

UNIVERSIDADE SANTO AMARO

**Programa de pós-graduação *stricto sensu* em
medicina veterinária nível mestrado**

Karina Camargo Arroyo dos Santos

Orientador: Dr. Guilherme José da Costa Silva

**CALIBRAÇÃO DA TAXA DE EVOLUÇÃO DO DNA BARCODE EM
PEIXES TELEOSTEI**

São Paulo

2021

Karina Camargo Arroyo dos Santos

**CALIBRAÇÃO DA TAXA DE EVOLUÇÃO DO DNA BARCODE EM
PEIXES TELEOSTEI**

Qualificação apresentada ao Programa de Pós-Graduação *Stricto Sensu* da Universidade Santo Amaro – UNISA, como requisito para defesa do título de Mestre em Medicina Veterinária e Bem Estar Animal
Orientador: Prof. Dr. Guilherme José da Costa Silva

São Paulo

2021

Karina Camargo Arroyo dos Santos

**CALIBRAÇÃO DA TAXA DE EVOLUÇÃO DO DNA BARCODE EM
PEIXES TELEOSTEI**

Qualificação apresentada ao Programa de Pós -Graduação *Stricto Sensu* da
Universidade Santo Amaro – UNISA, como requisito para defesa do título de Mestre
em Medicina Veterinária e Bem Estar Animal

Orientador: Prof. Dr. Guilherme José da Costa Silva

São Paulo, 17 de dezembro de 2021.

Prof. Dr. Guilherme José da Costa Silva

Prof. Dr. Luiz Henrique da Silva Nali

Prof. Dr. Luiz Henrique Garcia Pereira

Conceito final: _____

Dedico esta dissertação à minha filha Mia que diariamente me ensina, me motiva e me mostra o quão importante correr atrás dos meus sonhos é a melhor forma de cuidar e amar ela.

AGRADECIMENTO

Agradeço à Deus que sempre me guia e me conduz pelos desafios e novas aventuras, aos meus pais que estão sempre a postos para me apoiar, ao professor Guilherme que me propôs um desafio, me ajudou e guiou por todo processo com uma paciência infinita, demonstrando toda a essência de que um professor deve ser.

RESUMO

O presente estudo busca avaliar o gene Citocromo Oxidase C subunidade I (COI ou DNA Barcode) como potencial marcador de relógio molecular em Teleostei, tal como avaliar as interferências nas taxas de evolução nucleotídicas que esse gene está sujeito a depender das variáveis ecológicas e sistemáticas. Para estimar essa taxa de evolução foram buscadas na literatura árvores filogenéticas robustas nas quais foram empregadas técnicas de relógio molecular. Foram realizadas análises com dados de peixes marinhos e peixes dulcícolas coracoides. As sequências de COI das espécies foram obtidas no GenBank buscando a maior representatividade dos táxons das topologias referidas. Para ambas as análises o COI demonstrou estar saturado para grupos mais antigos, indicando que este tipo de análise deve ser feita para grupos mais recentes. Para obter as taxas do COI foi usado o programa Beast, sendo inserido neste as matrizes com as sequências, as árvores filogenéticas e pontos de calibração obtidos pelos artigos utilizados. Os resultados apontaram taxas de evolução com COI semelhante para os dois grupos de peixes, 1,17%/Ma em peixes marinhos e 1,05%/Ma em coracoides. Os valores encontrados foram validados em dois eventos geológicos conhecidos, a formação do Istmo do Panamá e da cachoeira das Sete-Quedas, relevando datas de divergência coerente com a literatura de 3,2 Ma e 1,6 Ma respectivamente. Não descartamos a possibilidade que fatores como temperatura e tamanho populacional possa influenciar nas taxas de evolução, visto que a taxa de evolução do COI em outros grupos pode ser bastante distintas. Os resultados encontrados nos dão indícios da possibilidade de se usar taxas evolutivas do COI para realizar relógios moleculares e trazer luz a eventos históricos pouco conhecidos e entender o processo de evolução, com a ressalva de se utilizar as taxas para linhagens próximas.

Palavras-chave: Relógio molecular. Citocromo Oxidase C Subunidade 1. COI. Taxa de substituição molecular.

ABSTRACT

The present study seeks to evaluate the gene Cytochrome Oxidase C subunit I (COI or DNA Barcode) as a potential universal molecular clock marker to Teleostei and assess the interferences in the nucleotide evolution rates that ecological and systematic variables influence this gene. To estimate this rate of evolution, robust phylogenetic trees were found in literature in which molecular clock techniques were employed. Analyses were performed with data from marine fish and coracoid freshwater fish. The COI sequences of the species were obtained from GenBank, seeking the most significant representation of the taxa of the referred topologies. The COI proved to be saturated for older groups for both analyses, indicating that this type of analysis should be done for more recent groups. To obtain the COI rates, the Beast program was used, inserting the matrices with the sequences, the phylogenetic trees and calibration points obtained by the articles used. The results showed similar COI evolution rates for both groups of fish, 1.17%/Ma in marine fish and 1.05%/Ma in coracoids. The values found were validated in two known geological events, the formation of the Panama Ístimo and the Sete-Quedas waterfall, showing divergence dates consistent with the literature of 3.2 Ma and 1.6 Ma, respectively. We do not rule out the possibility that temperature and population size factors may influence the evolution rates since the evolution rate of COI in other groups can be quite different. The results indicate the possibility of using COI evolutionary rates to make molecular clocks, shed light on little-known historical events, and understand the evolution process, except using rates for close lineages.

Key words: Molecular clock. Cytochrome Oxidase C Subunit 1. COI. Substitution molecular rate.

SUMÁRIO

1	INTRODUÇÃO	09
1.1	Históricos dos relógios moleculares	09
1.2	Formas de calibrar um relógio molecular	10
1.2.1	Fósseis	11
1.2.2	Eventos geológicos	12
1.2.3	Calibrações secundárias	12
1.3	Prós e contras do relógio molecular	13
1.3.1	Variação da taxa de evolução dos genes em diferentes grupos	14
1.4	Saturação	15
1.5	DNA Barcoding	15
1.6	Citocromo Oxidase C subunidade 1	15
1.7	Evolução dos peixes teleostei.....	16
2	OBJETIVO	18
2.1	Objetivos gerais.....	18
2.2	Objetivos específicos.....	18
3	MATERIAL E MÉTODO	18
4	RESULTADOS.....	22
5	DISCUSSÃO.....	26
	REFERÊNCIAS BIBLIOGRÁFICAS	29
	ANEXO A - Short communication enviado à revista BMC Evolutionary	
	ANEXO B – Sequencias do COI utilizadas e seus códigos de acesso no GenBank.	
	ANEXO C - Tolologia de peixes marinhos e coracoides calibradas.	
	ANEXO D - Testes das taxas de evolução do COI.	

1. INTRODUÇÃO

1.1 Histórico dos relógios moleculares.

Datar um evento evolutivo de uma espécie sempre foi um desafio para os estudiosos do tema. Ao longo dos séculos a ciência que sempre se destacou para esse fim foi a paleontologia com a datação de fósseis por meio de diferentes técnicas. Contudo sabe-se que o registro fóssil é falho e possui várias lacunas temporais e de representatividade de táxons. Embora muitas dessas lacunas tenham sido preenchidas ao longo das últimas décadas resta-se muito a descobrir e desvendar sobre o histórico evolutivo das espécies.

Uma alternativa pensada para auxiliar na datação de eventos evolutivos surgiu nas décadas de 1960 e 70 com a advento de pesquisas que sugeriram que um possível relógio molecular que fosse capaz de determinar o tempo de divergência de evolução entre espécies através da análise molecular de proteínas. Emile Zuckerkandl e Linus Pauling (1962) propuseram a possibilidade de conseguir se estimar de forma aproximada o tempo de divergência entre duas espécies desde que compartilharam um mesmo ancestral utilizando como marcador as sequências de hemoglobina. Esses autores observaram uma surpreendente taxa linear na quantidade de diferenças de aminoácidos durante o tempo evolutivo, o que é a base da teoria do relógio molecular. Para essa análise as informações utilizadas eram: o número de diferenças entre cadeias correspondentes em diferentes espécies e a idade geológica que o antepassado em comum das diferentes espécies em questão tivesse sido considerado vivo (Zuckerkandl e Pauling, 1962).

Seguindo essa linha de pesquisa, Kimura e Ohta (1971) sugeriram que a evolução dos aminoácidos em proteínas era causada principalmente pela fixação aleatória de mutações neutras (que não provocam alterações fenotípicas) do que por mutações pela seleção natural, explicando desta forma o conceito de uma taxa constante de evolução. Entende-se portanto, que a própria pressão seletiva sobre uma característica poderia distorcer as taxas de evolução de um marcador molecular.

De fato, desde o princípio, Zuckerkandl e Pauling (1962) já consideravam que existiam alguns fatores que poderiam afetar no cálculo do tempo de divergência entre espécies ancentrais, como mudanças de temperatura e da intensidade de irradiação ionizante, assim como o tamanho da população. Para demonstrar este último fator, Ohta e Kimura (1971) observaram que quanto menor for o tamanho da população e menor o tempo de geração, maior será a taxa de evolução.

Além das considerações acima, Dickerson (1971) concluiu que há diferenças na taxa de evolução entre diferentes proteínas em uma mesma espécie. Ao estudar o citocromo c, descobriram que este gene tem taxas de mutações menores quando comparadas com hemoglobinas e fibrinopeptídeos. Em outras palavras, regiões diferentes do genoma podem evoluir em taxas distintas. Desde então nas últimas cinco décadas, diferentes técnicas de datação molecular têm sido propostas, a fim de, principalmente, resolver a questão da variabilidade das taxas na datação molecular. Atualmente tem ficado cada vez mais claro que taxas constantes de evolução molecular podem na verdade ser a exceção e não a regra como se presumia pelo termo “relógio molecular” (Welch e Bromham, 2005). Diversos métodos analíticos buscam corrigir essas variações minimizando os desvios, porém comparar esses métodos é difícil porque existem muitas diferenças nas formas de se implementar cada técnica (Welch e Bromham, 2005), como iremos discutir os métodos mais adiante.

1.2 Formas de calibrar um relógio molecular.

Devida a natureza confusa das taxas de evolução e da análise dos tempos de divergência é preciso que o relógio molecular seja calibrado de forma independente usando informações da escala de tempo evolucionária, porém esta é uma ciência dinâmica e dependendo do grupo a ser estudado pode ser necessário uma combinação de métodos (Hipsley e Muller, 2014).

Para se ter uma idéia dos métodos de calibração mais utilizados, Hipsley e Muller (2014), analisaram 697 publicações entre os anos de 2008 e 2012 sobre o tema e pode constatar a seguinte distribuição: calibração por fósseis 52%, eventos

geológicos e calibrações secundárias 15% cada, taxa de substituição 12%, amostras 4%. Com fósseis e calibrações secundárias mostrando crescimento ao longo dos anos, eventos geológicos e taxa de substituição mostrando leve decréscimo, amostras mesmo mostrando baixo percentual nesses anos, no último ano da pesquisa (2011) teve aumento de 0 a 9%.

1.2.1 Fósseis

Como já citamos os fósseis são os mais usados para a calibração do relógio molecular. Dentro das calibrações fósseis temos dois tipos mais usados, a mais tradicional é a calibração nos nós da árvore filogenética e a mais recente e inovadora calibração nos ramos. A primeira usa a premissa que os fósseis informam a idade mínima do clado baseado na idade do fóssil mais antigo encontrado (Donoghue e Yang, 2016). Já a segunda faz uma combinação de informações moleculares e morfológicas de ambas espécies existentes e extintas (Pryon, 2011).

Seguindo a linha da calibração em ramos, um novo método descrito por Heath et al (2014), o “Fossilized birth-death” integra informações fósseis ao processo de diversificação baseado nas linhagens, superando limitações dos métodos de calibração por densidade.

Apesar das controvérsias, O’Reilly e Donoghue (2016), tentaram demonstrar que os dois métodos podem ser complementares e se usados juntos podem ser mais precisos do que quando usados sozinhos.

Pode-se concluir que a descoberta e descrição contínua de fósseis, assim como o desenvolvimento de métodos para se incorporar as calibrações nas análises, quando usados cuidadosamente, podem ser de valiosa contribuição para o conhecimento do processo evolutivo (Ho, 2014).

1.2.2 Eventos geológicos

A calibração por evento geológico assume que a especiação se baseia principalmente na vicariância, porém novas evidências indicam que a especiação por dispersão pode ser mais comum em vários grupos (Kodandaramaiah, 2011). Alguns autores como Kodandaramaiah (2011) sugere que calibrações baseadas somente em vicariância deveriam ser evitadas ou então ao se fazer uma combinação com calibrações fósseis deve-se reportar os resultados com e sem a inclusão dos eventos geológicos.

Porém um novo sistema de calibração se mostra promissor, baseado em mudanças geológicas no padrão de drenagem dos rios. Essa calibração se baseia no fato da vicariância em rios serem substancialmente mais rápidas do que as taxas tradicionalmente usadas para peixes, porém se mostram consistentes com outras inferências baseadas em pontos de calibrações geológicas mais recentes (Waters et al, 2007).

1.2.3 Calibrações secundárias

Calibrações secundárias derivam de calibrações utilizadas em publicações anteriores, o que por um lado pode facilitar os projetos e estudos sobre evolução, por outro lado pode induzir a equívocos. Usualmente as informações são usadas como fatos e usadas inapropriadamente, principalmente quando as referências usadas têm um único ponto de calibração. Essa ilusão de precisão quando alcançada através de conversão de estimativas estatísticas, leva a resultados errados (Graur e Martin, 2004) e deve ser usada com cautela.

Por isso é aconselhável consultar estudos com múltiplas calibrações e utilizar metodologias que considerem a variação da taxa e incerteza do prior. E deve-se reportar explicitamente intervalos de confiança para essas novas datas, para que os próximos a utilizarem possam utilizar essas informações de forma apropriada (Hispley e Muller, 2014).

1.3 Prós e contras do relógio molecular.

O relógio molecular trouxe uma nova forma de estudo da escala de tempo na história da vida e transformou a biologia evolucionária ao estimar os tempos de eventos evolucionários antigos e também ao explorar os mecanismos e processos de evolução (Bromham e Penny, 2003). Apesar da fantástica proposta, há muita controvérsia, principalmente porque batem de frente com estimativas de referências mais tradicionais como fósseis (Smith e Peterson, 2002).

A aplicação mais prática do relógio molecular está no estudos de doenças virais emergentes, diferentemente discutido anteriormente, para os vírus não se usa fósseis para calibração, mas sim com amostras de tecidos de pacientes que estejam armazenadas ou até mesmo com várias amostras coletas de tempos em tempos do mesmo paciente no caso dos retrovírus. Os relógios moleculares para vírus pode ajudar a reconstituir a história de epidemias a partir de amostras de vírus já extintos (Bromham, 2002).

Com o estudo dos relógios moleculares em vírus também se pode ter uma melhor compreensão de como diferentes taxa e genes se comportam, por exemplo os vírus da Dengue e da influenza possuem uma taxa de evolução constante, diferente de outros vírus, essas informações são importantes para compreender e lidar com a evolução dessas doenças (Twiddy et al., 2003; Fitch et al., 1991).

Para se utilizar o relógio molecular de forma adequada, há de se considerar diversos fatores que podem influenciar nos resultados como, a variação das taxas de evolução entre espécies e também entre genes, tipos de calibração, o tamanho da população, entre outros (Bromham e Penny, 2003). Destacamos a seguir a principais variantes que devemos nos atentar quando aplicarmos técnicas de relógio moleculares.

1.3.1 Variação da taxa de evolução dos genes em diferentes grupos.

Como já comentado, há vários fatores que influenciam sobre a taxa de evolução, podendo atuar sobre o tempo ou nas espécies, são variantes que podem levar a efeitos nas linhagens, que resulta em variação das taxas entre espécies.

O tamanho da população pode influenciar já que mutações deletérias podem passar despercebidas em grandes populações o que pode não ocorrer em populações menores, sendo mais fácil a fixação destas mutações (Otha 1987).

O próprio metabolismo pode influenciar, a taxa metabólica das espécies pode influenciar porque quanto mais acelerado é o metabolismo de determinada espécie, maior é estresse oxidativo, produção de radicais livres, consequentemente maior dano no DNA, aumentando a concentração de mutações, o que por fim pode influenciar na taxa (Rand, 1994)

A taxa de evolução é maior em espécies com tempos de geração mais curtos, por exemplo, ao comparar roedores e humanos, os roedores são capazes de ter novas gerações em apenas três meses, enquanto os humanos demoram anos para atingir sua maturidade reprodutiva (Martin e Palumbi, 1993).

Fatores intrínsecos da análise tanto como forma de calibração e o fato de que cada gene evolui em tempos diferentes, podem tal como influir numa datação por relógio molecular. Um gene de mutação rápida serviria para calcular eventos cladogenéticos mais recentes e genes de taxa lenta serviriam para calibrar relógios mais antigos. A escolha equivocada, por exemplo um marcador lento para um evento recente pode não ser resolutivo para as análises. Já um gene de evolução rápida para eventos antigos pode causar distorções na estimativa de tempo, sendo a mais sensível a saturação. (Frankham *et al.* 2004) portanto a taxa de evolução de cada gene é diferente em uma mesma espécie e por isso é importante encontrar um gene para analisar que melhor represente a evolução em geral de todas as espécies.

1.4 Saturação

As substituições utilizadas em relógio molecular ocorrem durante o tempo e de forma aleatória e aproximadamente constante, porém quando essas substituições ocorrem na mesma posição pode-se perder informações valiosas sobre a quantidade de mutações e a taxa de evolução de determinada espécie, quando isso acontece a sequência fica saturada e se for utilizada para análises de relógio molecular, provavelmente resultará em uma estimativa equivocada de tempo (Xia et al., 2003).

A saturação é um fenômeno probabilístico e estocástico, portanto, quanto maior o tempo de divergência dos táxons que compõem uma matriz, maior são as chances que os marcadores empregados estejam saturados. Desta forma é de extrema importância saber se uma matriz está demasiadamente saturada para ser analisada e para isso há métodos probabilísticos que estimam se a matriz de dados moleculares possui ou não sinal filogenético confiável (Xia et al., 2003).

1.5 DNA Barcoding

Com os avanços no estudos sobre o DNA, Herbert et al, (2003) propôs algo revolucionário, identificar táxons através de um “códigos de barras” da vida, através de um sistema que emprega as sequências de DNA na identificação de espécies. Para isso era preciso escolher um gene que servisse de base para um sistema de bioidentificação global dos animais e esse gene é o citocromo oxidase c subunidade 1 (COI).

1.6 Citocromo Oxidase C subunidade 1.

O COI é um gene mitocondrial, que codifica parte uma proteína de membrana mitocondrial que faz parte da cadeia transportadora de elétrons. É um gene de aproximadamente 1100 bp e está presente na cadeia leve do DNA Mitocondrial. Apesar de sua importante função metabólica esse gene ganhou grande notoriedade ao ser selecionado como marcador molecular do projeto multinacional Barcoding of Life. Esse projeto ambicioso busca por meio de um único marcador catalogar todas

espécies vivas utilizando o COI como o DNA barcode. A princípio esse gene foi o escolhido porque tem uma taxa de evolução rápida o suficiente para acumular mutações visíveis entre espécies distintas, mas ao mesmo tempo é lenta o suficiente para não mudar demais a sequência a ponto de as espécies perderem sua identidade genética. Evidentemente esse projeto demonstrou que um único marcador seria insuficiente para delimitar cada espécie vivente e em alguns grupos outros marcadores foram elencados para esse fim (Herbert et al., 2003).

Apesar disso, o COI é o gene dominante como DNA barcode e no ano de 2003 Herbert e colaboradores estimaram que uma diferenciação de 2% ou mais na sequência do COI, significa que estamos tratando de espécies distintas, isso com uma precisão de 95%. Apontando um novo potencial para esse marcador sendo largamente empregado em estudos taxonômicos, embora haja ressalvas.

A taxa de evolução desse marcador não é conhecida em peixes, tal como não é conhecida a variação dessas taxas em diferentes grupos de organismos. Por se tratar de um marcador altamente utilizado em grupos animais esse marcador também poderia se tornar em um elemento de relógio molecular universal, altamente utilizável em situações onde não se conhece pontos de calibração confiáveis.

1.7 Evolução dos peixes teleostei

A fauna de peixes da América do Sul foi formada a mais de 100 milhões de anos, a região tropical do continente foi a única do mundo a escapar das extinções em massa e reviravoltas bióticas associadas com o esfriamento climático Cenozóico, a formação de zonas boreais em altas latitudes e aridificação em vários locais de latitudes equatoriais. Esse grupo diverso de organismos conta só na região com mais de 10000 espécies distribuídas entre ambiente marinho e de água doce (Reis *et al.* 2016). O grande número de espécies da região deve-se sobretudo a um extenso histórico biogeográfico que promoveu muito isolamento e consequentemente especiação (Lundberg *et al.* 1998) (Albert & Reis 2011). Os Teleosteis datam de um tempo pretérito à quebra do supercontinente Gondwana e com a fragmentação desse

imenso bloco de continente impactou o histórico e diversificação de peixes tanto marinhos quanto dulcícolas (Ribeiro 2006) (Melo *et al.* 2021).

Os peixes de água doce em particular, devido à sua fisiologia não conseguem atravessar barreira físicas, portanto tem sua capacidade de dispersão limitada ficando restritos aos rios e lagos (Bizerril 1998). Por isso a historia da evolução de rios se equipara com a evolução de linhagens de peixes que neles habitam. Nesse sentido, o estudo do processo evolutivo dos corpos d'água pode auxiliar no estudo das evoluções dos peixes e vice versa. Para isso, fatores ecológicos como altitude, precipitação, temperatura, cobertura florestal, composição química da água também devem ser levados em consideração.

Ao contrário dos peixes de água doce, na qual a especiação ocorre pela separação por barreiras geológicas, nos peixes marinhos a especiação normalmente é simpátrica ou parapátrica que se dá pela adaptação ecológica de barreiras semi permeáveis (Thacker 2017). Tendo em vista a América do Sul pode se observar a grande distinção entre espécies dos Oceanos Pacífico e Atlântico, isso se dá tanto pelo histórico geológico quanto pela composição taxonômica, especialmente ao nível de espécies.

Levando-se em conta que o grupo dos peixes são excelentes modelos de estudos evolutivos, pois estão sujeitos a alterações ambientais e de mudanças de cursos d'água, o presente trabalho busca por meio de estudo de grupos de peixes, com sistemática bem conhecida, estimar a taxa evolutiva do COI buscando reconhecer o potencial desse marcador como relógio universal em Teleostei, tal como avaliar as interferências nas taxas de evolução nucleotídicas que esse gene está sujeito a depender das variáveis ecológicas e sistemáticas.

2. OBJETIVO

2.1. Objetivos gerais

Identificar a taxa de evolução do COI em diferentes linhagens de peixes Teleostei avaliando a eficiência desse marcador como um relógio universal desses organismos.

2.2 Objetivos específicos

Calcular a taxa de evolução do COI em Teleostei marinhos.

Calcular a taxa de evolução do COI em peixes Teleostei de água doce.

Testar a taxa de evolução obtida de peixes marinhos utilizando em relógio molecular para um evento geológico marinho conhecido.

Testar a taxa de evolução obtida de peixes Teleostei de água doce utilizando em relógio molecular para um evento geológico fluvial conhecido.

3. MATERIAL E MÉTODO

Coleta de dados e alinhamento e edição da matriz

Foi pesquisado artigos que trouxessem informações de árvores filogenéticas confiáveis e com informações sobre calibrações fósseis daquelas espécies estudadas. Para isso foram selecionados trabalhos baseados em filogenômica com excelente suporte estatístico dos nós internos.

Após escolha dos artigos concluída foi feita a busca das sequências de COI no GenBank (NCBI, 1988) em formato FASTA das espécies descritas nos artigos e todos esses dados colocados em uma planilha Excel com o nome da espécie, sequência COI, identificação da sequência no GenBank e referência. As espécies não encontradas foram excluídas da topologia. (dados das sequências são apresentados no anexo VI)

As sequências foram alinhadas no programa Muscle (Edgar, 2004) no servidor online (disponível no link <https://www.ebi.ac.uk/Tools/msa/muscle/>) sob parâmetros de default. Visto que o COI é um gene mitocondrial codificante, esse não apresenta íntrons, ou stop códons e dificilmente apresenta mutações frameshift. Sequências que apresentassem tais elementos foram retiradas das análises. No programa Bioedit (Hall, 1999) o alinhamento foi checado manualmente retirando sequências portadoras de incongruências tal como editando a ponta da matriz para que todas as sequências finais tivessem o mesmo comprimento. Todo esse procedimento foi realizado separadamente para peixes marinhos e Characoidei produzindo duas matrizes distintas.

O mesmo procedimento de coleta de sequências, alinhamento e edição foi feito para elaboração de outras duas matrizes, estas para serem utilizadas em testes que validem os relógios moleculares encontrados. Uma das matrizes com peixes do Alto Paraná/Bacia do Prata e outra com peixes do Pacífico/Caribe.

Análise de saturação

Utilizamos dois testes de saturação distintos para cada matriz, ambos no Programa DAMBE. O teste de Xia et al. (2001) foi realizado com parâmetros de default. Para isso a matriz foi importada ao programa e calculado os valores de *iss* (valor observado de substituições) e do *iss.c* (valor máximo de substituições possíveis), se o *iss* for maior que o *iss.c* significa que as sequências estão saturadas e se o *iss* for menor que o *iss.c* as sequências não estão saturadas e são úteis para análises. O teste gráfico da relação “Transition and Transversion vs genetic distance” foi calculado no DAMBE e o valor de R^2 das linhas de tendência das transições e transversões foram estimados no Excel. Nessa segunda análise interpreta-se que se

o valor de R^2 das transições e transversões foram muito discrepantes a matriz está com tendências a saturação (Page & Holmes 1988).

Topologias e pontos de calibração empregados

Para as estimativas do tempo de divergência e consequentemente a taxa evolutiva do COI foram utilizadas topologias de grupos consolidados, no caso baseadas em análises filogenômicas com elevada sustentação estatísticas das relações filogenéticas. Foram feitas análises de peixes marinhos, inspiradas na topologia de peixes marinhos de Alfaro et al., (2018) e Coracoides inspiradas na topologia de Melo et al., (2021).

O programa Mega X foi usado para montar as árvores baseadas nas sequências do COI obtidas pelo processo descrito acima, porém a árvore somente com as sequências do COI não são bem estruturadas, ao contrário das árvores presentes nos artigos que usam vários genes para sua base, sendo mais confiáveis. Por causa disso para que as topologias no presente estudo representem corretamente a filogenia, foram reproduzidas e editadas no TreeView 3.1 (Page, 1996) de acordo com o artigo relacionado.

Os pontos de calibração utilizados foram obtidos dos próprios artigos que foram utilizados como base respectivamente nos dois grupos estudos, são apresentados na tabela a seguir.

Quadro 1 - pontos de calibração para peixes marinho.

Nome do Fóssil	Ponto da calibração	Data do fóssil	Referência do fóssil
<i>Aipichthys minor</i>	P1	98.0 Ma.	Ogg, et al. (2012)
<i>Homonotichthys dorsalis</i>	P2	93.9 Ma.	Patterson (1964)
<i>Mcconichthys longipinnis</i>	P3	63.1 Ma.	Borden, et al. (2013)
<i>Eocoelopoma portentosum</i>	P5	54.17 Ma.	Alfaro et al. (2018)
<i>Rhamphexocoetus volans</i>	P6	49.0 Ma.	Alfaro et al. (2018)
<i>Eastmanalepes primaevus</i>	P9	49 Ma.	Alfaro et al. (2018)
<i>Chaetodontidae indet</i>	P10	29.62 Ma.	Alfaro et al. (2018)
<i>Tarkus squirei</i>	P11	49 Ma.	Alfaro et al. (2018)
<i>Calotomus preisli</i>	P12	11.9 Ma	Alfaro et al. (2018)

Quadro 2 - Pontos de calibração para peixes coracoides.

Nome do fóssil	Ponto de calibração	Data do fóssil	Referência do fóssil
Alestidae	P1	50 Ma.	de la Peña Zarzuelo (1996)
<i>Leporinus scalabrinii</i>	P2	7,5 Ma.	Bogan et al. (2012)
<i>Cyphocharax mosesi</i>	P3	23,8 Ma.	Travassos e Santos (1955)
<i>Brycon avus</i>	P4	23,8 Ma.	Woodward (1898)
Stervadiinae	P5	23,8 Ma.	Weiss et al. (2012)

Estimativa da taxa de evolução nucleotídica.

O pacote do programa Beast 1.8.1 (Drummond e Rambaut, 2007) foi utilizado para rodar as análises de relógio molecular. Primeiramente no Beaut, importando a matriz final, a árvore filogenética e também os pontos de calibração fósseis. No modelo de evolução nucleotídica utilizado foi o modelo GTR de substituição e os fatores de correção Gamma e invariant sites. O tipo de relógio escolhido foi o relógio relaxado e o tipo de distribuição lognormal. Nas configurações da árvore foi escolhido

o modelo de especiação Yule Process. O tipo de operação foi com topologia fixa (predefinida por filogenômicas como descrito anteriormente) e finalmente o MCMC escolhido foi de 20 milhões no comprimento da cadeia. Depois de salvo essas informações fez-se a análise no Beast e por último analisado os resultados pelo Tracer onde buscou-se identificar a taxa de evolução nucleotídica tal como a variância encontrada.

Validação da taxa de evolução

Para testar a taxa de evolução obtida dos peixes marinhos e do coracoides, foi rodado um relógio molecular no programa Beast 1.8.1 (Drummond e Rambaut, 2007) com a taxa de cada grupo, foi usado dois conhecidos eventos de vicariação como base da validação. 1) A separação das populações pelo surgimento da cachoeira das Sete-Quedas que isolou as populações de peixes do alto Paraná do resto da baía do Prata (Costa-Silva et al., 2018), foi utilizada a taxa encontrada para coracoides no presente estudo; 2) Separação entre populações pelo fechamento do istmo do Panamá, que separou os peixes do Caribe e do Pacífico Oriental, foi utilizada a taxa encontrada para peixes marinhos no presente estudo (Thacker 2017). Como programação foi escolhido o relógio estrito e taxa de evolução molecular constante.

4. RESULTADOS

No GenBank foram encontradas sequências de COI e após alinhamento foram excluídas sequências que apresentavam incongruências restando uma matriz de 102 táxons de peixes marinhos (presentes em Alfaro et al., 2018) e outra com 166 referentes à peixes coracoides (presentes em Melo et al. 2021). Características finais de cada matriz obtida são apresentadas na tabela 1:

Tabela 1 - dados merísticos das matrizes até então analisadas

Característica da matriz	Peixes marinhos	Peixes coracoides
Número de pares de base	422	369
Número de caracteres variáveis	219	204
Conteúdo A	22%	23%
Conteúdo C	30%	30%
Conteúdo G	17%	16%
Conteúdo T	29%	31%

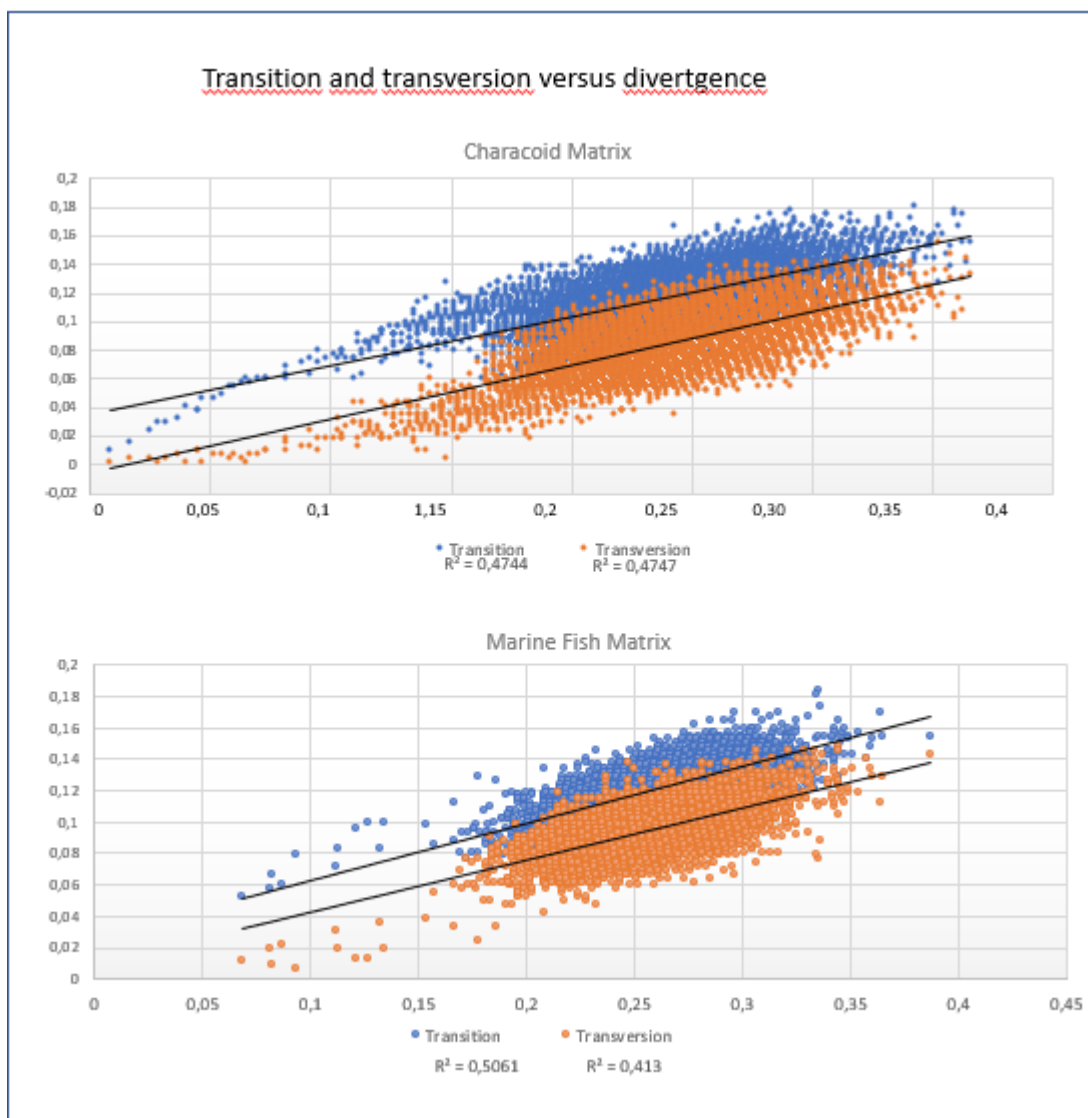
Quanto à análise de saturação segundo Xia et al 2001, feito no programa DAMBE, o valor do índice de saturação em ambas as matrizes foi menor do que o valor máximo de substituições possíveis. Os valores estão demonstrados na tabela 2.

Tabela 2 - resultado das análises de saturação das duas matrizes analisadas.

Peixes marinhos			Coracoide		
NumOTU	ISS	ISS:C	NumOTU	ISS	ISS:C
4	0,284	0,790	4	0,266	0,786
8	0,285	0,744	8	0,260	0,739
16	0,289	0,706	16	0,260	0,695
32	0,294	0,693	32	0,260	0,688

Porém na análise gráfica da saturação há indicação de que para grupos mais antigos o COI está saturado, com R^2 abaixo de 0,7 conforme observado na Figura 1, abaixo:

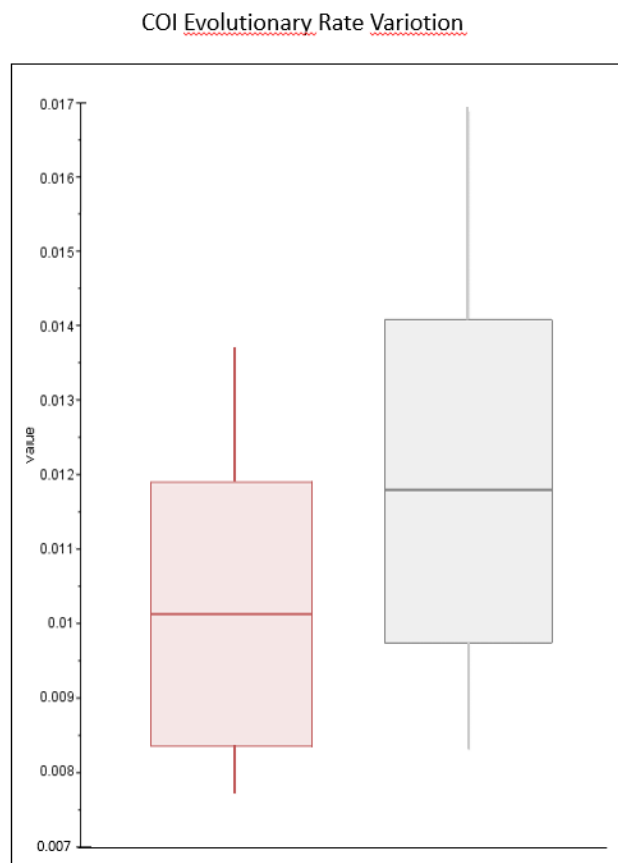
Figura 1 - Resultado em gráficos do teste de saturação, acima a saturação de peixes coracoides e abaixo peixes marinhos, revelando-se saturado em ambos os grupos com R^2 abaixo de 0,7.



Taxa de evolução do COI.

O resultados das análises no programa Beast revelou taxa de evolução de 1,17% a cada milhão de anos para os peixes marinhos (figura 1) e de 1,05% a cada milhão de anos para os peixes coracoides (Figura 2).

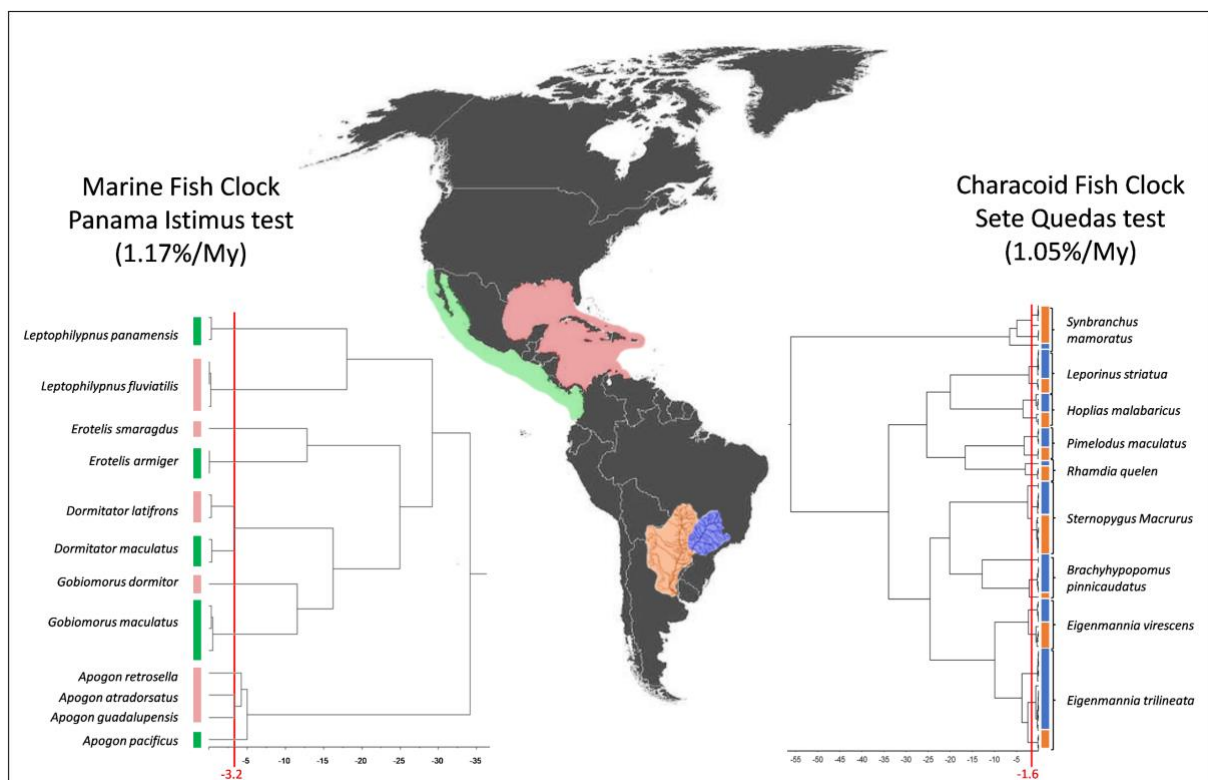
Figura 2 - resultados das taxas evolutivas do COI analisadas pelo programa Beast 1.8.1, o box em vermelho representa peixes coracoides com 1,05%/Ma de taxa média e o box cinza representa peixes marinhos com 1,17%/Ma de taxa evolutiva média.



Validação das taxas de evolução do COI

O teste para validação demonstrou que as taxas encontradas no presente trabalho trazem datas de cladogênese consistentes com conhecidos eventos geológicos. Com a taxa de 1,17%/Ma dos peixes marinhos calculamos através de relógio molecular o valor de 3,2 milhões de anos como tempo divergência para separação do istmo do Panamá. Já para a taxa de 1,05%/Ma para peixes coracoides foi possível chegar a um tempo de divergência de 1,6 milhões de anos para as espécies dos rios alto Paraná e baixo Paraná/Paraguai.

Figura 3 - Figura ilustra o conjunto dos resultados e sua validação, no mapa pode-se observar em rosa e verde a área das populações de peixes separados pelo fechamento do istmo do Panamá, à esquerda indicada a taxa de evolução do COI para peixes marinhos de 1,17%/Ma, taxa que foi utilizada para determinar o tempo de divergência de 3,2 milhões de anos através de relógio molecular, indicada pela linha vermelha na topologia (à esquerda), demonstrando a separação das espécies do Panamá. A cores azul e laranja do mapa representam os rios alto Paraná e baixo Paraná/Paraguai e suas populações que foram separadas pelo surgimento da cachoeira das Sete-Quedas, à direita indicada a taxa de evolução do COI para peixes coracoides de 1,05%/Ma, taxa que foi utilizada para determinar o tempo de divergência de 1,6 milhões de anos através do relógio molecular, indicada pela linha vermelha na topologia à direita, demonstrando a separação da espécie da bacia do Prata.



5. DISCUSSÃO

A taxa de evolução do COI pode estar sujeita a diferentes fatores que podem variar entre linhagens, como tamanho da população, metabolismo, tempo de geração, diferenças de taxa entre diferentes proteínas, saturação e calibração. Não se pode descartar a possibilidades de fatores que possam influenciar a taxa de evolução como

temperatura e tamanho da população, embora Orton et al. (2018) sugere que a taxa de evolução do COI não seja significantemente divergente nas diferentes latitudes.

É importante observar a grande variação da taxa evolutiva do COI entre espécies animais, por exemplo, foi encontrada taxa de 1,21%/Ma em bivalves (Marko, 2002), 1,4%/Ma em camarões (Knowlton e Weigt, 1998), 3,54% em insetos (Papadopoulou et al., 2010). Estes fatos associados com a saturação observada em nossas análises indica a importância de calcular os relógios em linhagens filogeneticamente próximas.

Outros genes mitocondriais como o citocromo B foram analisados em peixes da família ciprídeos e resultaram em uma taxa de aproximadamente 0,76%/Ma (Zardoya e Doadrio, 1999) e do DNA mitocondrial de animais pecilotérmicos com taxa de evolução de 0,7% (Martin e Palumbi, 1993), sendo portando aproximadamente a metade da taxa encontrada para o COI. Esse dado demonstra que nem todo o genoma mitocondrial está sujeito às mesmas taxas de evolução.

Quanto aos testes de validação foram encontrados resultados consistentes aos eventos geológico já conhecidos, usando a taxa de 1,05%/Ma para peixes marinhos chegou-se ao valor de 3,2 milhões de anos como ponto de divergência, o que condiz com o que se é suposto pela estimativa geológica de aproximadamente 3 milhões de anos da formação do Ístimo do Panamá (O'Dea et al., 2016). E o valor de divergência obtido para a Bacia da Plata de 1,6 também se mostra consistente com o descrito sobre o surgimento da cachoeira das Sete-Quedas há aproximadamente 1,5 milhões de anos (Barotelli, 2004). Esses resultados de validação são de extrema importância já que ilustram de forma fidedigna dois eventos geológicos reconhecidos e fundamentados, podendo assim ser usado para trazer luz a eventos pouco conhecidos, entender se em determinados locais espécies estão sofrendo processo de vicariação, funcionando portanto como um relógio molecular universal, ao menos em eventos geológicos recentes.

Ao fim desta pesquisa tivemos a satisfação de alcançar os objetivos traçados, primeiramente ao determinar taxas de evolução do COI em dois grupos de peixes teleosteis, peixes marinhos e coracoides, e posteriormente validando essas taxas encontradas para determinar o tempo de divergência em eventos geológicos conhecidos, estes resultados são de extrema relevância podendo-se concluir que

essas taxas possam ser utilizadas para linhagens próximas, nos desperta o interesse em expandir esta pesquisa, para aprofundar e entender ainda mais o comportamento do COI como marcador de relógio molecular.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERT, J.S.; REIS, R. (2011) Historical Biogeography of Neotropical Freshwater Fishes. U. O. C. PRESS (Ed). London, 408 pp. Available from: <https://market.android.com/details?id=book-HbLcMtCFdtwC%5Cnhttps://books.google.com/books?hl=en&lr=&id=HbLcMtCFdtwC&oi=fnd&pg=PP1&dq=Historical+Biogeography+of+Neotropical+Freshwater+Fishes&ots=20tCk2mugl&sig=kKfHxL6C5lGJwgpCrPc5fNqkGzI> LB - IRAV
- ALFARO, Michael E., et al. Explosive diversification of marine fishes at the Cretaceous-Palaeogene boundary. **Nature ecology and evolution**. v.2, n.4, p. 688-696, 2018.
- ARCILA, Dahiana et al. An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). **Molecular Phylogenetics and Evolution**. V.82, p.131-145, 2015.
- BIZERRIL, C. A Ictiofauna da Bacia do Rio Paraíba do Sul. Biodiversidade e Padrões Biogeográficos. **Brazilian Archives of Biology and Technology** 41, 67–81, 1998.
- BOGAN S., SIDLAUSKAS B., VARI R.P., AGNOLIN F. *Arrhinolemur scalabrinii* Ameghino, 1898, of the late Miocene—a taxonomic journey from the Mammalia to the Anostomidae (Ostariophysi: Characiformes). **Neotrop. Ichthyol.** 10:555–560, 2012.
- BORDEN, W. C., GRANDE, T. & SMITH, W. L. Comparative osteology and myology of thecaudal fin in the Paracanthopterygii (Teleostei: Acanthomorpha). Mesozoic fishes 419–455(2013). Grande, T., Borden, W. C. & Smith, W. L. Limits and relationships of Paracanthopterygii: A molecular framework for evaluating past morphological hypotheses. Mesozoic fishes 5, 385–418 2013.
- BROMHAM, L. The human zoo: endogenous retroviruses in the human genome. **Trends Ecol. Evol.** 17, 91–97 (2002).
- BROMHAM, Lindell; PENNY, David. The modern molecular clock. **Nature Reviews Genetics**. V.4, n.3, p.216-224, mar. 2003.
- DE LA PEÑA ZARZUELO A. 1996. Characid teeth from the Lower Eocene of the Ager Basin (Lérida, Spain): paleobiogeographical comments. Copeia, 746–750, 1996.
- DICKERSON, Richard E. The structure of cytochrome c and the rates of molecular evolution. **Journal of Molecular Evolution**. California, v.1, n.1, p. 26-45, jan. 1971.
- DONOGHUE, Philip C.J.; YANG, Ziheng. The evolution of methods for establishing evolutionary timescales. **Philosophical Transactions of the Royal Society B: Biological Sciences**. v.371, n.1699, p.1-11, mar. 2016.
- DRUMMOND A, RAMBAUT A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC Evol Biol.**, 8(7): 214, 2007.

EDGAR, R.C. Muscle: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**. 5:113, 2004.

FITCH, W. M.; LEITER, J. M.; LI, X. Q.; PALESE, P. Positive Darwinian evolution in human influenza A viruses. **Proc. Natl Acad. Sci. USA** 88, 4270–4274 (1991).

FRANKHAM, R., BALLOU, J.D., BRISCOE, D.A. & McINNES, K.H. A Primer of Conservation Genetics. Cambridge University Press. Available from: <https://www.cambridge.org/core/product/identifier/9780511817359/type/book>, 2004.

GRAUR, Dan; MARTIN, Willian. Reading the entrails of chickens: Molecular timescales of evolution and the illusion of precision. **Trends in Genetics**. V.20, n.2, p. 80-86, fev. 2004.

HALL, T.A. Bioedit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. **Nucleic Acids Symposium Series**. V. 41, p. 95-98, 1999.

HEATH, Tracy; HUELSENBECK, John P.; STADLER, Tanja. The fossilized birth-death process for coherent calibration of divergence-time estimates. **Proceedings of the National Academy of Sciences of the United States of America**. Seattle, v.111, n.29, p.2957-2966, jul. 2014.

HEBERT, Paul et al. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. **Proceedings of the Royal Society B: Biological Sciences**. Canadá, v.270, n. suppl1, p.96-99, mai. 2003.

HEBERT, Paul et al. Biological Identifications through DNA barcodes. **Proceedings of the Royal Society B: Biological Sciences**. Canadá, v.270, n.1512, p.313.321, jan. 2003.

HIPSLEY, Christy A; MULLER, Jahannes. Beyond fossil calibrations: Realities of molecular clock practices in evolutionary biology. **Frontiers in Genetics**. V.5, p. 1-11, mai. 2014.

HO, Simon Y.W.; DUCHENE, Sebastian. Molecular-clock methods for estimating evolutionary rates and timescales. **Molecular Ecology**. Australia, v. 23, n. 24, p. 5947-5965, set. 2014.

KIMURA, Motoo; OHTA, Tomoko. On the rate of molecular dating. **Journal of Molecular Evolution**. Japão, v.1, n.1, p. 1-17, set. 1971.

KODANDARAMAIAH, Ullasa. Tectonic calibrations in moleceuar dating. **Current Zoology**. Suécia, v.57, n.1, p. 116-124, 2011.

LUNDBERG, J.G., MARSHALL, L.G., GUERREO, J., HORTON, B., MALABARBA, M.C.S.L. & WESSELINGH, F. The Stage for Neotropical Fish Diversification: A History of Tropical South American Rivers. In: L. R. Malabarba, R. E. Reis, R. P. Vari, C. A. S. Lucena, and Z. M. Lucena (Eds), Phylogeny and classification of Neotropical Fishes. EDIPUCRS, Porto Alegre, p. 603, 1998.

MARTIN, Andrew P.; PALUMBI, Stephen R. Body size, metabolic rate, generation time, and the molecular clock. **Proceedings of the National Academy of Sciences**. v. 90, n.9, p.4087-4091, mai. 1993.

National Center for Biotechnology Information (NCBI). **GenBank**.1988 Disponível em: <https://www.ncbi.nlm.nih.gov/> Acesso em: 01 set. 2021.

OGG, J. G., HINNOV, L. A. & HUANG, C. Cretaceous. in The Geological Time Scale 2012 (eds. Gradstein, F. M., Ogg, J. G., Schmitz, M. D. & Ogg, G. M.) 793–853, 2012.

O'REILLY, Joseph; DONOGHUE, Philip C.J.; Tips and nodes are complementary not competing approaches to the calibration of molecular clocks. **Biology Letters**. V.12, n. 4, mar. 2016.

OHTA, Tomoko. Very slightly deleterious mutations and the molecular clock. **Journal of Molecular Evolution**. v.25, p.1-6, nov. 1987.

ORTON, Mathew G. et al. Is molecular evolution faster in the tropics? **The genetics society**. V. 122, p. 513-524, set. 2019.

PAGE, R.D. & HOLMES, E.C. 1 . Molecular evolution: a phylogenetic approach. John Wiley & Sons, 2009. **Molecular Evolution** — A Phylogenetic Approach. Blackwell Science, Oxford, 417 pp. Available from: <http://www.nature.com/doifinder/10.1046/j.1365-2540.2000.0696a.x>, 1988.

PAGE, Roderic D.M. Tree View: An application to display phylogenetic trees on personal computers. **Bioinformatics**. v. 12, n. 4, p.357-358, ago. 1996.

PATTERSON, C. A review of Mesozoic acanthopterygian fishes, with special reference to those of the English chalk. **Philos. Trans. R. Soc. Lond. B Biol. Sci.** 247, 213–482 (1964).

PRYON, R. Alexander. Divergence time estimation using fossils as terminal taxa and the origin of lissamphibia. **Systematic Biology**. v.60, n.4, p.466-481, mai. 2011.

RAND, David M. Thermal habit, metabolic rate and the evolution of mitochondrial DNA. **Trends in Ecology & Evolution**. v.9, n.4, p.125-131, abr. 1994.

REIS, R.E., ALBERT, J.S., DI DARIO, F., MINCARONE, M.M., PETRY, P. & ROCHA, L.A. Fish biodiversity and conservation in South America. **Journal of fish biology** 89, 12–47. <https://doi.org/10.1111/jfb.13016>, 2016.

RIBEIRO, A.C. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: An example of faunal evolution associated with a divergent continental margin. **Neotropical Ichthyology** 4, 225–246. <https://doi.org/10.1590/s1679-62252006000200009>, 2006.

ROXO, Fábio F., et al. Phyllogenomic reappraisal of the Neotropical catfish family Loricariidae (teleostei: Siluriformes) using ultraconserved elements. **Molecular Phylogenetics and Evolution**. Botucatu-SP, v.135, p. 148-165, fev. 2019.

THACKER, C.E. Patterns of divergence in fish species separated by the Isthmus of Panama. **BMC Evolutionary Biology** 17, 1–14. <https://doi.org/10.1186/s12862-017-0957-4>, 2017.

TRAVASSOS H., SANTOS R.S. Caracídeos fósseis da bacia do Paraíba. **Anais Academia Brasileira Ciências**. 27:313–321, 1955.

TWIDDY, S. S., HOLMES, E. C. & RAMBAUT, A. Inferring the rate and time-scale of dengue virus evolution. **Mol. Biol. Evol.** 20, 122–129 (2003)

WATERS, Jonathan et al. Geological dates and molecular rates: Rapid divergence of rivers and their biotas. **Systematic Biology**. v.56, n.2, p.271-282, mai. 2007.

WEISS F.E., MALABARBA L.R., MALABARBA M.C. Phylogenetic relationships of Paleotetra, a new characiform fish (Ostariophysi) with two new species from the Eocene-Oligocene of south-eastern Brazil. **J Syst. Palaeontol.** 10:73–86, 2012.

WELCH, John J; BROMHAM, Lindell. Molecular dating when rates vary. **Trends in ecology and evolution**. Grã Bretanha, v. 20, n.6, p. 320-327, jun.2005.

WOODWARD A.S. Considerações sobre alguns peixes Terciários dos schistos de Taubaté, Estado de S. Paulo, Brazil. **Rev. Mus. Paulista**. 3:63–70, 1898.

XIA X, XIE Z. DAMBE: Data analysis in molecular biology and evolution. **J Hered.**, 92, 371-373, 2001.

XIA, Xuhua et al. An index of substitution saturation and its application. **Molecular Phylogenetics and Evolution**. V.26, n.1, p.1-7, fev. 2003.

ZARDOYA, Rafael; DOADRIO, Ignacio. Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. **Journal of molecular evolution**. V. 49, p.227-237, mar. 1999.

ZUCKERKANDL, Emile; PAULING, Linus. Molecular Disease, Evolution, and Genic Heterogeneity. In: KASHA, M.; PULLMAN, B. **Horizons in Biochemistry**. New York. Academic Press, 1962.

ANEXO A - Short communication enviado à revista BMC Evolutionary

Calibrating the evolutionary rate of DNA barcode in Teleostei fishes

Author: Santos, Karina Camargo Arroyo¹; Roxo Fábio Fernandes²; Oliveira, Claudio²; Costa-Silva; Guilherme José *

Filiation:

1: Pós Graduação em Manejo e Conservação da Fauna Silvestre, Universidade Santo Amaro, São Paulo, SP, Brazil.

2: Instituto de Biociências, Universidade Estadual Paulista, Botucatu, SP, Brazil

Correspondin author: GJCS, costasilvagj@profunisa.br

Av. Prof. Enéas de Siqueira Neto, 340 - Jardim das Imbuías, São Paulo - SP, 04829-300.

Abstract

DNA barcoding is the best known molecular marker in animals and can be used as a universal molecular clock marker. Through bioinformatics analysis, we calculate its evolution rate in marine and freshwater teleosts fish, reaching values of 1.17% and 1.05% per million years, respectively. The saturation level found is high, which discourages the use of this marker for analyzing old events. The results are consistent with geological data when calculating the vicariance date of known recent geological events, suggesting that the universal molecular clock may be helpful for dating recent lineages separation.

Communication

The DNA Barcoding project proposed by Hebert *et al.* (2003) aims to identify all animal species by a single and standard sequence. For this purpose, the researcher selected the cytochrome c oxidase subunit 1 (COI), and since then, the global database has been increased. Today, the project assembles, only in Teleostei fish, almost a half-million analyzed specimens sampled worldwide (Bold System - Ratnasingham & Hebert 2007). This vast data bank provides new possibilities for exploring biodiversity, evolution, and distribution patterns. But, with this intention, it is necessary to understand some aspects of COI evolution, among them the evolutionary rate of the gene. The fishes are excellent models for evolutionary study, as they are subject to changes in water bodies (Albert & Reis 2011). Identifying the COI evolution rate can be an essential tool to estimate the divergence time between populations/species, working as a universal molecular clock marker. In this sense, the main idea of this communication is estimating the COI evolutionary rate in marine and freshwater Teleostei fish and testing the rates found with known vicariant events.

We obtained all COI sequences in GenBank and Bold System (matrixes on supplementary material 1) and performed the molecular clock analysis in Beast 1.7 (Drummond *et al.* 2012) (supplementary material 2: details of analysis criteria). The topology according to Alfaro *et al.* (2018) was studied for the marine fish analysis, with nine fossils calibration points. Meanwhile, when considering the Characoid (largest

freshwater teleosts group), the topology used was based on Melo *et al.* (2021), with five fossils calibration points. We also test the evolution rates obtained in two known vicarious events. 1) The separation of populations by the rise of the Sete Quedas waterfall, which isolated fish populations from the upper Paraná from the rest of the La-Prata basin (Costa-silva *et al.* 2018), dated at least 1.5 MY ago (Barotelli 2004); 2) the separation between populations by the closing of the Panama isthmus, that separated fish between the Caribbean and the Pacific Eastern (Thacker 2017), dated from approximately 3 MY (O'Dea *et al.* 2016).

The COI evolution rate is slightly different between marine fish (1.17%/MY) and Characoid (1.05%/MY) (Figure 1A). The graphical analysis of saturation indicates that, for old groups, the COI is saturated (figure 1B). The found rates test pointed that the Panama isthmus vicariance isolates the marine fish species at a minimum of 3.2 MY (Figure 1c), and the Sete Quedas vicariance that isolates the characoids fish between the upper Paraná river and the remaining of the Prata Basin dates back to 1.6 MY.

The evolution rate of COI evolution in a lineage may be subjected to different factors, such as population size, metabolism, generation time, variation of rates between protein sequences, and saturation (Frankham *et al.* 2004). Thus, it is possible to observe the wide variation of this rate among animals group, e.g. 1.21%/MY in bivalves (Marko 2002), 1.4%/MY in shrimps (Knowlton & Weigt 1998), and 3.54% in insects (Papadopoulou *et al.* 2010). These facts, associated with the saturation observed in our analyses, indicate the importance of calculating the clocks in phylogenetically close lineages.

Our tests show that the observed rates rescue cladogenesis dates that are very consistent with known recent geological events, whether in the closing of the Pamaná Isthmus (3 MY of geological estimating Vs 3.2 MY of molecular estimating) or in the rise of Sete-Quedas waterfall (1.5MY of geological estimating Vs 1.6 MY of molecular estimating), and could become a useful tool to calculate others cladogenetic events in the absence of geological data, especially in recent events. The estimated rate could help develop new horizons in evolutionary studies by identifying groups that suffer changes depending on lineages and habitats. Besides, it can also be used for preservation, through the recognition os ancient lineages that could be at threat of extinction.

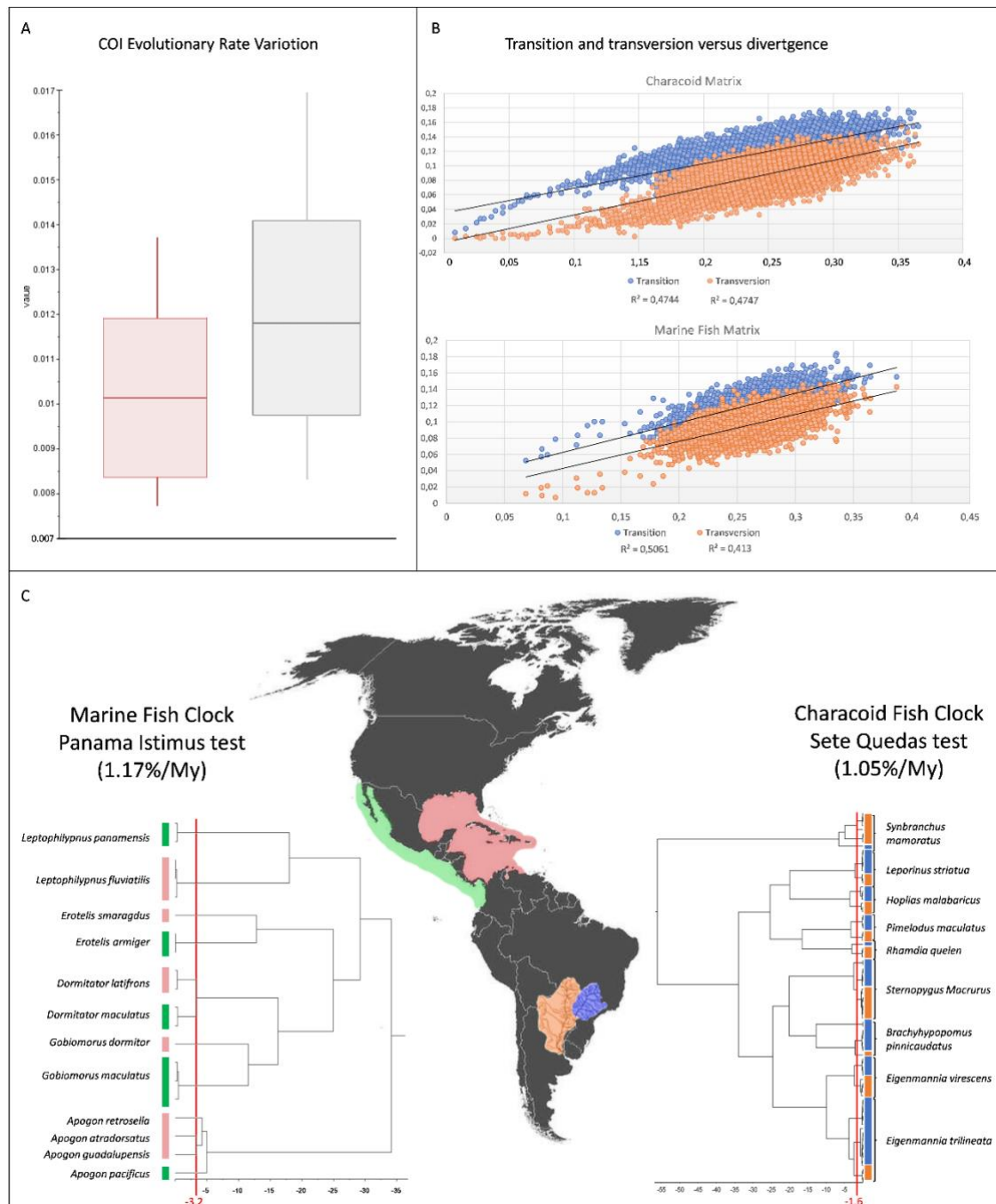


Figure 1: A) Evolution rate of COI gene in Teleostei: Red Box is variation in characoid fish (mean = 0,0105 or 1.05%/My). Gray Box variation in marine fishes (mean = 0,0117 or 1.17%/My). B) Transition and transversion versus divergence with tendency to saturate (R^2 lower than 0.75). C) Left: Genetic comparison between populations from Caribbean (red) and Eastern Pacific (Green), according to marine fish clock. Right: Genetic comparison between populations from the upper Paraná river (Blue) and lower Paraná/Paraguay river basins (Orange) according to characoid clock, and genetics. Redline marks represent 3.2 and 1.6 millions years in respective tests.

Bibliography

- Albert, J.S. & Reis, R. (2011) *Historical Biogeography of Neotropical Freshwater Fishes*. U. O. C. PRESS (Ed). London, 408 pp. Available from: <https://market.android.com/details?id=book-HbLcMtCFdtwC%5Cnhttps://books.google.com/books?hl=en&lr=&id=HbLcMtCFdtwC&oi=fnd&pg=PP1&dq=Historical+Biogeography+of+Neotropical+Freshwater+Fishes&ots=20tCk2mugl&sig=kKfHxL6C5lGJwgpCrPc5fNqkGzI> LB - IRAv
- Barotelli, A. (2004) Origem das grandes cachoeiras do Planalto Basáltico da bacia do Paraná: Evolução quaternária e geomorfologia. In: *Geologia do continente sul americano: Evolução da obra de Fernando Flávio Marques de Almeida*. Beca, p. 570.
- Bizerril, C. (1998) A Ictiofauna da Bacia do Rio Paraíba do Sul. In: *Biodiversidade e Padrões Biogeográficos*. Brazilian Archives of Biology and Technology 41, pp. 67–81.
- Costa-silva, G.J., Ashikaga, F.Y., Kioko, C., Dias, S., Henrique, L., Pereira, G. & Foresti, F. (2018) DNA barcoding techniques used to identify the shared ichthyofauna between the Pantanal floodplain and Upper Parana River. *Mitochondrial DNA Part A* 29, 1063–1072. <https://doi.org/10.1080/24701394.2017.1404046>
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29, 1969–1973. <https://doi.org/10.1093/molbev/mss075>
- Frankham, R., Ballou, J.D., Briscoe, D.A. & McInnes, K.H. (2004) *A Primer of Conservation Genetics*. Cambridge University Press. Available from: <https://www.cambridge.org/core/product/identifier/9780511817359/type/book>
- Hebert, P.D.N., Cywinska, A., Ball, S.L. & DeWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270, 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Knowlton, N. & Weigt, L.A. (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences* 265, 2257–2263. <https://doi.org/10.1098/rspb.1998.0568>
- Lundberg, J.G., Marshall, L.G., Guerreiro, J., Horton, B., Malabarba, M.C.S.L. & Wesselingh, F. (1998) The Stage for Neotropical Fish Diversification: A History of Tropical South American Rivers. In: L. R. Malabarba, R. E. Reis, R. P. Vari, C. A. S. Lucena, and Z. M. Lucena (Eds), *Phylogeny and classification of Neotropical Fishes*. EDIPUCRS, Porto Alegre, p. 603.
- Marko, P.B. (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution* 19, 2005–2021. <https://doi.org/10.1093/oxfordjournals.molbev.a004024>
- Melo, B.F., Sidlauskas, B.L., Near, T.J., Roxo, F.F., Ghezelayagh, A., Ochoa, L.E., Stiassny, M.L.J., Arroyave, J., Chang, J., Faircloth, B.C., MacGuigan, D.J., Harrington, R.C., Benine, R.C., Burns, M.D., Hoekzema, K., Sanches, N.C.,

- Maldonado-Ocampo, J.A., Castro, R.M.C., Foresti, F., Alfaro, M.E. & Oliveira, C. (2021) Accelerated Diversification Explains the Exceptional Species Richness of Tropical Characoid Fishes. *Systematic Biology* 0, 1–15. <https://doi.org/10.1093/sysbio/syab040>
- O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., De Queiroz, A., Farris, D.W., Norris, R.D., Stallard, R.F., Woodburne, M.O., Aguilera, O., Aubry, M.P., Berggren, W.A., Budd, A.F., Cozzuol, M.A., Coppard, S.E., Duque-Caro, H., Finnegan, S., Gasparini, G.M., Grossman, E.L., Johnson, K.G., Keigwin, L.D., Knowlton, N., Leigh, E.G., Leonard-Pingel, J.S., Marko, P.B., Pyenson, N.D., Rachello-Dolmen, P.G., Soibelzon, E., Soibelzon, L., Todd, J.A., Vermeij, G.J. & Jackson, J.B.C. (2016) Formation of the Isthmus of Panama. *Science Advances* 2, 1–12.
- Page, R.D. & Holmes, E.C. (1988) 1 . Molecular evolution: a phylogenetic approach. John Wiley & Sons, 2009. *Molecular Evolution — A Phylogenetic Approach*. Blackwell Science, Oxford, 417 pp. Available from: <http://www.nature.com/doifinder/10.1046/j.1365-2540.2000.0696a.x>
- Papadopoulou, A., Anastasiou, I. & Vogler, A.P. (2010) Revisiting the insect mitochondrial molecular clock: The mid-aegean trench calibration. *Molecular Biology and Evolution* 27, 1659–1672. <https://doi.org/10.1093/molbev/msq051>
- Reis, R.E., Albert, J.S., Di Dario, F., Mincarone, M.M., Petry, P. & Rocha, L.A. (2016) Fish biodiversity and conservation in South America. *Journal of fish biology* 89, 12–47. <https://doi.org/10.1111/jfb.13016>
- Ribeiro, A.C. (2006) Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: An example of faunal evolution associated with a divergent continental margin. *Neotropical Ichthyology* 4, 225–246. <https://doi.org/10.1590/s1679-62252006000200009>
- Thacker, C.E. (2017) Patterns of divergence in fish species separated by the Isthmus of Panama. *BMC Evolutionary Biology* 17, 1–14. <https://doi.org/10.1186/s12862-017-0957-4>

ANEXO B – Sequências do COI utilizadas e seus códigos de acesso no GenBank.

Tabela SUP 1 - Sequências do COI encontradas no GenBank de peixes marinho presente no artigo de Alfaro et al., 2018.

Taxon	GenBank Acesso	Taxon	GenBank Acesso
Acanthurus bahianus	GU224639.1	Stylephorus chordatus	MH777708.1
Acanthurus bariene	JN311611.1	Syngnathus fuscus	MT456195.1
Acanthurus pyroferus	MK566781.1	Taenianotus triacanthus	MK657247.1
Alectis indica	MK777146.1	Taractes asper	GU440550.1
Alepes kleinii	MN623875.1	Thalassoma ballieui	MG816739.1
Alepisaurus ferox	KY984978.1	Valenciennea strigata	MH049306.1
Anoplogaster cornuta	MT323253.1	Xenentodon cancila	JX983514.1
Antennarius striatus	MK777416.1	Zenopsis conchifer	MN883250.1
Antigonia capros	MH980019.1	Zeus faber	JN312803.1
Aphredoderus sayanus	MT456238.1	Zu elongatus	MN123529.1
Apogon maculatus	EU381025.1	Centropomus sp.	KJ641479.1
Assurger anzac	EU263815.1	Haemulon sp.	HM379879.1
Aulostomus maculatus	KY066073.1	Hypoplectrus sp.	JN312340.1
Balistes capriscus	MT455280.1	Lutjanus sp.	EU595219.1
Bathygobius soporator	MG450288.1	Ogcocephalus sp.	GU225412.1
Bothus pantherinus	MT076518.1	Scarus sp.	MN733654.1
Brama japonica	LC494448.1	Takifugu sp.	MK843763.1
Canthigaster figueiredoi	KT832864.1	Ostorhinchus nigrofasciatus	MK658291.1

Canthigaster margaritata	MT590658.1	Pareques acuminatus	JN313819.1
Canthigaster rostrata	JN311772.1	Parupeneus multifasciatus	MN870573.1
Caranx melampygus	MW034014.1	Pempheris schomburgkii	MF999165.1
Carapus bermudensis	MT323644.1	Percopsis omiscomycus	MT577174.1
Cephalopholis argus	MH707305.1	Periophthalmus barbarus	AF391339.1
Ceratoscopelus warmingii	MT806149.1	Pholidichthys leucotaenia	KT883612.1
Chaetodon kleinii	MW034078.1	Poecilopsetta plinthus	MK617168.1
Chaetodon ocellatus	MT456100.1	Polydactylus sexfilis	MG816719.1
Cheimarrichthys fosteri	AY722120.1	Polymixia lowei	KF930292.1
Chiasmodon niger	MT323552.1	Pomacanthus paru	KF930295.1
Chromis enchrysur	JQ707139.1	Pomatomus saltatrix	MT456177.1
Citharoides macrolepis	MK617139.1	Psenes cyanophrys	AB205446.1
Coryphaena hippurus	MT456259.1	Psenopsis anomala	AB205441.1
Coryphaenoides rupestris	KY033895.1	Pseudanthias squamipinnis	MH049292.1
Cryptopsaras couesii	HQ956180.1	Pseudochromis flavivertex	MN560855.1
Dactyloptena orientalis	MK777794.1	Pseudupeneus maculatus	KY402369.1
Epibulus insidiator	MW033994.1	Ptereleotris evides	MK657939.1

Erotelis smaragdus	HQ987875.1	Rachycentron canadum	MH559419.1
Evorthodus minutus	AY077607.1	Regalecus glesne	MN123480.1
Gadus morhua	MT456169.1	Rondeletia loricata	MN123485.1
Gramma loreto	HM379877.1	Ruvettus pretiosus	LN907525.1
Haemulon album	JN313823.1	Scartella cristata	MT456057.1
Haemulon parra	MK829330.1	Scarus trispinosus	MF999162.1
Halichoeres poeyi	MF999181.1	Scomber scombrus	MT456185.1
Halichoeres radiatus	JN313614.1	Scomberomorus maculatus	MT456078.1
Hippoglossina oblonga	MT455880.1	Seriola zonata	MT455069.1
Istiophorus platypterus	MK472814.1	Sphyræna putnamae	MK777494.1
Kurtus gulliveri	KF930008.1	Myripristis leiognathus	JX390743.1
Labrisomus nuchipinnis	HQ571157.1	Myripristis violacea	MW034077.1
Lampris guttatus	JF931919.1	Naso unicornis	MK658419.1
Lepophidium profundorum	MN883256.1	Odontobutis obscura	JX679046.1
Micropogonias undulatus	MT456167.1	Ophiocara porocephala	MW322096.1
Micropterus salmoides	MT456067.1	Mogurnda adspersa	KJ669528.1

Tabela SUP2 - Sequências do COI encontradas no GenBank de peixes coracoides presente no artigo de Melo et al., 2021.

Taxon	GenBank Acesso	Taxon	GenBank Acesso
Abramites hypselonotus	MN853761.1	Steindachnerina bimaculata	MH537329.1
Acestrocephalus sardina	MT754767.1	Steindachnerina hypostoma	MH537346.1
Acestrorhynchus falcatus	MZ051620.1	Stygichthys typhlops	HM405019.1
Acestrorhynchus microlepis	MZ052059.1	Tetragonopterus argenteus	MH003293.1
Acinocheiroduon melanogramma	HQ600807.1	Tetragonopterus chalceus	MZ051966.1
Acrobrycon ipanquianus	KF210031.1	Tetragonopterus georgiae	MZ051918.1
Alestes inferus	JF800982.1	Thayeria obliqua	MH411323.1
Alestes liebrechtsii	MK073967.1	Thoracocharax stellatus	KU289060.1
Alestopetersius compressus	KX186526.1	Tometes trilobatus	KX868698.1
Anodus elongatus	MW377178.1	Triportheus auritus	MF182364.1
Anodus orinocensis	MW377181.1	Tytocharax madeirae	MH003295.1
Anostomus brevior	MZ051960.1	Anostomoides atrianalis	MT427907.1
Apareiodon ibitiensis	HM405052.1	Hoplias curupira	MG699478.1
Argonectes robertsi	MW377191.1	Mimagoniates microlepis	KF210240.1
Argopleura chocoensis	KF210032.1	Moenkhausia gracilima	KU532828.1
Astyanax aeneus	GU806993.1	Moenkhausia oligolepis	MW916371.1

Astyanax fasciatus	KY268280.1	Myleus setiger	MG752165.1
Astyanax jordani	KY268281.1	Myloplus asterias	MG752256.1
Attonitus ephimeros	KF210034.1	Myxiops aphos	KY327453.1
Bivibranchia velox	MW377210.1	Nannocharax ansorgii	KF541816.1
Boehlkea fredcochui	MH411306.1	Nannocharax elongatus	HM880217.1
Boulengerella cuvieri	MH411355.1	Nannostomus trifasciatus	MH411342.1
Brachypetersius huloti	JN710387.1	Nematobrycon palmeri	KU568937.1
Brittanichthys axelrodi	MN974149.1	Nematocharax venustus	MN011542.1
Brycon falcatus	MZ051888.1	Odontostilbe fugitiva	MT823252.1
Brycon pesu	MZ051612.1	Odontostoechus lethostigmus	KF210247.1
Brycon petrosus	MG936814.1	Oligosarcus hepsetus	HM065009.1
Bryconadenos thanaotoros	KF210038.1	Orthospinus franciscensis	HM405169.1
Bryconaethiops microstoma	MK074013.1	Ossubtus xinguense	MG752465.1
Bryconamericus exodon	MN879963.1	Paracheiroidon axelrodi	KU568958.1
Caenotropus maculosus	MZ052048.1	Parodon hilarii	HM405179.1
Caenotropus mestomorgmatos	KF562416.1	Phago boulengeri	MK074594.1
Carlana eigenmanni	MG496117.1	Phenacogaster wayana	MZ051920.1

Carnegiella strigata	MH411380.1	Phenagoniates macrolepis	MG937101.1
Catoprion mento	MT519743.1	Piabarchus analis	MH003279.1
Chalceus macrolepidotus	MZ052030.1	Piabina argentea	HM144107.1
Characidium etheostoma	MK861670.1	Pimelodus argenteus	MZ345630.1
Characidium fasciatum	MK464151.1	Planaltina myersi	MK464051.1
Chilodus gracilis	KF562422.1	Poptella brevispina	MZ052067.1
Chilodus punctatus	MH411373.1	Pristella maxilaris	KU568983.1
Citharinus congicus	HM418208.1	Probolodus heterostomus	KY327456.1
Compsura heterura	HM404972.1	Prochilodus nigricans	MN996701.1
Copionodon pecten	KY857929.1	Protocheirodon pi	MH411310.1
Corydoras aeneus	MN879941.1	Psectrogaster ciliata	MH537308.1
Creagrutus peruanus	MH003085.1	Psectrogaster rutiloides	MG911822.1
Ctenobrycon spilurus	KY348907.1	Psellogrammus kennedyi	MN879945.1
Ctenolucius hujeta	JF800932.1	Pseudanos trimaculatus	MH411358.1
Curimata incompta	MH537234.1	Pseudocheirodon terrabae	MG496209.1
Curimata roseni	MH537240.1	Pseudocorynopoma doriae	JX983442.1
Curimata vittata	MH537242.1	Pseudocurimata lineopunctata	MH537322.1
Curimatella meyeri	MG911808.1	Pseudopimelodus mangurus	EU179816.1

Curimatopsis macrolepis	KX086740.1	Pterobrycon myrnae	MG496212.1
Cyanocharax alburnus	KF210124.1	Pygocentrus piraya	MG752589.1
Cyphocharax magdalenae	MG496127.1	Pyrrhulina filamentosa	MZ052021.1
Cyphocharax plumbeus	MH537285.1	Rachoviscus crassiceps	HM562858.1
Cyprinus carpio	JX983284.1	Rhoadsia altipinna	KU252647.1
Delturus carinotos	GQ225388.1	Rhytidodus argenteofuscus	KF569004.1
Deuterodon iguape	KY327421.1	Roeboides xenodon	HM405215.1
Distichodus affinis	MK074183.1	Salminus brasiliensis	MF969175.1
Distichodus decemmaculatus	HM418095.1	Schizodon fasciatus	MZ051441.1
Distichodus fasciolatus	MK074186.1	Semaprochilodus insignis	MH991717.1
Distichodus hypostomatus	MN935444.1	Semaprochilodus kneri	KX086783.1
Distichodus notospilus	MN935443.1	Semaprochilodus varii	MZ051939.1
Distichodus sexfasciatus	HM880208.1	Serrapinnus piaba	HM405229.1
Eretmobrycon emperador	MH003146.1	Serrasalmus rhombeus	MZ052051.1
Erythrinus erythrinus	MZ052053.1	Spintherobolus leptoura	MG967588.1
Galeocharax humeralis	KU288921.1	Steatogenys elegans	KT166880.1

Gasteropelecus sternichla	MZ051889.1	Jupiaba polylepis	FJ749068.1
Glandulocauda melanopleura	MH036053.1	Knodus cf delta	KT248103.1
Hasemania hansenii	MK464168.1	Knodus moenkhausii	MN879953.1
Hemibrycon metae	MH003201.1	Kolpotocheiroduon theloura	MK464161.1
Hemigrammus rodwayi	KY348914.1	Laemolyta proxima	MT427928.1
Hemiodus quadrimaculatus	MT949027.1	Lepidocharax burnsi	KF210231.1
Hepsetus cuvieri	KX911921.1	Leporellus vittatus	KF568980.1
Hepsetus kingsleyae	KX911953.1	Leporinus agassizii	MF664283.1
Hepsetus lineata	MK074345.1	Leporinus fasciatus	MZ051768.1
Hollandichthys multifasciatus	GU701423.1	Leporinus striatus	KX086761.1
Hoplerethrinus unitaeniatus	MG911835.1	Lophiobrycon weitzmani	KF210233.1
Hoplias intermedius	HM405123.1	Loricaria proluxa	HM899913.1
Hoplias microlepis	MG937000.1	Metynnis luna	MN853770.1
Hydrolicus scomberoides	EU514497.1	Micralestes acutidens	MK074483.1
Hyphessobrycon compressus	HQ564603.1	Microstomatichthy oborus bashforddeani	KF541788.1
Jupiaba abramoides	MZ051806.1	Jupiaba acanthogaster	KC618768.1

Tabela SUP3 - Sequências de COI encontradas no GenBank e seus códigos de acesso, de espécies do istmo do Panamá, baseadas no artigo de Thacker (2017).

Taxon	GenBank Acesso	Taxon	GenBank Acesso
Dormitator latifrons	AY722142	Leptophilypnus fluviatilis	AY722130
Dormitator maculatus	AY722143	Leptophilypnus panamensis	AY722127
Erotelis armiger	AY722167	Apogon retrosella	AB890018.1
Eleotris smaragdus	AF391499.1	Apogon pacificus	AB890015.1
Gobiomorus dormitor	AY722144	Apogon guadalupensis	AB890013.1
Gobiomorus maculatus	AY722132	Apogon atradorsatus	AB890008.1

ANEXO C - Topologia de peixes marinhos e coracoides calibradas.

Figura SUP 1 - Topologia de peixes coracoides obtida após análise com as sequências de COI, indicando os tempos de divergência entre os nós.

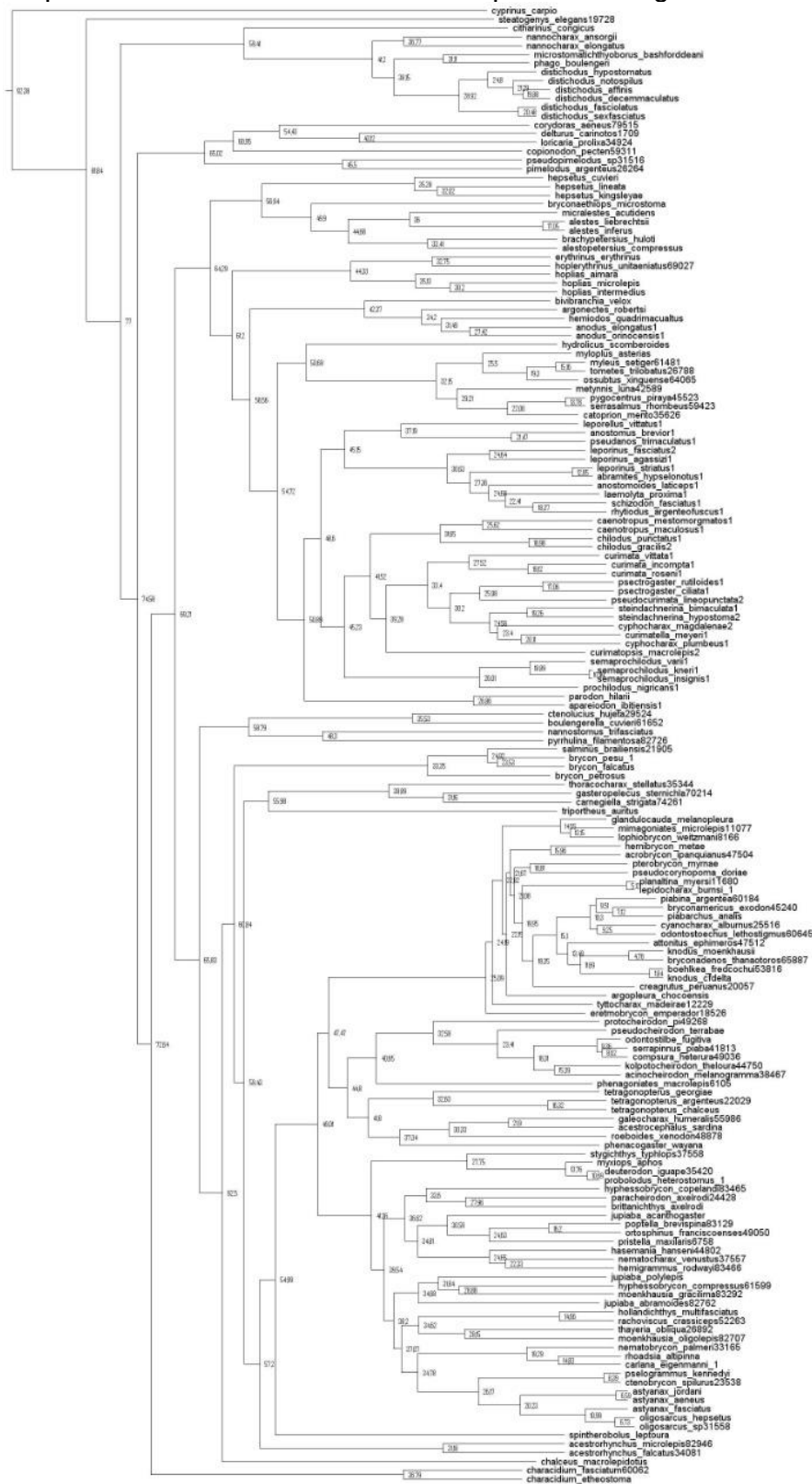


Figura SUP2 - Topologia de peixes marinhos obtida após análise com as sequências de COI, indicando os tempos de divergência entre os nós.

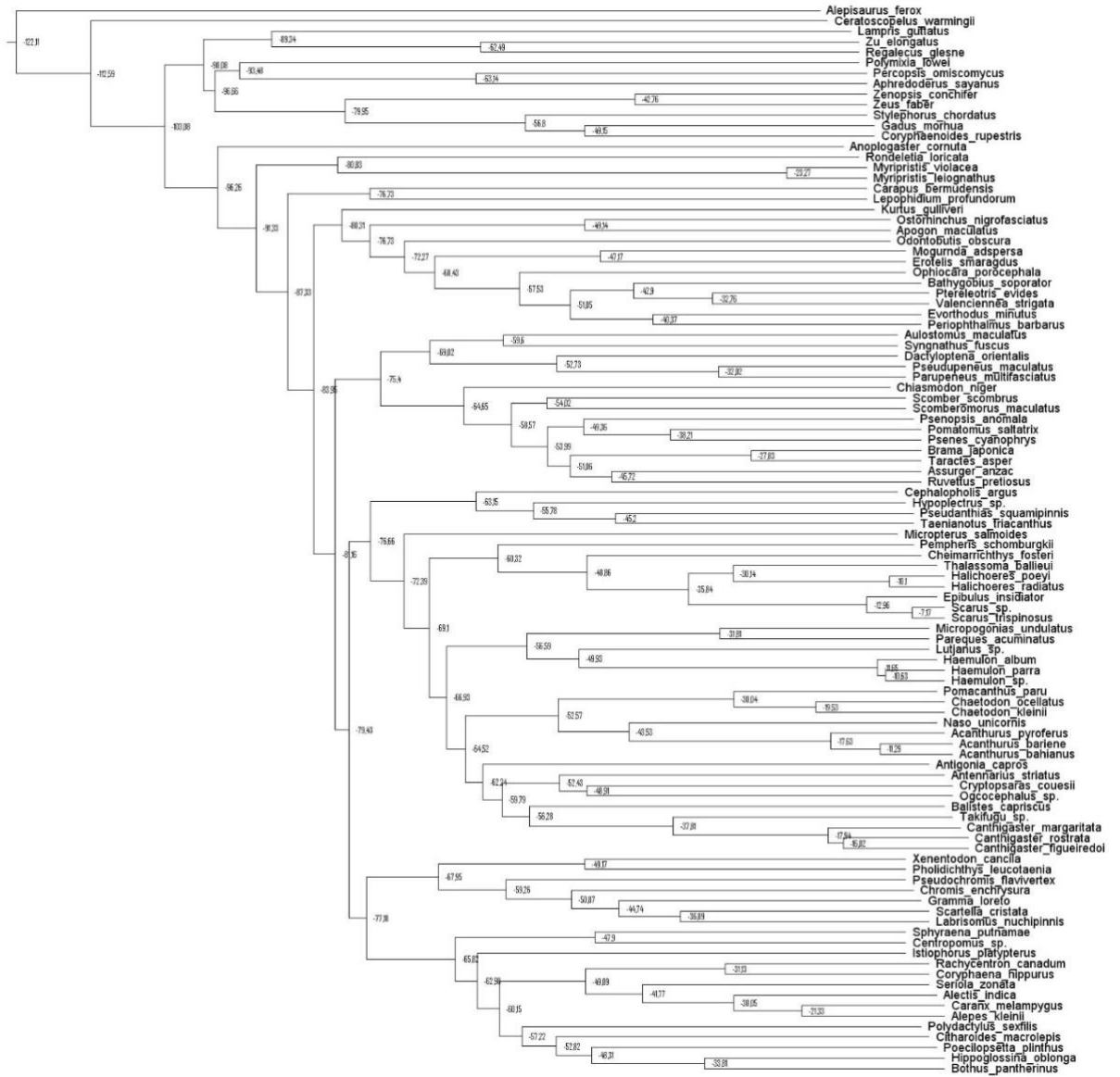


Figura SUP4 - Topologia baseada na taxa de evolução do COI para peixes coracoides. Essa topologia marca a relação entre as espécies separadas pelo surgimento da cachoeira das Sete-Quedas.

