



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS AGRÁRIAS**  
**DEPARTAMENTO DE ENGENHARIA DE PESCA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA DE RECURSOS**  
**NATURAIS**

**JOÃO PEDRO FERNANDES QUEIROZ**

**EVOLUÇÃO DAS GLOBINAS NOS ANFÍBIOS**

**FORTALEZA**

**2021**

JOÃO PEDRO FERNANDES QUEIROZ

EVOLUÇÃO DAS GLOBINAS NOS ANFÍBIOS

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia de Recursos Naturais da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Biotecnologia de Recursos Naturais. Área de concentração: Purificação e caracterização de biomoléculas.

Orientador: Prof. Dr. Bruno Anderson Matias da Rocha.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

Q44e

Queiroz, João Pedro Fernandes.

Evolução das globinas nos anfíbios / João Pedro Fernandes Queiroz. – 2021.  
51 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-Graduação em Biotecnologia de Recursos Naturais, Fortaleza, 2021.

Orientação: Prof. Dr. Bruno Anderson Matias da Rocha.

1. Duplicação gênica. 2. Neofuncionalização. 3. Anotação de genes. I. Título.

CDD 660.6

---

JOÃO PEDRO FERNANDES QUEIROZ

EVOLUÇÃO DAS GLOBINAS NOS ANFÍBIOS

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia de Recursos Naturais da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Biotecnologia de Recursos Naturais. Área de concentração: Purificação e caracterização de biomoléculas.

Aprovada em: 28/04/2021.

BANCA EXAMINADORA

---

Prof. Dr. Bruno Anderson Matias da Rocha (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Nicholas Costa Barroso Lima  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. José Hélio Costa  
Universidade Federal do Ceará (UFC)

Ao Billy, melhor companheiro, nas anotações  
de genes e na hora da pizza

## AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

Ao Prof. Dr. Bruno Anderson Matias da Rocha, pela orientação e confiança no meu projeto.

Aos professores participantes da banca examinadora Nicholas Costa Barroso Lima e José Hélio Costa pelo tempo, pelas valiosas colaborações e sugestões.

Aos colegas do GBME, pelo apoio e companheirismo.

Ao Prof. Dr. Nicholas Costa Barroso Lima, pela orientação nas análises de genomas.

Ao Dr. Eduardo Henrique Salviano Bezerra pela paciência de escutar meus monólogos sobre globinas e pelo apoio.

Ao meu amigo Daniel Matias, pelas caminhadas e pelas horas jogando *Red Dead Redemption 2*, que ajudaram a sossegar minha mente das anotações de genes e análises filogenéticas. Estou ansioso pela campanha de *Resident Evil: Village* nestas férias.

Ao meu amigo Mateus Linhares, pelas pedaladas noturnas até o Cumbuco que tornaram a vida pós-pandemia mais suportável e me ajudaram a perder peso, uma tendência inédita desde que ingressei no ensino superior.

Ao meu primo, Carlos Henrique, que me emprestou o notebook para eu fazer as correções da dissertação, pois esta dissertação é o último documento que escrevi espontaneamente usando Word, e não tenho mais o Office instalado no meu notebook. Até o fim dos tempos só usarei R Markdown.

À minha mãe, Maria Avani Gonçalves Fernandes, que sempre me apoiou.

“A evolução é um tear encantado repleto de códigos de DNAs recombinantes cujos padrões evanescentes tecem, à medida que bailam com seus parceiros através de profundos tempos geológicos, um banco de dados volumoso de sabedoria ancestral, uma descrição digitalmente codificada de mundos ancestrais e das condições necessárias para habitá-los.” (DAWKINS, 1998, p. 352).

## RESUMO

O repertório de genes de globinas nos vertebrados gnatostomados é ditado pela perda e retenção diferencial de nove tipos de genes parálogos: *androglobina*, *neuroglobina*, *globina X*, *citoglobina*, *globina Y*, *mioglobina*, *globina E*, e as globinas dos tipos  $\alpha$  e  $\beta$ . Nós descrevemos o repertório de genes de globinas nas três ordens de anfíbios modernos: Anura, Caudata e Gymnophiona. Combinando análise filogenética e de sintenia conservada, mostramos que *mioglobina* e *globina E* foram perdidas no clado Batrachia, mas retidas em Gymnophiona. Os principais grupos de anfíbios também retiveram diferentes parálogos de *globina X*. Nenhum anfíbio apresentou o gene de globina do tipo  $\alpha^D$ . Todavia, dois clados de globinas do tipo  $\beta$  estão presentes em todos os anfíbios, indicando que o ancestral dos anfíbios possuía dois parálogos do tipo  $\beta$ . Nós também mostramos que ortólogos do gene que codifica uma hemoglobina monomérica que foi encontrada no coração de *Rana catesbeiana* estão presentes em espécies analisadas pertencentes aos grupos Neobatrachia e Pelobatoidea. Nós sugerimos que tais genes realizem funções análogas à mioglobina e globina E. Nós concluímos que o repertório de globinas em anfíbios é ditado pela perda e retenção dos tipos de genes parálogos citados acima, assim como pelo surgimento de um novo tipo de gene de globina através da cooptação de uma globina do tipo  $\alpha$ , possivelmente facilitada por um evento de transposição prévio.

**Palavras-chave:** duplicação gênica; neofuncionalização; anotação de genes.



## ABSTRACT

The globin gene repertoire of gnathostome vertebrates is dictated by differential retention and loss of nine paralogous gene types: *androglobin*, *neuroglobin*, *globin X*, *cytoglobin*, *globin Y*, *myoglobin*, *globin E*, and the  $\alpha$ - and  $\beta$ -globins. We report the globin gene repertoire of three orders of modern amphibians: Anura, Caudata, and Gymnophiona. Combining phylogenetic and conserved synteny analysis, we show that *myoglobin* and *globin E* were lost only in the Batrachia clade, but retained in Gymnophiona. The major amphibian groups also retained different paralogous copies of *globin X*. None of the amphibian presented  $\alpha^D$ -globin gene. Nevertheless, two clades of  $\beta$ -globins are present in all amphibians, indicating that the amphibian ancestor possessed two paralogous proto  $\beta$ -globins. We also show that orthologs of the gene coding for the monomeric hemoglobin found in the heart of *Rana catesbeiana* are present in Neobatrachia and Pelobatoidea species we analyzed. We suggest that these genes might perform myoglobin- and globin E-related functions. We conclude that the repertoire of globin genes in amphibians is dictated by both retention and loss of the paralogous genes cited above and the rise of a new globin gene through co-option of an  $\alpha$ -globin, possibly facilitated by a prior event of transposition.

**Keywords:** gene duplication; neofunctionalization; gene annotation.

## LISTA DE ABREVIATURAS E SIGLAS

Adgb	Androglobina
aHb	Hemoglobina (agnatos)
aMb	Mioglobina (agnatos)
b%	Porcentagem bootstrap
Cygb	Citoglobina
GbA	Globina A
GbE	Globina E
GbX	Globina X
GbY	Globina Y
Hb	Hemoglobina (gnatostomados)
Hba	Globina do tipo $\alpha$
Hbb	Globina do tipo $\beta$
Mb	Mioglobina (gnatostomados)
NCBI	National Center for Biotechnology Information
Ngb	Neuroglobina
ppb	Probabilidade posterior bayesiana
UCSC	University of California, Santa Cruz

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>10</b>
<b>2</b>	<b>EVOLUÇÃO DAS GLOBINAS DE VERTEBRADOS</b> .....	<b>11</b>
<b>2.1</b>	<b>Mecanismos evolutivos na origem das hemoglobinas</b> .....	<b>11</b>
<b>2.2</b>	<b>Tempo e modo na diversificação das globinas <math>\alpha</math> e <math>\beta</math> dos tetrápodes</b> .....	<b>12</b>
<b>2.3</b>	<b>Anfíbios: lacunas no conhecimento das globinas de vertebrados</b> .....	<b>13</b>
<b>3</b>	<b>OBJETIVOS</b> .....	<b>15</b>
<b>4</b>	<b>MATERIAIS E MÉTODOS</b> .....	<b>16</b>
<b>4.1</b>	<b>Anotação e curadoria manual dos genes de globinas</b> .....	<b>16</b>
<b>4.2</b>	<b>Reconstrução filogenética</b> .....	<b>17</b>
<b>4.3</b>	<b>Análise de sintenia conservada</b> .....	<b>19</b>
<b>5</b>	<b>RESULTADOS</b> .....	<b>20</b>
<b>5.1</b>	<b>O repertório de globinas nos anfíbios</b> .....	<b>20</b>
<b>5.2</b>	<b>Filogenia dos genes de globina dos tipos <math>\alpha</math> e <math>\beta</math></b> .....	<b>25</b>
<b>5.3</b>	<b>Sintenia dos genes de globina dos tipos <math>\alpha</math> e <math>\beta</math></b> .....	<b>29</b>
<b>6</b>	<b>DISCUSSÃO</b> .....	<b>31</b>
<b>6.1</b>	<b>Discordância antiga na árvore dos genes</b> .....	<b>31</b>
<b>6.2</b>	<b>Distribuição filética das globinas em vertebrados</b> .....	<b>31</b>
<b>6.3</b>	<b>Hemoglobina nos anfíbios</b> .....	<b>33</b>
<b>6.4</b>	<b>Globina A</b> .....	<b>34</b>
<b>7</b>	<b>CONCLUSÃO</b> .....	<b>36</b>
	<b>REFERÊNCIAS</b> .....	<b>37</b>

## 1 INTRODUÇÃO

Nove membros da superfamília de genes de globinas estão presentes nos gnatostomados: *androglobina* (*Adgb*), *globina X* (*GbX*), *neuroglobina* (*Ngb*), *citoglobina* (*Cygb*), *globina Y* (*GbY*), *mioglobina* (*Mb*), *globina E* (*GbE*) e as globinas dos tipos  $\alpha$  e  $\beta$  (*Hba* e *Hbb*, respectivamente), que codificam as cadeias da proteína hemoglobina (Hb), um tetrâmero  $\alpha_2\beta_2$  (BURMESTER; HANKELN, 2014; KEPPNER *et al.*, 2020; STORZ, 2019). O domínio globina é a estrutura terciária conservada através de todas as globinas de vertebrados. Ele compreende entre sete e oito hélices  $\alpha$  (nomeadas de A até H de acordo com a estrutura cristalográfica da Mb) (KENDREW *et al.*, 1958; PERUTZ *et al.*, 1960). Um grupo prostético composto por protoporfirina IX está confinado num bolsão hidrofóbico do polipeptídeo de globina. No centro do grupo prostético existe um átomo de ferro sendo coordenado por quatro nitrogênios pirrólicos coplanares. Um resíduo de histidina (His F8 ou o oitavo resíduo da hélice  $\alpha$  F, ou ainda histidina proximal) também é responsável por coordenar o ferro do grupo heme em um dos sítios axiais. Um resíduo de fenilalanina (Phe CD1 ou o primeiro resíduo da região entre as hélices  $\alpha$  C e D) mantém o grupo heme numa posição estável. His F8 e Phe CD1 são conservadas em todas as globinas de vertebrados. O sexto sítio de coordenação do ferro pode ser ou não coordenado por um resíduo de histidina ou glutamina (His/Gln E7), também na posição axial. O sexto sítio de coordenação é responsável pela ligação de ligantes gasosos, tais como  $O_2$  e NO (GELL, 2018). As globinas pentacoordenadas [cujo sexto sítio de coordenação não está sendo coordenado pela cadeia lateral de um resíduo de aminoácido da proteína] como Hb, Mb e GbE, desempenham funções respiratórias no transporte e armazenamento de  $O_2$  (BLANK *et al.*, 2011a; LÜDEMANN *et al.*, 2019; MAIRBÄURL; WEBER, 2012; WEBER; FAGO, 2004). Globinas hexacoordenadas [cujo sexto sítio de coordenação está sendo coordenado pela cadeia lateral de um resíduo de aminoácido da proteína] como Adgb, Ngb, GbX e Cygb vêm sendo caracterizadas estruturalmente e funcionalmente (CORTI *et al.*, 2016; DE SANCTIS *et al.*, 2004; DEWILDE *et al.*, 2001; SAWAI *et al.*, 2003; TRENT III; WATTS; HARGROVE, 2001). Funções propostas para as globinas hexacoordenadas incluem regulação redox, função de sensor de  $O_2$  e funções enzimáticas, tal como atividade de redução de nitrito, que estão envolvidas em mecanismos de sinalização celular ou na limpeza de espécies reativas de oxigênio e nitrogênio (BLANK *et al.*, 2011b; BURMESTER; HANKELN, 2009, 2014; HANKELN *et al.*, 2005; KAKAR *et al.*, 2010).

## 2 EVOLUÇÃO DAS GLOBINAS DE VERTEBRADOS

As globinas encontradas em vertebrados possuem origens duplicativas distintas. Por exemplo, *Adgb*, *Ngb* e *GbX* evoluíram a partir de eventos de duplicação antigos que ocorreram antes da separação entre protostômios e deuterostômios (BLANK; BURMESTER, 2012; BURMESTER *et al.*, 2000; DRÖGE *et al.*, 2012; EBNER *et al.*, 2010; HOFFMANN *et al.*, 2012; HOOGEWIJS *et al.*, 2012). A classificação mais recente da superfamília das globinas em metazoários apresenta cinco clados (SONG *et al.*, 2020), onde estão incluídas todas as globinas presentes nos vertebrados, exceto *Adgb*. A *Ngb* pertence ao clado V, enquanto que a *GbX* de vertebrados pertence ao clado III. No clado I encontramos todas as globinas específicas de vertebrados, que possivelmente se originaram a partir de um gene ancestral do tipo *Cygb* (SONG *et al.*, 2020). O grupo das globinas específicas de vertebrados (*Cygb*, *GbY*, *Mb*, *GbE*, *Hba* e *Hbb*) também inclui as hemoglobinas e mioglobinas encontradas nos agnatos (*aHb* e *aMb*), que são produtos de duplicações específicas na linhagem dos agnatos e não são ortólogas aos genes *Hba/Hbb* ou *Mb* dos gnatostomados (FAGO *et al.*, 2018; SCHWARZE *et al.*, 2014). Portanto, podemos dizer que o repertório de globinas do último ancestral comum dos vertebrados incluiu um gene de globina que deu origem às globinas específicas de vertebrados através de duas rodadas de duplicação do genoma inteiro acompanhadas por duplicações em tandem, assim como cópias de *Adgb*, *Ngb*, e pelo menos quatro parálogos de *GbX* (GOODMAN; MOORE; MATSUDA, 1975; HOFFMANN; OPAZO; STORZ, 2012; OPAZO *et al.*, 2015a; STORZ; OPAZO; HOFFMANN, 2013). Um terceiro evento de duplicação do genoma inteiro específico da linhagem dos peixes teleósteos forneceu material genético para especialização dos genes de globinas nesse grupo numeroso em espécies (DAANE *et al.*, 2020; GALLAGHER; MACQUEEN, 2017; OPAZO *et al.*, 2013).

### 2.1 Mecanismos evolutivos na origem das hemoglobinas

Como explicado acima, as hemoglobinas de agnatos e gnatostomados evoluíram de forma independente durante o Período Cambriano (430 a 540 Ma) a partir de genes parálogos distintos. No entanto, ambas convergiram num padrão de expressão específico em eritrócitos (ROHLFING *et al.*, 2016), o que nos permite questionar se a expressão específica em eritrócitos evoluiu duas vezes independentemente. Usando cuidadosas análises bioquímicas e genéticas para esclarecer os dados da genômica comparativa, Miyata *et al.*

(2020) mostraram que o ancestral comum dos vertebrados veio a possuir um gene *NPRL3* contendo um forte elemento regulatório ligado a pelo menos dois diferentes genes de globina. Um dos genes de globina era similar à *Cygb* e foi cooptado para o transporte de  $O_2$  nos agnatos enquanto que o outro gene, mais similar à *Hba* e *Hbb* dos gnatostomados, foi cooptado para o transporte de  $O_2$  nos gnatostomados. Dessa forma, as Hbs de gnatostomados e de agnatos evoluíram a partir de genes diferentes, mas convergiram no mesmo padrão de expressão pelo fato de terem herdado o mesmo elemento regulatório localizado à montante dos genes de globina presentes no ancestral comum (HARDISON, 2020; MIYATA *et al.*, 2020).

As Hbs de agnatos e gnatostomados funcionam bem no transporte de  $O_2$  através do sistema circulatório porque possuem a capacidade de ligar-se cooperativamente ao  $O_2$ . Porém, as Hbs de agnatos e gnatostomados evoluíram o efeito cooperativo através de mecanismos estruturais distintos. No caso dos agnatos, a Hb é monomérica quando oxigenada, mas forma dímeros ou tetrâmeros quando é desoxigenada. A Hb dos gnatostomados é um tetrâmero  $\alpha_2\beta_2$  que é capaz de sofrer mudanças conformacionais durante a oxigenação ou desoxigenação. Em ambos os casos, os mecanismos estruturais são capazes de mudar a afinidade por  $O_2$  e produzir em maior ou menor grau um efeito cooperativo (INGERMANN, 2011; PERUTZ, 1978; ROYER, 1992; WEBER; FAGO, 2004).

Recentemente, Pillai *et al.* (2020) usaram uma abordagem de bioquímica evolutiva, consistindo na reconstrução e ressuscitação dos polipeptídeos ancestrais da Hb dos gnatostomados e caracterização de suas propriedades estruturais e funcionais, a fim de elucidar o processo evolutivo que originou o tetrâmero cooperativo a partir de ancestrais monoméricos e não cooperativos. Em resumo, após o evento de duplicação que deu origem à Mb e ao ancestral monomérico da Hb, surgiu uma interface na Hb monomérica ancestral que permitiu a formação de homodímeros. A duplicação da Hb homodimérica ancestral gerou os ancestrais das globinas  $\alpha$  e  $\beta$ , que eram capazes de formar heterodímeros. Uma nova interface surgiu na globina  $\beta$ , capaz de unir os heterodímeros e produzir o heterotetrâmero. O estudo demonstrou que estruturas e funções moleculares complexas podem evoluir por meio de mecanismos genéticos simples que recrutam caracteres biofísicos existentes para formar arquiteturas de ordem mais elevada, consistindo na primeira descrição do processo evolutivo de uma proteína multimérica (BERENBRINK, 2020; PILLAI *et al.*, 2020).

## **2.2 Tempo e modo na diversificação das globinas $\alpha$ e $\beta$ dos tetrápodes**

Os vertebrados frequentemente possuem grupamentos com múltiplos parálogos de *Hba* e *Hbb* que são subfuncionalizados de forma a regular a expressão de diferentes isoformas de Hb durante a ontogenia (ALEV *et al.*, 2009; DAMSGAARD *et al.*, 2013; FORGET; HARDISON, 2009; GRISPO *et al.*, 2012; JOHNSON *et al.*, 2006; JOHNSON; GUMUCIO; GOODMAN, 2002; OPAZO *et al.*, 2015b; RUNCK; MORIYAMA; STORZ, 2009; STORZ *et al.*, 2015; STORZ, 2016; STORZ *et al.*, 2011). Durante a evolução, a ligação dos grupamentos de *Hba* e *Hbb* observada em peixes (CHAN *et al.*, 1997; GILLEMANS *et al.*, 2003; MARINO *et al.*, 2007; OPAZO *et al.*, 2013) foi conservada na linhagem dos anfíbios (HOSBACH; WYLER; WEBER, 1983; JEFFREYS *et al.*, 1980; PATIENT *et al.*, 1980), enquanto que na linhagem dos amniotas um gene de *Hbb* foi transposto para outro cromossomo, produzindo um grupamento separado de parálogos de *Hbb* (HARDISON, 2012; JEFFREYS *et al.*, 1980; OPAZO; HOFFMANN; STORZ, 2008a; PATEL *et al.*, 2008, 2010; PATEL; DEAKIN, 2010; WHEELER *et al.*, 2001, 2004). O tempo e o modo na evolução da regulação ontogenética dos grupamentos de *Hba* e *Hbb* são diferentes. Enquanto a divisão de trabalho entre genes ontogeneticamente expressos cedo e tarde no grupamento de *Hba* aconteceu no ancestral comum dos tetrápodes, a regulação da expressão de *Hbb* durante o desenvolvimento evoluiu independentemente nas diferentes linhagens de tetrápodes que herdaram uma cópia do mesmo tipo de *Hbb* (GAUDRY *et al.*, 2014; HOFFMANN *et al.*, 2010, 2018; HOFFMANN; OPAZO; STORZ, 2008a, 2008b; OPAZO; HOFFMANN; STORZ, 2008a, 2008b; ZHANG *et al.*, 2014). O arranjo ancestral do grupamento de *Hba* pode ser representado da seguinte maneira: 5'- $\alpha^E$ - $\alpha^D$ - $\alpha^A$ -3'. Duas hipóteses alternativas tentam explicar as origens dos parálogos de *Hba*. De acordo com uma das hipóteses, a globina  $\alpha^E$  e o gene progenitor das globinas  $\alpha^A/\alpha^D$  se originaram na linhagem basal dos tetrápodes via duplicação em tandem, seguida de outra duplicação em tandem do gene proto- $\alpha^{A/D}$  na linhagem basal dos amniotas, produzindo a topologia ( $\alpha^E$  ( $\alpha^D$ ,  $\alpha^A$ )). De acordo com a hipótese alternativa e também mais aceita, duas duplicações em tandem sucessivas do gene proto-*Hba* na linhagem basal dos tetrápodes deram origem, primeiro, ao gene de globina  $\alpha^A$  e ao gene proto- $\alpha^{E/D}$ , e em seguida às globinas  $\alpha^E$  e  $\alpha^D$  a partir do gene proto- $\alpha^{E/D}$ , produzindo a topologia (( $\alpha^E$ ,  $\alpha^D$ )  $\alpha^A$ ) (COOPER *et al.*, 2006; HOFFMANN; STORZ, 2007).

### 2.3 Anfíbios: lacunas no conhecimento das globinas de vertebrados

Fuchs, Burmester e Hankeln (2006) analisaram a versão inicial do genoma de *Xenopus tropicalis* (Anura: Pipidae) e descobriram a *GbY*. Eles também notaram a ausência de *Mb* em *Xenopus*. Desde então, o cenário onde a *Mb* e a *GbE* foram deletadas na linhagem dos anfíbios anuros é a hipótese mais aceita (HOFFMANN; OPAZO; STORZ, 2011). A expressão de *Cygb* no músculo esquelético e cardíaco e também nos fotorreceptores de segmentos internos da retina sugere que a *Cygb* pode compensar pela falta de *Mb* e *GbE* em *Xenopus* (XI *et al.*, 2007). Não obstante, uma hemoglobina monomérica foi purificada a partir de extratos do coração de *Rana catesbeiana* (Anura: Neobatrachia) e sua sequência de aminoácidos apresentou similaridade a cadeias  $\alpha$  de Hb expressas em estágios larvais de anfíbios (MAEDA; FITCH, 1982). A partir dos estudos citados e do conhecimento sobre a evolução dos genes de globinas em vertebrados, podemos notar que o repertório de globinas de anfíbios ainda é pouco compreendido. A primeira limitação está no fato de *Xenopus* fazer parte do grupo dos Pipidae, que por ser pouco numeroso em espécies e por ter divergido cedo na filogenia dos anuros, é pouco representativo dos anfíbios modernos. Além disso, o repertório de globinas de anfíbios precisa também levar em conta a distribuição filética das globinas no grupo das salamandras (Caudata) e dos caecilianos (Gymnophiona). Ademais, o fato de anfíbios serem um grupo irmão dos amniotas é algo importante para compreendermos em profundidade a evolução dos genes de *Hba* e *Hbb*, pois uma amostragem apropriada de anfíbios adicionaria informação filogenética em regiões da filogenia das globinas ainda não totalmente esclarecidas. Por exemplo, apesar da hipótese mais aceita para a origem da globina  $\alpha^D$  sugerir que ela surgiu antes da divergência dos anfíbios, devemos notar que nenhum gene de globina  $\alpha^D$  foi ainda caracterizado em anfíbios. Por outro lado, conclusões sobre a origem dos genes de *Hbb* a partir de um único gene ancestral em anfíbios, da mesma forma que ocorreu em diferentes linhagens de amniotas, são fortemente limitadas pelo fato de que os estudos anteriores não utilizaram uma amostragem taxonômica apropriada de anfíbios. E devemos lembrar que após quase 40 anos da caracterização da hemoglobina monomérica no coração da *Rana catesbeiana*, ainda não temos nenhuma descrição do gene que codifica essa proteína, nem sabemos a sua distribuição filética em anfíbios ou sua relação evolutiva com outros genes de globina.



### **3 OBJETIVOS**

#### **3.1 Objetivo geral**

Determinar o repertório de globinas nas três ordens de anfíbios modernos: Anura, Caudata e Gymnophiona.

#### **3.2 Objetivos específicos**

- Anotar os genes de globinas em genomas e transcriptomas de representantes das três ordens de anfíbios modernos (Anura, Caudata e Gymnophiona).
- Reconstruir a filogenia dos genes de globina através da região codificante.
- Determinar as relações entre os genes de globinas (ortologia e paralogia) contrastando a árvore dos genes com a árvore das espécies.
- Analisar a sintenia da região genômica dos genes de globinas e comparar com a sintenia observada em peixes e amniotas.
- Identificar linhagens e genes que estejam sofrendo rápida evolução.

## 4 MATERIAIS E MÉTODOS

### 4.1 Anotação e curadoria manual dos genes de globinas

Nossa abordagem para a anotação de genes de globinas em genomas de anfíbios usou evidências de homologia fornecidas por alinhamentos de um conjunto representativo de sequências de aminoácidos de globinas conhecidas contra genomas e transcriptomas de anfíbios, buscando prever os genes através de regiões com suficiente similaridade e identidade para posterior inspeção e anotação, conforme as boas práticas de anotação de genomas de eucariotos (YANDELL; ENCE, 2012). Realizamos a curadoria manual dos genes usando o conhecimento disponível na literatura sobre a estrutura dos genes de globinas, como a quantidade de éxons e íntrons relatados para cada tipo de gene de globina, assim como elementos indispensáveis da estrutura proteica das globinas, tais como os resíduos Phe CD1 e His F8. Adicionalmente, buscamos evidências da expressão dos genes nos transcriptomas ou em bancos de dados de ESTs de cada espécie analisada. Também verificamos a presença de domínios conservados (LU *et al.*, 2020). Usamos o software *Exonerate* para alinhar proteínas contra os genomas e transcriptomas (SLATER; BIRNEY, 2005) e polimos os genes com auxílio de browsers de genoma como UCSC e o software GBiB (<https://genome.ucsc.edu/index.html>) (HAEUSSLER *et al.*, 2015; ZWEIG *et al.*, 2008), ou o software *Artemis* (CARVER *et al.*, 2012). Obtivemos os genomas e transcriptomas a partir do NCBI (AGARWALA *et al.*, 2018) e outras fontes mencionadas abaixo, juntamente com as espécies com genomas e transcriptomas disponíveis até o momento.

Genomas de anfíbios disponíveis na base de dados do NCBI que analisamos incluem: *Xenopus tropicalis* (v9.1) (HELLSTEN *et al.*, 2010); o alotetraplóide *Xenopus laevis* (v2), no qual o conjunto de cromossomos é dividido em dois subgenomas, L e S, que correspondem aos cromossomos de cada espécie ancestral (SESSION *et al.*, 2016); *Rana catesbeiana* (RCv2.1) (HAMMOND *et al.*, 2017); *Pyxicephalus adspersus* (Pads\_1.0) (DENTON *et al.*, 2018); *Nanorana parkeri* (ASM93562v1) (SUN *et al.*, 2015); *Rhinella marina* (RM170330) (EDWARDS *et al.*, 2018); *Ambystoma mexicanum* (ASM291563v2) (NOWOSHILOW *et al.*, 2018; SMITH *et al.*, 2018); *Microcaecilia unicolor* (aMicUni1.1); *Rhinatrema bivittatum* (aRhiBiv1.1); *Geotrypetes seraphini* (aGeoSer1.1); *Leptobranchium leishanense* (ASM966780v1) (LI *et al.*, 2019a); *Limnodynastes dumerilii* (BGI\_Ldum\_1.0) (LI *et al.*, 2020); *Oophaga pumilio* (GCA\_009801035.1 UCB\_Opum\_1.0) (ROGERS *et al.*, 2014); *Rana temporaria* (UCB\_Rtem\_1.0); *Spea multiplicata* (usc\_Smult\_1.0), *Spea*

*bombifrons* (usc\_Sbombifrons\_0.1), *Scaphiopus holbrookii* (usc\_Sholbrookii\_0.1), *Scaphiopus couchii* (usc\_Scouchii\_0.1) (SEIDL *et al.*, 2019). O genoma de *Leptobrachium ailaonicum* foi obtido a partir do repositório GigaGD (LI *et al.*, 2019b).

Transcriptomas de anfíbios disponíveis na base de dados do NCBI que analisamos incluem: *Rana clamitans* (GAEG01.1), *Pseudacris regilla* (GAEI01.1) (ROBERTSON; CORNMAN, 2014); *Hynobius chinensis* (GAQK01.1) (CHE *et al.*, 2014); *Pelophylax nigromaculatus* (GBET01.1) (XU *et al.*, 2015); *Rana catesbeiana* (GDDO01.1, GFBS01.1, GFBS01.2, GFBS01.3) (BIROL *et al.*, 2015); *Bufo viridis* (GDRL01.1) (GERCHEN *et al.*, 2016); *Ambystoma mexicanum* (GEBK01.1, GFBM01.1, GFBM01.2, GFBM01.3, GFZP01.1) (BRYANT *et al.*, 2017; HALL; EISTHEN; WILLIAMS, 2016); *Microhyla fissipes* (GECV01.1) (ZHAO *et al.*, 2016); *Rhacophorus omeimontis* (GEGF01.1), *Rhacophorus dennysi* (GEGG01.1), *Polypedates megacephalus* (GEGH01.1), *Pelophylax nigromaculatus* (GEGI01.1), *Odorrana margaretae* (GEGJ01.1), *Leptobrachium boringii* (GEGK01.1), *Xenophrys sangzhiensis* (G EGL01.1) (HUANG *et al.*, 2016); *Hyla cinerea* (GENE01.1); *Tylotriton wenxianensis* (GESS01.1); *Ambystoma tigrinum* (GFLI01.1), *Ambystoma texanum* (GFLJ01.1), *Ambystoma laterale* (GFLO01.1) (MCELROY *et al.*, 2017); *Rhinella marina* (GFMT01.1) (RICHARDSON *et al.*, 2018); *Ambystoma maculatum* (GFMY01.1) (BURNS *et al.*, 2017); *Oreobates cruralis* (GFNJ01.1) (MONTERO-MENDIETA *et al.*, 2017); *Caecilia tentaculata* (GFOD01.1), *Microcaecilia dermatophaga* (GFOE01.1), *Typhlonectes compressicauda* (GFOH01.1) (TORRES-SÁNCHEZ *et al.*, 2018); *Odorrana tormota* (GGLB01.1) (SHU *et al.*, 2018); *Rana temporaria* (GGNS01.1) (MA *et al.*, 2018); *Mantella betsileo* (GGTL01.1) (CIVITELLO *et al.*, 2019); *Anaxyrus baxteri* (GGUQ01.1, GGUR01.1, GGUS01.1); *Pelobates cultripes* (GHBH01.1); *Scaphiopus couchii* (GHBO01.1); *Rhinella arenarum* (GHCG01.1); *Batrachuperus yenyuanensis* (GHDZ01.1) (XIONG *et al.*, 2019); *Taricha granulosa* (GHKF01.1) (GLASS; MELIN; VAMOSI, 2019); *Bolitoglossa valleculea* (GHME01.1) (ARENAS GÓMEZ *et al.*, 2020); *Odorrana tormota* (GICS01.1); *Salamandra salamandra* (GIKK01.1) (GOEDBLOED *et al.*, 2017); *Hypsiboas pugnax* (GINY01.1, GINY01.2, GINY01.3); *Bombina bombina* (HADQ01.1, HADU01.1), *Bombina variegata variegata* (HADR01.1, HADV01.1), *Bombina variegata scabra* (HADS01.1), *Bombina orientalis* (HADT01.1) (NÜRNBERGER *et al.*, 2016); *Pyxicephalus adspersus* (ICLD02.1, ICPN01.1) (YOSHIDA; KAITO, 2019). O transcriptoma de *Rana pipiens* (versão 001) foi obtido em (<https://www.davislab.net/rana/>) (CHRISTENSON *et al.*, 2014).

## 4.2 Reconstrução filogenética

Fizemos alinhamentos múltiplos das regiões codificantes de três conjuntos de genes de globinas: (1) todas as globinas de anfíbios exceto *Adgb* devido ao seu domínio globina permutado; (2) *Hba* de anfíbios; (3) *Hbb* de anfíbios. Em cada conjunto de dados incluímos uma amostragem taxonômica adequada de representantes de outros grupos de vertebrados para comparação. Utilizamos diferentes softwares para alinhamento de sequências, como MUSCLE (EDGAR, 2004), ClustalW (LARKIN *et al.*, 2007), os modelos G-INS-1, E-INS- i, FFT-NS-i, G-INS-i, e L-INS-i do MAFFT disponíveis no servidor web (<https://mafft.cbrc.jp/alignment/server/>) (KATOHI; ROZEWICKI; YAMADA, 2019) e o M-coffee (<http://tcoffee.crg.cat/apps/tcoffee/do:mcoffee>) (MORETTI *et al.*, 2007; WALLACE *et al.*, 2006). Inspecionamos manualmente os alinhamentos e depois utilizamos o software MUMSA (LASSMANN; SONNHAMMER, 2005) disponível no servidor web (<https://msa.sbc.su.se/cgi-bin/msa.cgi>) (LASSMANN; SONNHAMMER, 2006) para verificar e comparar automaticamente a qualidade dos diferentes alinhamentos múltiplos de sequências obtidos. Em todos os casos, o modelo G-INS-i do MAFFT produziu os melhores alinhamentos de acordo com a pontuação do MUMSA. Usamos os alinhamentos múltiplos de sequências selecionados na etapa anterior para estimar por máxima verossimilhança os melhores modelos de substituição de nucleotídeos de acordo com as menores pontuações no *Bayesian Information Criterion* (BIC), conforme implementado em softwares como MEGA X (KUMAR *et al.*, 2018) ou o ModelFinder do servidor IQ-Tree (KALYAANAMOORTHY *et al.*, 2017). Realizamos a reconstrução filogenética usando máxima verossimilhança pelo software IQ-Tree (versão 1.6.12) a partir do servidor web (<http://iqtree.cibiv.univie.ac.at/>) (NGUYEN *et al.*, 2015; TRIFINOPOULOS *et al.*, 2016), calculando os valores de *bootstrap* pelo procedimento ultrarrápido do IQ-Tree com 1000 pseudoreplicas (HOANG *et al.*, 2018). Também reconstruímos a filogenia das globinas usando o método bayesiano pelo software MrBayes (versão 3.2.7) (RONQUIST *et al.*, 2012), configurando duas corridas independentes com quatro cadeias simultâneas durante  $2 \times 10^7$  gerações, coletando amostras a cada 1000 gerações. *Average standard deviation of split frequencies* foi inferior a 0,01 antes das corridas convergirem. Avaliamos a convergência pelo *potential scale reduction factor* estimado, que foi igual ou próximo a 1,0. Estimamos as probabilidades posteriores a partir das últimas 5000 árvores usando *majority rule of consensus*.

Inferimos os eventos de duplicação e especiação na árvore dos genes com auxílio de uma filogenia conhecida das espécies de anfíbios e de outros grupos de vertebrados incluídos para comparação. Usamos algoritmos do pacote *forester* para adicionar informação

taxonômica à árvore dos genes no formato phyloXML e então mapear os eventos de duplicação e especiação (HAN; ZMASEK, 2009; ZMASEK; EDDY, 2001, 2002). Como o algoritmo requer uma árvore das espécies que seja confiável a fim de contrastar com a árvore dos genes, usamos a árvore consenso das famílias de anfíbios disponível no website AmphibiaWeb (<https://amphibiaweb.org>), baseada numa literatura recente (BLACKBURN; WAKE, 2011; BROWN *et al.*, 2015; FENG *et al.*, 2017; HEINICKE *et al.*, 2018; JETZ; PYRON, 2018; KAMEI *et al.*, 2012; PELOSO *et al.*, 2016; SAN MAURO *et al.*, 2014; STREICHER *et al.*, 2018; TU *et al.*, 2018; VIEITES *et al.*, 2011; YAN *et al.*, 2016; YUAN *et al.*, 2019; ZHANG; WAKE, 2009), para reconstruir a filogenia das espécies usadas no estudo. Primeiro, usamos os dados e as ferramentas fornecidos pelo website VertLife.org (<http://vertlife.org/>) (JETZ; PYRON, 2018) para reconstruir a filogenia das espécies selecionadas em nosso estudo, e em seguida corrigimos a filogenia das espécies de acordo com a árvore consenso das famílias de anfíbios fornecida pelos especialistas do AmphibiaWeb. Para visualizar os filogramas e cladogramas, usamos os pacotes *ape* (versão 5.4-1) (PARADIS; SCHLIEP, 2019) e *ggtree* (versão 2.2.4) (YU, 2020; YU *et al.*, 2017) disponíveis na linguagem R (R CORE TEAM, 2020). Cada etapa da reconstrução filogenética é descrita em detalhes em outro lugar (KAPLI; YANG; TELFORD, 2020). Avaliamos o suporte para os ramos de acordo com os seguintes critérios. Porcentagens bootstrap (b%) entre 90 e 100% foram consideradas como fortes, 70 e 89% moderadas, 50 e 69% fracas, e valores abaixo de 50% como ausência de suporte. Consideramos probabilidades posteriores bayesianas (ppb) iguais ou acima de 0.95 como confiáveis e abaixo desse valor como não confiáveis.

### 4.3 Análise de sintenia conservada

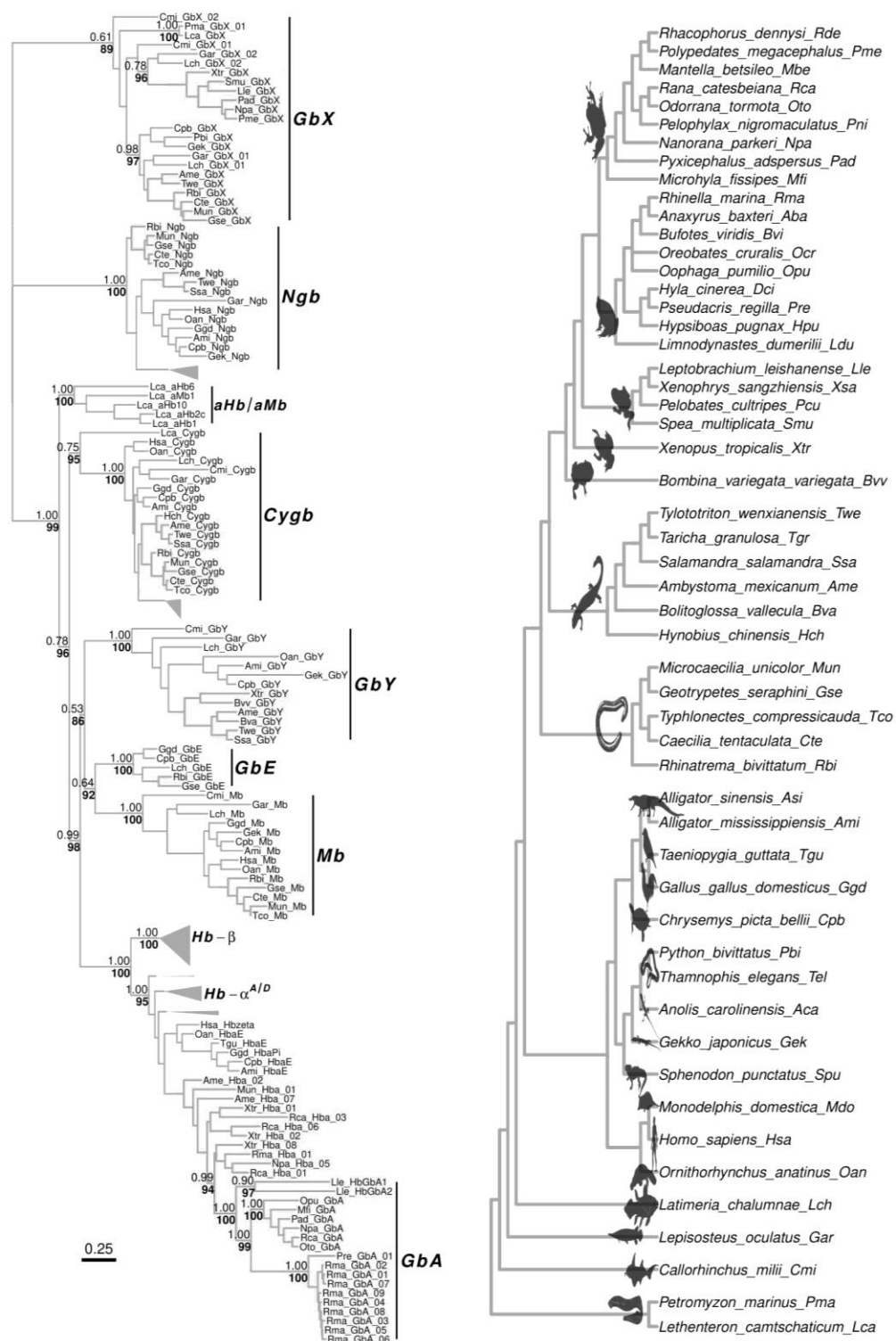
Usamos o browser Genome Data Viewer do NCBI (AGARWALA *et al.*, 2018) para mapear a posição dos genes de globinas e outros genes sintênicos nos cromossomos ou *scaffolds* de cada espécie de anfíbio analisada que possuía anotação de referência disponível. Analisando a sintenia conservada buscamos revelar inversões, duplicações em pequena ou em grande escala, e transposições de genes na região genômica dos genes de globinas, de forma a complementar e em alguns casos esclarecer as inferências filogenéticas. Utilizamos o software IBS 1.0.3 (<http://ibs.biocuckoo.org/>) (LIU *et al.*, 2015) para desenhar modelos de nucleotídeos das regiões genômicas dos genes de globinas.

## **5 RESULTADOS**

### **5.1 O repertório de globinas nos anfíbios**

Nossa análise filogenética recuperou todos os clados de globinas de vertebrados, como mostrado na Figura 1.

Figura 1 – Filograma de máxima verossimilhança das globinas de anfíbios

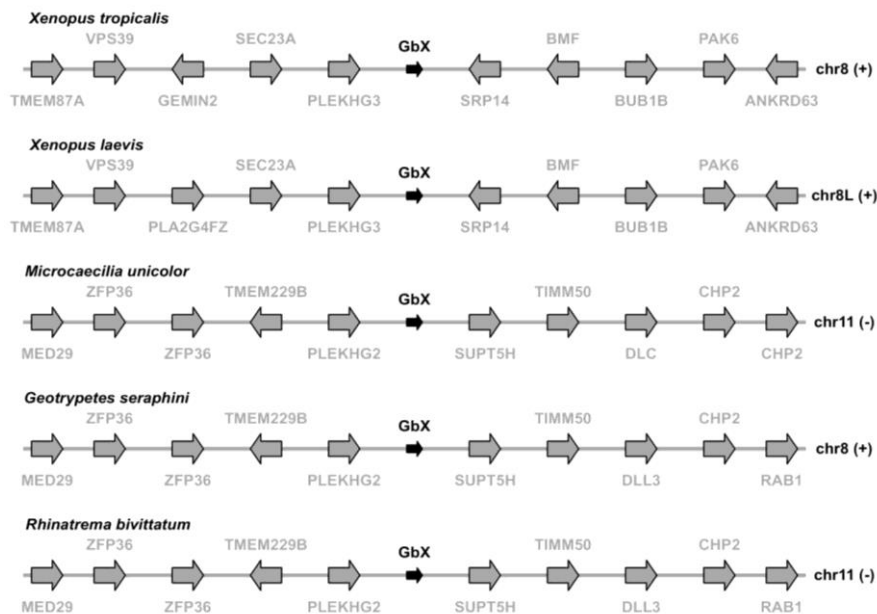


Fonte: elaborado pelo autor.

A árvore das globinas foi reconstruída por máxima verossimilhança no software IQ-Tree usando o modelo de substituição de nucleotídeos GTR + F + I +  $\Gamma_4$ . Os números em negrito abaixo dos nós são os valores de porcentagem do *bootstrap*. Os números acima dos nós são os valores de probabilidades posteriores bayesianas. A análise bayesiana foi realizada usando o modelo de substituição de nucleotídeos GTR +  $\Gamma$  + I. As sequências de *GbX* e *Ngb* foram usadas como *outgroup*. Retângulos em cinza destacam os genes de anfíbios. Alguns ramos foram colapsados (triângulos em cinza) para facilitar a visualização. Entre eles incluímos os clados *Ngb* e *Cygb* de anuros. O cladograma à direita representa a filogenia das espécies esperada.

Por terem surgido antes da separação entre protostômios e deuterostômios, *GbX* e *Ngb* foram usadas como *outgroup*. Nós usamos um conjunto representativo de parálogos de *GbX* de outros vertebrados para verificar quais estão presentes em anfíbios. Todos os parálogos de *GbX* formaram um clado com suporte moderado, embora a probabilidade posterior seja baixa ( $b\% = 89$ ,  $ppb = 0,61$ ). O clado de *GbX* recuperou os quatro clados de parálogos de *GbX*, em concordância com Opazo *et al.* (2015a). Mas observamos que os genes das diferentes ordens de anfíbios foram agrupados em diferentes clados. *GbX* de caecilianos e salamandras fazem parte do clado contendo *GbX1* do Gar manchado (*Lepisosteus oculatus*) e do celacanto (*Latimeria chalumnae*), assim como a *GbX* de amniotas. Por outro lado, genes de *GbX* de anuros fazem parte do clado contendo *GbX2* do Gar manchado e do celacanto. A sintenia da região de genômica da *GbX* (Figura 2) confirma os resultados da análise filogenética, onde podemos notar que os ortólogos de *GbX1* de caecilianos são flanqueados por *TIMM50* e *SUPT5H*, enquanto que os ortólogos de *GbX2* em anuros são flanqueados por *PLEKHG3* e *SRP14*, em concordância com Opazo *et al.* (2015a). Não analisamos a sintenia da região de *GbX* de salamandras, pois nenhum genoma de referência para esse grupo estava disponível no browser *Genome Data Viewer* do NCBI.

Figura 2 – Sintenia da região genômica da *GbX* de anfíbios

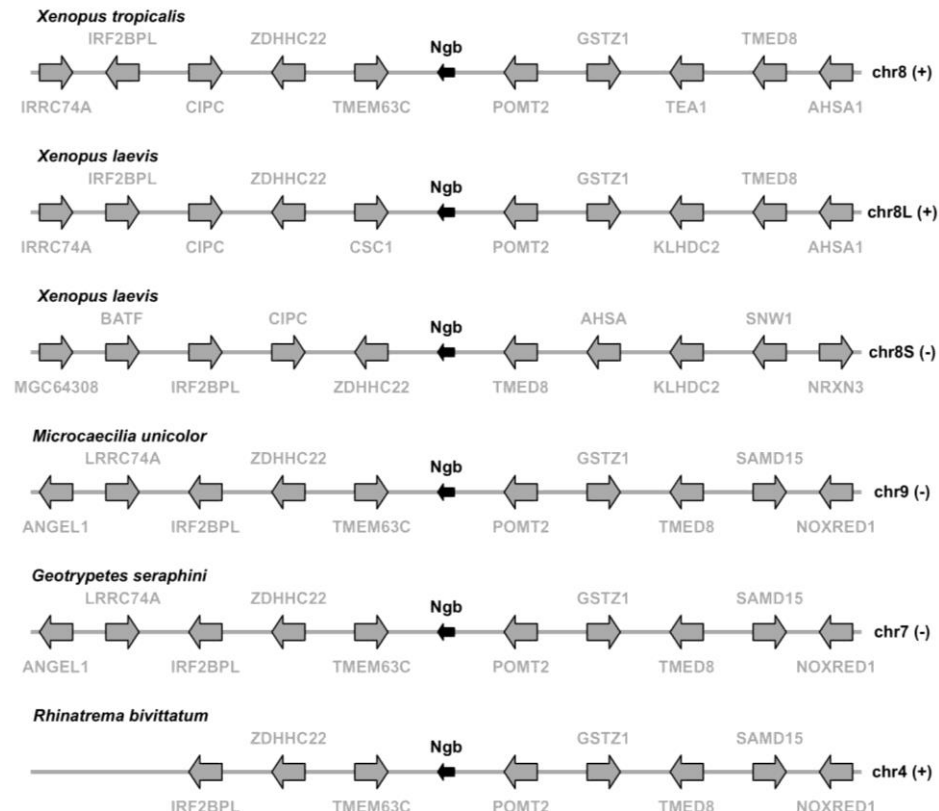


Fonte: elaborado pelo autor.

Todos os genes de *Ngb* formaram um clado com suporte forte ( $b\%/ppb = 100/1,00$ ). Apesar do clado de *Ngb* desviar da filogenia dos organismos esperada, a sintenia é conservada em caecilianos e anuros (*TMEM63C-Ngb-POMT2*), como podemos observar na Figura 3, sugerindo que a árvore dos genes desviou da árvore das espécies devido a alguma faceta do processo de evolução que iremos discutir mais adiante.



Figura 3 – Sintenia da região genômica da *Ngb* de anfíbios



Fonte: elaborado pelo autor.

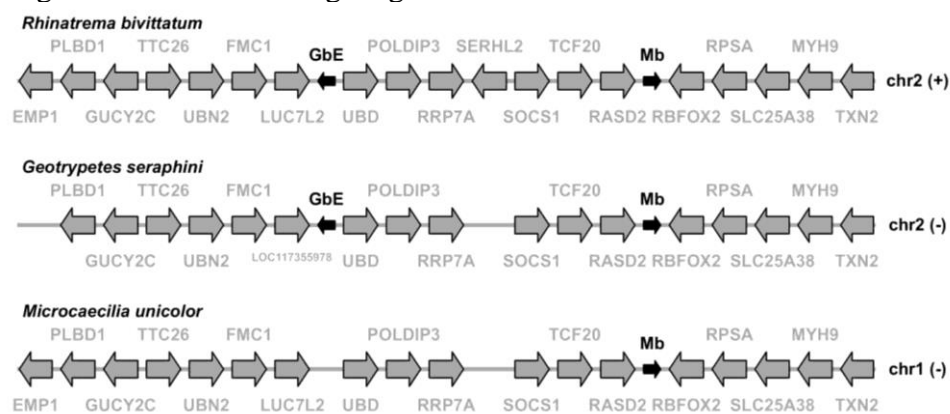
As globinas específicas de vertebrados formaram um clado com suporte forte ( $b\%/ppb = 99/1,00$ ). Divergindo primeiro, o clado formado pelos genes *aHb* e *aMb* dos ciclóstomos teve suporte forte ( $b\%/ppb = 100/1,00$ ). A topologia sugere que ele é um clado irmão de todas as outras globinas específicas de vertebrados, mas não estamos certos dessa relação porque o clado de globinas específicas de vertebrados menos *aHb* e *aMb* não teve probabilidade posterior confiável ( $b\%/ppb = 96/0,78$ ). A *Cygb* do agnato *Lethenteron camtschaticum* foi colocada como ortóloga da *Cygb* dos gnatostomados com suporte forte, mas probabilidade posterior não confiável ( $b\%/ppb = 95/0,75$ ). A *Cygb* dos gnatostomados formou um clado bem suportado, mas com desvios da filogenia dos organismos esperada, em especial a posição dos genes de *Cygb* de mamíferos. Pensamos inicialmente que tal discordância ocorreu devido à presença de um quarto éxon que se originou na linhagem dos mamíferos e que não é homólogo a qualquer parte de genes de *Cygb* em não mamíferos (BURMESTER *et al.*, 2002). Porém, a topologia permaneceu a mesma quando repetimos as análises depois de remover o quarto éxon dos genes de *Cygb* de mamíferos.

Os cladogramas *GbE*, *Mb*, *GbY* e (*Hba*, *Hbb*) foram suportados fortemente ( $b\%/ppb = 100/1,00$ ). Todos foram colocados juntos em um clado de globinas específicas de gnatostomados que é irmão do clado *Cygb* com suporte forte ( $b\%/ppb = 98/0,99$ ).

Observamos que *Mb* e *GbE* estão presentes em caecilianos, mas não *GbY*, e o clado (*Mb*, *GbE*) é colocado como irmão do clado *GbY* com forte suporte, embora a probabilidade posterior não seja confiável ( $b\%/ppb = 92/0,64$ ). Cada clado exibiu algumas discordâncias da filogenia dos organismos esperada. Não obstante, o clado (*GbY* (*GbE*, *Mb*)) é colocado como sendo irmão do clado (*Hba*, *Hbb*) com suporte moderado e probabilidade posterior não confiável ( $b\%/ppb = 86/0,53$ ).

Nós verificamos a sintonia da região genômica da *Mb* e *GbE* em caecilianos (Figura 4). Todos os dados de caecilianos analisados, incluindo os transcriptomas, mostram a presença de pelo menos uma cópia de *Mb*. Por outro lado, *GbE* parece ser restrita a *Geotrypetes seraphini* e *Rhinatrema bivittatum*, sendo deletada em *Microcaecilia unicolor*. Nenhum dos transcriptomas apresentou *GbE*. Curiosamente, existem duas cópias de *Mb* distintas no transcriptoma de *Microcaecilia dermatophaga*.

Figura 4 – Sintonia da região genômica da *Mb* e *GbE* de caecilianos



Fonte: elaborado pelo autor.

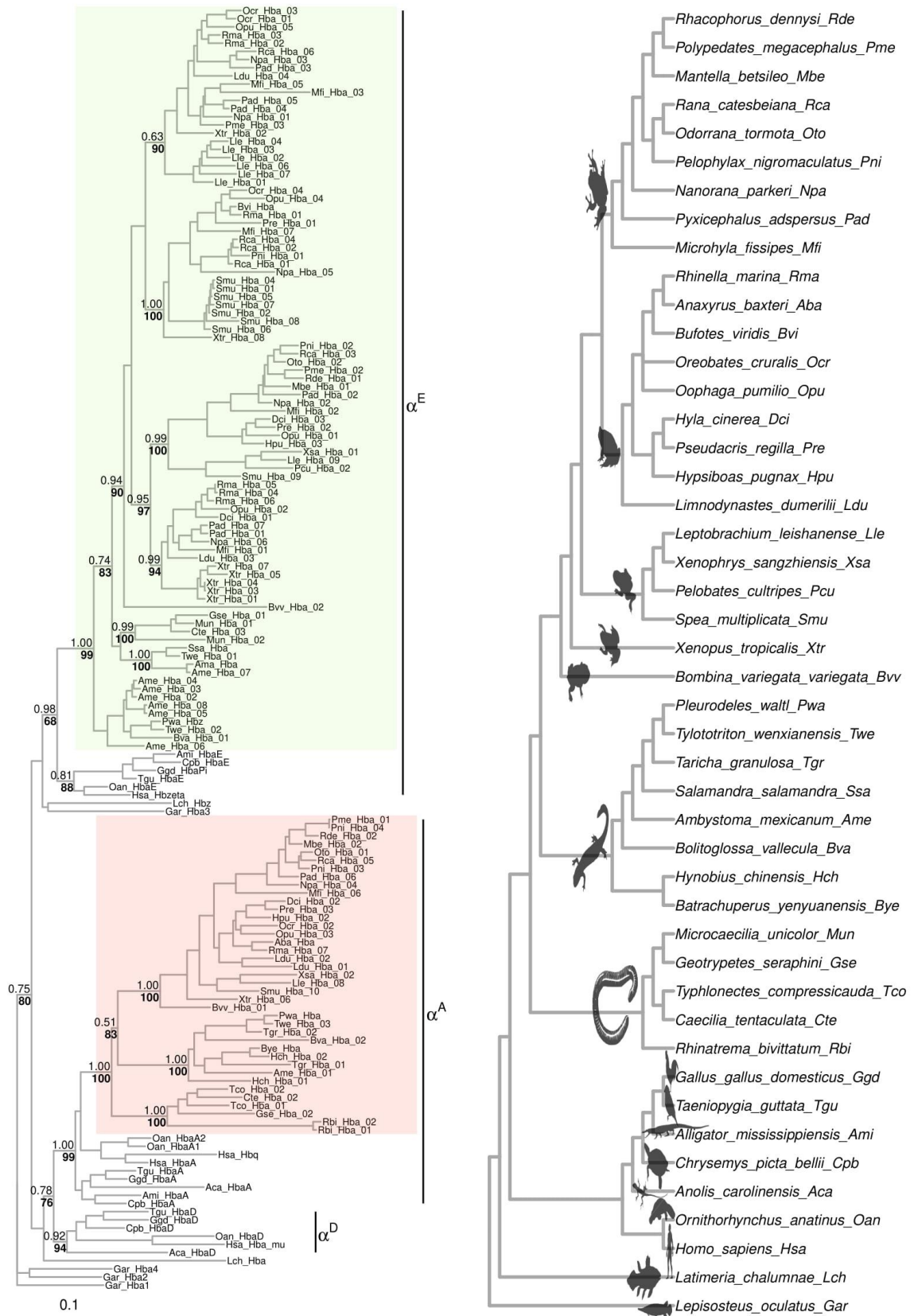
Pelo fato de que realizamos análises filogenéticas de conjuntos de genes de *Hba* e *Hbb* separadamente, incluímos apenas uma fração desses genes na Figura 1. Então, nos concentraremos por enquanto apenas no clado *GbA*. Esse clado inclui todos os genes que codificam proteínas similares à hemoglobina monomérica que foi encontrada no coração de *Rana catesbeiana* (MAEDA; FITCH, 1982). Encontramos vários genes do tipo *GbA* em espécies do grupo Neobatrachia, comumente na forma de várias cópias em indivíduos da família Bufonidae. Também encontramos dois genes do tipo *GbA* em *Leptobranchium leishanense*, membro do grupo Pelobatoidea, um grupo irmão dos Neobatrachia. Todos os genes de *GbA* formam um clado fortemente suportado ( $b\%/ppb = 100/1,00$ ) e concordam totalmente com a filogenia das espécies quando levamos em conta os eventos de duplicação nos Bufonidae. Além disso, o clado *GbA* é relacionado às globinas do tipo  $\alpha^E$  com forte suporte ( $b\%/ppb = 94/0,99$ ). Genes de *GbA* foram localizados no mesmo cromossomo que a

*Adgb* em algumas espécies, mas não pudemos analisar a sintonia da região da *GbA*, exceto nos *scaffolds* do genoma de *Nanorana parkeri*. Observamos que a *GbA* é flanqueada por *ANKRD9*, *YPEL5* e *LBH* de um lado, e por *ALK* e *CLIP4* do outro lado. Todos os genes que flanqueiam *GbA* em *Nanorana parkeri* foram anotados nos genomas de caecilianos e, com exceção de *ANKRD9*, localizamos todos no mesmo cromossomo da *Adgb*.

## 5.2 Filogenia dos genes de globina dos tipos $\alpha$ e $\beta$

A filogenia dos genes de globinas do tipo  $\alpha$  (Figura 5) mostra que os anfíbios possuem genes do tipo  $\alpha^E$  ou  $\alpha^A$ , mas não o tipo  $\alpha^D$ . O clado de globinas  $\alpha^A$  dos tetrápodes recebeu suporte forte (b%/ppb = 99/1,00), mas notamos alguns desvios da filogenia das espécies nas sequências de amniotas. O clado  $\alpha^A$  dos anfíbios, assim como o clado  $\alpha^A$  de caecilianos, salamandras e anuros, todos receberam suporte forte (b%/ppb = 100/1,00). O clado de globinas  $\alpha^A$  dos Batrachia (Anura, Caudata) recebeu suporte forte, mas a probabilidade posterior não foi confiável (b%/ppb = 90/0,51). De modo geral, os genes recuperaram a filogenia das espécies e ocorreram poucos eventos de duplicação, os últimos restritos à linhagem das salamandras. O clado de globinas  $\alpha^E$  dos tetrápodes recebeu suporte fraco, embora a probabilidade posterior seja confiável (b%/ppb = 68/0,98). O ramo dos amniotas concordou com a filogenia das espécies em geral. Por outro lado, o clado de globinas  $\alpha^E$  dos anfíbios é marcado por muitos eventos de duplicação e recebeu suporte forte (b%/ppb = 99/1,00). Alguns dos ramos internos de globinas  $\alpha^E$  também receberam bom suporte, mas algumas divergências profundas não foram bem resolvidas.

Figura 5 – Filograma de máxima verossimilhança das globinas do tipo  $\alpha$



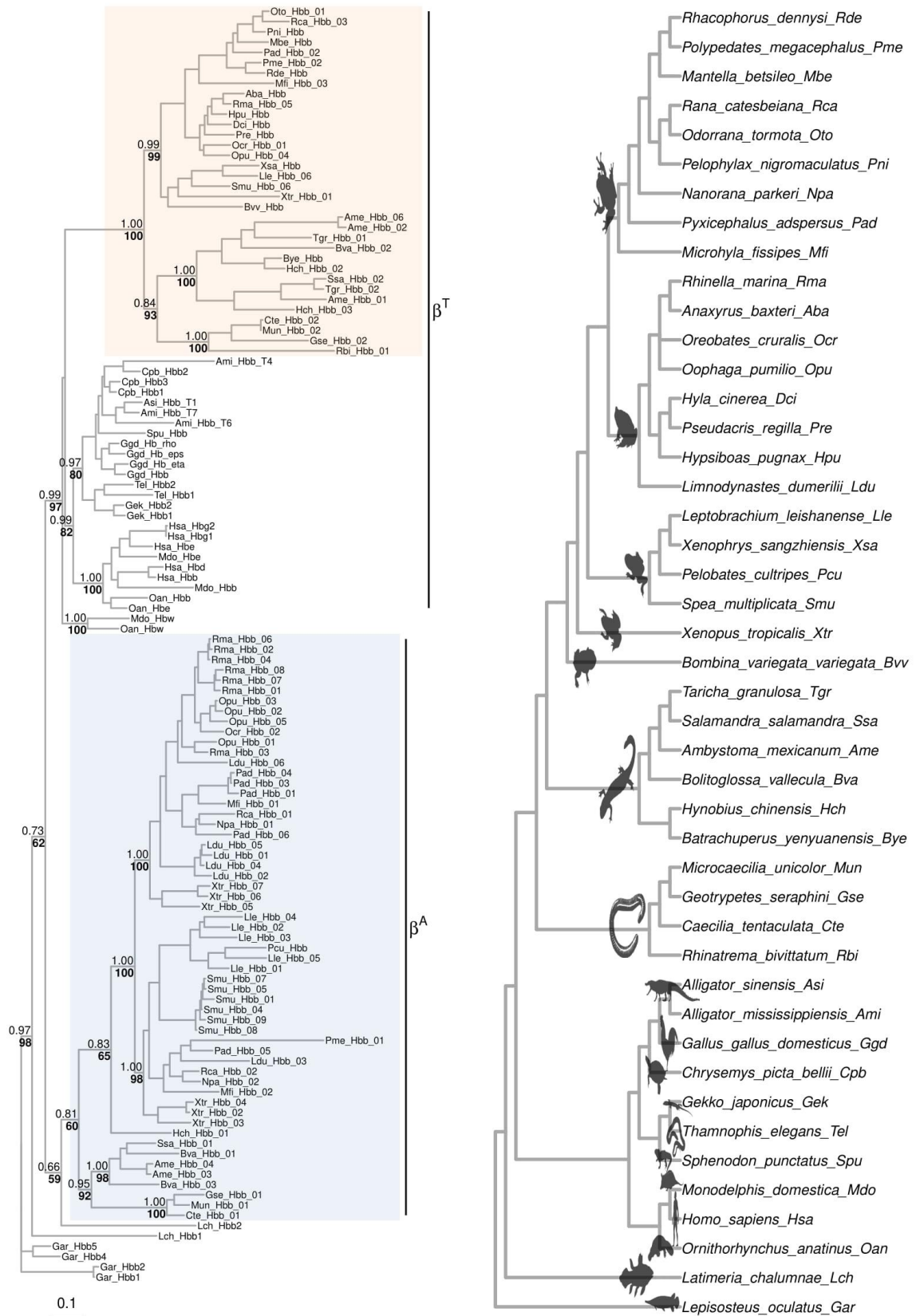
Fonte: elaborado pelo autor.

A árvore das globinas do tipo  $\alpha$  foi reconstruída por máxima verossimilhança no software IQ-Tree usando o modelo de substituição de nucleotídeos SYM + I +  $\Gamma_4$ . Os números em negrito abaixo dos nós são os valores

de porcentagem do *bootstrap*. Os números acima dos nós são os valores de probabilidades posteriores bayesianas. A análise bayesiana foi realizada usando o modelo de substituição de nucleotídeos HKY +  $\Gamma$  + I. A sequência de *Hba1* do gar manchado (*Lepisosteus oculatus*) foi usada como *outgroup*. Retângulos coloridos destacam os genes de anfíbios. O cladograma à direita indica a filogenia das espécies esperada.

As relações filogenéticas dos genes de *Hbb* de tetrápodes são complexas. Supreendentemente, os anfíbios possuem dois clados de globinas do tipo  $\beta$  (Figura 6). Um dos clados é irmão do clado de globinas do tipo  $\beta$  de amniotas, e por isso nomeamos o clado de  $\beta^T$ , referente a globinas  $\beta$  de tetrápodes. O outro clado foi específico de anfíbios e por isso denominados  $\beta^A$ , referente a globinas  $\beta$  de anfíbios. Apesar do clado  $\beta^T$  ter sido colocado como irmão das globinas  $\beta$  de amniotas, o relacionamento entre eles não foi suportado em ambas as análises, por máxima verossimilhança e pelo método bayesiano, à exclusão dos genes de globinas  $\beta$  órfãos encontrados no grupamento de *Hba* de marsupiais e monotremados, referidos aqui como *Hbw* de acordo com Aguileta, Bielawski e Yang (2006). Na análise bayesiana temos uma politomia ( $\beta^T$  de anfíbios,  $\beta^T$  de amniotas, *Hbw* de marsupiais e monotremados) com suporte forte (b%/ppb = 97/0,99). O clado  $\beta^A$  dos anfíbios foi colocado como sendo um clado irmão de um dos genes *Hbb* do celacanto, mas o suporte foi fraco (b%/ppb = 59/0,66). O clado  $\beta^T$  de anfíbios, assim como o clado  $\beta^T$  de caecilianos e salamandras, ambos receberam suporte forte (b%/ppb = 100/1,00). No entanto, os clados  $\beta^T$  de caecilianos e salamandras agruparam juntos, desviando da topologia esperada para o clado Batrachia com suporte forte, mas probabilidade posterior não confiável (b%/ppb = 93/0,84). O clado  $\beta^T$  de anuros recebeu suporte forte (b%/ppb = 99/0,99). O clado  $\beta^A$  de anfíbios junto com *Hbb2* do celacanto recebeu suporte fraco, mas alguns ramos internos foram bem suportados. É possível que adições futuras de outros taxa, como representantes dos peixes pulmonados, ajudem a resolver as relações filogenéticas desse novo clado de globinas  $\beta$ .

Figura 6 – Filograma de máxima verossimilhança das globinas do tipo  $\beta$



Fonte: elaborado pelo autor.

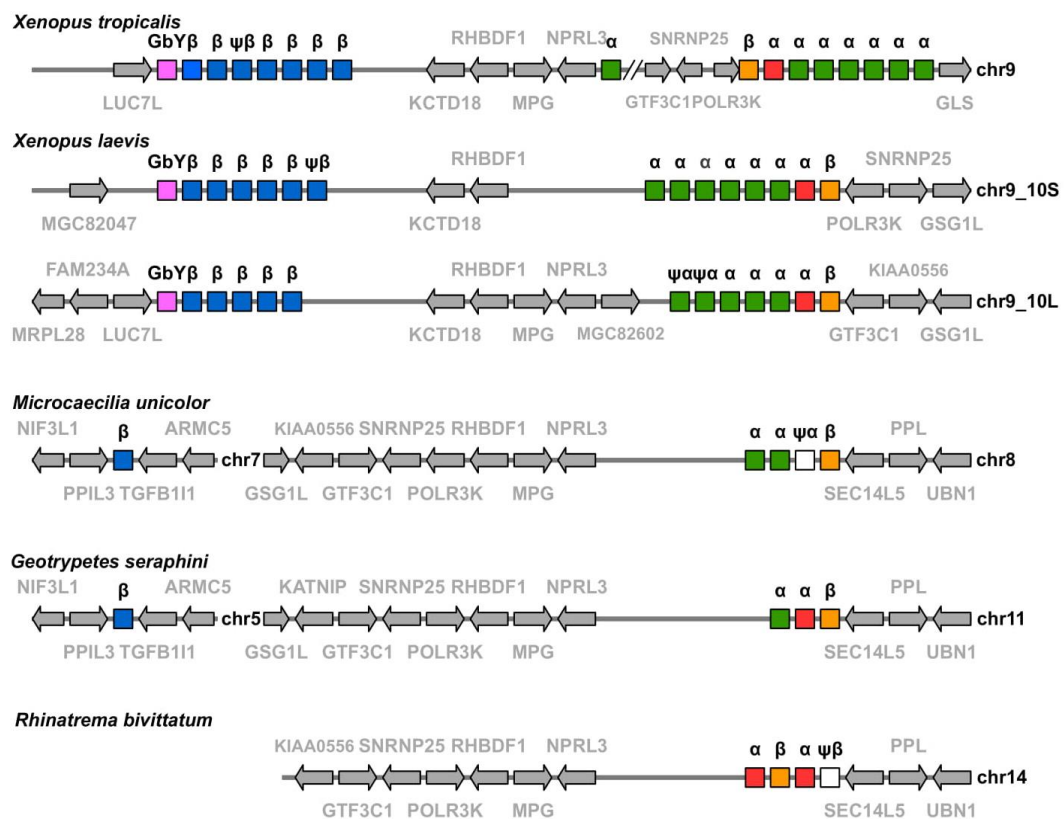
A árvore das globinas do tipo  $\beta$  foi reconstruída por máxima verossimilhança no software IQ-Tree usando o modelo de substituição de nucleotídeos TIM3e + I +  $\Gamma_4$ . Os números em negrito abaixo dos

nós são os valores de porcentagem do *bootstrap*. Os números acima dos nós são os valores de probabilidades posteriores bayesianas. A análise bayesiana foi realizada usando o modelo de substituição de nucleotídeos GTR +  $\Gamma$  + I. As sequências do gar manchado (*Lepisosteus oculatus*) foram usadas como *outgroup*. Retângulos coloridos destacam os genes de anfíbios. O cladograma à direita indica a filogenia das espécies esperada.

### 5.3 Sintenia dos genes de globina dos tipos $\alpha$ e $\beta$

A região genômica dos genes de *Hba* e *Hbb* de anfíbios é apresentada na Figura 7.

Figura 7 – Representação sem escala do grupamento de globinas  $\alpha$  e  $\beta$  de anfíbios



Fonte: elaborado pelo autor.

Representamos os genes de globina com quadrados coloridos:  $GbY$  (rosa);  $\alpha^E$  (verde);  $\alpha^A$  (vermelho);  $\beta^T$  (laranja);  $\beta^A$  (azul). Os genes que flanqueiam os genes de globina são representados por setas.  $\Psi$  denota os pseudogenes e quadrados em branco são pseudogenes cuja ortologia não conseguimos inferir. O símbolo // se refere a segmentos do genoma que foram cortados da representação por clareza.

Podemos observar que os genes de globina de *Xenopus* estão arranjados em dois grupamentos no mesmo cromossomo (o alotetraplóide *Xenopus laevis* possui dois cromossomos parálogos). Um dos grupamentos possui várias cópias de globinas  $\beta^A$  ligadas à  $GbY$  e flanqueadas por *LUC7L* e *KCTD18*. O segundo grupamento possui várias cópias de globinas  $\alpha^E$  ligadas a cópias únicas de globina  $\alpha^A$  e  $\beta^T$ , sendo flanqueado por *NPRL3*. Por que o grupamento de *Hba* é frequentemente flanqueado por *NPRL3* (e/ou *RHBDF1*) e *LUC7L* em tubarão e em amniotas (PHILIPSEN; HARDISON, 2018), podemos inferir que o primeiro

grupamento é produto de um rearranjo específico da linhagem dos anfíbios. Apesar de não ser mostrado na Figura 7, o mesmo arranjo ocorreu em outros genomas de anuros, como observamos nas montagens a nível de *scaffolds* de *Nanorana parkeri* e *Rhinella marina*, e a montagem a nível de cromossomos de *Pyxicephalus adspersus*. O rearranjo nos impediu de representar o grupamento de *Hba* e *Hbb* nos moldes do esquema convencional, onde o gene *NPRL3* é representado na extremidade 5' do grupamento. Caecilianos também apresentaram o grupamento  $\alpha^E$ - $\alpha^A$ - $\beta^T$  ligado ao *NPRL3*, mas a globina  $\beta^A$  foi encontrada como uma única cópia num cromossomo diferente onde é flanqueada por *PPIL3* e *TGFB1L1*. Isto se deve provavelmente a um evento de transposição específico na linhagem dos caecilianos. Contrário à diversificação das globinas  $\alpha^E$  e  $\beta^A$  em anuros e salamandras, os caecilianos apresentaram um número reduzido destes genes. De fato, *Rhinatrema bivittatum* parece ter perdido ambas as globinas,  $\alpha^E$  e  $\beta^A$ .



## 6 DISCUSSÃO

### 6.1 Discordância antiga na árvore dos genes

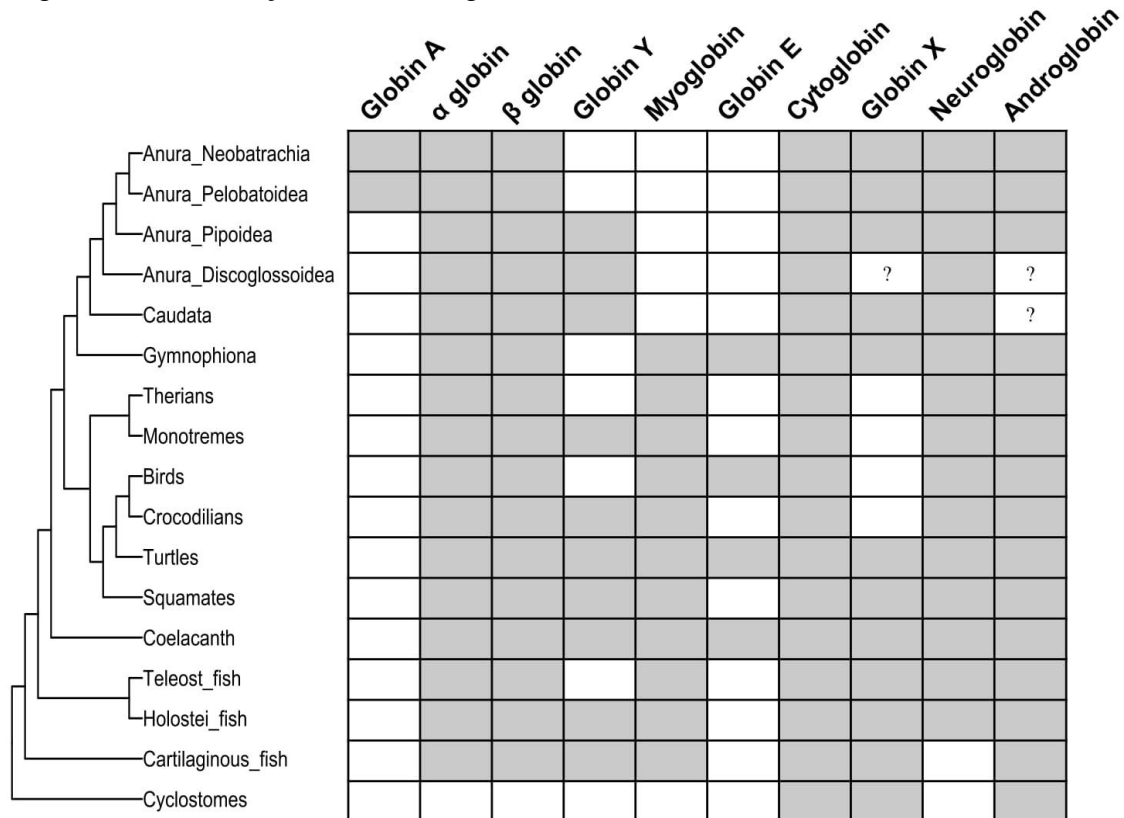
Em todos os clados de globinas que podemos observar conflitos protuberantes com a filogenia dos organismos, como nos clados *Ngb*, *Cygb*,  $\alpha^E$ ,  $\beta^A$  e  $\beta^T$ , os conflitos geralmente ocorrem nas relações mais profundas entre as ordens de anfíbios, ou entre anfíbios e amniotas. Quando olhamos para as relações dentro de cada ordem ou família de anfíbios, a árvore dos genes frequentemente concorda com a filogenia dos organismos esperada. Nós pensamos que tais discordâncias observadas nas partes interiores da árvore dos genes se devem provavelmente ao sorteio incompleto das linhagens (também conhecido como coalescência profunda), um fenômeno mais provável de ocorrer quando os ramos internos da árvore das espécies são curtos (KAPLI; YANG; TELFORD, 2020; MADDISON, 1997). De fato, Hime *et al.* (2020) realizaram uma análise filogenômica de anfíbios que revelou discordâncias antigas na árvore de genes que compõe a árvore da vida dos anfíbios. Apesar de suportarem fortemente a hipótese Batrachia (Gymnophiona (Anura, Caudata)), os autores observaram também que diferentes subconjuntos dos genomas suportavam hipóteses alternativas. A partir de 195 loci incluindo pelo menos um membro de cada ordem dos anfíbios, os autores encontraram 89 loci suportando a hipótese Batrachia, 51 suportando Procera (Anura, (Caudata, Gymnophiona)), e 55 suportando Acauda (Caudata (Anura, Gymnophiona)). Os autores sugerem que muitos loci são insuficientemente informativos para resolver os ramos mais profundos da filogenia dos anfíbios. Nós pensamos que esse é provavelmente o caso com o nosso conjunto de dados, pois globinas são proteínas pequenas que tipicamente possuem um tamanho de 200 aa ou menos, limitando o número de sítios informativos em potencial e aumentando as incertezas inerentes ao processo de estimativa filogenética, tais como ambiguidades nos alinhamentos de sequências, dúvidas sobre os modelos de substituição mais adequados, e o poder dos algoritmos usados para a reconstrução filogenética (STORZ; OPAZO; HOFFMANN, 2011).

### 6.2 Distribuição filética das globinas em vertebrados

Analisando genomas e transcriptomas de 55 espécies, compreendendo 26 das 110 famílias de anfíbios, nós relatamos pela primeira vez a distribuição filética das globinas nas três ordens de anfíbios modernos e contrastamos o repertório de anfíbios com o de outros

vertebrados para entender melhor os padrões de evolução de genes de globinas (Figura 8). Como nosso foco foi nos padrões de evolução de globinas em grande escala, incluímos apenas uma espécie representativa por gênero nas análises filogenéticas.

Figura 8 – Distribuição filética das globinas de vertebrados



Fonte: elaborado pelo autor.

A presença de cada tipo de gene de globina é sinalizada por retângulos preenchidos em cinza. Genes que podem não ter sido encontrados devido à cobertura dos genomas e transcriptomas são sinalizados por “?”. A ilustração é uma atualização de uma figura similar apresentada em Storz (2019) usando nosso conjunto de dados mais representativo de anfíbios.

Como podemos ver na Figura 8, *Adgb* é encontrada em todos os principais grupos de vertebrados, mas não encontramos genes putativos de *Adgb* em Caudata e Discoglossoidea (Anura). Devido ao grande tamanho da *Adgb* e ao seu padrão de expressão restrito (HOOGEWIJS *et al.*, 2012), podemos esperar que genomas e transcriptomas com pouca cobertura falhem em capturar a sequência desse gene de globina. A *Ngb* (monomérica) e a *Cygb* (homodimérica) são expressas principalmente em neurônios e fibroblastos, respectivamente. Enquanto *Cygb* é encontrada em todos os principais taxa de vertebrados, exceto peixes pulmonados, *Ngb* foi perdida na linhagem dos peixes pulmonados, peixes cartilagineos e lampreias (LÜDEMANN *et al.*, 2020; OPAZO *et al.*, 2015a; SCHWARZE *et al.*, 2014). Nós confirmamos a presença de *Ngb* e *Cygb* em todos os grupos de anfíbios. *GbX*, uma globina monomérica ligada à membrana e possuindo sítios de acilação no N-terminal, foi provavelmente perdida em mamíferos, aves e crocodilianos (BLANK; BURMESTER, 2012).

Nós encontramos *GbX* em todos os anfíbios, exceto Discoglossoidea, provavelmente devido à cobertura dos genomas e transcriptomas. Mas as diferentes linhagens de anfíbios retiveram parálogos diferentes de *GbX*. Outrossim, como o celacanto e as tartarugas possuem o repertório completo de globinas de vertebrados (SCHWARZE; BURMESTER, 2013; SCHWARZE; SINGH; BURMESTER, 2015), sabemos que a distribuição filética dos principais tipos de genes parálogos de globinas nas demais linhagens de vertebrados é ditada pelo ganho de genes via duplicação e pela perda de genes via deleção ou inativação (STORZ, 2019). Tendo em conta esse fato, *GbA* é uma nova adição à família gênica das globinas de vertebrados, estando presente nos grupos de anuros Pelobatoidea e Neobatrachia.

*GbY* foi perdida independentemente em peixes teleósteos, aves e mamíferos térios, e nós mostramos aqui que a *GbY* também foi perdida em Gymnophiona, Pelobatoidea e Neobatrachia, mas permaneceu em Caudata, Discoglossoidea e Pipoidea (Figura 8). Enquanto *GbE* tem uma distribuição filética restrita a celacanto, aves, tartarugas e peixes pulmonados (KUGELSTADT *et al.*, 2004; LÜDEMANN *et al.*, 2019; SCHWARZE; BURMESTER, 2013; SCHWARZE; SINGH; BURMESTER, 2015), *Mb* é encontrada em todos os gnatostomados, exceto seis espécies da família de peixes-gelo da Antártida, Channichthyidae (FERRERAS *et al.*, 2016; KOCH *et al.*, 2016; LÜDEMANN *et al.*, 2020; SIDELL; O'BRIEN, 2006). Por que *Mb* desempenha papéis importantes no transporte difusivo de O<sub>2</sub> nos tecidos de vertebrados, não possuir *Mb* pode ser maladaptativo. De fato, todas as 16 espécies atuais de peixes-gelo da Antártida também perderam a Hb. A maior [O<sub>2</sub>] em água e a menor taxa metabólica proporcionadas pelo ambiente gélido, juntamente com um maior débito cardíaco devem explicar como as espécies de Channichthyidae compensam a falta de Hb e *Mb* no transporte convectivo e difusivo de O<sub>2</sub> (SIDELL; O'BRIEN, 2006; STORZ, 2019). Nós observamos que apenas Gymnophiona reteve *Mb* e *GbE* no clado dos anfíbios (Figura 8). Como os anuros e salamandras que analisamos parecem ambos ter perdido *Mb* e *GbE*, e outra salamandra, *Cynops pyrrhogaster*, não apresentou *Mb* (XI *et al.*, 2007), nós sugerimos a hipótese de que a perda de *Mb* e *GbE* ocorreu no último ancestral comum do clado Batrachia que viveu por volta de 297 Ma (KUMAR *et al.*, 2017). A perda dos genes curiosamente coincide com o aumento da concentração de O<sub>2</sub> atmosférica no final do Carbonífero e começo do Permiano (BERNER; VANDENBROOKS; WARD, 2007; KRAUSE *et al.*, 2018), mas na falta de evidência adicional sobre a relação entre tais eventos, o modelo nulo da evolução molecular coloca que a perda dos genes deve ter sido aproximadamente neutra.

### 6.3 Hemoglobina nos anfíbios

A nossa análise do grupamento de genes codificantes da Hb de anfíbios revelou pistas sobre rearranjos que parecem ter ocorrido na região cromossômica onde os genes estavam presentes. A partir dos grupamentos de Hb de anuros e caecilianos modernos (Figura 7), e levando em conta os padrões de evolução dos loci de Hb já conhecidos (PHILIPSEN; HARDISON, 2018), inferimos que o arranjo ancestral dos parálogos de *Hba*, *Hbb* e a cópia única de *GbY* no ancestral dos anfíbios deve ter sido o seguinte: 5'-*NPRL3*- $\alpha^E$ - $\alpha^A$ - $\beta^T$ - $\beta^A$ -*GbY*-*LUC7L*-3'. A sintenia dos genes na extremidade 3' do grupamento foi logo quebrada por um rearranjo, que gerou dois grupamentos no mesmo cromossomo: 5'-*LUC7L*-*GbY*- $\beta^A$ -3' e 5'-*NPRL3*- $\alpha^E$ - $\alpha^A$ - $\beta^T$ -3'. Caecilianos perderam *GbY* e têm um gene de globina  $\beta^A$  em outro cromossomo, sugerindo que uma transposição ocorreu na linhagem dos caecilianos. Podemos ver nos filogramas das Figuras 5 e 6 que globinas  $\alpha^E$  e  $\beta^A$  se diversificaram em anuros e salamandras, enquanto globinas  $\alpha^A$  e  $\beta^T$  geralmente permaneceram como cópias únicas. Ao checarmos o padrão de expressão, notamos que globinas  $\alpha^A$  e  $\beta^T$  são geralmente expressas em estágios adultos. Uma interessante exceção foi o *Xenopus tropicalis*, onde também pudemos observar outro rearranjo (Figura 7) que provavelmente inverteu a expressão ontogenética de genes ligados ao *NPRL3*. A diversidade de globinas  $\alpha^E$  e  $\beta^A$  deve estar relacionada à variação na expressão de isoformas de Hb (isoHbs) observada na ontogenia e filogenia de muitas espécies de anuros (AGGARWAL; RIGGS, 1969; ATHA *et al.*, 1979; CARDELLINI; SALA, 1978, 1979b, 1979a, 1983, 1987; CLEMENTI *et al.*, 2007; LIMA *et al.*, 1983, 1985; MEIRELLES *et al.*, 1979; VAN AARDT; WEBER, 2010; VIEIRA *et al.*, 1982) e salamandras (BONAVENTURA *et al.*, 1977; CONDÒ *et al.*, 1989; FLAVIN; DUPRAT; ROSA, 1979; KOUSSOULAKOS; KAPAROS; STATHAKOS, 1986; MACLEAN; JURD, 1971; TAKETA; NICKERSON, 1973a, 1973b; WEBER; WELLS; ROSSETTI, 1985). Encontramos apenas um estudo relatando as isoHbs expressas em um caeciliano, *Typhlonectes compressicauda*, no qual o sangue maternal e fetal apresentaram o mesmo padrão consistindo de um componente majoritário e um minoritário (GARLICK *et al.*, 1979). Tal padrão é consistente com o número de globinas que codificam as cadeias da Hb que encontramos em caecilianos. Estudos bioquímicos futuros podem revelar o papel dos diferentes parálogos de *Hba* e *Hbb* na expressão de isoHbs, assim como revelar os mecanismos genéticos e bioquímicos das adaptações respiratórias em anfíbios.

#### 6.4 Globina A

Encontramos genes do tipo *GbA* em espécies de anuros pertencentes aos grupos Neobatrachia e Pelobatoidea. Várias cópias de *GbA* estão presentes na família Bufonidae, indicando uma possível evolução por duplicação gênica seguida de subfuncionalização nesse grupo. Maeda e Fitch (1982) sugeriram que a proteína monomérica codificada pela *GbA* de *Rana catesbeiana* e expressa no tecido cardíaco surgiu por volta do tempo do evento de duplicação que originou *Hba* e *Hbb*. Nossa análise filogenética (Figura 1) nos permite inferir uma origem mais recente para a *GbA*, que ocorreu por meio da cooptação de uma globina  $\alpha^E$  no último ancestral comum do clado que compreende os grupos Neobatrachia e Pelobatoidea, por volta de 186 Ma (KUMAR *et al.*, 2017). Os genes de *GbA* estão localizados no mesmo cromossomo da *Adgb* em *Pyxicephalus adspersus* e *Leptobranchium leishanense*. Por isso, nós sugerimos a hipótese de que ocorreu um evento de transposição de uma globina  $\alpha^E$  para o mesmo cromossomo da *Adgb* antes do surgimento da *GbA*. Pelo fato da *GbA* de *Rana catesbeiana* codificar uma proteína monomérica (MAEDA; FITCH, 1982), nós podemos inferir que a neofuncionalização da globina  $\alpha^E$  que deu origem à *GbA* envolveu a perda de resíduos que permitem cadeias  $\alpha$  formarem homodímeros, um caráter que já estava presente no polipeptídeo ancestral das cadeias  $\alpha$  e  $\beta$  da Hb (PILLAI *et al.*, 2020). De fato, Maeda e Fitch (1982) observaram que enquanto os resíduos envolvidos no contato com o heme foram conservados, os resíduos envolvidos no contato entre as subunidades não foram conservados. Estudos futuros devem esclarecer a estrutura e as propriedades funcionais da *GbA*, assim como revelar seu papel fisiológico e os mecanismos genéticos e bioquímicos responsáveis pela evolução dessa pequena globina monomérica (~132 aa) em anfíbios.

## 7 CONCLUSÃO

Os principais grupos de anfíbios retiveram diferentes parálogos de *GbX*. *Gymnophiona* reteve *Mb* e *GbE*, com algumas deleções de *GbE* em linhagens específicas, enquanto *Anura* e *Caudata* perderam ambos os genes. *GbY*, por sua vez, foi perdida em *Gymnophiona*, mas retida em *Caudata* e em algumas linhagens dos *Anura*. Os anfíbios modernos não apresentam o gene de globina  $\alpha^D$ , mas apresentam dois tipos de parálogos de globinas  $\beta$ , indicando que pelo menos dois genes de *Hbb* já estavam presentes no ancestral dos anfíbios. *GbA* é um novo tipo de gene de globina que evoluiu em anfíbios através da cooptação de uma globina  $\alpha^E$  que foi transposta previamente para outra região genômica. Os demais tipos de globinas presentes em gnatostomados (*Adgb*, *Ngb* e *Cygb*) foram retidos nos anfíbios.

## REFERÊNCIAS

- AGARWALA, R. *et al.* Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, [s.l.], v. 46, n. D1, p. D8–D13, 2018.
- AGGARWAL, S. J.; RIGGS, A. F. The hemoglobins of the bullfrog, *Rana catesbeiana* I. Purification, amino acid composition, and oxygen equilibria. **Journal of Biological Chemistry**, Estados Unidos da América, v. 244, n. 9, p. 2372–2383, 1969.
- AGUILETA, G.; BIELAWSKI, J. P.; YANG, Z. Proposed standard nomenclature for the  $\alpha$ - and  $\beta$ -globin gene families. **Genes and Genetic Systems**, [s.l.], v. 81, n. 5, p. 367–371, 2006.
- ALEV, C. *et al.* Genomic organization of zebra finch alpha and beta globin genes and their expression in primitive and definitive blood in comparison with globins in chicken. **Development Genes and Evolution**, [s.l.], v. 219, n. 7, p. 353–360, 2009.
- ARENAS GÓMEZ, C. M. *et al.* A de novo reference transcriptome for *Bolitoglossa valleculea*, an Andean mountain salamander in Colombia. **Data in Brief**, [s.l.], v. 29, p. 105256, 2020.
- ATHA, D. H. *et al.* Hemoglobins of the tadpole of the bullfrog, *Rana catesbeiana*: pH dependence of ligand binding and subunit dissociation equilibria and kinetics. **Journal of Biological Chemistry**, Estados Unidos da América, v. 254, n. 9, p. 3393–3400, 1979.
- BERENBRINK, M. Evolution of a molecular machine. **Nature**, [s.l.], v. 581, p. 388–389, 2020.
- BERNER, R. A.; VANDENBROOKS, J. M.; WARD, P. D. Oxygen and evolution. **Science**, [s.l.], v. 316, n. 5824, p. 557–558, 2007.
- BIROL, I. *et al.* De novo transcriptome assemblies of *Rana (Lithobates) catesbeiana* and *Xenopus laevis* tadpole livers for comparative genomics without reference genomes. **PLoS ONE**, [s.l.], v. 10, n. 6, p. e0130720, 2015.
- BLACKBURN, D. C.; WAKE, D. B. Class Amphibia Gray, 1825. In: ZHANG, Z.-Q. (ed.). **Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness**. [s.l.], v. 3148p. 39–55, 2011.
- BLANK, M. *et al.* Oxygen supply from the bird's eye perspective: globin E is a respiratory protein in the chicken retina. **Journal of Biological Chemistry**, Estados Unidos da América, v. 286, n. 30, p. 26507–26515, 2011a.
- BLANK, M. *et al.* A membrane-bound vertebrate globin. **PLoS ONE**, [s.l.], v. 6, n. 9, p. e25292, 2011b.
- BLANK, M.; BURMESTER, T. Widespread occurrence of N-terminal acylation in animal globins and possible origin of respiratory globins from a membrane-bound ancestor. **Molecular Biology and Evolution**, Oxford, v. 29, n. 11, p. 3553–3561, 2012.
- BONAVENTURA, C. *et al.* Anion modulation of the negative Bohr effect of haemoglobin

from a primitive amphibian. **Nature**, [s.l.], v. 265, n. 5593, p. 474–476, 1977.

BROWN, R. M. *et al.* Multilocus phylogeny and a new classification for Southeast Asian and Melanesian forest frogs (family Ceratobatrachidae). **Zoological Journal of the Linnean Society**, Londres, v. 174, n. 1, p. 130–168, 2015.

BRYANT, D. M. *et al.* A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. **Cell Reports**, [s.l.], v. 18, n. 3, p. 762–776, 2017.

BURMESTER, T. *et al.* A vertebrate globin expressed in the brain. **Nature**, [s.l.], v. 407, n. 6803, p. 520–523, 2000.

BURMESTER, T. *et al.* Cytoglobin: a novel globin type ubiquitously expressed in vertebrate tissues. **Molecular Biology and Evolution**, Oxford, v. 19, n. 4, p. 416–421, 2002.

BURMESTER, T.; HANKELN, T. What is the function of neuroglobin? **Journal of Experimental Biology**, [s.l.], v. 212, n. 10, p. 1423–1428, 2009.

BURMESTER, T.; HANKELN, T. Function and evolution of vertebrate globins. **Acta Physiologica**, Escandinávia, v. 211, n. 3, p. 501–514, 2014.

BURNS, J. *et al.* Transcriptome analysis illuminates the nature of the intracellular interaction in a vertebrate-algal symbiosis. **eLife**, [s.l.], v. 6, p. 1–32, 2017.

CARDELLINI, P.; SALA, M. Metamorphic variations in the hemoglobins of *Hyla arborea* L. **Comparative Biochemistry and Physiology Part B: Comparative Biochemistry**, Grã-Bretanha, v. 61, n. 1, p. 21–24, 1978.

CARDELLINI, P.; SALA, M. Hemoglobin transition in the toad *Bufo viridis laurenti* during development. **Italian Journal of Zoology**, Itália, v. 46, n. 1–2, p. 51–56, 1979a.

CARDELLINI, P.; SALA, M. Metamorphic variations in the hemoglobins of *Bombina variegata* (L.). **Comparative Biochemistry and Physiology Part B: Comparative Biochemistry**, Grã-Bretanha, v. 64, n. 1, p. 113–116, 1979b.

CARDELLINI, P.; SALA, M. Developmental time of the hemoglobin transition in the anuran *Bombina orientalis*. **Comparative Biochemistry and Physiology Part B: Comparative Biochemistry**, Grã-Bretanha, v. 75, n. 2, p. 259–262, 1983.

CARDELLINI, P.; SALA, M. Hemoglobin transition in the anuran *Pelodytes punctatus*. **Comparative Biochemistry and Physiology Part A: Comparative Physiology**, Grã-Bretanha, v. 86, n. 1, p. 85–89, 1987.

CARVER, T. *et al.* Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. **Bioinformatics**, Oxford, v. 28, n. 4, p. 464–469, 2012.

CHAN, F.-Y. *et al.* Characterization of adult  $\alpha$ - and  $\beta$ -globin genes in the zebrafish. **Blood**, [s.l.], v. 89, n. 2, p. 688–700, 1997.



- CHE, R. *et al.* Transcriptomic analysis of endangered Chinese salamander: identification of immune, sex and reproduction-related genes and genetic markers. **PLoS ONE**, [s.l.], v. 9, n. 1, p. e87940, 2014.
- CHRISTENSON, M. K. *et al.* De novo assembly and analysis of the northern leopard frog *Rana pipiens* transcriptome. **Journal of Genomics**, [s.l.], v. 2, p. 141–149, 2014.
- CIVITELLO, M. L. *et al.* Activation of the bile acid pathway and no observed antimicrobial peptide sequences in the skin of a poison frog. **G3: Genes, Genomes, Genetics**, [s.l.], v. 9, n. 2, p. 581–589, 2019.
- CLEMENTI, M. E. *et al.* Oxygen transport in Amphibia: the functional properties of hemoglobins from *Bufo bufo* and *Bufo viridis*. **Journal of Biological Sciences**, [s.l.], v. 7, n. 5, p. 786–790, 2007.
- CONDÒ, S. G. *et al.* The functional properties of amphibian hemoglobin: the case of *Salamander salamander* and *Hydromantes genei*. **Comparative Biochemistry and Physiology Part A: Physiology**, Grã-Bretanha v. 93, n. 2, p. 319–325, 1989.
- COOPER, S. J. B. *et al.* The mammalian  $\alpha^D$ -globin gene lineage and a new model for the molecular evolution of  $\alpha$ -globin gene clusters at the stem of the mammalian radiation. **Molecular Phylogenetics and Evolution**, [s.l.], v. 38, n. 2, p. 439–448, 2006.
- CORTI, P. *et al.* Globin X is a six-coordinate globin that reduces nitrite to nitric oxide in fish red blood cells. **Proceedings of the National Academy of Sciences**, Estados Unidos da América, v. 113, n. 30, p. 8538–8543, 2016.
- DAANE, J. M. *et al.* Adaptations to environmental change: globin superfamily evolution in Antarctic fishes. **Marine Genomics**, [s.l.], v. 49, p. 100724, 2020.
- DAMSGAARD, C. *et al.* Hemoglobin isoform differentiation and allosteric regulation of oxygen binding in the turtle, *Trachemys scripta*. **American Journal of Physiology-Regulatory, Integrative and Comparative Physiology**, Estados Unidos da América, v. 305, n. 8, p. R961–R967, 2013.
- DAWKINS, R. **A escalada do monte improvável**. São Paulo: Companhia das Letras, 1998.
- DE SANCTIS, D. *et al.* Crystal structure of cytoglobin: the fourth globin type discovered in man displays heme hexa-coordination. **Journal of Molecular Biology**, [s.l.], v. 336, n. 4, p. 917–927, 2004.
- DENTON, R. D. *et al.* The African Bullfrog (*Pyxicephalus adspersus*) genome unites the two ancestral ingredients for making vertebrate sex chromosomes. **bioRxiv**, [s.l.], p. 329847, 2018.
- DEWILDE, S. *et al.* Biochemical characterization and ligand binding properties of neuroglobin, a novel member of the globin family. **Journal of Biological Chemistry**, Estados Unidos da América, v. 276, n. 42, p. 38949–38955, 2001.
- DRÖGE, J. *et al.* Comparative genomics of neuroglobin reveals its early origins. **PLoS ONE**,

[*s.l.*], v. 7, n. 10, p. e47972, 2012.

EBNER, B. *et al.* The globin gene family of the cephalochordate amphioxus: implications for chordate globin evolution. **BMC Evolutionary Biology**, [*s.l.*], v. 10, n. 370, 2010.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, [*s.l.*], v. 32, n. 5, p. 1792–1797, 2004.

EDWARDS, R. J. *et al.* Draft genome assembly of the invasive cane toad, *Rhinella marina*. **GigaScience**, Oxford, v. 7, n. 9, p. giy095, 2018.

FAGO, A. *et al.* Functional diversification of sea lamprey globins in evolution and development. **Biochimica et Biophysica Acta - Proteins and Proteomics**, [*s.l.*], v. 1866, n. 2, p. 283–291, 2018.

FENG, Y. J. *et al.* Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. **Proceedings of the National Academy of Sciences of the United States of America**, Estados Unidos da América, v. 114, n. 29, p. E5864–E5870, 2017.

FERRERAS, J. M. *et al.* Insight into the phylogenetic relationship and structural features of vertebrate myoglobin family. **International Journal of Biological Macromolecules**, [*s.l.*], v. 93, p. 1041–1050, 2016.

FLAVIN, M.; DUPRAT, A. M.; ROSA, J. Ontogenetic changes in the hemoglobins of the salamander, *Pleurodeles waltlii*. **Cell Differentiation**, Irlanda, v. 8, p. 405–410, 1979.

FORGET, B. G.; HARDISON, R. C. The normal structure and regulation of human globin gene clusters. In: STEINBERG, M. H. *et al.* (eds.). **Disorders of hemoglobin: genetics, pathophysiology, and clinical management**. 2<sup>a</sup> ed. Cambridge: Cambridge University Press, 2009. p. 46–61.

FUCHS, C.; BURMESTER, T.; HANKELN, T. The amphibian globin gene repertoire as revealed by the *Xenopus* genome. **Cytogenetic and Genome Research**, [*s.l.*], v. 112, n. 3–4, p. 296–306, 2006.

GALLAGHER, M. D.; MACQUEEN, D. J. Evolution and expression of tissue globins in ray-finned fishes. **Genome Biology and Evolution**, Oxford, v. 9, n. 1, p. 32–47, 2017.

GARLICK, R. L. *et al.* A fetal-maternal shift in the oxygen equilibrium of hemoglobin from the viviparous caecilian, *Typhlonectes compressicauda*. **Comparative Biochemistry and Physiology Part A: Physiology**, Grã-Bretanha, v. 62, n. 1, p. 239–244, 1979.

GAUDRY, M. J. *et al.* Repeated evolution of chimeric fusion genes in the  $\beta$ -globin gene family of laurasiatherian mammals. **Genome Biology and Evolution**, Oxford, v. 6, n. 5, p. 1219–1233, 2014.

GELL, D. A. Structure and function of haemoglobins. **Blood Cells, Molecules, and Diseases**, [*s.l.*], v. 70, p. 13–42, 2018.

- GERCHEN, J. F. *et al.* A single transcriptome of a green toad (*Bufo viridis*) yields candidate genes for sex determination and -differentiation and non-anonymous population genetic markers. **PLoS ONE**, [s.l.], v. 11, n. 5, p. e0156419, 2016.
- GILLEMANS, N. *et al.* Functional and comparative analysis of globin loci in pufferfish and humans. **Blood**, [s.l.], v. 101, n. 7, p. 2842–2849, 2003.
- GLASS, H. C.; MELIN, A. D.; VAMOSI, S. M. De novo transcriptome analysis of dermal tissue from the rough-skinned newt, *Taricha granulosa*, enables investigation of tetrodotoxin expression. **bioRxiv**, [s.l.], p. 653238, 2019.
- GOEDBLOED, D. J. *et al.* Parallel habitat acclimatization is realized by the expression of different genes in two closely related salamander species (genus *Salamandra*). **Heredity**, [s.l.], v. 119, n. 6, p. 429–437, 2017.
- GOODMAN, M.; MOORE, G. W.; MATSUDA, G. Darwinian evolution in the genealogy of haemoglobin. **Nature**, [s.l.], v. 253, n. 5493, p. 603–608, 1975.
- GRISPO, M. T. *et al.* Gene duplication and the evolution of hemoglobin isoform differentiation in birds. **Journal of Biological Chemistry**, Estados Unidos da América, v. 287, n. 45, p. 37647–37658, 2012.
- HAEUSSLER, M. *et al.* Navigating protected genomics data with UCSC Genome Browser in a Box. **Bioinformatics**, Oxford, v. 31, n. 5, p. 764–766, 2015.
- HALL, K. W.; EISTHEN, H. L.; WILLIAMS, B. L. Proteinaceous pheromone homologs identified from the cloacal gland transcriptome of a male axolotl, *Ambystoma mexicanum*. **PLoS ONE**, [s.l.], v. 11, n. 2, p. e0146851, 2016.
- HAMMOND, S. A. *et al.* The North American bullfrog draft genome provides insight into hormonal regulation of long noncoding RNA. **Nature Communications**, [s.l.], v. 8, n. 1, p. 1433, 2017.
- HAN, M. V.; ZMASEK, C. M. PhyloXML: XML for evolutionary biology and comparative genomics. **BMC Bioinformatics**, [s.l.], v. 10, p. 356, 2009.
- HANKELN, T. *et al.* Neuroglobin and cytoglobin in search of their role in the vertebrate globin family. **Journal of Inorganic Biochemistry**, Estados Unidos da América, v. 99, n. 1, p. 110–119, 2005.
- HARDISON, R. C. Evolution of hemoglobin and its genes. **Cold Spring Harbor Perspectives in Medicine**, [s.l.], v. 2, n. 12, p. a011627, 2012.
- HARDISON, R. C. A Cambrian origin for globin gene regulation. **Blood**, [s.l.], v. 136, n. 3, p. 261–262, 2020.
- HEINICKE, M. P. *et al.* Phylogenomic support for evolutionary relationships of New World direct-developing frogs (Anura: Terraranae). **Molecular Phylogenetics and Evolution**, [s.l.], v. 118, p. 145–155, 2018.

HELLSTEN, U. *et al.* The genome of the western clawed frog *Xenopus tropicalis*. **Science**, [s.l.], v. 328, n. 5978, p. 633–636, 2010.

HIME, P. M. *et al.* Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. **Systematic Biology**, [s.l.], 2020.

HOANG, D. T. *et al.* UFBoot2: improving the ultrafast bootstrap approximation. **Molecular Biology and Evolution**, Oxford, v. 35, n. 2, p. 518–522, 2018.

HOFFMANN, F. G. *et al.* Lineage-specific patterns of functional diversification in the  $\alpha$ - and  $\beta$ -globin gene families of tetrapod vertebrates. **Molecular Biology and Evolution**, Oxford, v. 27, n. 5, p. 1126–1138, 2010.

HOFFMANN, F. G. *et al.* Evolution of the globin gene family in deuterostomes: lineage-specific patterns of diversification and attrition. **Molecular Biology and Evolution**, Oxford, v. 29, n. 7, p. 1735–1745, 2012.

HOFFMANN, F. G. *et al.* Gene turnover and diversification of the  $\alpha$ - and  $\beta$ -globin gene families in sauropsid vertebrates. **Genome Biology and Evolution**, Oxford, v. 10, n. 1, p. 344–358, 2018.

HOFFMANN, F. G.; OPAZO, J. C.; STORZ, J. F. Rapid rates of lineage-specific gene duplication and deletion in the  $\alpha$ -globin gene family. **Molecular Biology and Evolution**, Oxford, v. 25, n. 3, p. 591–602, 2008a.

HOFFMANN, F. G.; OPAZO, J. C.; STORZ, J. F. New genes originated via multiple recombinational pathways in the  $\beta$ -globin gene family of rodents. **Molecular Biology and Evolution**, Oxford, v. 25, n. 12, p. 2589–2600, 2008b.

HOFFMANN, F. G.; OPAZO, J. C.; STORZ, J. F. Differential loss and retention of cytoglobin, myoglobin, and globin-E during the radiation of vertebrates. **Genome Biology and Evolution**, Oxford, v. 3, n. 1, p. 588–600, 2011.

HOFFMANN, F. G.; OPAZO, J. C.; STORZ, J. F. Whole-genome duplications spurred the functional diversification of the globin gene superfamily in vertebrates. **Molecular Biology and Evolution**, Oxford, v. 29, n. 1, p. 303–312, 2012.

HOFFMANN, F. G.; STORZ, J. F. The  $\alpha^D$ -globin gene originated via duplication of an embryonic  $\alpha$ -like globin gene in the ancestor of tetrapod vertebrates. **Molecular Biology and Evolution**, Oxford, v. 24, n. 9, p. 1982–1990, 2007.

HOOGEWIJS, D. *et al.* Androglobin: a chimeric globin in metazoans that is preferentially expressed in mammalian testes. **Molecular Biology and Evolution**, Oxford, v. 29, n. 4, p. 1105–1114, 2012.

HOSBACH, H. A.; WYLER, T.; WEBER, R. The *Xenopus laevis* globin gene family: chromosomal arrangement and gene structure. **Cell**, [s.l.], v. 32, n. 1, p. 45–53, 1983.

HUANG, L. *et al.* Comparative transcriptome analyses of seven anurans reveal functions and adaptations of amphibian skin. **Scientific Reports**, [s.l.], v. 6, p. 24069, 2016.

INGERMANN, R. L. Vertebrate hemoglobins. In: TERJUNG, R. (ed.). **Comprehensive Physiology**, 2011, p. 357–408.

JEFFREYS, A. J. *et al.* Linkage of adult  $\alpha$ - and  $\beta$ -globin genes in *X. laevis* and gene duplication by tetraploidization. **Cell**, [s.l.], v. 21, n. 2, p. 555–564, 1980.

JETZ, W.; PYRON, R. A. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. **Nature Ecology and Evolution**, [s.l.], v. 2, n. 5, p. 850–858, 2018.

JOHNSON, R. M. *et al.* Phylogenetic comparisons suggest that distance from the locus control region guides developmental expression of primate  $\beta$ -type globin genes. **Proceedings of the National Academy of Sciences of the United States of America**, Estados Unidos da América, v. 103, n. 9, p. 3186–3191, 2006.

JOHNSON, R. M.; GUMUCIO, D.; GOODMAN, M. Globin gene switching in primates. **Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology**, Grã-Bretanha, v. 133, n. 3, p. 877–883, 2002.

KAKAR, S. *et al.* Structure and reactivity of hexacoordinate hemoglobins. **Biophysical Chemistry**, [s.l.], v. 152, n. 1–3, p. 1–14, 2010.

KALYAANAMOORTHY, S. *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. **Nature Methods**, [s.l.], v. 14, n. 6, p. 587–589, 2017.

KAMEI, R. G. *et al.* Discovery of a new family of amphibians from northeast India with ancient links to Africa. **Proceedings of the Royal Society B: Biological Sciences**, Londres, v. 279, n. 1737, p. 2396–2401, 2012.

KAPLI, P.; YANG, Z.; TELFORD, M. J. Phylogenetic tree building in the genomic age. **Nature Reviews Genetics**, [s.l.], v. 21, p. 428–444, 2020.

KATOH, K.; ROZEWICKI, J.; YAMADA, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. **Briefings in Bioinformatics**, [s.l.], v. 20, n. 4, p. 1160–1166, 2019.

KENDREW, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. **Nature**, [s.l.], v. 181, p. 662–666, 1958.

KEPPNER, A. *et al.* Lessons from the post-genomic era: globin diversity beyond oxygen binding and transport. **Redox Biology**, [s.l.], v. 37, p. 101687, 2020.

KOCH, J. *et al.* Unusual diversity of myoglobin genes in the lungfish. **Molecular Biology and Evolution**, Oxford, v. 33, n. 12, p. 3033–3041, 2016.

KOUSSOULAKOS, S.; KAPAROS, G.; STATHAKOS, D. Multiple hemoglobins in *Triturus cristatus*: their study by analytical isoelectrofocusing. **Comparative Biochemistry and Physiology**, Grã-Bretanha, v. 83, n. 2, p. 475–481, 1986.

- KRAUSE, A. J. *et al.* Stepwise oxygenation of the Paleozoic atmosphere. **Nature Communications**, [s.l.], v. 9, p. 4081, 2018.
- KUGELSTADT, D. *et al.* Neuroglobin, cytoglobin, and a novel, eye-specific globin from chicken. **Biochemical and Biophysical Research Communications**, [s.l.], v. 325, n. 3, p. 719–725, 2004.
- KUMAR, S. *et al.* TimeTree: a resource for timelines, timetrees, and divergence times. **Molecular Biology and Evolution**, Oxford, v. 34, n. 7, p. 1812–1819, 2017.
- KUMAR, S. *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. **Molecular Biology and Evolution**, Oxford, v. 35, n. 6, p. 1547–1549, 2018.
- LARKIN, M. A. *et al.* Clustal W and Clustal X version 2.0. **Bioinformatics**, Oxford, v. 23, n. 21, p. 2947–2948, 2007.
- LASSMANN, T.; SONNHAMMER, E. L. L. Automatic assessment of alignment quality. **Nucleic Acids Research**, [s.l.], v. 33, n. 22, p. 7120–7128, 2005.
- LASSMANN, T.; SONNHAMMER, E. L. L. Kalign, Kalignvu and Mumsa: web servers for multiple sequence alignment. **Nucleic Acids Research**, [s.l.], v. 34, n. WEB. SERV. ISS., p. 596–599, 2006.
- LI, J. *et al.* Genomic and transcriptomic insights into molecular basis of sexually dimorphic nuptial spines in *Leptobrachium leishanense*. **Nature Communications**, [s.l.], v. 10, n. 1, p. 5551, 2019a.
- LI, Q. *et al.* A draft genome assembly of the eastern banjo frog *Limnodynastes dumerilii* (Anura: Limnodynastidae). **bioRxiv**, [s.l.], 2020.
- LI, Y. *et al.* Chromosome-level assembly of the mustache toad genome using third-generation DNA sequencing and Hi-C analysis. **GigaScience**, Oxford, v. 8, p. 1–10, 2019b.
- LIMA, A. A. B. *et al.* Separation and functional characterization of *Leptodactylus labyrinthicus* hemoglobin components. **Comparative Biochemistry and Physiology Part A: Physiology**, Grã-Bretanha, v. 76, n. 1, p. 123–125, 1983.
- LIMA, A. A. B. *et al.* Allosteric effect of protons and adenosine triphosphate on hemoglobins from aquatic amphibia. **Journal of Comparative Physiology Part B: Comparative Biochemistry**, Grã-Bretanha, v. 155, n. 3, p. 353–355, 1985.
- LIU, W. *et al.* IBS: an illustrator for the presentation and visualization of biological sequences. **Bioinformatics**, Oxford, v. 31, n. 20, p. 3359–3361, 2015.
- LU, S. *et al.* CDD/SPARCLE: the conserved domain database in 2020. **Nucleic Acids Research**, [s.l.], v. 48, n. D1, p. D265–D268, 2020.
- LÜDEMANN, J. *et al.* Globin E is a myoglobin-related, respiratory protein highly expressed in lungfish oocytes. **Scientific Reports**, [s.l.], v. 9, n. 280, p. 1–11, 2019.

LÜDEMANN, J. *et al.* Genetic and functional diversity of the multiple lungfish myoglobins. **FEBS Journal**, [s.l.], v. 287, n. 8, p. 1598–1611, 2020.

MA, W.-J. *et al.* Evolutionary and developmental dynamics of sex-biased gene expression in common frogs with proto-Y chromosomes. **Genome Biology**, [s.l.], v. 19, n. 156, 2018.

MACLEAN, N.; JURD, R. D. Electrophoretic analysis of the haemoglobins of *Ambystoma mexicanum*. **Comparative Biochemistry and Physiology Part B: Comparative Biochemistry**, Grã-Bretanha, v. 40, n. 3, p. 751–755, 1971.

MADDISON, W. P. Gene trees in species trees. **Systematic Biology**, [s.l.], v. 46, n. 3, p. 523–536, 1997.

MAEDA, N.; FITCH, W. M. Isolation and amino acid sequence of a monomeric hemoglobin in heart muscle of the bullfrog, *Rana catesbeiana*. **Journal of Biological Chemistry**, Estados Unidos da América, v. 257, n. 6, p. 2806–2815, 1982.

MAIRBÄURL, H.; WEBER, R. E. Oxygen transport by hemoglobin. **Comprehensive Physiology**, Estados Unidos da América, v. 2, n. 2, p. 1463–1489, 2012.

MARINO, K. *et al.* Organisation of the *Hb 1* genes of the Antarctic skate *Bathyraja eatonii*: new insights into the evolution of globin genes. **Gene**, [s.l.], v. 406, n. 1–2, p. 199–208, 2007.

MCELROY, K. E. *et al.* Genome expression balance in a triploid trihybrid vertebrate. **Genome Biology and Evolution**, Oxford, v. 9, n. 4, p. 968–980, 2017.

MEIRELLES, N. C. *et al.* Some larval properties of *Pipa carvalhoi* adult hemoglobins. **Comparative Biochemistry and Physiology Part A: Physiology**, Grã-Bretanha, v. 62, n. 4, p. 859–862, 1979.

MIYATA, M. *et al.* An evolutionary ancient mechanism for regulation of hemoglobin expression in vertebrate red cells. **Blood**, [s.l.], v. 136, n. 3, p. 269–278, 2020.

MONTERO-MENDIETA, S. *et al.* A practical guide to build de-novo assemblies for single tissues of non-model organisms: the example of a Neotropical frog. **PeerJ**, [s.l.], v. 5, p. e3702, 2017.

MORETTI, S. *et al.* The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. **Nucleic Acids Research**, [s.l.], v. 35, n. SUPPL.2, p. 645–648, 2007.

NGUYEN, L. T. *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. **Molecular Biology and Evolution**, Oxford, v. 32, n. 1, p. 268–274, 2015.

NOWOSHILOW, S. *et al.* The axolotl genome and the evolution of key tissue formation regulators. **Nature**, [s.l.], v. 554, n. 7690, p. 50–55, 2018.

NÜRNBERGER, B. *et al.* Para-allopatry in hybridizing fire-bellied toads (*Bombina bombina* and *B. variegata*): inference from transcriptome-wide coalescence analyses. **Evolution**, [s.l.],

v. 70, n. 8, p. 1803–1818, 2016.

OPAZO, J. C. *et al.* Whole-genome duplication and the functional diversification of teleost fish hemoglobins. **Molecular Biology and Evolution**, Oxford, v. 30, n. 1, p. 140–153, 2013.

OPAZO, J. C. *et al.* Ancient duplications and expression divergence in the globin gene superfamily of vertebrates: insights from the elephant shark genome and transcriptome. **Molecular Biology and Evolution**, Oxford, v. 32, n. 7, p. 1684–1694, 2015a.

OPAZO, J. C. *et al.* Gene turnover in the avian globin gene families and evolutionary changes in hemoglobin isoform expression. **Molecular Biology and Evolution**, Oxford, v. 32, n. 4, p. 871–887, 2015b.

OPAZO, J. C.; HOFFMANN, F. G.; STORZ, J. F. Genomic evidence for independent origins of  $\beta$ -like globin genes in monotremes and therian mammals. **Proceedings of the National Academy of Sciences of the United States of America**, Estados Unidos da América, v. 105, n. 5, p. 1590–1595, 2008a.

OPAZO, J. C.; HOFFMANN, F. G.; STORZ, J. F. Differential loss of embryonic globin genes during the radiation of placental mammals. **Proceedings of the National Academy of Sciences of the United States of America**, Estados Unidos da América, v. 105, n. 35, p. 12950–12955, 2008b.

PARADIS, E.; SCHLIEP, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. **Bioinformatics**, Oxford, v. 35, n. 3, p. 526–528, 2019.

PATEL, V. S. *et al.* Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals. **BMC Biology**, [s.l.], v. 6, n. 34, 2008.

PATEL, V. S. *et al.* Globin gene structure in a reptile supports the transpositional model for amniote  $\alpha$ - and  $\beta$ -globin gene evolution. **Chromosome Research**, [s.l.], v. 18, n. 8, p. 897–907, 2010.

PATEL, V. S.; DEAKIN, J. E. The evolutionary history of globin genes: insights from marsupials and monotremes. *In*: DEAKIN, J. E.; WATERS, P. D.; GRAVES, J. A. M. (eds.). **Marsupial Genetics and Genomics**. Dordrecht: Springer, 2010. p. 415–433.

PATIENT, R. K. *et al.* Internal organization of the major adult  $\alpha$ - and  $\beta$ -globin genes of *X. laevis*. **Cell**, [s.l.], v. 21, n. 2, p. 565–573, 1980.

PELOSO, P. L. V. *et al.* The impact of anchored phylogenomics and taxon sampling on phylogenetic inference in narrow-mouthed frogs (Anura, Microhylidae). **Cladistics**, [s.l.], v. 32, n. 2, p. 113–140, 2016.

PERUTZ, M. F. *et al.* Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. **Nature**, [s.l.], v. 185, n. 4711, p. 416–422, 1960.

PERUTZ, M. F. Hemoglobin structure and respiratory transport. **Scientific American**, Estados Unidos da América, v. 239, n. 6, p. 92–125, 1978.



PHILIPSEN, S.; HARDISON, R. C. Evolution of hemoglobin loci and their regulatory elements. **Blood Cells, Molecules, and Diseases**, [s.l.], v. 70, p. 2–12, 2018.

PILLAI, A. S. *et al.* Origin of complexity in haemoglobin evolution. **Nature**, [s.l.], v. 581, p. 480–485, 2020.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: Austria, 2020. Disponível em: <https://www.r-project.org/>. Acesso em: 22 nov. 2020.

RICHARDSON, M. F. *et al.* Improving amphibian genomic resources: a multitissue reference transcriptome of an iconic invader. **GigaScience**, Oxford, v. 7, n. 1, p. 1–7, 2018.

ROBERTSON, L. S.; CORNMAN, R. S. Transcriptome resources for the frogs *Lithobates clamitans* and *Pseudacris regilla*, emphasizing antimicrobial peptides and conserved loci for phylogenetics. **Molecular Ecology Resources**, [s.l.], v. 14, n. 1, p. 178–183, 2014.

ROGERS, R. L. *et al.* Genomic takeover by transposable elements in the strawberry poison frog. **Molecular Biology and Evolution**, Oxford, v. 35, n. 12, p. 2913–2927, 2014.

ROHLFING, K. *et al.* Convergent evolution of hemoglobin switching in jawed and jawless vertebrates. **BMC Evolutionary Biology**, [s.l.], v. 16, p. 30, 2016.

RONQUIST, F. *et al.* Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. **Systematic Biology**, [s.l.], v. 61, n. 3, p. 539–542, 2012.

ROYER, W. E. Structures of red blood cell hemoglobins. *In*: **Blood and tissue oxygen carriers**. Berlin, Heidelberg: Springer, 1992. p. 87–116.

RUNCK, A. M.; MORIYAMA, H.; STORZ, J. F. Evolution of duplicated  $\beta$ -globin genes and the structural basis of hemoglobin isoform differentiation in *Mus*. **Molecular Biology and Evolution**, Oxford, v. 26, n. 11, p. 2521–2532, 2009.

SAN MAURO, D. *et al.* Life-history evolution and mitogenomic phylogeny of caecilian amphibians. **Molecular Phylogenetics and Evolution**, [s.l.], v. 73, n. 1, p. 177–189, 2014.

SAWAI, H. *et al.* Characterization of the heme environmental structure of cytoglobin, a fourth globin in humans. **Biochemistry**, [s.l.], v. 42, n. 17, p. 5133–5142, 2003.

SCHWARZE, K. *et al.* The globin gene repertoire of lampreys: convergent evolution of hemoglobin and myoglobin in jawed and jawless vertebrates. **Molecular Biology and Evolution**, Oxford, v. 31, n. 10, p. 2708–2721, 2014.

SCHWARZE, K.; BURMESTER, T. Conservation of globin genes in the “living fossil” *Latimeria chalumnae* and reconstruction of the evolution of the vertebrate globin family. **Biochimica et Biophysica Acta - Proteins and Proteomics**, [s.l.], v. 1834, n. 9, p. 1801–1812, 2013.

SCHWARZE, K.; SINGH, A.; BURMESTER, T. The full globin repertoire of turtles provides insights into vertebrate globin evolution and functions. **Genome Biology and Evolution**, Oxford, v. 7, n. 7, p. 1896–1913, 2015.

- SEIDL, F. *et al.* Genome of *Spea multiplicata*, a rapidly developing, phenotypically plastic, and desert-adapted spadefoot toad. **G3: Genes, Genomes, Genetics**, [s.l.], v. 9, p. 3909–3919, 2019.
- SESSION, A. M. *et al.* Genome evolution in the allotetraploid frog *Xenopus laevis*. **Nature**, [s.l.], v. 538, n. 7625, p. 336–343, 2016.
- SHU, Y. *et al.* Integrated analysis of mRNA and miRNA expression profiles reveals muscle growth differences between adult female and male Chinese concave-eared frogs (*Odorrana tormota*). **Gene**, [s.l.], v. 678, p. 241–251, 2018.
- SIDELL, B. D.; O'BRIEN, K. M. When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. **Journal of Experimental Biology**, [s.l.], v. 209, n. 10, p. 1791–1802, 2006.
- SLATER, G. S. C.; BIRNEY, E. Automated generation of heuristics for biological sequence comparison. **BMC Bioinformatics**, [s.l.], v. 6, n. 1, p. 31, 2005.
- SMITH, J. J. *et al.* A chromosome-scale assembly of the axolotl genome. **bioRxiv**, [s.l.], p. 373548, 2018.
- SONG, S. *et al.* Globins in the marine annelid *Platynereis dumerilii* shed new light on hemoglobin evolution in bilaterians. **BMC Evolutionary Biology**, [s.l.], v. 20, p. 165, 2020.
- STORZ, J. F. *et al.* Developmental regulation of hemoglobin synthesis in the green anole lizard *Anolis carolinensis*. **Journal of Experimental Biology**, [s.l.], v. 214, n. 4, p. 575–581, 2011.
- STORZ, J. F. *et al.* Oxygenation properties and isoform diversity of snake hemoglobins. **American Journal of Physiology-Regulatory, Integrative and Comparative Physiology**, Estados Unidos da América, v. 309, n. 9, p. R1178–R1191, 2015.
- STORZ, J. F. Gene duplication and evolutionary innovations in hemoglobin-oxygen transport. **Physiology**, [s.l.], v. 31, n. 3, p. 223–232, 2016.
- STORZ, J. F. **Hemoglobin**: insights into protein structure, function, and evolution. 1<sup>a</sup> ed. New York: Oxford University Press, 2019.
- STORZ, J. F.; OPAZO, J. C.; HOFFMANN, F. G. Phylogenetic diversification of the globin gene superfamily in chordates. **IUBMB Life**, [s.l.], v. 63, n. 5, p. 313–322, 2011.
- STORZ, J. F.; OPAZO, J. C.; HOFFMANN, F. G. Gene duplication, genome duplication, and the functional diversification of vertebrate globins. **Molecular Phylogenetics and Evolution**, [s.l.], v. 66, n. 2, p. 469–478, 2013.
- STREICHER, J. W. *et al.* Evaluating methods for phylogenomic analyses, and a new phylogeny for a major frog clade (Hyloidea) based on 2214 loci. **Molecular Phylogenetics and Evolution**, [s.l.], v. 119, p. 128–143, 2018.

SUN, Y.-B. *et al.* Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. **Proceedings of the National Academy of Sciences**, Estados Unidos da América, v. 112, n. 11, p. E1257–E1262, 2015.

TAKETA, F.; NICKERSON, M. A. Comparative studies on the hemoglobins of representative salamanders of the families Cryptobranchidae, Proteidae and Hynobiidae. **Comparative Biochemistry and Physiology Part B: Comparative Biochemistry**, Grã-Bretanha, v. 45, n. 3, p. 549–556, 1973a.

TAKETA, F.; NICKERSON, M. A. Hemoglobin of the aquatic salamander, *Cryptobranchus*. **Comparative Biochemistry and Physiology Part A: Physiology**, Grã-Bretanha, v. 46, p. 583–591, 1973b.

TORRES-SÁNCHEZ, M. *et al.* Multi-tissue transcriptomes of caecilian amphibians highlight incomplete knowledge of vertebrate gene families. **DNA Research**, Oxford, v. 26, n. 1, p. 13–20, 2018.

TRENT III, J. T.; WATTS, R. A.; HARGROVE, M. S. Human neuroglobin, a hexacoordinate hemoglobin that reversibly binds oxygen. **Journal of Biological Chemistry**, Estados Unidos da América, v. 276, n. 32, p. 30106–30110, 2001.

TRIFINOPOULOS, J. *et al.* W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. **Nucleic Acids Research**, [s.l.], v. 44, n. W1, p. W232–W235, 2016.

TU, N. *et al.* A large-scale phylogeny of Microhylidae inferred from a combined dataset of 121 genes and 427 taxa. **Molecular Phylogenetics and Evolution**, [s.l.], v. 126, p. 85–91, 2018.

VAN AARDT, W. J.; WEBER, R. E. Respiration and hemoglobin function in the giant African bullfrog *Pyxicephalus adspersus tschudi* (Anura: Pyxicephalidae) during rest, exercise and dormancy. **African Journal of Herpetology**, [s.l.], v. 59, n. 2, p. 173–190, 2010.

VIEIRA, H. F. *et al.* Some functional and structural properties of *Bufus paracnemis* and *Pipa pipa* hemoglobins. **Comparative Biochemistry and Physiology Part A: Physiology**, Grã-Bretanha, v. 73, n. 2, p. 197–200, 1982.

VIEITES, D. R. *et al.* A multigenic perspective on phylogenetic relationships in the largest family of salamanders, the Plethodontidae. **Molecular Phylogenetics and Evolution**, [s.l.], v. 59, n. 3, p. 623–635, 2011.

WALLACE, I. M. *et al.* M-Coffee: combining multiple sequence alignment methods with T-Coffee. **Nucleic Acids Research**, [s.l.], v. 34, n. 6, p. 1692–1699, 2006.

WEBER, R. E.; FAGO, A. Functional adaptation and its molecular basis in vertebrate hemoglobins, neuroglobins and cytoglobins. **Respiratory Physiology & Neurobiology**, [s.l.], v. 144, n. 2–3, p. 141–159, 2004.

WEBER, R. E.; WELLS, R. M. G.; ROSSETTI, J. E. Adaptations to neoteny in the salamander, *Necturus maculosus*. Blood respiratory properties and interactive effects of pH, temperature and ATP on hemoglobin oxygenation. **Comparative Biochemistry and**

**Physiology**, Grã-Bretanha, v. 80, n. 4, p. 495–501, 1985.

WHEELER, D. *et al.* An orphaned mammalian  $\beta$ -globin gene of ancient evolutionary origin. **Proceedings of the National Academy of Sciences of the United States of America**, Estados Unidos da América, v. 98, n. 3, p. 1101–1106, 2001.

WHEELER, D. *et al.* Linkage of the  $\beta$ -like  $\omega$ -globin gene to  $\alpha$ -like globin genes in an Australian marsupial supports the chromosome duplication model for separation of globin gene clusters. **Journal of Molecular Evolution**, [s.l.], v. 58, n. 6, p. 642–652, 2004.

XI, Y. *et al.* Gene expression and tissue distribution of cytoglobin and myoglobin in the Amphibia and Reptilia: possible compensation of myoglobin with cytoglobin in skeletal muscle cells of anurans that lack the myoglobin gene. **Gene**, [s.l.], v. 398, n. 1, p. 94–102, 2007.

XIONG, J. *et al.* The first transcriptome assembly of yenyuan stream salamander (*Batrachuperus yenyuanensis*) provides novel insights into its molecular evolution. **International Journal of Molecular Sciences**, [s.l.], v. 20, p. 1529, 2019.

XU, W. *et al.* Low concentrations of dihydrotestosterone induce female-to-male sex reversal in the frog *Pelophylax nigromaculatus*. **Environmental Toxicology and Chemistry**, [s.l.], v. 34, n. 10, p. 2370–2377, 2015.

YAN, F. *et al.* The Australasian frog family Ceratobatrachidae in China, Myanmar and Thailand: discovery of a new Himalayan forest frog clade. **Zoological Research**, [s.l.], v. 37, n. 1, p. 7–14, 2016.

YANDELL, M.; ENCE, D. A beginner's guide to eukaryotic genome annotation. **Nature Reviews Genetics**, [s.l.], v. 13, n. 5, p. 329–342, 2012.

YOSHIDA, N.; KAITO, C. De novo transcriptome of the African bullfrog *Pyxicephalus adspersus* for molecular analysis of aestivation. **bioRxiv**, [s.l.], 2019.

YU, G. *et al.* GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. **Methods in Ecology and Evolution**, [s.l.], v. 8, n. 1, p. 28–36, 2017.

YU, G. Using ggtree to visualize data on tree-like structures. **Current Protocols in Bioinformatics**, [s.l.], v. 69, n. 1, p. 1–18, 2020.

YUAN, Z. Y. *et al.* Natatanuran frogs used the Indian Plate to step-stone disperse and radiate across the Indian Ocean. **National Science Review**, Oxford, v. 6, n. 1, p. 10–14, 2019.

ZHANG, G. *et al.* Comparative genomics reveals insights into avian genome evolution and adaptation. **Science**, [s.l.], v. 346, n. 6215, p. 1311–1321, 2014.

ZHANG, P.; WAKE, D. B. Higher-level salamander relationships and divergence dates inferred from complete mitochondrial genomes. **Molecular Phylogenetics and Evolution**, [s.l.], v. 53, n. 2, p. 492–508, 2009.

ZHAO, L. *et al.* Transcriptome profiles of metamorphosis in the ornamented pygmy frog *Microhyla fissipes* clarify the functions of thyroid hormone receptors in metamorphosis. **Scientific Reports**, [s.l.], v. 6, p. 27310, 2016.

ZMASEK, C. M.; EDDY, S. R. A simple algorithm to infer gene duplication and speciation events on a gene tree. **Bioinformatics**, Oxford, v. 17, n. 9, p. 821–828, 2001.

ZMASEK, C. M.; EDDY, S. R. RIO: analyzing proteomes by automated phylogenomics using resampled inference of orthologs. **BMC Bioinformatics**, [s.l.], v. 3, p. 14, 2002.

ZWEIG, A. S. *et al.* UCSC genome browser tutorial. **Genomics**, [s.l.], v. 92, n. 2, p. 75–84, 2008.