAS MOLECULARES

### INFERRING PHYLOGENETIC RELATIONSHIPS FROM CONSERVATION OF THE HOMEOTIC PROTEIN HOXA-1 OF DIFFERENT VERTEBRATES BY MOLECULAR EVOLUTIONARY GENETIC ANALYSES

Luisa Cunha OLIANI<sup>1</sup>, Jane Eyre GABRIEL<sup>2</sup>

1- Graduanda da Universidade Federal do Vale do São Francisco UNIVASF, Petrolina, PE.

2- Professora da Universidade Federal do Vale do São Francisco UNIVASF, Petrolina, PE.

Autor para correspondência: [jane.gabriel@univasf.edu.br](mailto:jane.gabriel@univasf.edu.br)

#### RESUMO:

O objetivo do presente estudo foi investigar as relações filogenéticas entre vertebrados a partir de diferenças na natureza química dos resíduos de aminoácidos da proteína homeótica HoxA por análises comparativas de alinhamentos múltiplos. Após a busca e seleção das sequências da proteína HoxA de 18 animais em um banco de dados de informação biológica, alinhamentos múltiplos de seus resíduos foram efetuados empregando o programa Clustal Omega, seguido pela construção de um dendograma para inferir as relações filogenéticas entre as espécies selecionadas. Análises comparativas dos alinhamentos múltiplos revelaram alta conservação estrutural com valor de identidade máxima de 35,593% entre os aminoácidos da proteína HoxA de distintos vertebrados. Além disso, o dendograma das relações filogenéticas demonstrou o agrupamento dos vertebrados selecionados dispostos em diferentes sub-ramos, onde animais com alta proximidade evolutiva foram intimamente agrupados em alguns sub-ramos, tais como: camundongo e ratazana (0,02) e homem e chimpanzé (0,01). Entretanto, divergências evolutivas em animais com alta proximidade evolutiva também foram observadas com o agrupamento de peixe-zebra (0,46), tubarão chifre (0,17) e celacanto (0,15) em ramos mais divergentes. Assim, conservação estrutural e proximidade filogenética descritas no presente estudo fornecem subsídios acerca da história evolutiva de vertebrados com base nas relações filogenéticas geradas a partir de alinhamentos múltiplos de aminoácidos da proteína HoxA de diferentes animais.

**Palavras-chave:** Alinhamentos múltiplos; Conservação proteica; Relações filogenéticas

#### ABSTRACT:

The objective of this study was to investigate the phylogenetic relationships among vertebrates from differences in the chemical nature of amino acid residues of the homeotic protein HoxA by comparative analyses of multiple alignments. Search and selection of the sequences of interest of 18 animals were performed from biological information database. Amino acid residues were aligned using program Clustal Omega, followed by constructing dendogram demonstrating phylogenetic relationships among animals selected. Comparative alignments of the protein HoxA revealed a maximum identity of 35.593% among different vertebrates. Moreover, the dendogram of phylogenetic relationships showed the grouping of vertebrates arranged in different sub-branches, where animals with high evolutionary proximity were closely grouped in branches, such as: mouse and rat

(0.02), and man and chimpanzee (0.01). Nevertheless, evolutionary divergences were also observed in animals with closest evolutionary proximity by clustering zebrafish (0.46), horn shark (0.17) and coelacanth (0.15) in more divergent branches. Thus, structural conservation and phylogenetic proximity described in current study provide insights into the evolutionary history of vertebrates based on phylogenetic relationships generated from multiple amino acid alignments of the protein HoxA of different animals.

**Keywords:** Multiple alignments; Phylogenetic relationships; Protein conservation

## 1. INTRODUÇÃO

Genes homeóticos são reguladores chaves capazes de codificar proteínas essenciais para o desenvolvimento de determinados segmentos ou estruturas do eixo corporal durante a embriogênese. A maioria dos genes homeóticos codifica fatores de transcrição contendo um homeodomínio típico, cuja ativação envolve uma cascata de genes reguladores essenciais para promover a síntese tardia de proteínas durante o desenvolvimento embrionário (MALLO & ALONSO, 2013). Os genes Hox em vertebrados determinam o eixo ântero-posterior do corpo, incluindo o tipo de vértebra a ser formada em cada região, além de serem também importantes para a morfogênese dos membros anteriores/superiores e posteriores/inferiores, para o estabelecimento do trato digestivo e do sistema reprodutor e para a organização do sistema nervoso (LAPPIN et al, 2006).

Estudos de mutagênese dirigida em células-tronco embrionárias para produzir mutantes carreando supressão na expressão de genes HoxA vem sendo conduzidos a fim de examinar o papel desses genes durante a proliferação e diferenciação celular no desenvolvimento animal. A maioria das mutações homeóticas com efeitos fenotípicos marcantes em mutantes de *Drosophila* apresentam efeitos disruptivos tão grandes no desenvolvimento, capazes de resultar em acentuada redução de suas viabilidades (FUTUYMA, 1997). Embriões mutantes para os genes Hox1, Hox5, Hox7 e Hox8 apresentavam fenótipos alterados relacionados com o desenvolvimento e a disposição das costelas e do esterno (MCINTYRE et al, 2007). Ainda, a análise de embriões mutantes para os genes Hox resultou em costelas extras nas vértebras lombares e ligação das costelas ao esterno nas vértebras mais posteriores (MCINTYRE et al, 2007, VINAGRE et al, 2010).

Embora um número expressivo de relatos na literatura tem desvendado as interações estruturais e funcionais dos clusters contendo sequências de nucleotídeos dos genes HoxA em diferentes animais, poucos relatos na literatura tem desvendado o grau de conservação dos produtos proteicos codificados por tais genes e suas implicações na área do conhecimento básico em biologia evolutiva. Nesse contexto, o objetivo do presente estudo foi inferir relações filogenéticas entre vertebrados a partir do grau de conservação e

---

da natureza química dos resíduos de aminoácidos da proteína homeótica HoxA-1 por análises genéticas evolutivas moleculares.

## 2. MATERIAIS E MÉTODOS

### 2.1 Busca de sequências de interesse em bancos de dados biológicos

As sequências de aminoácidos da proteína homeótica HoxA-1 foram obtidas por meio de buscas em banco de dados de informação biológica UniProt knowledgebase (CHEN et al, 2017). Durante as análises de busca e seleção nesse banco de dados foram identificadas sequências da proteína homeótica HoxA-1 expressas em 18 animais vertebrados: homem (P49639 *Homo sapiens*), peixe-zebra (Q90423, *Danio rerio*), camundongo (P09022 *Mus musculus*), ratazana (O08656 *Rattus norvegicus*), tubarão chifre (Q9IA19 *Heterodontus francisci*), canguru walabi (G9HPQ6 *Macropus eugenii*), chimpanzé (H2QUB3 *Pan troglodytes*), boi (E1B918 *Bos taurus*), celacanto (H3BCH5 *Latimeria chalumnae*), tartaruga-de-carapaça-mole-chinesa (K7F451 *Pelodiscus sinensis*), panda gigante (D2GVC2 *Ailuropoda melanoleuca*), gato (M3W8N7 *Felis catus*, diabo da tasmânia (G3WWK9 *Sarcophilus harrisi*), cavalo (F6YY78 *Equus caballus*), morcego preto (L5KNV4 *Pteropus alecto*), coelho europeu (B7NZT1 *Oryctolagus cuniculus*), morcego marrom (G1PQ54 *Myotis lucifugus*), cão (E2R5Z1 *Canis familiaris*). Embora esse banco de dados disponibilize dezenas de sequências de aminoácidos da proteína HoxA-1 nos mais variados organismos, foi estabelecido como critério de seleção que todas as sequências de interesse escolhidas para tais análises apresentassem número similar de resíduos com aproximadamente 300 aminoácidos.

### 2.2 Alinhamentos múltiplos dos resíduos de aminoácidos da proteína Hoax

Sequências totais dos segmentos peptídicos HoxA-1 de distintas espécies animais foram submetidas a alinhamentos múltiplos pela comparação direta entre os resíduos apresentados em formato FASTA pelo programa Clustal Omega (SIEVERS et al, 2011), disponíveis no programa computacional Expasy Proteomics Server. Ainda, baseado nos resultados dos alinhamentos múltiplos das sequências selecionadas foi estabelecido o grau de identidade e de similaridade entre os resíduos de aminoácidos.

### **2.3 Construção de dendograma das relações filogenéticas entre vertebrados selecionados a partir dos alinhamentos múltiplos da sequência da proteína HoxA**

Um dendograma das relações evolutivas entre as espécies selecionadas foi construído a partir de sequências de interesse empregando o programa Mega versão 6.0 (TAMURA et al, 2013), empregando o método de Neighbor-Joining (NJ) e de Kimura parâmetro-2 (SAITOU & NEI, 1987). Os ensaios experimentais envolveram a análise de 18 sequências de aminoácidos em um total de 350 posições. Os tempos de divergência para todos os pontos de ramificação foram calculados pelo método RelTime (TAMURA et al, 2012) a partir dos comprimentos dos ramos contidos na árvore inferida. Barras ao redor de cada nó representam intervalos de confiança de 95%, os quais foram computados pelo método descrito por TAMURA et al (2013).

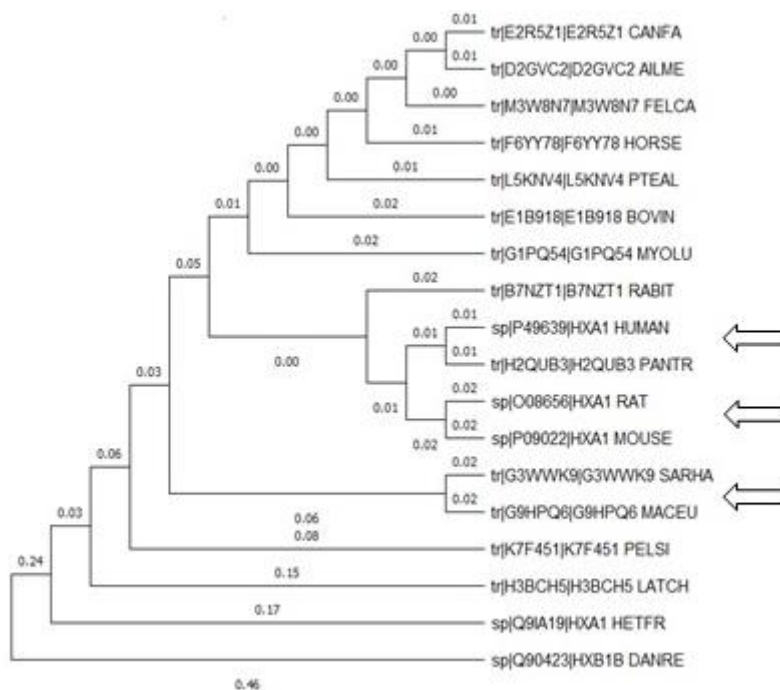
## **3. RESULTADOS**

Alinhamentos comparativos dos resíduos de aminoácidos da proteína HoxA-1 de distintos vertebrados revelaram uma identidade máxima de 35,593%, demonstrando que em 126 posições dessa proteína os resíduos de aminoácidos permaneceram conservados em todas as espécies selecionadas, ou seja, posições idênticas nas quais não foram observadas quaisquer alterações da natureza química dos aminoácidos (Figura 1, posições indicadas por asterisco na linha inferior do alinhamento). Além disso, 76 posições foram identificadas como sendo posições similares nas análises de alinhamentos múltiplos da proteína HoxA-1, ou seja, quando foi observada a alteração de natureza química entre resíduos de aminoácidos em uma mesma posição, sem que necessariamente tenha havido alterações nas suas propriedades bioquímicas (seta apresentada na Figura 1, demonstrando a presença dos aminoácidos treonina (T) ou serina (S) na mesma posição de alinhamento, sendo ambos os aminoácidos classificados como polares neutros contendo grupos químicos que tendem a formar ligações de hidrogênio).



**FIGURA 1.** ALINHAMENTOS MÚLTIPLOS PARCIAL DOS RESÍDUOS 1 A 180 DAS SEQUÊNCIAS DE RESÍDUOS DE AMINOÁCIDOS DA PROTEÍNA HOMEÓTICA HOXA-1 EM DIFERENTES VERTEBRADOS. OS SÍMBOLOS ASTERISCO \*, PONTO. E DOIS PONTOS: (APRESENTADOS NA LINHA INFERIOR EM CADA SÉRIE DE ALINHAMENTO) INDICAM MÁXIMA IDENTIDADE ENTRE TODAS AS SEQUÊNCIAS, ALTERAÇÃO DE UM RESÍDUO DE AMINOÁCIDO OU ALTERAÇÃO DE DOIS AMINOÁCIDOS, RESPECTIVAMENTE. A AUSÊNCIA DE SÍMBOLOS NA LINHA INFERIOR EM CADA SÉRIE INDICA A PRESENÇA DE PELO MENOS TRÊS AMINOÁCIDOS DIFERENTES ENTRE AS SEQUÊNCIAS ANALISADAS.

O dendograma das relações filogenéticas entre os vertebrados analisados demonstrou o agrupamento dos animais selecionados dispostos em diferentes sub-ramos (0,00, 0,01 e 0,02). Contudo, divergências acentuadas com menor grau de conservação da proteína HoxA-1 foram observadas a partir do agrupamento do peixe-zebra (0,46), tubarão chifre (0,17) e celacanto (0,15), que foram dispostos em ramos mais divergentes nesse dendograma (Figura 2). Interessantemente, alguns vertebrados foram intimamente agrupados em alguns sub-ramos, tais como: camundongo e ratazana (0,02); homem e chimpanzé (0,01); e canguru walabi e diabo da tasmânia (0,02) (setas apresentadas na Figura 2). Estimativas da divergência evolutiva inferidas a partir das sequências de aminoácidos demonstraram que as sequências da HoxA-1 humana e de peixe-zebra foram as mais diferenciadas (0,88403), enquanto as proteínas HoxA-1 humana e de panda gigante foram as menos diferenciadas (0,00908).



**FIGURA 2.** DENDOGRAMA ILUSTRATIVO DAS RELAÇÕES FILOGENÉTICAS ENTRE DIFERENTES VERTEBRADOS A PARTIR DA CONSERVAÇÃO DE RESÍDUOS DE AMINOÁCIDOS DA PROTEÍNA HOMEÓTICA HOXA-1. OS ESPÉCIMES ANALISADOS NESSE ESTUDO FORAM: HOMEM (P49639 *HOMO SAPIENS*), PEIXE-ZEBRA (Q90423, *DANIO RERIO*), CAMUNDONGO (P09022 *MUS MUSCULUS*), RATAZANA (O08656 *RATTUS NORVEGICUS*), TUBARÃO CHIFRE (Q9IA19 *HETERODONTUS FRANCISCI*), CANGURU WALABI (G9HPQ6 *MACROPUS EUGENII*), CHIMPANZÉ (H2QUB3 *PAN TROGLODYTES*), BOI (E1B918 *BOS TAURUS*), CELACANTO (H3BCH5 *LATIMERIA CHALUMNAE*), TARTARUGA-DE-CARAPAÇA-MOLE-CHINESA (K7F451 *PELODISCUS SINENSIS*), PANDA GIGANTE (D2GVC2 *AILUROPODA MELANOLEUCA*), GATO (M3W8N7 *FELIS CATUS*, DIABO DA TASMÂNIA (G3WWK9 *SARCOPHILUS HARRISII*), CAVALO (F6YY78 *EQUUS CABALLUS*), MORCEGO PRETO (L5KNV4 *PTEROPUS ALECTO*), COELHO EUROPEU (B7N2T1 *ORYCTOLAGUS CUNICULUS*), MORCEGO MARROM (G1PQ54 *MYOTIS LUCIFUGUS*), CÃO (E2R5Z1 *CANIS FAMILIARIS*). SETAS INDICAM TRÊS SUB-RAMOS APRESENTANDO ANIMAIS INTIMAMENTE AGRUPADOS.

#### 4. DISCUSSÃO

As descobertas descritas no presente estudo apontam alta identidade dos resíduos de aminoácidos da proteína HoxA-1 em animais evolutivamente distantes (Figura 1).

---

SANTINI et al (2003), comparando sequências de DNA entre animais com genomas evolutivamente distantes, identificaram que os genes Hox são altamente conservados durante a evolução em vertebrados, estruturalmente organizados em clusters, sem que haja interrupção de outros genes. Ainda, as regiões regulatórias intergênicas situadas entre tais genes são aparentemente mais conservadas evolutivamente do que as regiões localizadas nas extremidades terminais dos clusters dos genes Hox em vertebrados (SANTINI et al, 2003). Além disso, MCGINNIS & KRUMLAUF (1992) demonstraram que os homeodomínios individuais das proteínas Hox geralmente exibem maior similaridade com os homeodomínios de outras espécies do que com proteínas codificadas por genes adjacentes dentro do próprio cluster Hox. Assim, os resultados descritos no presente estudo indicam alta conservação dos segmentos proteicos da HoxA-1, mesmo entre espécies com vastas distâncias evolutivas.

Baseado nos alinhamentos múltiplos de resíduos de aminoácidos da proteína HoxA-1 em diferentes vertebrados por análises moleculares evolutivas é possível sugerir que agrupamentos de animais evolutivamente próximos em sub-ramos pouco divergentes podem sugerir acentuada proximidade filogenética entre os vertebrados analisados (Figura 2). De acordo com YANG & BOURNE (2009), eventos associados com alterações das características e das ocorrências de um domínio proteico no genoma de diferentes organismos são de suma importância para o entendimento das relações filogenéticas e da consequente história biológica evolutiva. Ainda, OLIANI et al (2015) também descreveram resultados consistentes com descobertas apresentadas no presente estudo, tendo em vista que membros do grupo monofilético Ruminantia (boi, búfalo e ovelha) foram agrupados mais proximamente a partir de análises de alinhamentos múltiplos de sequências de aminoácidos dos fatores transcricionais associados à miogênese MyoD and MyoG. Nesse contexto, KIM et al (2000) sequenciaram e compararam clusters de Hox em diferentes vertebrados, demonstrando que o cluster HoxM de tubarão chifre é ortólogo ao cluster HoxA de mamífero, além de demonstrar alta similaridade estrutural ao cluster de anfioxo. Essa persistência de um cluster HoxA identificável entre animais com uma divergência de 800 milhões de anos na história evolutiva demonstra que os clusters de genes Hox são entidades genéticas altamente integradas e estruturadas, com função essencial em motivos genéticos associados ao controle do processo de desenvolvimento (KIM et al, 2000).

Ainda, animais com alta proximidade evolutiva e pertencentes aos clados Euteleostomi (peixe-zebra e celacanto) e Chondrichthyes (tubarão-chifre) foram dispostos em ramos mais distantes e com maior divergência em comparação aos agrupamentos

observados para os demais vertebrados analisados (Figura 2). Interessantemente, LUTZ et al (1996) demonstraram a acentuada conservação filogenética funcional das proteínas Hox, quando um gene ortólogo ao gene HoxA de galinha foi expresso perfeitamente em mutantes nulos para esse gene de *Drosophila melanogaster*, sugerindo que a função dos genes Hox é conservada filogeneticamente. Além disso, BODENMILLER et al (2002) compararam 11 sequências do gene HoxA-11 de galinha com seus ortólogos humanos e de camundongo, revelando múltiplas regiões de 80% ou homologias superiores entre estas espécies por análises filogenéticas. Comparações das sequências de nucleotídeos do transcrito HoxA-11 de camundongo, galinha e homem e seus respectivos homeodomínios indicam uma forte pressão seletiva em vertebrados contra mutações que resultam em alterações no padrão de codificação desses genes. CHIU et al (2000) demonstraram aceleradas taxas de evolução e de substituições acumuladas ao analisar sequências de genes HoxA-11A e HoxA-11B e de suas proteínas em diferentes mamíferos, embora alta conservação dos mesmos tenha sido observada entre celacanto e tetrápodes nessas condições experimentais.

A reconstrução da história evolutiva a partir das origens dos clusters de Hox e da natureza química dos produtos finais de seus homeodomínios fornece subsídios sobre seu significado evolutivo na filogenia de vertebrados, bem como seu papel nas origens dos vários planos corporais de vertebrados. Em conclusão, as descobertas apresentadas no presente estudo fornecem subsídios acerca da história evolutiva de vertebrados baseado nas relações filogenéticas geradas a partir de alinhamentos múltiplos de aminoácidos da proteína HoxA-1 de diferentes animais.

## 5. AGRADECIMENTOS

Os autores agradecem o Grupo de Estudos em Biologia Computacional e Bioinformática, designado "BIO in BYTES", pela cooperação na execução das análises computacionais apresentadas no presente estudo.

## 6. REFERÊNCIAS

BODENMILLER DM, BAXTER CS, HANSEN DV, *et al.* 2002. Phylogenetic analysis of HoxA 11 sequences reveals absence of transposable elements, conservation of transcription factor binding sites, and suggests antisense coding function. *DNA Sequence*, 13, 77-83.



---

CHEN C, HUANG H, WU CH. 2017. Protein Bioinformatics Databases and Resources. *Methods in Molecular Biology*, 1558, 3-39.

CHIU CH, NONAKA D, XUE L, *et al.* 2000. Evolution of Hoxa-11 in lineages phylogenetically positioned along the fin-limb transition. *Molecular Phylogenetics and Evolution*, 17, 305-316.

FUTUYMA DJ. *Biologia Evolutiva*. Ribeirão Preto: Sociedade Brasileira de Genética, 1997. 631p.

KIM CB, AMEMIYA C, BAILEY W, *et al.* 2000. Hox cluster genomics in the horn shark, *Heterodontus francisci*. *Proceedings of the National Academy of Sciences of the USA*, 97, 1655-1660.

LAPPIN TRJ, GRIER DG, THOMPSON A, *et al.* 2006. HOX GENES: seductive science, mysterious mechanisms. *Ulster Medical Journal*, 75, 23–31.

LUTZ B, LU HC, EICHELE G, *et al.* 1996. Rescue of *Drosophila* labial null mutant by the chicken ortholog Hoxb-1 demonstrates that the function of Hox genes is phylogenetically conserved. *Genes & Development*, 10,176–184.

MALLO M, ALONSO CR. 2013. The regulation of Hox gene expression during animal development. *Development*, 140, 3951-3963.

MCGINNIS W, KRUMLAUF R. 1992. Homeobox genes and axial patterning. *Cell*, 68, 283-302

MCINTYRE DC, RAKSHIT S, YALLOWITZ AR, *et al.* 2007. Hox patterning of the vertebrate rib cage. *Development*, 134, 2981-2999.

OLIANI LC, LIDANI KC, GABRIEL JE. 2015. Differentiated evolutionary relationships among chordates from comparative alignments of multiple sequences of MyoD and MyoG myogenic regulatory factors. *Genetics and Molecular Research*, 14, 12561-12566.

SAITOU N, NEI M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4, 406-425.

SANTINI S, BOORE JL, MEYER A. 2003. Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. *Genome Research*, 13, 1111-1122.

SIEVERS F, WILM A, DINEEN DG, *et al.* 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539.

TAMURA K, BATTISTUZZI FU, BILLING-ROSS P, *et al.* 2012. Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences of the USA*, 109, 19333-19338.

TAMURA K, STECHER G, PETERSON D, *et al.* 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725-2729.

VINAGRE T, MONCAUT N, CARAPUÇO M, *et al.* 2010. Hox-modulated regional expression of Myf5/Myf6 specifies global vertebral domains. *Developmental Cell*, 18, 655-661.

YANG S, BOURNE PE. 2009. The evolutionary history of protein domains viewed by species phylogeny. *PLoS One*, 4, e8378.