

UNIVERSIDADE FEDERAL DO PARANÁ

ERASTO VILLA BRANCO JUNIOR

ESTUDO FILOGENÉTICO DO GENE *RAPHI*

CURITIBA
2013

ERASTO VILLA BRANCO JUNIOR

ESTUDO FILOGENÉTICO DO GENE *RAPH1*

Tese de doutoramento apresentada ao Programa
de Pós-Graduação em Genética do Departamento
de Genética do Setor de Ciências Biológicas da
Universidade Federal do Paraná

Orientador: Prof. Dr. Ricardo Lehtonen
Rodrigues de Souza

CURITIBA
2013

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Rosilei Vilas Boas – CRB/9-939).

Branco Junior, Erasto Villa.
Estudo filogenético do gene RAPH1. / Erasto Villa Branco Junior. –
Curitiba, 2013.
140 f. : il.

Orientador: Ricardo Lehtonen Rodrigues de Souza.

Tese (Doutorado) – Universidade Federal do Paraná, Setor de
Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Filogenia. 2. Butirilcolinesterase. 3. Colinesterase. I. Título. II. Souza,
Ricardo Lehtonen Rodrigues de. III. Universidade Federal do Paraná. Setor
de Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 575



PARECER

Os abaixo-assinados, membros da Banca Examinadora da defesa de tese de Doutorado, a qual se submeteu **ERASTO VILLA BRANCO JUNIOR**, para obtenção do título de Doutor em Genética pela Universidade Federal do Paraná, no Programa de Pós-Graduação em Genética, são de parecer que se confira ao candidato o conceito "B".

Secretaria da Coordenação do Programa de Pós-Graduação em Genética do Setor de Ciências Biológicas da Universidade Federal do Paraná.

Curitiba, 31 de janeiro de 2013

Doutora Kelly Nunes

Universidade de São Paulo – Membro Titular

Doutor Pablo Sandro Carvalho Santos

Universidade Federal do Paraná – Membro Titular

Doutora Patrícia Dalzoto

Universidade Federal do Paraná – Membro Titular

Doutora Nina Amália Brancia Pagnan

Universidade Federal do Paraná – Membro Titular

Doutor Ricardo Lehtonen Rodrigues de Souza

Universidade Federal do Paraná – Orientador e Presidente da Banca

Visto

Professora Doutora Enilze Maria de Souza Fonseca Ribeiro
Vice-Coordenadora do Programa de Pós-Graduação em Genética

1 RESUMO

Foram analisadas árvores filogenéticas construídas com base em 53 sequências de DNA e 53 sequências de aminoácidos, correspondentes ao gene RAPH1, de 48 diferentes espécies de animais, incluindo nematódeo, peixes, anfíbio, réptil, mamíferos e aves. Foram construídas árvores com as sequências inteiras e com segmentos delas correspondentes aos domínios proteicos GM, RA, PH e EVH1 e aos sítios de ligação de proteínas que a compõem. Foram utilizados métodos de construção de filogenia baseados em distância, na Máxima Parcimônia e na Máxima Verossimilhança. Em algumas espécies de peixe o gene *RAPH1* (*Ras association (RalGDS/AF-6) and pleckstrin homology domains 1*) apresenta-se duplicado, com duas formas semelhantes, porém não idênticas. Esta análise, que incluiu os genes de cinco espécies, deixou evidente que os dez genes se reúnem, por semelhança, nitidamente em dois grupos, cada um dos quais incluindo cinco genes, um de cada uma das cinco espécies, o que sugere que a duplicação do gene tenha ocorrido em época anterior aos momentos de especiação que separaram as cinco espécies. A análise evidencia, também, que genes de espécies que apresentam cópia única, entre alguns peixes, anfíbio, réptil, aves e mamíferos, são, todos eles, mais estreitamente aparentados com apenas um destes dois grupos. Isto deixou evidente que a duplicação do gene RAPH1 ocorreu em um ancestral comum a todas as espécies analisadas e não em um ancestral comum às espécies que apresentam dois genes RAPH1.

Evolução. Filogenia. Lamelipodina. Butirilcolinesterase. Colinesterase.

2 ABSTRACT

Phylogenetic trees constructed based on 53 DNA sequences and 53 amino acid sequences, corresponding to the RAPH1 gene, from 48 different animal species, including nematodes, fish, amphibians, reptiles, mammals and birds, were analyzed. Trees were constructed with the entire sequences and with segments corresponding to the protein domains GM, RA, PH and EVH1 and the protein binding sites that comprise it. Phylogeny construction methods based on distance, Maximum Parsimony and Maximum Likelihood were used. In some species of fish, the RAPH1 gene (Ras association (RalGDS / AF-6) and pleckstrin 1 homology domains) is duplicated, with two similar, but not identical, forms. This analysis, which included the genes of five species, made it evident that the ten genes come together, by similarity, clearly in two groups, each of which includes five genes, one from each of the five species, which necessarily means that the duplication of the gene occurred before the speciation moments that separated the five species. The analysis also shows that the genes of species that have a single copy, among species fish, amphibian, reptile, birds and mammals, are all more closely related to only one of these two groups. This made it evident that the duplication of the RAPH1 gene occurred in an ancestor common to all analyzed species and not in an ancestor common to species that have two RAPH1 genes.

Evolution. Phylogeny. Lamellipodin. Butyrylcholinesterase. Cholinesterase.

3 SUMÁRIO

1	RESUMO	5
2	ABSTRACT	6
3	SUMÁRIO	7
4	INTRODUÇÃO	9
5	REVISÃO BIBLIOGRÁFICA	17
6	OBJETIVOS	24
6.1	OBJETIVO GERAL.....	24
6.2	OBJETIVOS ESPECÍFICOS	24
7	JUSTIFICATIVA	25
8	CAPÍTULO 1 - ANÁLISE DO GENE RAPH1 COM USO DE FILOGENIAS	27
8.1	RESUMO	27
8.2	INTRODUÇÃO.....	27
8.3	MATERIAL E MÉTODOS.....	32
8.4	RESULTADOS E DISCUSSÃO	36
8.4.1	<i>Árvores de cDNA de 53 sequências</i>	41
8.4.2	<i>Árvores de cDNA de 39 sequências</i>	43
8.4.3	<i>Árvores de Aminoácidos de 53 sequências</i>	45
8.4.4	<i>Árvores de Aminoácidos de 39 sequências</i>	48
8.4.5	<i>Domínios</i>	50
8.5	CONCLUSÕES.....	51
8.6	REFERÊNCIAS	53
9	CAPÍTULO 2 - ANÁLISE EVOLUTIVA DO GENE RAPH1 EM ALGUMAS ESPÉCIES DE PEIXE	55
9.1	RESUMO	55
9.2	INTRODUÇÃO.....	55
9.3	MATERIAL E MÉTODOS.....	56
9.4	RESULTADOS E DISCUSSÃO	57
9.5	CONCLUSÕES.....	63
9.6	REFERÊNCIAS	64
10	DISCUSSÃO	65
	REFERÊNCIAS	70
	APÊNDICES	73
	APÊNDICE 1 - MATERIAL E MÉTODOS	73
	APÊNDICE 2 – ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO GENE RAPH1	81
	APÊNDICE 3 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE cDNA DO GENE RAPH1	91
	APÊNDICE 4 – ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO RA DO GENE RAPH1	101
	APÊNDICE 5 – ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO RA DO GENE RAPH1	106
	APÊNDICE 6 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO GM DO GENE RAPH1	111
	APÊNDICE 7 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO PBS DO GENE RAPH1	116
	APÊNDICE 8 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO EVH1 1 DO GENE RAPH1	118

APÊNDICE 9 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO EVH1 2 DO GENE RAPH1	120
APÊNDICE 10 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO EVH1 3 DO GENE RAPH1	122
APÊNDICE 11 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO EVH1 4 DO GENE RAPH1	124
APÊNDICE 12 – ÁRVORES AMINOACÍDICAS DE GENES PARÁLOGOS E ORTÓLOGOS RAPH1	129
APÊNDICE 13 – TABELAS DE COMPOSIÇÃO AMINOACÍDICA DO GENE RAPH1	134

4 INTRODUÇÃO

O que se denomina, em eucariontes, de células excitáveis são quase sempre células que apresentam dentro e fora de si uma grande diferença na concentração dos íons Na^+ e K^+ . Como regra geral, esta diferença é de cerca de 30 vezes mais potássio que sódio no interior e o inverso, cerca de 30 vezes mais sódio que potássio, no meio externo.

O início da excitação, impulso nervoso, por exemplo, é um grande aumento da permeabilidade da membrana plasmática ao sódio, tornando-se cerca de quinhentas vezes mais permeável ao sódio que ao potássio. Isto provoca uma intensa e espontânea entrada do sódio na célula por aquele ponto de início. Em seguida, esta alteração da permeabilidade da membrana, eletricamente detetável, propaga-se ao longo dela, enquanto que a situação original é restabelecida no ponto de origem. Este impulso, atingindo uma região de sinapse, provoca a liberação dos neurotransmissores, que dão continuidade a ele.

Sinapse é um contato funcional entre duas células excitáveis, como por exemplo, dois neurônios. O simples fato de duas células estarem com suas membranas plasmáticas juntas não permitiria que a despolarização da membrana de uma iniciasse a despolarização da membrana da outra, mas havendo uma sinapse, o impulso elétrico produz uma transmissão química do próprio impulso, que assim induz o início de um outro na célula vizinha. Esta transmissão química se dá pela liberação, na fenda sináptica, do conteúdo de vesículas da primeira das células. Este conteúdo são os neurotransmissores, que sendo produzidos em uma célula e liberados na fenda sináptica, têm a capacidade de desencadear o impulso nervoso em outra célula.

Um dos nutrientes essenciais para a manutenção da saúde, a colina é de presença obrigatória nos alimentos de seres humanos em quantidade relativamente grande: cerca de meio grama ao dia para adultos. É uma amina, hidrossolúvel, agrupada com as vitaminas do complexo B, que tem como uma de suas principais funções no organismo eucarionte a formação de fosfoglicerídios, que, juntamente com esfingolipídios e esteroides, são os principais lipídeos formadores da membrana celular. Além disto, a colina é formadora de um dos principais neurotransmissores, que é a acetilcolina.

Acetilcolina é um neurotransmissor das sinapses colinérgicas, tanto do sistema nervoso central quanto do periférico. É caracteristicamente encontrada nas sinapses neuromusculares dos vertebrados. Ela é sintetizada nas terminações dos axônios a partir da colina e da acetilcoenzima A, numa única reação enzimática, promovida pela enzima colina-acetiltransferase. A reação produz acetilcolina e libera coenzima A.

Quando a acetilcolina é liberada por um neurônio na fenda sináptica, é induzida a formação de impulso nervoso no neurônio seguinte, em uma cadeia de neurônios que estão formando a trajetória de um impulso nervoso. E, quando isto ocorre na placa motora, a acetilcolina estimula as células musculares esqueléticas à contração. Em ambos os casos, a acetilcolina necessita ser eliminada rapidamente, da fenda sináptica ou da placa motora, depois de cumprida uma destas suas funções. Esta eliminação é realizada nestes locais pela enzima acetilcolinesterase que degrada a acetilcolina ao convertê-la em metabólitos inativos de colina e ácido acético.

A ação da acetilcolinesterase é fundamental nestes casos, uma vez que na sua ausência o impulso nervoso torna-se ininterrupto, o que inviabiliza todas as funções coordenadas por contração muscular e todo o controle nervoso do organismo. Inibições mesmo parciais da acetilcolinesterase provocam, por exemplo, paralisia dos músculos

que realizam a respiração e parada dos batimentos cardíacos, ocasionadas pelo acúmulo de acetilcolina nas placas motoras.

A butirilcolinesterase (BChE), também denominada colinesterase do soro, é uma enzima muito similar à acetilcolinesterase (AChE), também denominada colinesterase verdadeira, tanto na estrutura terciária quanto na composição nucleotídica do gene que a determina.

Esta enzima é muito conhecida e de grande interesse fisiológico em seres humanos, o que gerou grande quantidade de estudos fisiológicos e genéticos a seu respeito. As semelhanças entre as enzimas fizeram com que a butirilcolinesterase fosse descoberta, tendo sido, em seguida, caracterizada por suas diferenças. Esta descoberta ocorreu como consequência do uso cirúrgico da succinilcolina (suxametônio) como relaxante muscular pré-operatório, na ocasião em que se percebeu que alguns variantes da butirilcolinesterase faziam com que seus portadores, quando homozigotos, fossem extremamente sensíveis ao relaxante muscular, a ponto de produzir neles o óbito por apneia.

Embora a homozigose para este variante da butirilcolinesterase fosse letal nesta situação, estes indivíduos eram indistinguíveis, em termos de vigor e sobrevivência, dos demais portadores de outros alelos.

Estudos em seres humanos demonstraram sua ação enzimática específica, bem como evidenciaram a existência de muitos outros variantes alélicos, sendo cinco deles comuns e a grande maioria, cerca de 50, bastante rara.

Outros estudos demonstraram a existência de variantes gênicos desta enzima em outros mamíferos e em outros seres vivos, o que leva a pensar em uma ação enzimática

estabelecida há muito tempo, bem como em uma ação de relevância para o grupo analisado de seres vivos.

No entanto, a despeito de tantos estudos, o papel da butirilcolinesterase na fisiologia humana não é ainda bem compreendido. De muitos estudos, depreende-se desempenhar um papel significativo, visto que se sabe estar a enzima relacionada a importantes atividades celulares e sistêmicas.

Está, em seres humanos, demonstrado que a intensidade de sua atividade é relacionada, entre outras coisas, à obesidade e seus variantes alélicos, estão relacionados ao índice de massa corporal e ao peso.

Estando muito provavelmente relacionada ao metabolismo de lipídios, acredita-se que ela participe do desenvolvimento embrionário, tanto na proliferação, quanto na diferenciação celular.

Colinesterases são enzimas de dois tipos fundamentais: a acetilcolinesterase e a butirilcolinesterase. A acetilcolinesterase também é denominada de colinesterase dos glóbulos vermelhos, colinesterase verdadeira, colinesterase eritrocitária ou ainda, oficialmente, acetilcolina-acetil-hidrolase. A butirilcolinesterase é denominada de colinesterase do soro, pseudocolinesterase, colinesterase sérica ou ainda, tecnicamente, acilcolina-acil-hidrolase. Encontra-se, a primeira, principalmente nas hemácias, nas sinapses e na musculatura estriada. A segunda, principalmente no fígado, no intestino delgado, no pâncreas e no plasma.

Ambas promovem a degradação do neurotransmissor acetilcolina, convertendo-o em colina e ácido acético, o que faz cessar a ação do neurotransmissor na placa motora e na fenda sináptica. Embora ambas promovam esta reação, a acetilcolinesterase

apresenta maior afinidade pela acetilcolina e conseqüentemente maior velocidade de reação.

A butirilcolinesterase tem como substrato de preferência a butirilcolina, que, embora apresente estrutura similar à acetilcolina, é sintética e não é encontrada no organismo humano, sendo a sua degradação uma das maneiras de distingui-las. A butirilcolinesterase, no plasma, é responsável pela degradação por hidrólise de muitos ésteres de colina e também da procaína e da cocaína, sendo o principal responsável pela eliminação desta do organismo.

As colinesterases podem ser inibidas em suas atividades por substâncias de origem não-humana, como o veneno de algumas serpentes e por outras de fabricação humana, dentre as quais estão principalmente alguns gases neurotóxicos, utilizados como armas químicas e inseticidas organofosforados, de uso mais intenso como defensivos agrícolas. O bloqueio ou a inibição das funções das colinesterases provocam, como sintomas, aumento da movimentação intestinal, bradicardia, hipotensão arterial e redução da pressão dos globos oculares, dilatação pupilar, salivação abundante e sufocação provocada pela constrição dos brônquios, tendo a morte como consequência.

Uma das funções atribuídas à butirilcolinesterase é a de realizar proteção da acetilcolinesterase, por competição, contra a ação destes inibidores de colinesterases, impedindo ou diminuindo sua ação sobre ela, o que pode chegar a ser letal. Sua reação com os inibidores provocaria com eles ligações permanentes, diminuindo a quantidade de inibidor disponível para ligar-se com a acetilcolinesterase.

Os efeitos mais marcantes da deficiência de butirilcolinesterase são o acúmulo, por retardo na eliminação, de procaína, succinilcolina e cocaína. A succinilcolina é utilizada como relaxante muscular em cirurgias, agindo como competidor da acetilcolina na placa motora, impedindo a contração muscular e provocando uma

paralisia muscular flácida, o que é muito conveniente em muitos procedimentos cirúrgicos. A maior parte da succinilcolina administrada é eliminada pela butirilcolinesterase, que a converte em ácido succínico antes mesmo que atinja a placa motora, mas parte dela produz o efeito desejado. Quando há deficiência na ação da butirilcolinesterase, como ocorre no caso de indivíduos homocigotos para alguns alelos de BchE, ou no caso de sua inibição, a succinilcolina permanece no organismo por muito mais tempo que o habitual, bloqueando a placa motora e impedindo a contração muscular. Apneia prolongada e a morte são as principais consequências deste relaxamento muscular que atinge toda a musculatura esquelética.

Inicialmente, quando ainda só se realizavam estudos eletroforéticos e de intensidade de ação enzimática, verificou-se que, independentemente do alelo para o gene BCHE que o indivíduo possuísse, ele poderia ser C5+ (cerca de 10% da população humana) ou C5- (os demais 90%).

Caracterizavam-se os indivíduos C5+ através de um padrão eletroforético: presença de banda diferente da banda presente nos indivíduos C5-. Estes indivíduos com a banda C5+ apresentam mais intensa ação da enzima BChE e apresentam, na média, 5 quilogramas a menos na massa corporal.

Posteriormente, explicou-se esta característica como sendo devida a um polipeptídeo ligado ao tetrâmero da BChE, que seria formado por um outro gene, diferente do gene BCHE e não ligado a ele, então denominado de gene CHE2, que apresentaria dois alelos: C5+ e C5-.

Um grupo de pesquisa isolou e identificou o polipeptídeo que se liga ao tetrâmero BChE e descobriu que é idêntico a um polipeptídeo rico em prolina determinado pelo gene RAPH1, que forma a lamelipodina. Lamelipodina é uma proteína que participa da formação do lamelipódio, que é uma projeção, plana e fina, da

membrana plasmática de algumas células. Esta projeção laminar é estruturada internamente por uma rede de filamentos de actina e de muitas outras proteínas que, inter-relacionadas complexa e dinamicamente, dão-lhes sustentação e, simultaneamente, mobilidade, sendo responsável pelo deslocamento destas células, tanto ao longo de um substrato qualquer, quanto por entre outras células.

Outro grupo de pesquisa, localizando o gene RAPH1, verificou que está no mesmo local, no braço longo do cromossoma 2, em que já se sabia estar, por trabalhos de ligação, o gene CHE2.

A conclusão foi que ambos são um único gene e que os “alelos do gene CHE2” são na verdade dois “alelos” do gene RAPH1, sendo que nos indivíduos com C5+ o polipeptídio é mais longo, provavelmente por haver no DNA (e no RNA) um códon de término alterado, o que faz que o polipeptídio seja traduzido até um outro códon de término a jusante, ou, alternativamente, que haja no DNA (e no RNA) uma diferença em relação ao outro alelo que faça com que um corte deixe de ocorrer nas modificações pós-traducionais.

Estudos de Evolução apresentam interesses variados e utilizam estratégias muito diversificadas. Um grupo destas é denominado de análise filogenética, que permite a realização de inferências tomando por base as semelhanças e diferenças apresentadas por grupos reprodutivos, principalmente espécies.

Os caracteres dos vivos utilizados para fazer estas inferências são também muito diversos: morfológicos, fisiológicos e comportamentais, entre outros.

A partir das décadas de 1960 e de 1970 foram obtidos resultados muito significativos através da comparação das estruturas macromoleculares, principalmente

proteínas, de grupos de espécies. Posteriormente, o uso do DNA passou a fornecer informações ainda mais consistentes e com um grau de refinamento ainda superior.

O presente trabalho tem como propósito comparar um grupo de espécies, no que diz respeito aos seus genes determinantes da lamelipodina, com o propósito de tentar aclarar os acontecimentos evolutivos que envolveram estes genes, nas espécies analisadas. (Souza, 2001) (Arpagus, Kott, Vatsis, Bartels, Ladu, & Lockridge, 1990) (Nunes, 2007)

5 REVISÃO BIBLIOGRÁFICA

Filogenias anteriormente construídas com informações morfológicas ou fisiológicas, como as imunológicas, tiveram grande acréscimo de acuidade com o uso de sequências de nucleotídios. (Brimijoin, Rakonczay, & Mintz, 1986)

Como se sabe, algumas informações seguras sobre a filogenia de um grupo podem permitir ou facilitar a conclusão de estudos, ao produzir uma intensa redução no número de filogenias possíveis. A estrutura do gene ACHE, juntamente com outro, foi utilizada para permitir a escolha de uma dentre muitas possíveis árvores filogenéticas em ascídias. (Jeffery, Swalla, Ewing, & Kusakabe, 1998) (Cleto, et al., 2003)

Há indícios de ter ocorrido seleção natural na determinação de códons do gene da ACHE. Estudos em mais de 500 íntrons e éxons de genes inativos de camundongos revelaram ação de seleção natural na determinação dos códons utilizados. Embora o teste não tenha sido feito especificamente para os genes BCHE e ACHE, este se encontrava entre os mais de 100 genes examinados, o que sugere a possibilidade de seleção natural ter estado em ação na determinação de códons deste gene. (Chamary & Hurst, 2004)

Sabe-se que a BChE está relacionada ao metabolismo de lipídios, e, provavelmente, de alguma maneira, à síntese de membrana plasmática. Sabe-se também que muitas doenças resultantes de degeneração de sistema nervoso, como mal de Parkinson, mal de Alzheimer, quase todas elas tendo ainda causas desconhecidas, são influenciadas pela maneira pela qual se dá a mielinização dos neurônios. Por estes motivos, existem especulações no sentido haver relação entre as formas de BChE e a ocorrência destas formas de degeneração. (Bartzokis, 2004)

Em virtude do ainda pouco conhecimento sobre as funções da BChE, estão sendo realizados trabalhos que tentam elucidar suas relações com a AChE, como o que descobriu que a inibição da BChE, com inibidores seletivos, em ratos, produz o aumento na liberação de AChE na córtex cerebral. (Cerbai, Giovannini, Melani, Enz, & Pepeu, 2007)

Diferentes testes de ação enzimática da AChE e da BChE, em peixe-bruxa, demonstraram ação preferencial da AChE sobre a acetilcolina, como ocorre em outros vertebrados. Verificação de sequência de bases nitrogenadas do DNA demonstrou a presença de sítio ativo da enzima similar também ao de outros vertebrados. (Sanders, Mathews, Sutherland, Soong, Giles, & Pezzementi, 1996)

A resistência de mosquito a inseticida dos mais utilizados, inibidor da AChE, deve-se, em *Culex pipiens*, à presença de alelos que conferem diferentes capacidades de resistir aos inseticidas. Estes alelos estão em genes duplicados, originados em três ou mais processos de duplicação independentes, e mantidos na espécie recentemente por seleção favorecendo a presença de mais de uma cópia do gene. (Labbé, et al., 2007)

Há grande quantidade de esterases em eucariontes, sendo grande parte delas classificadas como carboxilesterases. Estas são subdivididas em três grupos, conforme a ação inibitória de organofosfatos sobre elas. A um dos três grupos pertencem as colinesterases, dentre as quais estão a acetilcolinesterase e a butirilcolinesterase.

As proteases serina-peptidases são enzimas filogeneticamente próximas umas das outras e as esterases também apresentam entre si esta proximidade. Algumas proteases apresentam ação de esterase, enquanto que muitas esterases apresentam ação proteolítica. Mas, embora funcionalmente relacionadas com a família das serina-endopeptidases, as esterases não apresentam com elas muita semelhança na estrutura

molecular. Quanto às sequências de nucleotídeos, também não há muita semelhança entre ambos os grupos enzimáticos.

Apesar, portanto, das semelhanças funcionais, provavelmente ambos os grupos não têm origem comum ou não tem origem comum muito próxima. (Myers, Richmond, & Oakeshott, 1988)

A BChE tem ação de aril-acilamidase conhecida há bastante tempo. Tem ação catalítica na hidrólise de muitos compostos, inclusive das acetanilidas, compostos de ação analgésica em nosso organismo, semelhante ao paracetamol. (Masson, Froment, Gillon, Nachon, Darvesh, & Schopfer, 2007)

Sabe-se que a butirilcolinesterase está, mesmo indiretamente, relacionada a muitas funções, mesmo considerando apenas sua ação de degradação da acetilcolina, que não é tão eficiente quanto à da acetilcolinesterase. A acetilcolina, além de neurotransmissor, age como sinalizador na comunicação entre as células e como hormônio, modulando muitos processos fisiológicos, como o desenvolvimento celular, envolvendo a sobrevivência, a proliferação e a diferenciação celular; o contato célula-célula e a adesão entre células; o controle do ciclo celular; locomoção; atividade elétrica da célula; organização do citoesqueleto; funções relacionadas à nutrição e à secreção; resposta fisiológica ao “stress”; apoptose e funções imunológicas. (Zheng-Xuan, 2008)

A butirilcolinesterase está relacionada, em associação com a acetilcolinesterase, à manutenção da temperatura corporal e da respiração, como demonstra um experimento no qual camundongos com o gene da acetilcolinesterase bloqueado apresentam diminuição do volume respiratório e queda na temperatura corporal quando a butirilcolinesterase é inibida. (Boudinot, et al., 2009)

Recentemente, com o propósito de verificar em que regiões cerebrais age a butirilcolinesterase, foi realizado em camundongos um experimento no qual foi injetada butirilcolinestase fluorescente em camundongos que não a produziam por terem o gene BCHE bloqueado. Verificou-se que houve maior concentração de BChE nos ventrículos cerebrais e na córtex, especificamente no giro denteado do hipocampo. Sabe-se que o giro denteado do hipocampo está relacionado com a memória, principalmente com a conversão da memória recente em memória mantida por longo prazo e, mais recentemente, descobriu-se que lá ocorre a multiplicação de neurônios. Pensava-se até recentemente que ela não ocorresse em adultos, mas agora se sabe que sim, em situações relacionadas com a aprendizagem. (Johnson, Duysen, & Lockridge, 2009)

Um experimento com drogas utilizadas no tratamento contra a doença de Alzheimer teve o objetivo de verificar se os pacientes humanos com alelos pouco eficientes de BCHE seriam tolerantes a este tratamento tradicional. Camundongos com o gene BCHE bloqueado que recebem estas drogas mencionadas morrem com sintomas de problemas neurológicos, o que indica que a BChE desempenha função relacionada à neurotransmissão. (Duysen, Li, Darvesh, & Lockridge, 2007)

Baseados na sequência de nucleotídeos do gene da BCHE de humanos foi feita uma árvore filogenética relacionando 25 espécies, incluindo plantas, nematódeos, cordados e bactérias. (Sridhar, Lakshmi, & Rao, Phylogenetic Tree Construction of Butyrylcholinesterase Sequences in Life Forms, 2006) (Sridhar, et al., 2005)

No intuito de melhorar o conhecimento sobre as bem estudadas funções da AChE e de descobrir algo que ajude a esclarecer as mal conhecidas funções da BChE, foram medidas as atividades enzimáticas de AChE e BChE em cérebro e plasma de 16 espécies de peixes teleósteos dulcícolas, e estas se mostraram muito diferentes de espécie para espécie. A atividade no cérebro é maior que a atividade no plasma, mas a

atividade da BChE não existe no plasma de algumas espécies, sendo inexistente apenas em alguns indivíduos de outras. No entanto, a atividade de BChE tem, nestes casos todos, grande correlação com a atividade da AChE. (Chuiko, Podgornaya, & Zhelmin, 2003)

AChE e BChE apresentam-se no plasma humano em 12 diferentes formas, (Juul, 1968) demonstradas em bandas de eletroforese em gel de poliacrilamida. Estas formas correspondem às formas moleculares das enzimas, homólogas ou heterólogas. As formas homólogas são monômeros, dímeros, trímeros e também tetrâmeros que são a forma mais abundante. Quanto às formas heterólogas, a AChE ou a BChE ligam-se à albumina, à alfa2-macroglobulina, provavelmente ao colágeno (Frederick, et al., 2008), e a muitas outras ainda não identificadas. (Alcântara, 2000) (Nunes, 2007)

Uma função conhecida da BChE em seres humanos é a eliminação da cocaína da corrente sanguínea, realizada nos hepatócitos, a hidrólise catalítica da cocaína. (Gao & Zhan, 2006)

O besouro *Schizaphis graminum* apresenta um único gene da ACHE, que, no entanto, é derivado de uma cópia ancestral diferente daquela que originou a maioria dos genes ACHE dos demais insetos. (Gao, Kambhampati, & Zhu, 2002)

Já há bastante tempo foram feitos estudos enzimáticos e imunológicos da BChE, que forneceram informações sobre a história evolutiva de símios, incluindo bonobo, chimpanzé, gibão, siamango, gorila, homem e orangotango. Foram utilizados anticorpos monoclonais contra a enzima BChE isolada do sangue humano. As enzimas mostraram grande diversidade de atividade entre as espécies e também entre os indivíduos de uma mesma espécie. Os anticorpos monoclonais humanos reagiram com as BChE de todas as espécies, tendo apresentado a seguinte ordem decrescente de afinidade: homem,

chimpanzé, bonobo, gorila, orangotango siamango e gibão. (Mintz, Weinshilboum, & Brimijoin, 1984)

Há muitas evidências de a BChE ser responsável pela destruição, no organismo do homem e de outros seres, de substâncias de alta nocividade. Os inibidores de colinesterase são alguns deles produzidos pelo homem, como alguns inseticidas, tais quais os organofosforados e os carbamatos, e outros são substâncias que ocorrem naturalmente em alguns seres vivos, como, por exemplo, a eserina, no eserê (*Physostigma venenosum*), a cocaína, na coca (*Erythoxylum coca*) e a fasciculina, no veneno de algumas cobras. Estes inibidores afetam tanto a AChE, quanto a BChE e outras colinesterases, e o fato de a BChE ser encontrada em maiores concentrações nos pulmões e no fígado é compatível com uma das funções atribuídas a ela, de detoxicação. (Jbilo, Bartels, Chatonnet, Toutant, & Lockridge, 1994)

Em insetos, existem espécies com formas parálogas do gene da ACHE: no pulgão do algodão (*Aphis gossypii*) foram identificados dois destes genes. (Li & Han, 2002)

Foi verificada, em nematóides, grande preservação de estrutura gênica do gene ACHE, entre espécies de mesmo gênero e entre gêneros diferentes proximalmente aparentados, com manutenção, principalmente, dos resíduos de aminoácidos responsáveis pela conservação da estrutura e da função da enzima. (Piotte, Arthaud, Abad, & Rosso, 1999)

Os cefalocordados anfioxos *Branchiostoma lanceolatum* e *Branchiostoma floridae* foram investigados, no que diz respeito à existência dos genes ACHE e BCHE, tanto enzimaticamente, quanto com uso de sequenciamento de bases nitrogenadas do DNA, e ficou evidente a existência de dois genes, um similar ao ACHE e outro similar ao BCHE, o que faz supor um evento de duplicação gênica que antecedeu ao momento

em que houve a separação de cefalocordados dos demais cordados. (Sutherland, et al., 1997)

Suspeitava-se da existência de dois genes de ACHE em alguns mosquitos, pelo fato de indivíduos resistentes a inseticidas inibidores de colinesterase não apresentarem os alelos associados à resistência. Em experimento com quinze espécies de mosquito (Díptera), ficou evidenciado que há dois genes parálogos de ACHE, sendo que um deles é bastante similar, em sequência de bases nitrogenadas, ao único gene de ACHE das drosófila. (Weill, Fort, Berthomieu, Dubois, Pasteur, & Raymond, 2002)

Informações sobre os genes ACHE e BCHE, seu número e sua sequência de bases nitrogenadas, têm fornecido informações de ordem prática, como por exemplo, a descoberta da maneira pela qual surgiram insetos e carrapatos resistentes a inseticidas: o surgimento, por mutação, de alelo com base nitrogenada alterada, condicionando uma AChE com um aminoácido alterado, o que impede a inibição da enzima pelo inseticida. (Aiki, Kozaki, Mizuno, & Kono, 2005)

Comparações entre genes ACHE e BCHE têm permitido também a solução de problemas na montagem de filogenias. Semelhanças e diferenças na sequência de bases nitrogenadas e também número de genes indicaram proximidade filogenética e apontaram soluções em casos, por exemplo, de ascídias e drosófilas do grupo Repleta. (Jeffery, Swalla, Ewing, & Kusakabe, 1998) (O'Grady, Durando, Heed, Wasserman, Etges, & Desalle, 2002).

6 OBJETIVOS

6.1 OBJETIVO GERAL

O presente trabalho teve o propósito de, comparando um grupo de espécies, no que diz respeito aos seus genes determinantes da lamelipodina, tentar aclarar os acontecimentos evolutivos que envolveram estes genes, nas espécies analisadas.

6.2 OBJETIVOS ESPECÍFICOS

Obter as sequências de DNA, com íntrons e éxons, de cDNA e de aminoácidos do gene *RAPH1* e de seus transcritos em diferentes espécies.

Definir as regiões das sequências obtidas a serem utilizadas, para então proceder ao alinhamento e ao ajuste das mesmas.

Caracterizar os genes em termos de composição de aminoácidos de seus polipeptídios.

Construir e comparar filogenias entre si.

Entender a origem evolutiva do gene.

7 JUSTIFICATIVA

Inicialmente, quando ainda só se realizavam estudos eletroforéticos e de intensidade de ação enzimática, verificou-se, em seres humanos, que, independentemente do alelo para o gene BCHE que o indivíduo possuísse, ele poderia ser, para o gene CHE2, C5+ (menos de 10% da população humana) ou C5- (mais de 90% dela).

Os indivíduos C5+ caracterizam-se através de seu padrão eletroforético: presença da banda usual e de uma banda adicional, sendo que os indivíduos C5- apresentam a banda usual e não a adicional. Estes indivíduos com a banda adicional apresentam, entre outras diferenças, mais intensa ação da enzima BChE e apresentam, na média, 6 quilogramas a menos na massa corporal, em comparação com os indivíduos C5-.

A existências destes dois variantes eletroforéticos ficaram sendo explicados através da ação de dois alelos do gene CHE2, que determinam a presença das bandas usual e usual juntamente com a adicional. O gene CHE2 determina a formação de um polipeptídio que se liga ao tetrâmero da BChE, na proteína madura, sendo, portanto, BCHE e CHE2 dois genes, que outros trabalhos demonstraram não estar ligados.

O polipeptídio que se liga ao tetrâmero BChE foi isolado, identificado e posteriormente sequenciado, o que permitiu que se descobrisse que é idêntico a um polipeptídio rico em prolina determinado pelo gene RAPH1, que forma a lamelipodina. Posteriormente, localizando-se o gene RAPH1, verificou-se que está no mesmo local, situado no braço longo do cromossoma 2, em que já se sabia estar, por trabalhos de ligação, o gene CHE2.

Ambos são, portanto, um único gene e os “alelos do gene CHE2” são, na verdade, dois “alelos” do gene RAPH1, sendo que nos indivíduos com C5+ o polipeptídio é mais longo, provavelmente por haver no DNA (e no RNA) um códon de término alterado, o que faz que o polipeptídio seja traduzido até um outro códon de término à jusante, ou, alternativamente, que haja no DNA uma diferença em relação ao outro alelo que faça que um corte deixe de ocorrer nas modificações pós-traducionais.

Conhecer a origem evolutiva do gene RAPH1 pode aclarar o seu relacionamento com o gene BCHE e também permitir compreender os motivos pelos quais o polipeptídio desempenha duplo papel: na formação do lamelipódio e na associação com a BChE.

8 CAPÍTULO 1 - ANÁLISE DO GENE RAPH1 COM USO DE FILOGENIAS

8.1 RESUMO

Foram analisadas árvores filogenéticas construídas com base em 53 sequências de DNA e 53 sequências de aminoácidos, correspondentes ao gene RAPH1, de 48 diferentes espécies de animais, incluindo nematódeo, peixes, anfíbio, réptil, mamíferos e aves. Foram construídas árvores com as sequências inteiras e com segmentos delas correspondentes aos domínios proteicos GM, RA, PH e EVH1 e aos sítios de ligação de proteínas que a compõem. Foram utilizados métodos de construção de filogenia baseados em distância, na Máxima Parcimônia e na Máxima Verossimilhança.

8.2 INTRODUÇÃO

Um diagrama de árvore é uma representação gráfica, com formato que lembra a estrutura geral de uma árvore real e que tem por propósito mostrar relações entre coisas de várias naturezas. (Amorin, 1997) (Mount, 2004)

Muitas vezes estes diagramas mostram relações de semelhança e diferença, referentes geralmente à estrutura das coisas representadas, mas podem representar características das mais variadas naturezas.

Estes diagramas são muitas vezes usados para representar relações entre seres vivos, e mais frequentemente relações entre suas estruturas morfológicas.

A forma de organizar e classificar os seres vivos com base nestas semelhanças e diferenças é denominada Fenética e se utiliza muito frequentemente de diagramas de árvore.

Uma vez feita a assunção de algumas premissas, principalmente a origem única de todos os seres vivos considerados, a origem das espécies novas por subdivisão das espécies já existentes e a ocorrência de crescente divergência entre os seres vivos à medida que o tempo passa, é possível utilizar um diagrama de árvore para representar graficamente a história evolutiva de um determinado grupo de espécies.

O diagrama, então, pode representar as relações de semelhanças e diferenças gerais dos patrimônios genéticos, bem como as sequências de subdivisão de espécies ocorridas ao longo do tempo e ainda os tempos relativos que separam os eventos. Dizendo de outro modo, mostra a história das origens das espécies e seus graus de parentesco.

Uma forma diferente de organizar e classificar os seres vivos toma por base não apenas as semelhanças e diferenças morfológicas, mas leva em conta as histórias evolutivas dos grupos por classificar. Esta forma é denominada Cladística ou Sistemática Filogenética e utiliza-se igualmente de diagramas de árvore para descobrir suas inter-relações e propor modos de classificar. (Graur & Li, 2000)

Mesmo que houvesse um consenso sobre qual é a árvore, única, que expressa a história da vida, seria difícil utilizá-la para, a partir exclusivamente dela e desprezando todo o conhecimento anterior, construir uma nova classificação taxonômica. A extrema irregularidade da história dos seres vivos não permitiria a construção de uma classificação útil e prática, se se pretendesse, com a classificação taxonômica, refletir a história da vida. (Nei & Kumar, 2000)

Mas, mesmo assim, é uma tendência atual usar cada vez mais os conhecimentos da história evolutiva dos vivos para “corrigir” as classificações tradicionais, ou, dizendo de outro modo, a taxonomia está deixando de ser puramente fenética e passando progressivamente a ser cladística.

Árvores filogenéticas são construídas a partir de características dos vivos atuais, a partir de seres vivos já extintos ou de uma combinação de ambos.

A maior parte delas foi construída a partir de características morfológicas e, dentre elas, principalmente os esqueletos rígidos, no que diz respeito às formas extintas. E no que diz respeito às espécies atuais, principalmente a partir de características morfológicas, fisiológicas e embrionárias.

Filogenias desde muito foram reconstituídas através do uso de características morfológicas, mas surgiram na década de 1960 os primeiros métodos para construir filogenias considerando a morfologia, porém quantificando os caracteres morfológicos analisados. Usando exclusivamente valores numéricos, estatísticos, atribuídos aos caracteres, filogenias passaram a ser construídas, em uma tentativa de deixar menos subjetivos os procedimentos. Este método, ou conjunto de métodos, é denominado Taxonomia Numérica e foi proposto inicialmente por (Sokal & Sneath, 1973). Surgiram independentemente, na mesma década, 1964, outras formas de construir filogenias de forma numérica, que foram propostas inicialmente por (Cavalli-Sforza & Edwards, 1967), fazendo uso de frequências alélicas de vários sistemas genéticos para avaliar as distâncias relativas entre populações. (Diniz Filho, 2000) (Schneider, 2003)

A partir destes dois conjuntos de métodos numéricos, foram propostos muitos outros, com a mesma idéia básica de quantificar diferenças e semelhanças entre grupos de vivos, para então depois descobrir as prováveis relações filogenéticas entre eles.

Recentemente tem havido um crescimento extraordinário dos métodos para se fazer inferência de filogenias, e isto se deve principalmente ao aumento da disponibilidade de informação genética referente às macromoléculas, proteínas, DNA e RNA. Sequências de aminoácidos e de ácidos nucleicos são propícias à avaliação quantitativa proposta pelos métodos idealizados a partir da década de 1960, e, como

várias melhorias na fidelidade, preço e rapidez dos métodos de obtenção destas sequências aumentaram muito a informação disponível, houve também um aumento das proposições de novas metodologias.

Embora praticamente todos os métodos de reconstituição de filogenia usados atualmente estejam baseados na quantificação de diferenças, de semelhanças ou de outros fatores, os métodos através dos quais as filogenias são feitas podem ser classificados em três principais grupos, cada um deles tendo como norteador um diferente princípio geral.

Os Métodos de Distância, ou Geométricos, são métodos que se baseiam nas distâncias encontradas ao se fazerem comparações entre sequências homólogas de aminoácidos ou nucleotídios. Sequências homólogas de ácidos nucleicos ou proteínas pertencentes a diferentes grupos de seres vivos apresentam diferenças que são aproximadamente proporcionais ao tempo decorrido desde o momento em que os grupos de seres vivos aos quais elas pertencem se separaram. É um pressuposto que as diferenças que ocorrem por mutação e permanecem na sequência nucleotídica do gene o fazem pelo fato de as novas sequências serem funcionalmente equivalentes às anteriores, o que determina que o acúmulo de alterações se dê em um ritmo constante. Assim, a quantidade de diferença que duas sequências apresentam equivale à distância que as separa em sua história evolutiva, e esta é a base para a construção de filogenias neste grupo de métodos.

Os Métodos de Máxima Parcimônia baseiam-se em um princípio relativamente simples que consiste na minimização dos fenômenos evolutivos necessários para explicar a história dos grupos de seres vivos analisados. Quando são analisadas sequências, os fenômenos evolutivos considerados são substituições de nucleotídios ou

aminoácidos e os métodos optam por, dentre as filogenias possíveis, aquelas que apresentem o menor número delas.

Os Métodos Probabilísticos, ou de Máxima Verossimilhança, têm como base o princípio da maximização da probabilidade de ocorrência dos fenômenos necessários para originar as sequências de nucleotídios ou aminoácidos em estudo. Dizendo de outro modo, os métodos probabilísticos, dentre todas as filogenias possíveis para explicar a origem das sequências analisadas, opta por aquela que apresente a maior probabilidade de ter realmente ocorrido. É óbvio que, para que este cálculo probabilístico possa ser realizado, é necessária a adoção de um conjunto de modelos evolutivos, com as correspondentes probabilidades de ocorrência de cada um dos possíveis fenômenos evolutivos. Para a escolha da melhor árvore, estes métodos levam em conta todas as formas que os possíveis ancestrais comuns das sequências possam ter tido, em todos os momentos de subdivisão de grupos taxonômicos. Sendo calculadas todas as probabilidades de ocorrência de todas as formas ancestrais das sequências analisadas, sabe-se a probabilidade de ocorrência de toda a árvore, o que permite escolher a de maior probabilidade. Estes métodos além de determinarem a topologia, que é o “desenho” das árvores, forma ramos de diferentes comprimentos, determinados a partir das probabilidades calculadas.

As árvores mencionadas até agora se referem a representações gráficas que mostram relações entre grupamentos de seres vivos, sejam eles espécies, grupamentos taxonômicos de nível mais elevado ou populações de uma mesma espécie. Mas uma forma similar de representação, também com forma de árvore, lida diretamente com genes. Nestes diagramas, apenas um gene, com suas formas ortólogas, é analisado, sem levar em conta em que grupos de seres vivos estas formas estavam presentes. Quando posteriormente esta forma de árvore de genes é comparada com as anteriormente

mencionadas, são obtidas informações acerca de quando e onde as formas alélicas surgiram e acerca de seu desaparecimento ou manutenção.

O propósito deste trabalho foi obter informações sobre a história evolutiva das espécies utilizadas, no que se refere ao gene RAPH1 (Ras association (RalGDS/AF-6) and pleckstrin homology domains 1).

A construção de filogenias, através de várias metodologias, permite atingir o propósito de verificar se a história do gene, ao longo do tempo, nestas espécies, esteve em concordância com a história evolutiva dos demais genes delas, o que equivale a dizer, grosso modo, em concordância com a história evolutiva delas, inferida através de muitos outros recursos.

8.3 MATERIAL E MÉTODOS

O gene RAPH1 foi o objeto do presente trabalho. É também chamado informalmente de Gene da Lamelipodina e foi anteriormente nominado através de um grande número de termos ainda presentes em grande parte dos trabalhos referentes a ele. Através do HUGO Gene Nomenclature Committee foi fixado o nome do gene em Ras Association (RalGDS/AF-6) and Pleckstrin Homology Domains 1 (RAPH1), deixando como sinóníma os demais termos. Como código de nomenclatura foi fixado o termo HGNC 14436.

Esse gene está localizado no braço longo do cromossomo 2 e é constituído por 17 éxons (Eppert, Wunder, Aneliunas, Tsui, Scherer, & Andrulis, 2005).

Cinquenta e três sequências de aminoácidos e o mesmo número de sequências de nucleotídios, apresentadas nas tabelas 1, 2, 3, e 4, foram utilizadas para realizar as comparações que permitiram estabelecer suas relações filogenéticas.

Estas sequências foram obtidas através da Rede Internacional de Computadores (Internet), principalmente das páginas Ensembl, do European Bioinformatics Institute e do Wellcome Trust Sanger Institute, acessada através do Universal Resource Locator (URL) <<http://www.ensembl.org/index.html>> e a do National Center for Biotechnology Information (NCBI), acessada através do URL <<http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>>. (NCBI) (Ensembl)

As sequências encontradas nestes bancos de informação genética apresentam-se em vários graus de completude e sob denominações que variam muito entre as espécies, mesmo quando apresentam alto grau de similitude.

As espécies apresentadas na tabela 2 foram as 48 espécies das quais se obtiveram, dos bancos de sequências disponibilizados na rede, 53 sequências nucleotídicas, de DNA e de RNA, e 53 sequências aminoacídicas, que correspondessem a uma parte razoável da extensão do gene. E, ainda, que estivessem anotadas sob a denominação de RAPH1, ou sob quaisquer denominações equivalentes.

Nestes bancos de informação genética, as sequências nucleotídicas depositadas tanto correspondem a nucleotídeos de ácido desoxirribonucleico (DNA) do gene como um todo, incluindo íntrons e éxons, quanto a nucleotídeos de ácido ribonucleico (RNA) correspondentes principalmente a ácidos ribonucleicos mensageiros (RNAm), em uma ou mais das composições alternativas utilizadas pelo organismo.

Foi feito o alinhamento das sequências obtidas, para avaliar sua completude e a correspondência delas entre si. Para isto foi utilizado o programa alinhador de múltiplas sequências ClustalW, contido no programa Bioedit, da North Carolina State University, obtido de sua página na rede, acessada através do URL <<http://www.mbio.ncsu.edu/BioEdit/BioEdit>>. (BioEdit)

Foram utilizadas no trabalho também sequências de aminoácidos correspondentes às sequências nucleotídicas referidas. Estas correspondem aos RNAm identificados e sequenciados que tiveram suas sequências nucleotídicas convertidas em aminoácidos pelos próprios bancos fornecedores de sequências. Como no caso das sequências de nucleotídios, as sequências de aminoácidos também foram alinhadas para avaliar sua completude e a correspondência delas entre si. Para isto foi também utilizado o programa alinhador de múltiplas sequências ClustalW, contido no programa Bioedit, da North Carolina State University. (Clustalw)

A cada gene correspondem, muitas vezes, mais de um RNAm, em suas várias composições alternativas, e por isto, foi necessário descobrir a que sequência de aminoácidos corresponderia um RNAm que incluísse todos os éxons do gene, utilizando para isto a informação contida nas sequências de aminoácidos correspondentes a todos os RNAm de um determinado gene. E isto só foi possível com a realização do alinhamento entre todas as possíveis composições alternativas de cada um dos genes de cada uma das espécies, o que foi também realizado com o uso do programa ClustalW.

Para os alinhamentos inicialmente mencionados, foram utilizados os seguintes parâmetros: penalidade para abertura de lacuna de 18 e penalidade para extensão de lacuna de 0,3 e para os posteriores, os seguintes: penalidade para abertura de lacuna de 0,1 e penalidade para extensão de lacuna de 0,1. Todos os alinhamentos foram inspecionados e corrigidos, quando necessário, manualmente.

As sequências de aminoácidos das espécies que as apresentaram com maior grau de completude foram divididas, cada uma delas, de acordo com a correspondência de seus segmentos com a sequência de aminoácidos do gene humano. Ou seja, as proteínas de todas estas espécies foram divididas de acordo com a segmentação da proteína humana Lamelipodina em domínios proteicos. Portanto, com esta subdivisão, foi

possível adotar para cada um dos domínios da proteína, os procedimentos adotados para as proteínas inteiras.

Uma vez alinhadas, ajustadas, corrigidas, montadas e delimitadas, estas sequências todas, de aminoácidos e nucleotídios, serviram de base para a construção das árvores que expressam suas relações filogenéticas, permitindo analisar visualmente as semelhanças e diferenças entre elas.

Para estas construções, foram utilizados cinco métodos, baseados nos três mais importantes princípios norteadores da construção de árvores filogenéticas. Os métodos utilizados foram de Máxima Verossimilhança (ML), Neighbor Joining (NJ), de Evolução Mínima (ME), Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA) e de Máxima Parcimônia (MP), sendo que os métodos Neighbor Joining, de Evolução Mínima e Unweighted Pair-Group Method Using Arithmetic Averages são métodos de distância, ou também chamados métodos geométricos. O método de Máxima Verossimilhança é um método probabilístico, enquanto que o de Máxima Parcimônia é denominado pelo próprio princípio que o norteia. (Amorin, 1997)

Todos estes métodos foram utilizados através do programa MEGA, em sua versão 5.05, disponível através da URL <<http://www.megasoftware.net/mega.html>>.

Com a finalidade de tentar elucidar os fenômenos evolutivos ocorridos com o gene RAPH1, o gene foi tomado não só por inteiro, mas também por seus trechos correspondentes aos domínios das proteínas determinadas pelo gene e aos sítios de ligação de proteínas presentes nele. Portanto, cada uma das 53 sequências de aminoácidos foi segmentada em seus domínios proteicos, que no caso de RAPH1 são os domínios GM, RA, PH e EVH1, e nos sítios de ligação de proteínas, que no caso são o sítio de ligação da Profilina e o sítio de ligação da SH3.

Foram construídas árvores filogenéticas com os segmentos de aminoácidos correspondentes a cada um destes domínios e sítios de ligação, com todas as sequências disponíveis. E, para isto, foram utilizados métodos de construção de árvores baseados na distância, bem como um baseado no princípio de Máxima Parcimônia e outro, no princípio da Máxima Verossimilhança.

8.4 RESULTADOS E DISCUSSÃO

Todas as sequências polipeptídicas utilizadas foram comparadas no que diz respeito à sua composição de aminoácidos, o que pode ser visualizado na Tabela 6, Tabela 7, Tabela 8 e Tabela 9. Como exemplo foram mostradas, nas figuras 1, 2, 3, 4, 5, e 6, as composições aminoacídicas de polipeptídios de seis das espécies utilizadas, pertencentes a diferentes grupos taxonômicos.

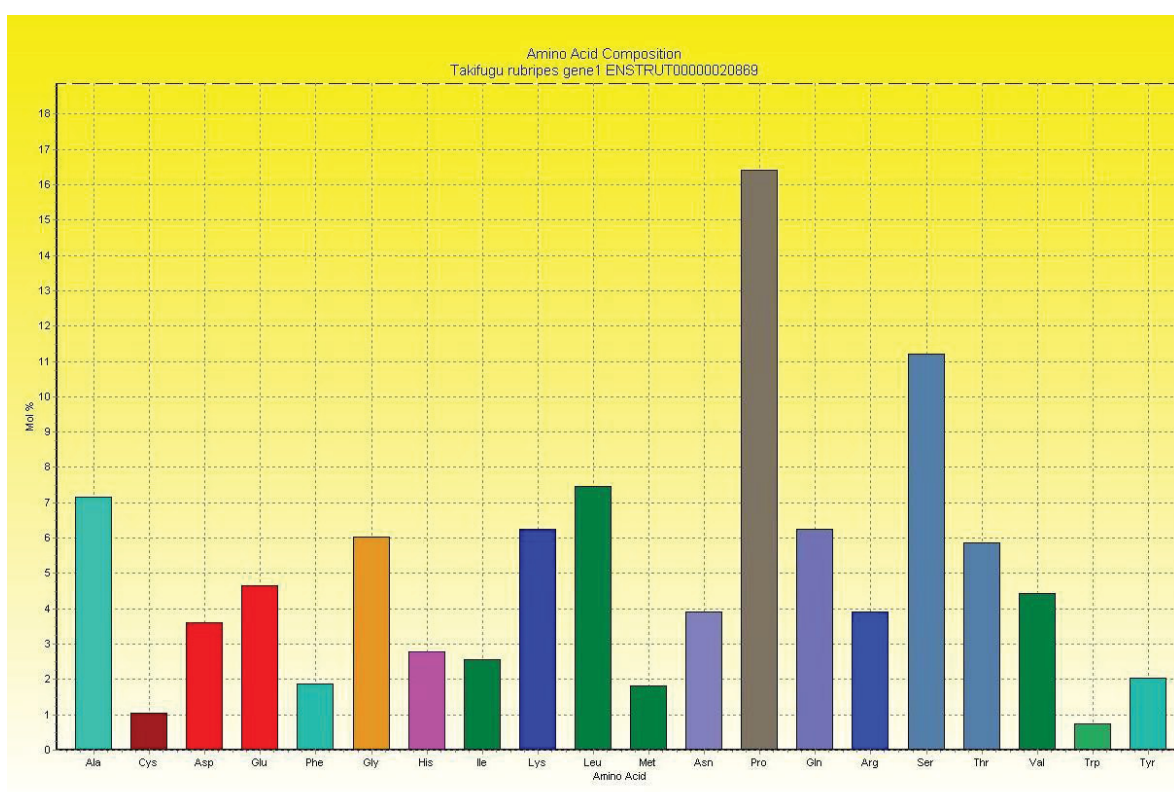


Figura 1 - Composição Aminoacídica do Gene 1 de *Takifugu rubripes*

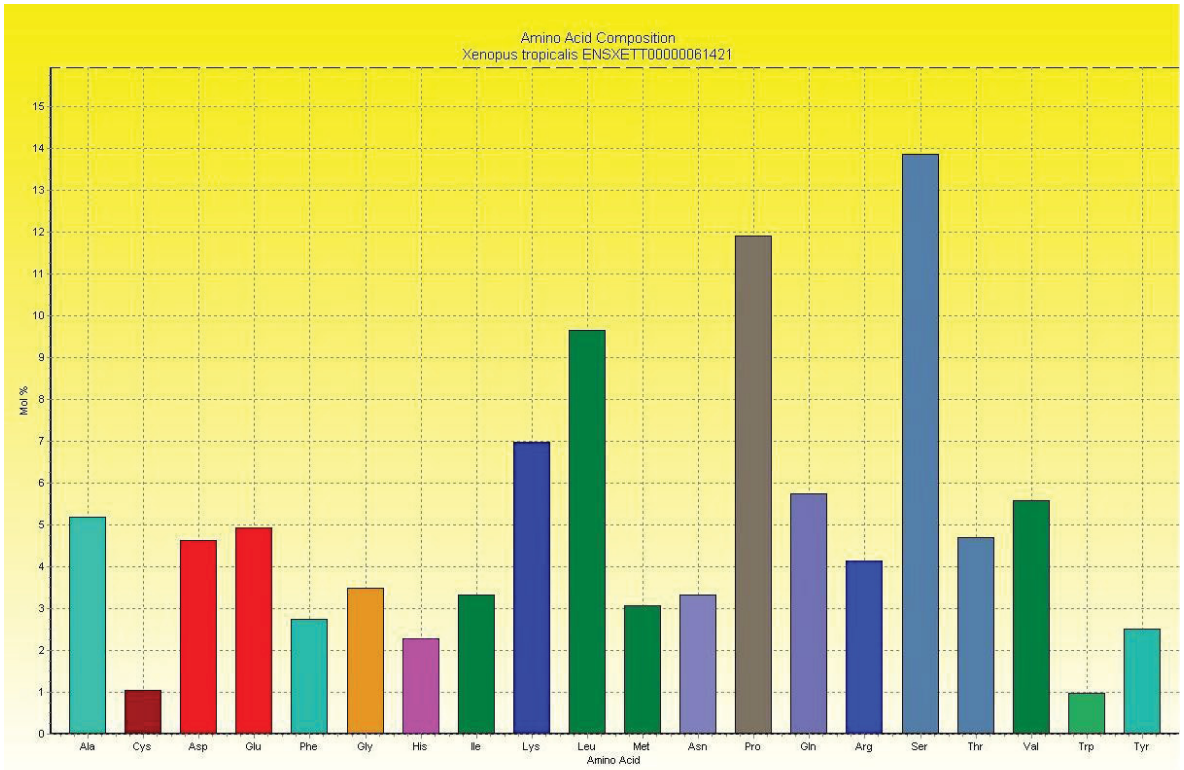


Figura 2 - Composição Aminoacídica de *Xenopus tropicalis*

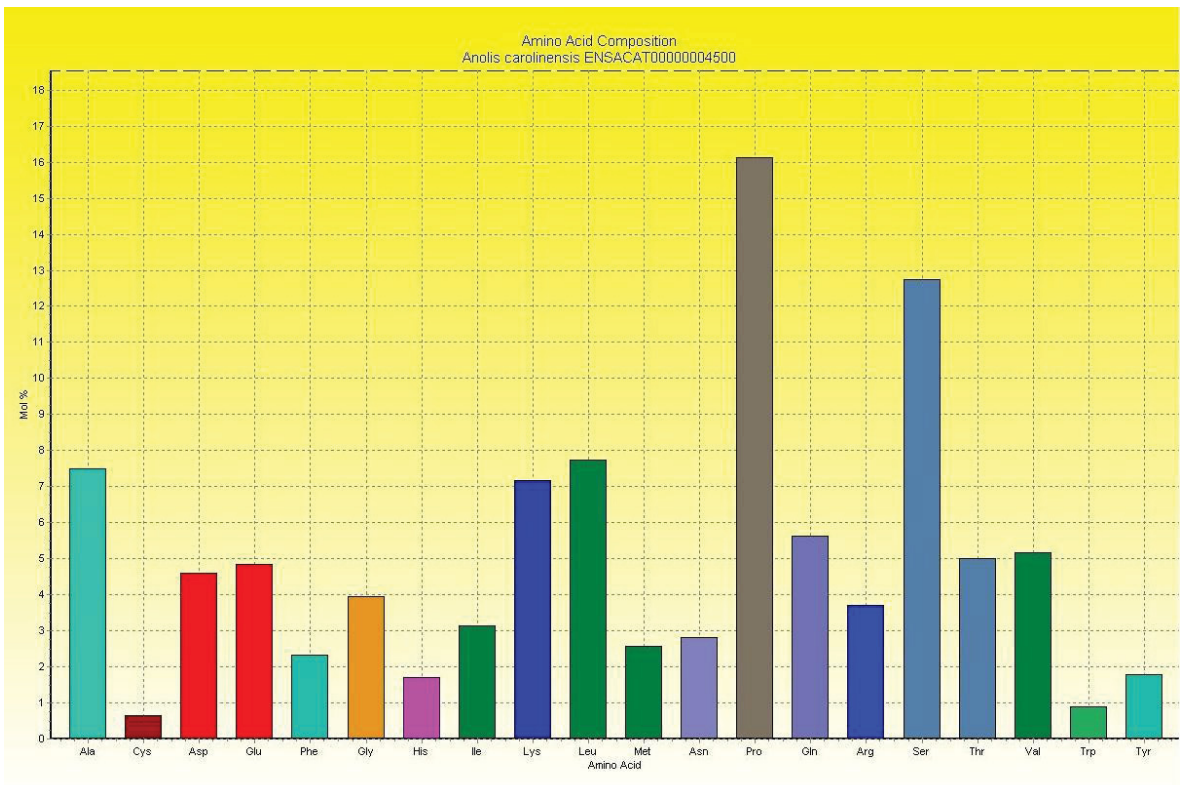


Figura 3 - Composição Aminoacídica de Anolis carolinensis

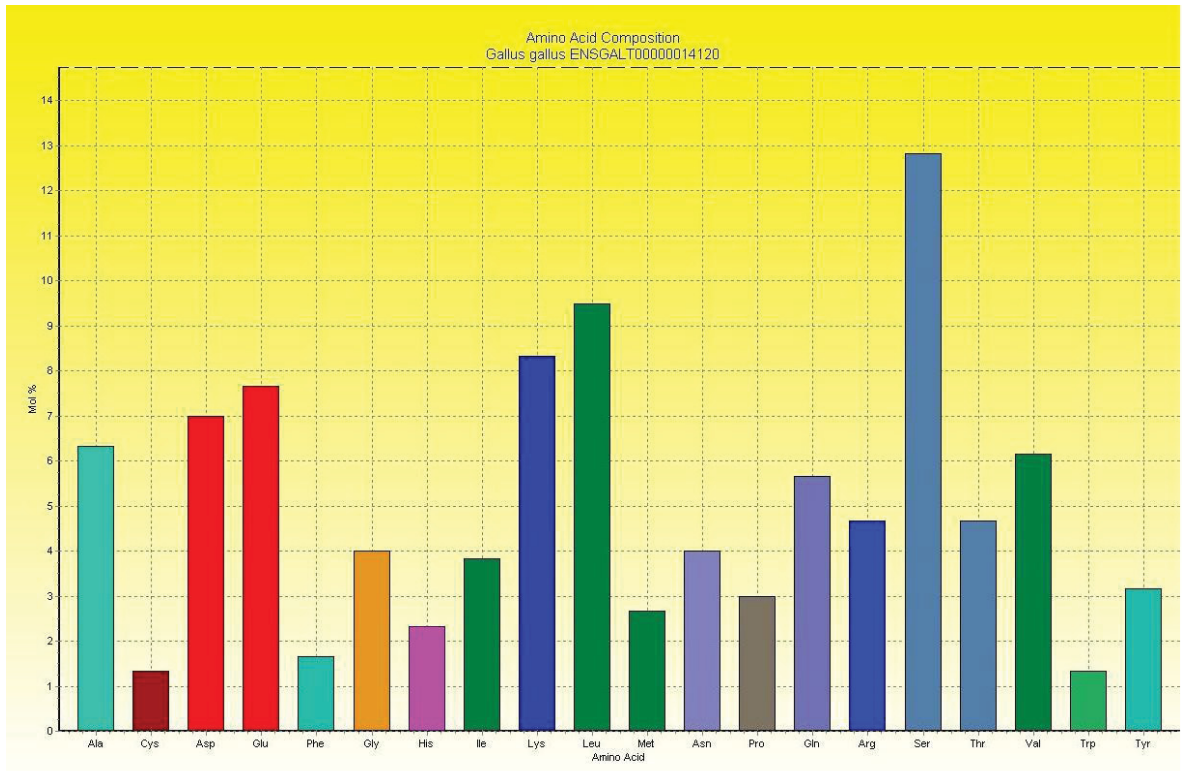


Figura 4 - Composição Aminoacídica de Gallus gallus

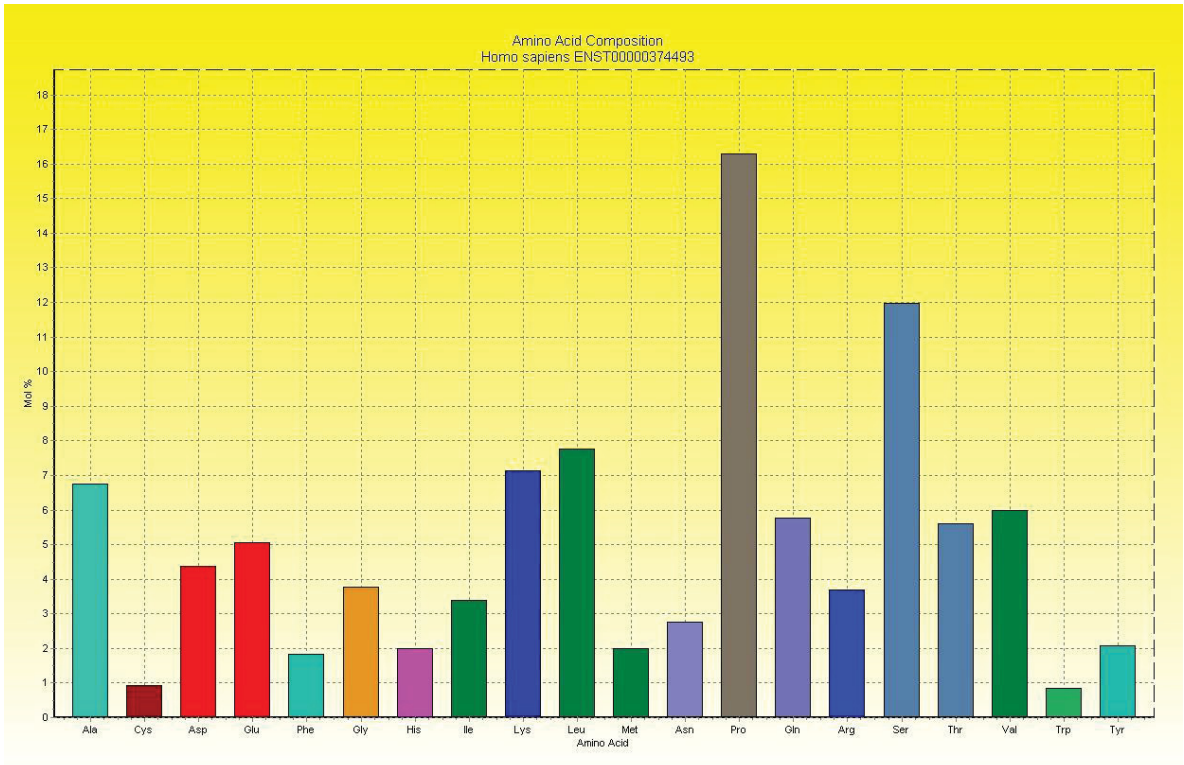


Figura 5 - Composição Aminoacídica de Homo Sapiens

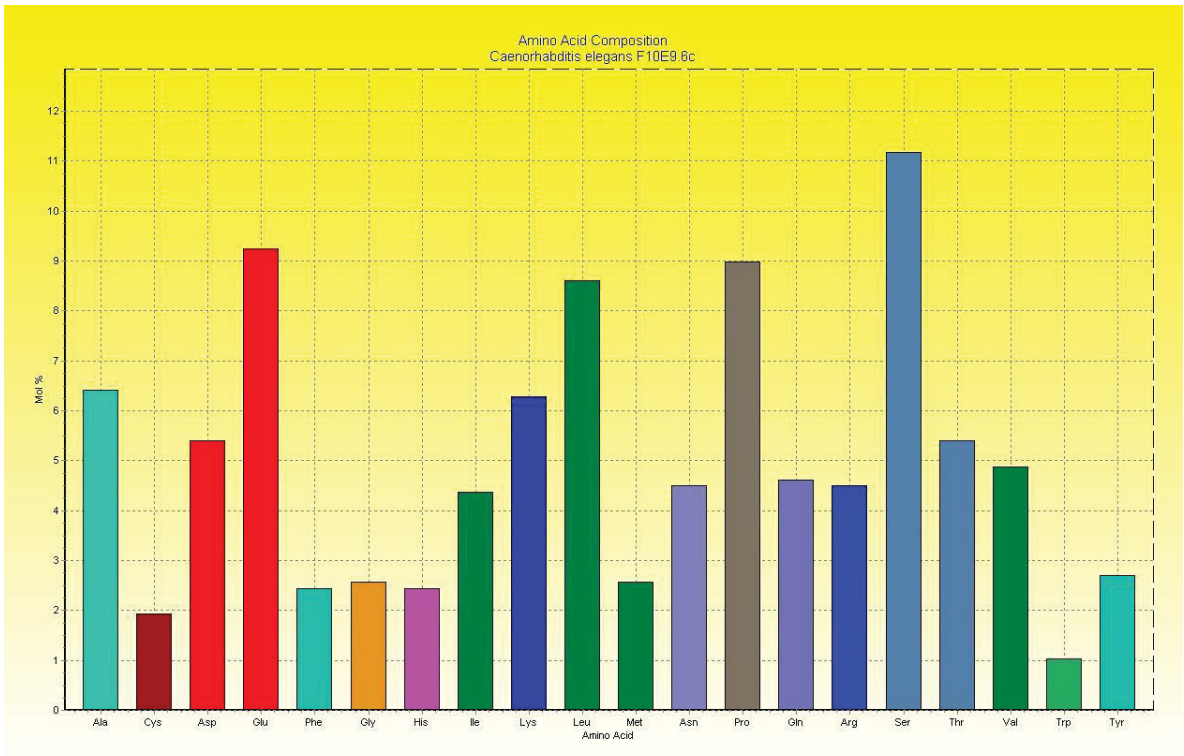


Figura 6 - Composição Aminoacídica de Caenorhabditis elegans

Como mostra a Figura 7, estes seis polipeptídios foram comparados, quanto à hidrofobicidade ao longo da sequência de aminoácidos, deixando evidente que as diferenças são grandes entres os polipeptídios das diferentes espécies, em alguns trechos deles, mas são muito pequenas em outros.

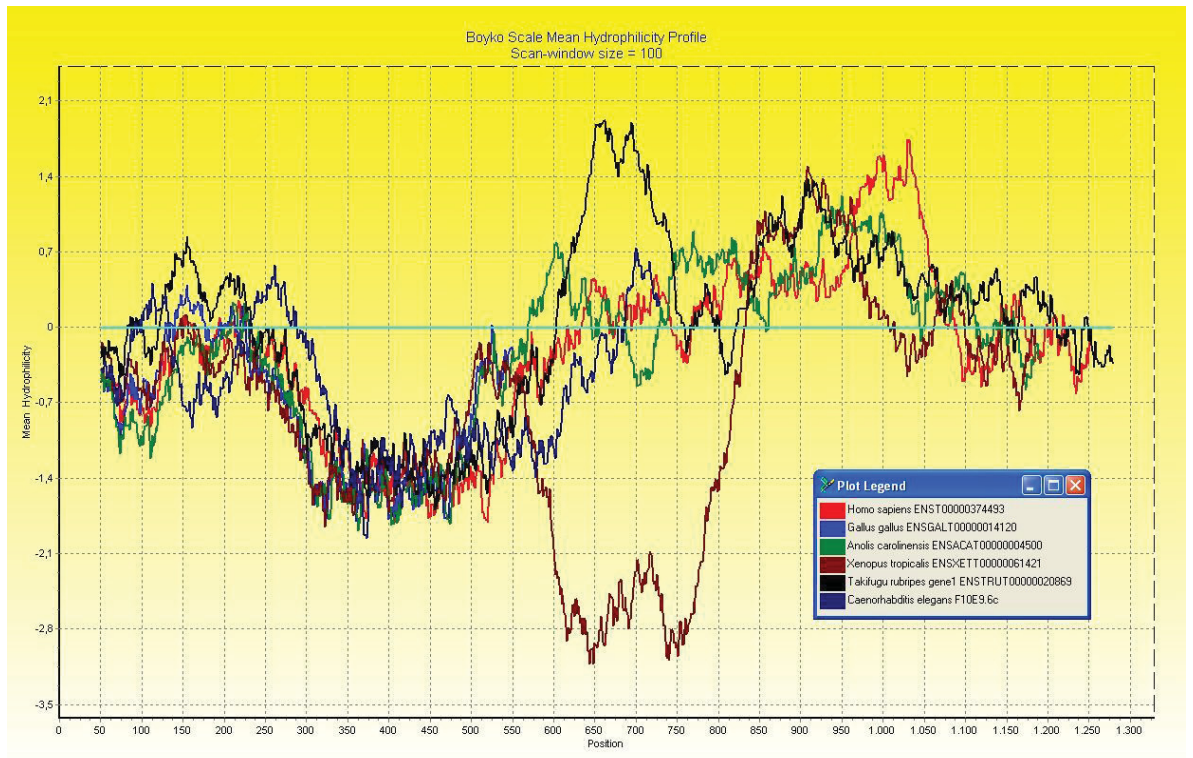


Figura 7 – Hidrofobicidade em RAPH1 de seis espécies

Os diferentes domínios funcionais do transcrito do gene RAPH1 humano, mostrado na figura 8 abaixo, permitem compreender em parte as diferenças de hidrofobicidade ao longo do transcrito do gene, entre as seis espécies comparadas acima.

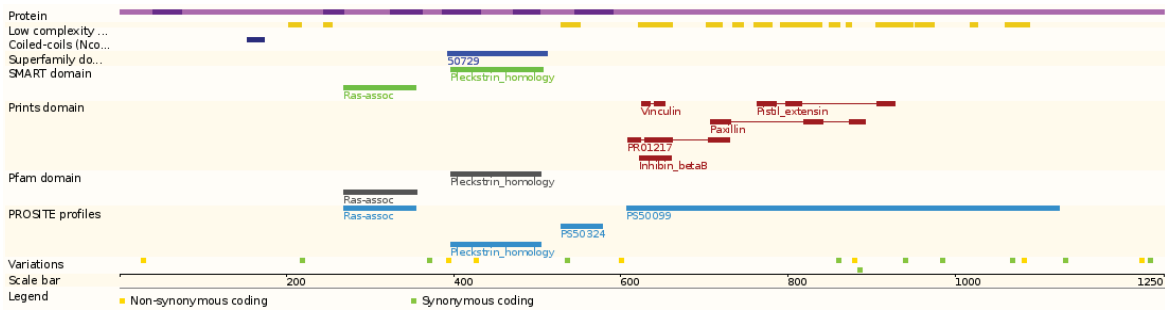


Figura 8 - Domínios funcionais do transcrito do gene RAPH1 em *Homo sapiens* (ENSP00000316543)

Fonte: http://www.ensembl.org/Homo_sapiens/Transcript/ProteinSummary

?db=core;g=ENSG00000173166; r=2:204259068-204400133; t=ENST00000319170

(Ensembl)

8.4.1 ÁRVORES DE CDNA DE 53 SEQUÊNCIAS

Utilizando o Método de Máxima Verossimilhança, a árvore filogenética construída com as 53 sequências de DNA disponíveis, de 48 diferentes espécies, mostrou muitas características que já se esperavam encontrar.

Os peixes todos formaram um único conjunto monofilético, próximo a *Xenopus* e com dois subconjuntos bem delimitados.

As aves encontraram-se juntas, formando um conjunto monofilético, com *Gallus* e *Meleagris* mais próximos entre si que de *Taeniopygia*, porém o conjunto todo das aves ficou afastado dos demais tetrápodos, permanecendo junto com *Caenorhabditis*, que, como era de se esperar, encontrou-se afastada de todas as demais espécies.

Os três marsupiais formaram também um grupo monofilético e os primatas formariam também, se não fossem exceções *Tarsius* e *Microcebus*, que, não ficando entre os primatas, ficaram mais próximos de roedores e lagomorfos, respectivamente. Mas, dentro dos primatas, os símios formaram, estes sim, um grupo monofilético, com *Homo* e *Pan* bem próximos e *Gorilla* junto a ambos.

As duas espécies de morcego utilizadas posicionaram-se muito distantes entre si: uma próxima aos primatas e outra junto a *Ailuropoda*.

Os carnívoros também não ficaram próximos entre si: *Felis* aproximou-se de roedores, *Ailuropoda* aproximou-se de um morcego e *Canis* ficou próximo, mas não muito, de um grupo bem diversificado, com artiodáctilos, perissodáctilos e primatas.

Não ocorreu a proximidade esperada entre *Bos* e *Sus*, como artiodáctilos, e nem sua proximidade com *Tursiops*, que se sabe pertencer a um grupo, cetáceos, que é filogeneticamente aproximado dos artiodáctilos.

Tupaia e *Sorex* ficaram, como se esperava, relativamente juntos, por pertencerem a grupos filogeneticamente próximos, Scandentia e Soricomorpha, que até algum tempo estavam sob uma única denominação: Insetívoros.

Os lagomorfos ficaram juntos e próximos a *Erinaceus*.

Nas construções das árvores com métodos de distância e de Máxima Parcimônia, observaram-se algumas diferenças em relação à árvore de Máxima Verossimilhança.

Nos métodos de distância, os peixes, em sua grande maioria, se mantiveram agrupados, porém com *Latimeria* e *Petromyzon* afastados.

Os lagomorfos e *Erinaceus* permaneceram juntos entre si e as aves também, porém elas, nos métodos de distância, ficaram mais próximas dos peixes, e não afastadas de todas as demais espécies, como se observou no método de Máxima Verossimilhança.

Os marsupiais formaram, como havia ocorrido no método de Máxima Verossimilhança, um grupo monofilético, e isto se deu tanto nos métodos de distância, quanto no de Máxima Parcimônia.

Algumas espécies que se esperava que formassem grupo monofilético na árvore de Máxima Parcimônia, realmente formaram, como ocorreu com os lagomorfos entre si, com os marsupiais, com os roedores e com os artiodáctilos.

Primatas teriam também constituído um grupo monofilético, se não fosse o caso de *Tarsius* e *Microcebus* ficarem afastados do restante do grupo.

8.4.2 ÁRVORES DE CDNA DE 39 SEQUÊNCIAS

Com Método de Máxima Verossimilhança, foi construída uma árvore filogenética com as 39 melhores sequências de DNA, no que diz respeito à sua completude, dentre todas as sequências disponíveis.

Os primatas quase formaram um grupo monofilético, isto só não ocorrendo pelo fato de *Microcebus* ter ficado fora do grupo, próximo a roedores, e também por ter ficado dentro um par de espécies formado por *Ailuropoda* e *Pteropus*. Esta proximidade de *Microcebus* com roedores foi observada na árvore construída com 53 sequências, e mantida nesta.

Mas, dentro dos primatas, os símios não formaram um grupo monofilético por terem ficado dentro *Callithrix* e o par de espécies acima mencionado.

Os peixes todos formariam um único conjunto monofilético, se *Latimeria* não tivesse ficado de fora do grupo, posicionada mais proximamente do tetrápodos, o que não é estranho pelo fato de esta espécie ser originária do grupo de peixes que os formou. Os demais peixes formaram os dois grupos já observados nas outras árvores filogenéticas.

As aves encontraram-se juntas, formando um conjunto monofilético, com *Gallus* e *Meleagris* mais próximos entre si que de *Taeniopygia*. Este conjunto ficou, como era

de se esperar, próximo aos demais tetrápodos, e não afastado deles como se observou na árvore formada com as 53 sequências.

Caenorhabditis ficou estranhamente muito mais próxima das aves do que de todos os outros cordados.

Anolis, por sua vez, ficou próxima ao grupo dos tetrápodos, porém ficou mais próxima às aves que aos demais membros do grupo.

Vicugna, que havia ficado afastada de *Bos*, na árvore de Máxima Verossimilhança de 53 sequências, permaneceu afastada nesta, porém, em ambas, ficou muito próxima de *Choloepus*.

As duas espécies de carnívoro ficaram, tanto na árvore de Máxima Verossimilhança de 53 sequências quanto nesta, bastante afastadas entre si.

As duas espécies marsupiais formaram um grupo monofilético, como também as duas espécies de roedores.

Oryctolagus, que havia formado grupo monofilético na árvore de Máxima Verossimilhança de 53 sequências com *Ochotona* ficou próxima de *Loxodonta*.

Nas construções das árvores com sequências de DNA com métodos de distância e de Máxima Parcimônia, observaram-se algumas diferenças em relação à árvore de Máxima Verossimilhança. Os peixes agruparam-se nas árvores construídas por métodos de distância, ficando, no entanto, *Latimeria* e *Petromyzon* afastados. Na árvore construída por Máxima Parcimônia o mesmo ocorreu, e em todos estes casos *Petromyzon* afastou-se de peixes e tetrápodos e *Latimeria*, afastando-se dos peixes aproximou-se dos tetrápodos.

As aves formaram grupo monofilético nas árvores formadas por quaisquer métodos, ficando sempre *Gallus* e *Meleagris* mais próximas que qualquer uma delas de *Taeniopygia*.

Oryctolagus ficou próxima de *Cavia*, como era de se esperar, nas árvores dos métodos de Máxima Verossimilhança, NJ e Máxima Parcimônia, ficando afastada, no entanto, na de UPGMA.

Quanto aos primatas, o par de espécies *Ailuropoda* e *Pteropus* impediu que se formasse um grupo monofilético nas árvores formadas por Máxima Verossimilhança e NJ, o afastamento de *Gorilla* fez o mesmo na de UPGMA, mas os primatas ficaram todos reunidos na árvore de Máxima Parcimônia, nas quais formaram grupo monofilético, com os símios formando também.

Os marsupiais formaram, como havia ocorrido no método de Máxima Verossimilhança, um grupo monofilético, e isto se deu tanto nos métodos de distância, quanto no de Máxima Parcimônia.

8.4.3 ÁRVORES DE AMINOÁCIDOS DE 53 SEQUÊNCIAS

Com as 53 sequências de aminoácidos disponíveis, de 48 diferentes espécies, foi construída uma árvore filogenética, utilizando o Método de Máxima Verossimilhança, e muitas características que já se esperavam encontrar, foram observadas.

Os peixes todos constituíram um único conjunto monofilético, com exceção de *Petromyzon*, que ficou afastada de todas as espécies, e de *Latimeria*, que, afastando-se dos peixes, ficou próxima de todos os tetrápodos. Os demais peixes ficaram reunidos, formando dois subconjuntos simétricos, constituídos pelos genes RAPH1 que a maior parte dos peixes estudados apresenta em duplicidade.

As aves encontraram-se juntas, formando um conjunto monofilético, com *Gallus* e *Meleagris* mais próximos entre si que de *Taeniopygia*, porém com *Xenopus* e *Anolis* mais próximas delas do que era de se esperar, uma vez que não há motivo para se esperar mais proximidade entre anfíbios e répteis com aves do que com qualquer outro grupo de tetrápodos.

Os três marsupiais formaram também um grupo monofilético, como também formaram os lagomorfos, que ainda mantiveram proximidade com *Cavia* e *Mus*, como era de se esperar, embora *Dipodomys* não tenha se posicionado proximamente.

As duas espécies de morcego utilizadas posicionaram-se separadamente uma de outra e o mesmo ocorreu entre *Sus* e *Bos*, e entre estes e *Tursiops*.

Quanto aos carnívoros, ficaram próximos entre si *Felis* e *Canis*, mas estes se distanciaram de *Ailuropoda*, que não manteve proximidade com nenhum grupo em particular.

Tupaia ficou próxima de *Bos* e não de *Sorex*, que por sua vez ficou próxima de *Equus*, carnívoros e primatas.

Quatro dos primatas formaram também um grupo monofilético, mas se posicionaram como exceções *Tarsius*, que ficou junto a roedores e lagomorfos, *Nomascus*, que ficou próximo a *Erinaceus*, *Microcebus*, que ficou junto a alguns carnívoros, *Otolemur* que se aproximou de *Procavia* e *Gorilla*, que ficou junto a *Equus*. Mas, dentro dos primatas, os símios, excetuando *Gorilla*, formaram, estes sim, um grupo monofilético, com *Homo* e *Pan* bem próximos e *Pongo* junto a ambos.

Observaram-se algumas diferenças, nas construções das árvores com métodos de distância, NJ e UPGMA, em relação à árvore de Máxima Verossimilhança.

Os peixes ficaram, naquelas árvores, divididos em dois grupos bastante simétricos, porém ficando um destes dois grupos mais próximo do grupo formado pelos tetrápodos do que do outro grupo de peixes. Na árvore formada por Máxima Parcimônia, ocorreu o mesmo, porém com a exceção de *Latimeria* e *Petromyzon*. Esta espécie ficou afastada de todas as espécies de peixes e tetrápodos e, também, *Latimeria* ficou afastada de peixes, mas mais proximamente posicionada dos tetrápodos, o que se justifica em função da proximidade filogenética que se sabe existir entre o grupo que originou os tetrápodos e aquele que originou os atuais *Latimeria*.

A simetria observada nestes casos ocorreu em função de alguns peixes apresentarem dois genes RAPH1, e a maior proximidade de um destes grupos de peixe com os tetrápodos se dá em função de uma das formas parálogas do gene ser mais proximamente aparentada com o gene presente nos tetrápodos.

Na árvore construída por NJ, *Gallus* e *Meleagris* ficaram juntas e *Taeniopygia* afastada delas e mais próxima de *Xenopus*, mas nas construídas por UPGMA e por Máxima Parcimônia elas ficaram todas juntas.

Nesta mesma árvore construída por NJ, as espécies de carnívoros, *Felis*, *Canis* e *Ailuropoda* ficaram separadas, mas na construída por UPGMA, as duas primeiras ficaram juntas e a última é que ficou distante delas. Já na construída por Máxima Parcimônia, os carnívoros formam um grupo monofilético, mas com *Equus* junto a eles.

Na árvore construída pelo método de Máxima Verossimilhança, os marsupiais formaram um grupo monofilético, e isto se deu também nas árvores formadas por NJ, UPGMA e por Máxima Parcimônia.

Os morcegos, na árvore construída pelo método de Máxima Verossimilhança, ficaram próximos e não formaram grupo monofilético, mas já, na árvore formada por

Máxima Parcimônia, formaram, embora tenham ficado afastados nas árvores de NJ e UPGMA.

Os Primatas teriam também constituído, na árvore formada por Máxima Parcimônia, um grupo monofilético, se não fosse o caso de *Tarsius* e *Microcebus* ficarem afastados do restante do grupo. O mesmo que foi observado aqui, entre os primatas, com as árvores de aminoácidos, foi observado quando foi construída a árvore com o cDNA das mesmas espécies. De maneira idêntica, os símios também, entre os primatas, formaram um grupo monofilético.

8.4.4 ÁRVORES DE AMINOÁCIDOS DE 39 SEQUÊNCIAS

Foi construída uma árvore filogenética, utilizando o Método de Máxima Verossimilhança, com as 39 melhores sequências de aminoácidos, no que diz respeito à sua completude, dentre todas as sequências, de diferentes espécies, disponíveis.

Tal como foi observado na árvore com 53 sequências, nesta árvore os peixes todos constituíram um único conjunto monofilético, com exceção de *Petromyzon*, que ficou afastada de todas as espécies, e de *Latimeria*, que, afastando-se dos peixes, ficou próxima de todos os tetrápodos. Nesta árvore também, os demais peixes ficaram reunidos, formando dois subconjuntos simétricos, constituídos cada um deles por um ou outro dos genes RAPH1 que a maior parte dos peixes estudados apresenta em duplicidade.

As aves encontraram-se juntas, formando um conjunto monofilético, com *Gallus* e *Meleagris* mais próximos entre si do que cada um deles de *Taeniopygia*.

Xenopus posicionou-se mais proximamente das aves do que era de se esperar, uma vez que não há motivo para se esperar mais proximidade de anfíbios com elas do

que com qualquer outro grupo de tetrápodos. Algo muito próximo disto foi observado na árvore de 53 sequências.

Formaram grupo monofilético, também, os marsupiais, e os primatas todos formariam também, se não fosse o fato de *Microcebus* ter ficado distante dos demais e próxima dos carnívoros, que, por sua vez, ficaram também juntos. Dentre os primatas, os símios também ficaram reunidos e formando dois subgrupos, estando em um deles *Pongo* e *Nomascus* e em outro *Homo*, *Pan* e *Gorilla*, o que já se tinha como esperado.

Os roedores, além de formarem grupo monofilético, ficaram junto a *Oryctolagus*, o que era de se esperar pela grande proximidade filogenética dos dois grupos que se sabe existir.

Os morcegos ficaram distantes entre si, tendo um deles ficado relativamente próximo aos primatas e o outro bem junto a roedores e lagomorfos.

Os carnívoros ficaram próximos entre si e os artiodáctilos, também, ficaram juntos e relativamente próximos aos primatas.

Nas árvores construídas com métodos de distância, NJ e UPGMA, observaram-se algumas diferenças em relação à árvore de Máxima Verossimilhança.

Os peixes ficaram, na árvore de UPGMA, divididos em dois grupos bastante simétricos, porém com um destes dois grupos mais próximo do grupo formado pelos tetrápodos do que do outro grupo de peixes. Nas árvores formadas por Máxima Parcimônia, NJ e Máxima Verossimilhança ocorreu também a subdivisão em dois grupos simétricos, porém sem que houvesse maior proximidade de um dos dois grupos com o grupo dos demais tetrápodos.

Como ocorreu nas árvores de 53 sequências, a árvore construída por NJ, apresentou *Gallus* e *Meleagris* juntas e *Taeniopygia* afastada delas e mais próxima de

Xenopus, mas nas construídas por Máxima Verossimilhança, UPGMA e por Máxima Parcimônia, elas ficaram todas juntas.

Os carnívoros ficaram sempre juntos, nas árvores construídas por todos os métodos.

Os marsupiais só não formaram um grupo monofilético na árvore formada por UPGMA, ficando juntos nas demais.

Os Primatas teriam constituído, na árvore formada por Máxima Verossimilhança, um grupo monofilético, se não fosse o caso de *Microcebus* ficar afastada do restante do grupo. Na árvore formada por Máxima Parcimônia, ocorreu algo similar, tendo ficado afastadas *Microcebus* e *Otolemur*, que, embora juntas, aproximaram-se mais dos artiodáctilos e carnívoros que dos demais primatas.

8.4.5 DOMÍNIOS

O gene RAPH1 foi tomado não só por inteiro, mas também por seus trechos correspondentes aos domínios das proteínas determinadas pelo gene e aos sítios de ligação de proteínas presentes nele. Portanto, cada uma das 53 sequências de aminoácidos foi segmentada em seus domínios proteicos, que no caso de RAPH1 são os domínios GM, RA, PH e EVH1, e nos sítios de ligação de proteínas, que no caso são os sítios de ligação da Profilina e os sítios de ligação da SH3. Foram construídas árvores filogenéticas com os segmentos de aminoácidos correspondentes a cada um destes domínios e sítios de ligação, com todas as sequências disponíveis.

Métodos de construção baseados na distância foram utilizados, bem como um baseado no princípio de Máxima Parcimônia e outro, no princípio da Máxima Verossimilhança.

Os sítios de ligação da SH3 são 8 sítios e a montagem de árvores filogenéticas com eles não se mostrou de nenhuma utilidade. A informação gerada foi desconstruída e mostrava óbvia falta de relação com a história evolutiva dos grupos estudados. Isto foi devido provavelmente às pequenas dimensões dos sítios, que têm comprimento variando de 18 a 21 nucleotídios.

Os sítios de ligação da proteína Profilina apresentaram um padrão de semelhanças e diferenças próximo ao do gene como um todo, com exceção de *Mus* e *Oryctolagus*, roedor e lagomorfo, e de *Monodelphis* e *Meleagris*, marsupial e ave, que se distanciaram bastante dos demais.

Os sítios RA, PH e GM mostram proximidade entre os peixes, e também entre as aves. Mostram separação de *Caenorhabditis* dos demais e também de *Petromyzon* em GM e PH.

Entre primatas esta proximidade não é tão pronunciada, com exceção dos símios que têm uma proximidade evidente. Os peixes apresentam-se próximos e com nítida separação das demais espécies, mas mostram um grau de diversidade maior do que a encontrada entre os tetrápodos.

8.5 CONCLUSÕES

Observando os resultados mostrados nas árvores filogenéticas construídas, podemos chegar a algumas conclusões. As árvores foram construídas tanto com todas as 53 sequências de nucleotídios, quanto com as 53 sequências de aminoácidos. Foram construídas também com as 39 melhores sequências de nucleotídios, como também com as 39 melhores sequências de aminoácidos. E, em cada um destes casos, as árvores foram construídas utilizando dois métodos de distância, NJ e UPGMA, um método de Máxima Parcimônia e um Método de Máxima Verossimilhança.

Embora tenham ocorrido muitas diferenças entre os resultados, o que já se sabia que ocorreria, em virtude das diferentes metodologias, que se baseiam em diferentes premissas, algumas características deles foram recorrentes. As aves permaneceram juntas muito proximamente. O mesmo ocorreu com os marsupiais e com os roedores. Os lagomorfos ficaram próximos aos roedores na maior parte dos casos. Caso oposto foram os morcegos, que raramente permaneceram juntos, aproximando-se ora de um grupo, ora de outro. Os artiodáctilos também não permaneceram juntos muito frequentemente. Os primatas apresentaram algum grau de coesão, mas dentre eles os símios ficaram bastante próximos.

Xenopus variou bastante de posição, raramente ficando entre os peixes e os demais cordados, que é onde se esperava que ficasse. Algo similar ocorreu com *Anolis*, que poucas vezes situou-se entre os peixes e os demais cordados, ou entre *Xenopus* e os demais cordados.

Caenorhabditis ficou afastada de todas as espécies, *Petromyzon* junto aos peixes, mas mais afastada de cada um deles do que eles entre si.

Latimeria algumas vezes ficou entre os peixes, mas sendo a mais distante deles, com exceção de *Petromyzon*, e outras vezes ficou entre os peixes e os tetrápodos. Mas em ambos os casos ficou em uma situação próxima da situação na qual se esperava encontrá-la em virtude da proximidade filogenética que se sabe existir entre o grupo que originou os tetrápodos e aquele que originou os atuais *Latimeria*.

Os peixes permaneceram juntos quase sempre, formando grupo monofilético, mas apresentando uma subdivisão com bastante simetria. Esta simetria se deve ao fato de vários peixes apresentarem o gene RAPH1 em duplicidade, e em decorrência disto, os peixes colocarem-se em posições semelhantes, em um e outro grupo. Cada um dos genes parálogos forma um dos subgrupos, que assim faz que haja, para cada espécie de

gene duplo, um gene em cada subgrupo. Em alguns casos um destes subgrupos aproximou-se mais dos tetrápodos do que do outro subgrupo, o que era de se esperar, uma vez que um dos genes parálogos é o gene mais proximamente aparentado com o gene presente nos tetrápodos.

Pode-se dizer, então, em termos gerais, mas com algumas exceções, que o gene RAPH1, nas espécies estudadas, tem sua história evolutiva muito próxima da história evolutiva de todo o restante do genoma. E infere-se que, com os diferentes graus de conservação que as diferentes porções do genoma apresentam, o gene RAPH1 manteve sua estrutura básica e acumulou diferenças de forma similar ao restante do genoma.

No que diz respeito ao gene em si, pode-se observar, da Figura 31 até a Figura 58, que os domínios proteicos GM, RA, PH e EVH1 apresentam um grau de alteração de sua estrutura muito inferior ao observado no gene inteiro, para o conjunto de espécies analisadas como um todo. Este fato, naturalmente, já se esperava, uma vez que os domínios são os principais responsáveis pelo desempenho fisiológico das proteínas, o que leva a um maior grau de preservação de suas estruturas ao longo da história evolutiva do gene.

8.6 REFERÊNCIAS

- Alcântara, V. M. (2000). Fenótipos da butirilcolinesterase e suas relações com dados antropométricos, bioquímico-hormonais e pressão arterial em obesos e na população geral de Curitiba, PR. Curitiba: Universidade Federal do Paraná (Doutorado em Ciências Biológicas - Genética - Setor de Ciências Biológicas).
- Amorin, D. (1997). Elementos Básicos de Sistemática Filogenética (2ª ed.). Ribeirão Preto: Holos Editora e Sociedade Brasileira de Entomologia.
- BioEdit. (s.d.). Fonte: <http://www.mbio.ncsu.edu/BioEdit/BioEdit>
- Blast N. (s.d.). Fonte: <http://www.ncbi.nlm.nih.gov/blast/blast.cgi>
- Clustalw. (s.d.). Fonte: <ftp://ftp.ebi.ac.uk/pub/software/clustalw2>
- Diniz Filho, J. F. (2000). Métodos Filogenéticos Comparativos. Ribeirão Preto: Holos Editora.
- DNAsp. (s.d.). Fonte: <http://www.ub.edu/DNAsp>
- Ensembl. (s.d.). Fonte: ensembl.org
- GeneBank. (s.d.). Fonte: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>
- Graur, D., & Li, W.-H. (2000). Fundamentals of Molecular Evolution (2ª ed.). Sunderland, Massachusetts, United States: Sinauer Associates Incorporation.
- Lewin, B. (2008). Genes IX. Sudbury: Jones and Bartlett.

Mega5. (s.d.). Fonte: Mega Software: <http://www.megasoftware.net/mega.html>

Mount, D. W. (2004). Bioinformatics: sequence and genome analysis (2ª ed.). New York: Cold Spring Harbor Laboratory Press.

Mulan. (s.d.). Fonte: mulan.dcode.org

Myers, M., Richmond, R. C., & Oakeshott, J. G. (1988). On the Origins of Esterases. *Molecular Biology and Evolution*, 5/2, 113-119.

NCBI. (s.d.). Fonte: <http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>

Nei, M., & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.

Schneider, H. (2003). *Métodos de Análise Filogenética* (2ª ed.). Ribeirão Preto: Holos Editora.

Souza, R. L. (2001). *Variação Fenotípica e Genotípica Da Butirilcolinesterase Humana e Sua Relação Com Altura e Índice De Massa Corporal*. Curitiba: Universidade Federal do Paraná (Setor de Ciências Biológicas - Doutorado em Ciências – Genética).

Tree of Life. (s.d.). Fonte: tolweb.org

9 CAPÍTULO 2 - ANÁLISE EVOLUTIVA DO GENE RAPH1 EM ALGUMAS ESPÉCIES DE PEIXE

9.1 RESUMO

Em algumas espécies de peixe o gene *RAPH1* (*Ras association (RalGDS/AF-6) and pleckstrin homology domains 1*) apresenta-se duplicado, com duas formas semelhantes, porém não idênticas. Esta análise, que incluiu os genes de cinco espécies, deixou evidente que os dez genes reúnem-se, por semelhança, nitidamente em dois grupos, cada um dos quais incluindo cinco genes, um de cada uma das cinco espécies, o que sugere que a duplicação do gene tenha ocorrido em época anterior aos momentos de especiação que separaram as cinco espécies. A análise evidencia, também, que genes de espécies que apresentam cópia única, entre alguns peixes, anfíbio, réptil, aves e mamíferos, são, todos eles, mais estreitamente aparentados com apenas um destes dois grupos. Isto deixou evidente que a duplicação do gene *RAPH1* ocorreu em um ancestral comum a todas as espécies analisadas e não em um ancestral comum às espécies que apresentam dois genes *RAPH1*.

9.2 INTRODUÇÃO

A sequência do gene *RAPH1* (*Ras association (RalGDS/AF-6) and pleckstrin homology domains 1*) foi identificada por (Nagase, Kikuno, Hattori, Kondo, Okumura, & Ohara, 2000) em um projeto desenvolvido para sequenciar cDNA humano de genes não identificados, obtidos de amostras do cérebro e hipocampo. Os autores obtiveram o sequenciamento completo de 100 cDNAs, e dentre estas sequências estava a do gene *RAPH1*, cuja localização foi determinada no cromossomo 2. Além disso, os autores

observaram que o produto do gene *RAPH1*, nomeado como KIAA1681, apresentava 1236 aminoácidos e 33% de identidade com a sequência da proteína Grb10 de camundongo. (Eppert, Wunder, Aneliunas, Tsui, Scherer, & Andrulis, 2005) refinaram a localização do gene *RAPH1* como 2q33 e também observaram homologia com as proteínas humanas Grb7, Grb10 e Grb14 (29% a 35% de identidade), levando os autores a sugerir um possível papel desta proteína na migração e sinalização celulares.

Genes *RAPH1* existem em cordados, tanto em peixes, anfíbios e répteis quanto em mamíferos e aves. Na grande maioria das espécies, este gene é único no genoma. No entanto, em algumas delas, apenas em algumas das espécies de peixe, existem dois genes *RAPH1*, não idênticos, mas semelhantes.

Para tentar elucidar os acontecimentos evolutivos envolvendo os genes *RAPH1* verificamos se estes genes existentes em dose dupla nestas espécies assemelham-se mais aos genes das outras espécies que ao outro gene *RAPH1* da mesma espécie.

9.3 MATERIAL E MÉTODOS

Foram utilizadas dez sequências de aminoácidos, obtidas do *Protein Database*, NCBI dos dez polipeptídios produzidos a partir dos dez genes *RAPH1* das cinco espécies de peixe que apresentam genes em duplicidade (*Danio rerio*, *Gadus morhua*, *Gasterosteus aculeatus*, *Takifugu rubripes* e *Tetraodon nigroviridis*) e submetidas a um alinhamento com o programa ClustalW, utilizando os seguintes parâmetros: penalidade para abertura de lacuna de 18 e penalidade para extensão de lacuna de 0,3. (NCBI) (Clustalw) (BioEdit)

Tabela 1 - Sequências de aminoácidos de genes parálogos

Danio rerio gene1	ENSDART00000061156 modif
Danio rerio gene2	ENSDART00000088290
Gadus morhua gene1	ENSGMOT00000007333
Gadus morhua gene2	ENSGMOT0000000606 modif

Gasterosteus aculeatus gene1	ENSGACT00000011348
Gasterosteus aculeatus gene2	ENSGACT00000020022
Takifugu rubripes gene1	ENSTRUT00000020869
Takifugu rubripes gene2	ENSTRUT00000040469
Tetraodon nigroviridis gene1	ENSTNIT00000017769
Tetraodon nigroviridis gene2	ENSTNIT00000014686

O alinhamento foi inspecionado e corrigido, quando necessário, manualmente.

Uma vez alinhadas, estas sequências serviram de base para a montagem de uma árvore filogenética que nos permitiu visualizar as semelhanças e diferenças entre genes e espécies. O método de montagem das árvores foi o de Máxima Verossimilhança utilizando o programa MEGA5.05, disponível em <<http://www.megasoftware.net/mega.html>>. (Mega4)

9.4 RESULTADOS E DISCUSSÃO

Uma árvore obtida está mostrada na Figura 66, e nela pode-se observar bem que as sequências reúnem-se nitidamente em dois grupos bem diferenciados, havendo uma sequência de aminoácidos de cada espécie em cada um dos dois grupos. O mesmo aparece mais nitidamente na Figura 67.

Um dos grupos (grupo A) é constituído pelos genes *Danio rerio* gene 1, *Gadus morhua* gene 2, *Gasterosteus aculeatus* gene 2, *Takifugu rubripes* gene 1 e *Tetraodon nigroviridis* gene 1, enquanto o outro grupo (grupo B) é formado por *Danio rerio* gene 2, *Gadus morhua* gene 1, *Gasterosteus aculeatus* gene 1, *Takifugu rubripes* gene 2 e *Tetraodon nigroviridis* gene 2.

A árvore mostra ainda perfeita simetria, com igual relacionamento filogenético entre as cinco espécies em ambos os grupos.

Mostra maior proximidade entre as espécies *Takifugu rubripes* e *Tetraodon nigroviridis*, o que era de se esperar, uma vez que ambas são, nas classificações tradicionais, posicionadas na mesma família Tetraodontidae, da ordem Tetraodontiformes, da superordem Acanthopterygii. Próximas a estas duas espécies, está *Gasterosteus aculeatus*, que pertence à família Gasterosteidae, da ordem Gasterosteiformes, da mesma superordem Acanthopterygii. Pouco mais afastada, ficou *Gadus morhua*, que pertence à família Gadidae, da ordem Gadiformes, de outra superordem, Paracanthopterygii. A espécie mais distante de todas as demais é *Danio rerio*, o que é compatível com as classificações tradicionais que colocam sua família Cyprinidae da ordem Cypriniformes em outra superordem, Ostariophysii, mais distante de todas as demais. (Helfman, Collette, Facey, & Bowen, 2009) (Lagler, Bardach, Miller, & Passino, 1977) (Amorin, 1997) (Moyle & Chech Jr, 2003)

Como pode ser visto na Figura 9, todas estas cinco espécies que apresentam genes *RAPH1* duplicados pertencem aos peixes ósseos, superclasse Osteichthyes, com nadadeiras com raios, classe Actinopterygii, infraclasse Teleostei.

Domínio Eucaria	
Reino Animalia	
Filo Chordata	
Infracilo Cyclostomata (peixes s mandíbula)	
Classe Myxini	peixe-bruxa
Classe Hyperoartia	lampreia
Infracilo Gnathostomata (peixes mandibulados)	
Classe Chondrichthyes (peixes cartilagosos)	
subclasse : Elasmobranchii	tubarão raia
subclasse Holocephali	quimera
Superclasse Osteichthyes (peixes ósseos)	
Classe Sarcopterygii (nadadeiras carnosas)	
Subclasse Actinistia	celacanto
Subclasse Dipnoi	peixes pulmonados piramboia
Classe Actinopterygii, (nadadeiras c raios)	
Infraclasse Holostei	Amia calva, Lepisosteus
Infraclasse Chondrostei	esturção
Infraclasse Teleostei	
Superordem Paracanthopterygii	
Ordem Gadiformes	
Família Gadidae	bacalhau
Superordem Acanthopterygii	
Ordem Beloniformes	
Família Adrianichthyidae	Oryzias
Ordem Tetraodontiformes	
Família Tetraodontidae	baiacu
Ordem Gasterosteiformes	
Família Gasterosteidae	esgana gato
Ordem Cyprinodontiformes	
Ordem Perciformes	
Superorder: Ostariophysii	
Ordem Cypriniformes	
Família Cyprinidae	paulistinha

Figura 9 – Taxonomia dos Peixes

Em todas as outras espécies analisadas, de aves, mamíferos, anfíbio e réptil, só há um gene *RAPH1*. Nas espécies de peixe *Latimeria chalumnae*, *Oryzias latipes* e *Petromyzon marinus*, parece haver também um único gene.

Oryzias latipes é uma espécie próxima das cinco que apresentam duplicação do gene, pois, sendo da família Adrianichthyidae, ordem Beloniformes, pertence também à superordem Acanthopterygii, sendo também da classe Actinopterygii. *Latimeria chalumnae* já é uma espécie mais distante, uma vez que não é uma espécie de peixe com nadadeiras com raios, mas com nadadeiras carnosas, classe Sarcopterygii da superclasse

dos peixes ósseos, Osteichthyes. *Petromyzon marinus* é ainda mais distante delas, uma vez que é uma espécie de peixe que pertence à classe Hyperoartia, e, portanto, é desprovida de mandíbula, como todas as pertencentes ao infrafilo Cyclostomata.

Esta forma de agrupamento sugere que houve duplicação do gene *RAPH1* bastante anterior às épocas em que as cinco espécies separaram-se umas das outras. Esta interpretação está de acordo com o fato de todas as dez sequências diferirem entre si, mas o grau de diferenciação ser sempre maior entre cada um dos genes e o outro gene da mesma espécie do que entre cada um dos genes e pelo menos um dos genes de qualquer uma das quatro outras espécies. Ou seja, esta interpretação é condizente com a existência de dois grupos de genes (grupos A e B), cada um deles com um gene de cada uma das cinco espécies. A composição e a disposição dos grupos permite dizer que houve duplicação gênica em um ancestral comum das cinco espécies, seguida de acúmulo, ao longo do tempo, de alterações em cada uma das duas cópias gênicas, o que fez que deixassem de ser idênticas e passassem a ser progressivamente diferentes. E, como mostram as diferenças existentes entre os cinco membros de cada grupo, após os fenômenos de especiação que separaram as cinco espécies, outras diferenças continuaram a se acumular, porém agora independentemente em cada um dos dez genes.

Tendo estas espécies duas cópias do gene *RAPH1* e as espécies de aves, mamíferos, réptil e anfíbio apenas uma, fica-se com dúvida a respeito do relacionamento filogenético que o gene destas últimas espécies teria com ambos os genes das espécies que o tem duplicados. Seria possível que um ou outro, de cada um dos dois genes das espécies *Danio rerio*, *Gadus morhua*, *Gasterosteus aculeatus*, *Takifugu rubripes* e *Tetraodon nigroviridis*, fosse o gene mais proximamente aparentado com os genes de cada uma das demais espécies. Ou seja, para cada espécie

seria interessante saber a qual dos dois grupos de genes duplos seu único gene mais se aproxima.

Para tirar esta dúvida, foram feitas filogenias, por método de Máxima Verossimilhança, com um de cada um dos genes cópia-única com as dez cópias de genes duplos das cinco espécies que os apresentam.

Como mostra a filogenia apresentada na Figura 59, o gene do réptil *Anolis carolinensis* é mais próximo filogeneticamente do primeiro dos grupos descritos acima. A Figura 60 mostra que este grupo é mais próximo também de gene do anfíbio *Xenopus laevis*. O primeiro dos grupos é o mais próximo também do mamífero marsupial *Monodelphis domestica*, Figura 63, e da ave *Taeniopygia guttata*, Figura 64.

Uma filogenia montada com todos os peixes disponíveis, Figura 66, mostra que *Latimeria chalumnae*, *Oryzias latipes* e *Petromyzon marinus* têm, todos os três, seus genes *RAPH1* mais proximamente relacionados com os genes que compõem o grupo 1.

A árvore filogenética mostrada na Figura 65, construída através de método de Máxima Verossimilhança, evidencia que o gene *RAPH1* humano tem mais proximidade filogenética também com o primeiro dos grupos acima mencionados.

Verificadas independentemente as proximidades de cada um dos genes das espécies de gene único com ambos os grupos formados pelos genes das espécies que os apresentam em duplicidade, fica demonstrado que todos os genes das espécies de gene único são mais estreitamente aparentados com um só dos grupos, que é o primeiro dos descritos acima.

Como as espécies que apresentam dois genes *RAPH1* são apenas peixes, sendo que réptil, anfíbio, mamíferos e ave, bem como alguns peixes, apresentam cópia única

do gene, seria possível optar por uma das duas seguintes explicações sobre as ocorrências evolutivas na história do gene.

Seria possível achar que tivesse ocorrido duplicação do gene no ancestral comum dos peixes que apresentam gene duplo e não tivesse ocorrido duplicação nos ancestrais das demais espécies (Hipótese A). Ou, alternativamente, seria possível achar que houvesse ocorrido duplicação no ancestral comum de todas as espécies analisadas e que tivesse ocorrido perda de uma das cópias do gene no ancestral comum das espécies que apresentam cópia única (Hipótese B).

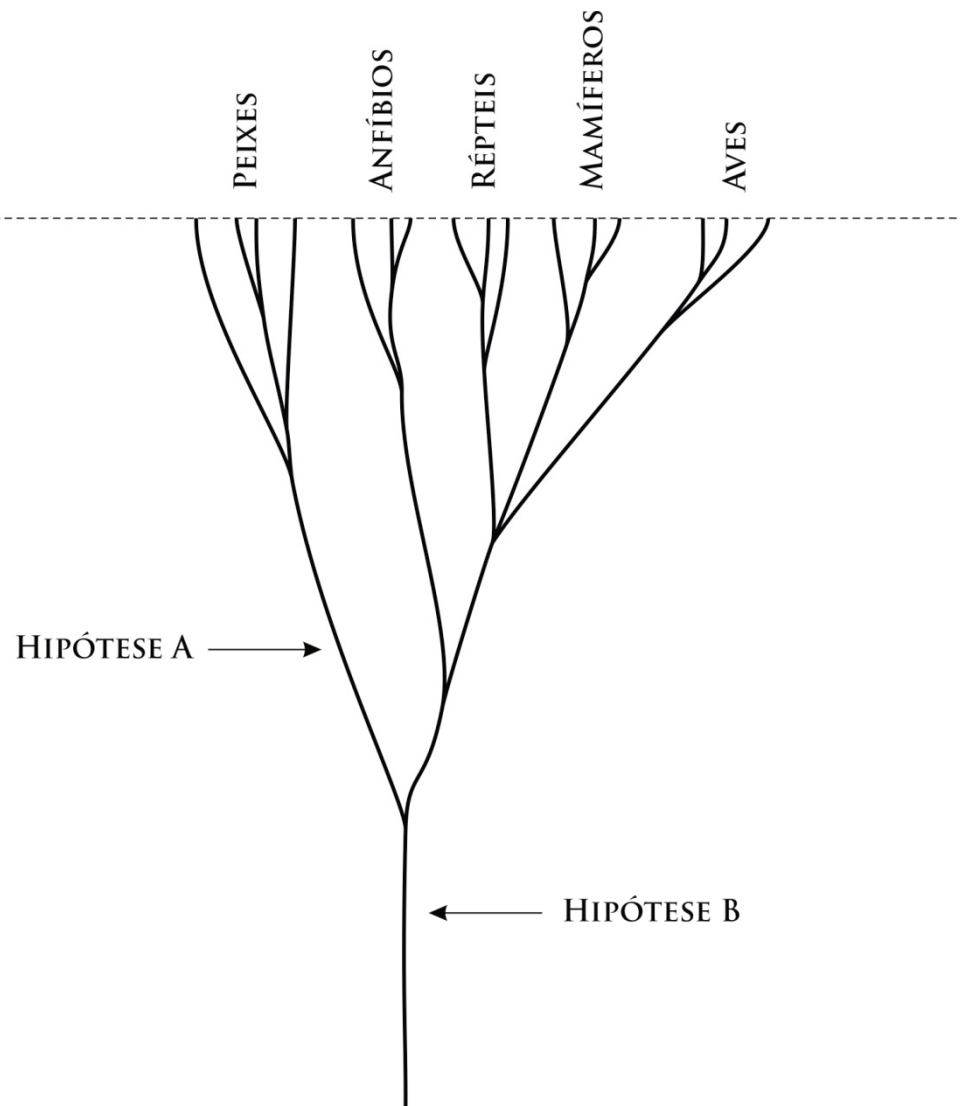


Figura 10 – Hipóteses A e B

No entanto, como pode ser observado nas filogenias da Figura 59, Figura 60, Figura 61, Figura 62, Figura 63, Figura 64 e Figura 65, as sequências de réptil, anfíbio, mamíferos e ave são sequências cuja maior proximidade filogenética ocorre com os genes todos de apenas um dos grupos de genes (grupo A) das espécies que apresentam dois genes RAPH1. Ou seja, os genes das espécies de cópia-única são mais próximos filogeneticamente de apenas um dos grupos de genes das espécies de gene duplo (grupo A).

9.5 CONCLUSÕES

Estes achados indicam fortemente que houve duplicação do gene RAPH1 no ancestral comum de todas as espécies analisadas e que, bastante posteriormente, houve perda de uma das cópias do gene em um, ou mais de um, ancestral comum das espécies que apresentam uma só cópia. A explicação alternativa exigiria que as espécies de cópia única tivessem seus genes proximamente aparentados indistintamente com ambas as cópias (grupos A e B) do gene das espécies de dois genes. Ou seja, dizendo de outra maneira, se a duplicação do gene tivesse ocorrido no ancestral comum das espécies de gene duplo, não haveria motivo para que todos os genes das espécies de cópia única do gene fossem mais próximos filogeneticamente de apenas um dos grupos de gene das espécies de gene duplo.

Portanto, sabe-se, com esta observação, que o gene RAPH1 teve sua duplicação em uma época em que vivia a espécie que originaria grupos de hoje tão diversificados quanto peixes e aves. E é possível concluir ainda que o gene RAPH1 duplicou há mais de 360 milhões de anos, que é o ponto da história da vida, no fim do período

Devoniano, a partir do qual se separaram um ancestral de alguns dos atuais peixes de outro ancestral dos demais atuais Chordatas.

Fica óbvio ainda, que esta duplicação gênica não se deu no exato momento em que se separaram estes ancestrais mencionados acima, uma vez que foi necessário que decorresse tempo suficiente para que as cópias originadas da duplicação acumulassem, independentemente, as diferenças que nos permitem hoje caracterizá-las. Sabemos, por isto, que a duplicação é bastante anterior ao fim do período Devoniano.

9.6 REFERÊNCIAS

- Amorin, D. (1997). Elementos Básicos de Sistemática Filogenética (2ª ed.). Ribeirão Preto: Holos Editora e Sociedade Brasileira de Entomologia.
- BioEdit. (s.d.). Fonte: <http://www.mbio.ncsu.edu/BioEdit/BioEdit>
- Blast N. (s.d.). Fonte: <http://www.ncbi.nlm.nih.gov/blast/blast.cgi>
- Clustalw. (s.d.). Fonte: <ftp://ftp.ebi.ac.uk/pub/software/clustalw2>
- Diniz Filho, J. F. (2000). Métodos Filogenéticos Comparativos. Ribeirão Preto: Holos Editora.
- DNAsp. (s.d.). Fonte: <http://www.ub.edu/DNAsp>
- Ensembl. (s.d.). Fonte: ensembl.org
- Eppert, K., Wunder, J. S., Aneliunas, V., Tsui, L. C., Scherer, S. W., & Andrulis, I. L. (2005). Altered Expression and Deletion of RMO1 in Osteosarcoma. *International Journal of Cancer*, 114/5, 738/746.
- GeneBank. (s.d.). Fonte: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>
- Graur, D., & Li, W.-H. (2000). *Fundamentals of Molecular Evolution* (2ª ed.). Sunderland, Massachusetts, United States: Sinauer Associates Incorporation.
- Helfman, G., Collette, B. B., Facey, D. E., & Bowen, B. W. (2009). *Diversity of Fishes, the: Biology, Evolution, and Ecology* (2ª ed.). Hoboken: Wiley-Blackwell.
- Lagler, K. F., Bardach, J. E., Miller, R. R., & Passino, D. R. (1977). *Ichthyology*. New York: John Willey.
- Lewin, B. (2008). *Genes IX*. Sudbury: Jones and Bartlett.
- Li, F., & Han, Z.-J. (2002). Two different genes encoding acetylcholinesterase existing in cotton aphid (*Aphis gossypii*). *Genome*, 45/6, 1134-1141.
- Mega5. (s.d.). Fonte: Mega Software: <http://www.megasoftware.net/mega.html>
- Mount, D. W. (2004). *Bioinformatics: sequence and genome analysis* (2ª ed.). New York: Cold Spring Harbor Laboratory Press.
- Moyle, P. B., & Chech Jr, J. J. (2003). *Fishes: An Introduction to Ichthyology* (5ª ed.). (P. Hall, Ed.)
- Mulan. (s.d.). Fonte: mulan.dcode.org
- Nagase, T., Kikuno, R., Hattori, A., Kondo, Y., Okumura, K., & Ohara, O. (2000). Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res*, 7/4, 347-355.
- NCBI. (s.d.). Fonte: <http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>
- Nei, M., & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.
- Schneider, H. (2003). *Métodos de Análise Filogenética* (2ª ed.). Ribeirão Preto: Holos Editora.
- Souza, R. L. (2001). *Variação Fenotípica e Genotípica Da Butirilcolinesterase Humana e Sua Relação Com Altura e Índice De Massa Corporal*. Curitiba: Universidade Federal do Paraná (Setor de Ciências Biológicas - Doutorado em Ciências – Genética).
- Tree of Life. (s.d.). Fonte: tolweb.org

10 DISCUSSÃO

A principal finalidade da execução deste trabalho foi tentar aclarar os acontecimentos evolutivos que envolveram os genes determinantes da lamelipodina.

Isto foi feito, principalmente, através das comparações realizadas entre os genes RAPH1 de 48 espécies de animais, em árvores filogenéticas construídas com suas sequências de aminoácidos e de nucleotídios. Com o mesmo propósito foram comparadas, também em árvores filogenéticas, sequências de aminoácidos de genes parálogos de 5 espécies de peixes que os apresentam.

As árvores foram construídas tanto com todas as 53 sequências de nucleotídeos correspondentes à reunião de todos os éxons dos genes, quanto com as 53 sequências de aminoácidos correspondentes a eles.

Foram elas construídas também com as 39 melhores sequências de nucleotídeos, escolhidas dentre as 53 disponíveis, como também com as 39 melhores sequências de aminoácidos e, em cada um destes casos, as árvores foram construídas utilizando dois métodos de distância, NJ e UPGMA, um método de Máxima Parcimônia e um Método de Máxima Verossimilhança.

Algumas conclusões puderam ser obtidas ao observar os resultados mostrados nas árvores filogenéticas construídas.

Os resultados diferiram quando comparadas as árvores filogenéticas entre si, o que já se sabia que ocorreria, em virtude das diferentes metodologias, que se baseiam em diferentes premissas, mas algumas características destes resultados foram recorrentes.

Os primatas apresentaram algum grau de coesão, mas dentre eles os símios ficaram bastante próximos. As aves permaneceram juntas muito proximamente. O mesmo ocorreu com os marsupiais e com os roedores. Os lagomorfos ficaram próximos aos roedores na maior parte dos casos.

Caso oposto foram os morcegos, que raramente permaneceram juntos, aproximando-se ora de um grupo, ora de outro. Os artiodáctilos também não permaneceram juntos muito frequentemente. O anfíbio *Xenopus* variou bastante de posição, raramente ficando entre os peixes e os demais cordados, que é onde se esperava que ficasse. Algo similar ocorreu com o réptil *Anolis*, que poucas vezes situou-se entre os peixes e os demais cordados, ou entre *Xenopus* e os demais cordados.

O nematelminto *Caenorhabditis* ficou afastado de todas as espécies, a lampreia, *Petromyzon*, junto aos peixes, mas mais afastada de cada um deles do que eles entre si.

O peixe de nadadeiras lobadas *Latimeria* algumas vezes ficou entre os peixes, mas sendo a espécie mais distante deles, com exceção de *Petromyzon*, e outras vezes ficou entre os peixes e os tetrápodos. Mas em ambos os casos ficou em uma situação próxima da situação na qual se esperava encontrá-la em virtude da proximidade filogenética que se sabe existir entre o grupo que originou os tetrápodos e aquele que originou os atuais peixes de nadadeira lobada.

Os peixes permaneceram juntos quase sempre, formando grupo monofilético, mas apresentando uma subdivisão com bastante simetria. Esta simetria se deve ao fato de vários peixes apresentarem o gene RAPH1 em duplicidade, e em decorrência disto, os peixes colocarem-se em posições semelhantes, em um e outro grupo.

Cada um dos genes parálogos forma um dos subgrupos, que assim faz que haja, para cada espécie de gene duplo, um gene em cada subgrupo. Em alguns casos um

destes subgrupos aproximou-se mais dos tetrápodos do que do outro subgrupo, o que era de se esperar, uma vez que um dos genes parálogos é o gene mais proximamente aparentado com o gene presente nos tetrápodos.

Como conclusão geral, embora com algumas exceções pode-se dizer, então, que o gene RAPH1, nas espécies estudadas, tem sua história evolutiva muito próxima da história evolutiva de todo o restante do genoma destas espécies. E infere-se que, com os diferentes graus de conservação que as diferentes porções do genoma apresentam, o gene RAPH1 manteve sua estrutura básica e acumulou diferenças de forma similar ao restante do genoma.

No que diz respeito ao gene em si, pode-se observar que os domínios proteicos GM, RA, PH e EVH1 apresentam um grau de alteração de sua estrutura muito inferior ao observado no gene inteiro, para o conjunto de espécies analisadas como um todo.

Naturalmente, este fato já se esperava, uma vez que os domínios são os principais responsáveis pelo desempenho fisiológico das proteínas, o que leva a um maior grau de preservação de suas estruturas ao longo da história evolutiva do gene.

Genes RAPH1 parálogos foram encontrados em 5 das 48 espécies analisadas e são todas elas espécies de peixes ósseos, com nadadeiras com raios, e, portanto, bastante próximas filogeneticamente. Com suas sequências aminoacídicas foram construídas filogenias que evidenciaram um agrupamento das 10 sequências em dois grupos de 5 sequências, com um gene de cada espécie em cada grupo. E esta forma de agrupamento sugere que houve duplicação do gene RAPH1 bastante anterior às épocas em que as cinco espécies separaram-se umas das outras. A composição e a disposição dos grupos permite dizer que houve duplicação gênica em um ancestral comum das cinco espécies, seguida de acúmulo, ao longo do tempo, de alterações em cada uma das duas cópias

gênicas, o que fez que deixassem de ser idênticas e passassem a ser progressivamente diferentes.

Restando a dúvida a respeito do relacionamento filogenético que os genes das demais espécies teriam com ambos os genes das espécies que o tem duplicados, foram feitas filogenias, por método de Máxima Verossimilhança, com genes de espécies com cópia-única com as dez cópias de genes duplos das cinco espécies que os apresentam.

As espécies que apresentam dois genes RAPH1 são apenas peixes, sendo que réptil, anfíbio, mamíferos e ave, bem como alguns peixes, apresentam cópia única do gene, portanto, seria possível achar que tivesse ocorrido duplicação do gene no ancestral comum dos peixes que apresentam gene duplo e não tivesse ocorrido duplicação nos ancestrais das demais espécies ou seria possível achar que houvesse ocorrido duplicação no ancestral comum de todas as espécies analisadas e que tivesse ocorrido perda de uma das cópias do gene no ancestral comum das espécies que apresentam cópia única.

Porém ficou demonstrado que todos os genes das espécies de gene único são mais estreitamente aparentados com um só dos grupos de genes das espécies de gene-duplo. E este fato indica fortemente que houve duplicação do gene RAPH1 no ancestral comum de todas as espécies analisadas e que, bastante posteriormente, houve perda de uma das cópias do gene em um, ou mais de um, ancestral comum das espécies que apresentam uma só cópia. Fica menos provável a explicação alternativa, pois exigiria que as espécies de cópia única tivessem seus genes mais proximamente aparentados indistintamente com ambas as cópias do gene das espécies de dois genes. Ou seja, dizendo de outra maneira, se a duplicação do gene tivesse ocorrido no ancestral comum das espécies de gene duplo, não haveria motivo para que todos os genes das espécies de cópia única fossem mais próximos filogeneticamente de apenas um dos grupos de gene das espécies de gene duplo. O trabalho evidencia, ainda, que esta duplicação gênica não

se deu no momento em que se separaram estes ancestrais mencionados acima, uma vez que foi necessário que decorresse tempo suficiente para que as cópias originadas por duplicação acumulassem, independentemente, as diferenças que nos permitem hoje caracterizá-las.

REFERÊNCIAS

- AIKI, Y.; KOZAKI, T.; MIZUNO, H.; & KONO, Y. Amino acid substitution in Ace paralogous acetylcholinesterase accompanied by organophosphate resistance in the spider mite *Tetranychus kanzawai*. **Pesticide Biochemistry and Physiology**, v.82, n.2, p.154-161. 2005.
- ALCÂNTARA, V. M. Fenótipos da butirilcolinesterase e suas relações com dados antropométricos, bioquímico-hormonais e pressão arterial em obesos e na população geral de Curitiba, PR. 2000. Tese (Doutorado em Ciências Biológicas – Genética) - Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2000.
- AMORIN, D. **Elementos Básicos de Sistemática Filogenética** Ribeirão Preto: Holos Editora e Sociedade Brasileira de Entomologia, 1997. 2ª edição.
- ARPAGUS, M.; KOTT, M.; VATSIS, K. P.; BARTELS, C. F.; LADU, B. N. & LOCKRIDGE, O. Structure of the gene for human butyrylcholinesterase: evidence for a single copy. **Biochemistry**, v.29, p.124-131. 1990.
- BARTZOKIS, G. Quadratic Trajectories of Brain Myelin Content: Unifying Constrict for Neuropsychiatric Disorders. **Neurobiology of Aging**, v.25, p.49-62. 2004.
- BIOEDIT. Disponível em: <http://www.mbio.ncsu.edu/BioEdit/BioEdit>. Acesso em:2013.
- BLAST N. Disponível em: <http://www.ncbi.nlm.nih.gov/blast/blast.cgi>. Acesso em:2013.
- BOUDINOT, E.; BERNARD, V.; CAMP, S.; TAYLOR, P.; CHAMPAGNAT, J.; KREJCI, E. Influence of differential expression of acetylcholinesterase in brain and muscle on respiration. **Respiratory Physiology and Neurobiology**, v.165, p.40-48. 2009.
- BRIMIJOIN, S.; RAKONCZAY, Z.; & MINTZ, K. Immunochemistry of a mammalian cholinesterases. **Fed. Proc.**, v.45, n.13, p.2960-2964. 1986.
- CAVALLI-SFORZA, L. & EDWARDS, A. F. Phylogenetic analysis: models and estimation procedures. **American Journal of Human Genetics**, v.19, p.233-257. 1967.
- CERBAI, F.; GIOVANNINI, M. G.; MELANI, C.; ENZ, A. & PEPEU, G. N Phenethyl-norcymserine, a Selective Butyrylcholinesterase inhibitor, Increases Acetylcholine Release in rat cerebral Cortex: A Comparison with Donepezil and Rivastigmine. **European Journal of Pharmacology**, v.572, p.142-150. 2007.
- CHAMARY, J.-V. & HURST, L. D. Similar Rates but Different Modes of Sequence Evolution in Introns and at Exonic Silent Sites in Rodents: Evidence for Selectively Driven Codon Usage. **Molecular Biology and Evolution**, v.21, n.6, p.1014-1023. 2004.
- CHUIKO, G. M.; PODGORNAYA, V. A. & ZHELNIN, Y. Y. Acetylcholinesterase and butyrylcholinesterase activities in brain and plasma of freshwater teleosts: cross-species and cross-family differences. **Biochemistry and Molecular Biology**, v.135B, n.1, p.55-61. 2003.
- CLETO, C. L.; VANDENBERGHE, A. E.; MACLEAN, D. W.; PANNUNZIO, P.; TORTORELLI, C.; MEEDEL, T. H. Ascidian Larva Reveals Ancient Origin of Vertebrate-Skeletal-Muscle Troponin I Characteristics in Chordate Locomotory Muscle. **Molecular Biology and Evolution**, v.20, n.12, p.2113-2122. 2003.
- CLUSTALW. Disponível em: <ftp://ftp.ebi.ac.uk/pub/software/clustalw2>. Acesso em:2013.
- DINIZ FILHO, J. F. **Métodos Filogenéticos Comparativos**. Ribeirão Preto: Holos Editora, 2000.
- DNASP. Disponível em: <http://www.ub.edu/DNAsp>. Acesso em:2013.
- DUYSEN, E. G.; LI, B., DARVESH, S. & LOCKRIDGE, O. Sensitive of butyrylcholinesterase knockout mice to huperzine A and donepezil suggests humans with butyrylcholinesterase deficiency may not tolerate these Alzheimer's disease drugs and indicates butyrylcholinesterase function in neurotransmission. **Toxicology**, v.233, p.60-69. 2007.
- ENSEMBL. Disponível em: <http://www.ensembl.org>. Acesso em: 2013.

- EPPERT, K.; WUNDER, J. S.; ANELIUNAS, V.; TSUI, L. C.; Scherer, S. W. & ANDRULIS, I. L. Altered Expression and Deletion of RMO1 in Osteosarcoma. **International Journal of Cancer**, v.114, n.5, p.738/746. 2005.
- FREDERICK, A.; TSIGELNY, I.; COHENOUR, F.; SPIKER, C.; KREJCI, E.; CHATONNET, A. Acetylcholinesterase from the invertebrate *Ciona intestinalis* is capable of assembling into asymmetric forms when co-expressed with vertebrate collagenic tail peptide. **FEBS Journal**, v.275, n.6, p.1309-1322. 2008.
- GAO, D. & ZHAN, C. G. Modeling evolution of hydrogen bonding and stabilization of transition states in the process of cocaine hydrolysis catalyzed by human butyrylcholinesterase. **Proteins**, v.62, n.1, p.99-100. 2006.
- GAO, J. R.; KAMBHAMPATI, S. & ZHU, K. Y. Molecular cloning and characterization of a greenbug (*Schizaphis graminum*) cDNA encoding acetylcholinesterase possibly evolved from a duplicate gene lineag. **Insect Biochemistry and Molecular Biology**, v.32, n.7, p.765-766. 2002.
- GENEBANK. Disponível em: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>. Acesso em:2013.
- GRAUR, D. & LI, W.-H. **Fundamentals of Molecular Evolution** Sunderland, Massachusetts, United States: Sinauer Associates Incorporation, 2000. 2a edição
- HELFMAN, G.; COLLETTE, B. B.; FACEY, D. E. & BOWEN, B. W. **The diversity of Fishes: Biology, Evolution, and Ecology** Hoboken: Wiley-Blackwell, 2009. 2a edição.
- JBILO, O.; BARTELS, C. F.; CHATONNET, A.; TOUTANT, J.-P. & LOCKRIDGE, O. Tissue distribution of human acetylcholinesterase and butyrylcholinesterase messenger RNA. **Toxicon**, v.32, n.11, p.1445-1457. 1994.
- JEFFERY, W. R.; SWALLA, B. J.; EWING, N. & Kusakabe, T. Evolution of the ascidian anural larva: evidences from embryos and molecules. **Molecular Biology and Evolution**, v.15, n.12, p.646-54. 1998.
- JOHNSON, N. D.; DUYSEN, E. G. & LOCKRIDGE, O. Intrathecal delivery of fluorescent labeled butyrylcholinesterase to the brains of butyrylcholinesterase knock-out mice: visualization and quantification of enzyme distribution in the brain. **Neurotoxicology**, v.30, p.386-392. 2009.
- JUUL, P. Human Plasma Cholinesterase Isoenzymes. **Clin. Chim. Acta**, v.19, p.205-213. 1968.
- LABBÉ, P.; BERTHOMIEU, A.; BERTICAT, C.; ALOUT, H., RAYMOND, M.; LENORMAND, T. Independent Duplications of the Acetylcholinesterase Gene Conferring Insecticide Resistance in the Mosquito *Culex pipiens*. **Molecular Biology and Evolution**, v.24, n.4, p.1056-1067. 2007.
- LAGLER, K. F.; BARDACH, J. E.; MILLER, R. R. & PASSINO, D. R. **Ichthyology**. New York: John Willey, 1977.
- LEWIN, B. **Genes IX**. Sudbury: Jones and Bartlett, 2008.
- LI, F. & HAN, Z.-J. Two different genes encoding acetylcholinesterase existing in cotton aphid (*Aphis gossypii*). **Genome**, v.45, n.6, p.1134-1141. 2002.
- MASSON, P.; FROMENT, M.-T.; GILLON, E.; NACHON, F.; DARVESH, S.; & SCHOPFER, L. M. Kinetic Analysis of Butyrylcholinesterase-catalyzed Hydrolysis of Acetanilides. **Biochimica et Biophysica Acta**, v.1774, p.1139-1147. 2007.
- MEGA5. Disponível em: Mega Software: <http://www.megasoftware.net/mega.html>. Acesso em:2013.
- MINTZ, K. P.; WEINSHILBOUM, R. M.; & BRIMIJOIN, W. A. Evolution of butyrylcholinesterase in higher primates: an immunological study. **Comparative Biochemistry and Physiology**, v.79, n.1, p.35-37. 1984.
- MOUNT, D. W. **Bioinformatics: sequence and genome analysis** New York: Cold Spring Harbor Laboratory Press, 2004. 2a edição.
- MOYLE, P. B. & CHECH Jr, J. J. **Fishes: An Introduction to Ichthyology** P. Hall Ed, 2003. 5a edição.
- MYERS, M.; RICHMOND, R. C. & OAKESHOTT, J. G. On the Origins of Esterases. **Molecular Biology and Evolution**, v.5, n.2, p.113-119. 1988.

- NAGASE, T.; KIKUNO, R.; HATTORI, A.; KONDO, Y.; OKUMURA, K. & OHARA, O. Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. **DNA Res**, v.7, n.4, p.347-355. 2000.
- NCBI. Disponível em: <http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>. Acesso em:2013.
- NEI, M. & KUMAR, S. **Molecular Evolution and Phylogenetics**. New York: Oxford University Press, 2000.
- O'GRADY, P. M.; DURANDO, C. M.; HEED, W. B.; WASSERMAN, M.; ETGES, W. & DESALLE, R. Genetic Divergence Within the *Drosophila* Mayaguana Subcluster, a Closely Related Triad of Caribbean Species In the Repleta Species Group. **Hereditas**, v.136, n.3, p.240-245. 2002.
- PIOTTE, C.; ARTHAUD, L.; ABAD, P. & ROSSO, M.-N. Molecular cloning of an acetylcholinesterase gene from the plant parasitic nematodes *Meloidogyne incognita* and *Meloidogyne javanica*. **Molecular and Biochemical Parasitology**, v.99, n.2, p.247-256. 1999.
- SANDERS, M.; MATHEWS, B.; SUTHERLAND, D.; SOONG, W.; GILES, H. & PEZZEMENTI, L. Biochemical and Molecular Characterization of Acetylcholinesterase from the Hagfish *Myxine glutinosa*. **Comparative Biochemistry and Physiology**, v.115B, n.1, p.97-109. 1996.
- SCHNEIDER, H. **Métodos de Análise Filogenética**. Ribeirão Preto: Holos Editora, 2003. 2ª edição.
- SOKAL, R. R. & SNEATH, P. H. **Numerical Taxonomy: the principals and practice of numerical classification**. San Francisco: W H Freeman and Company, 1973.
- SOUZA, R. L. Variação Fenotípica e Genotípica Da Butirilcolinesterase Humana e Sua Relação Com Altura e Índice De Massa Corporal. 2001. Tese (Doutorado em Ciências – Genética) - Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2001.
- SRIDHAR, G. R.; LAKSHMI, P. V. & RAO, A. A. Phylogenetic Tree Construction of Butyrylcholinesterase Sequences in Life Forms. **J Assoc Physicians India**, v.54, p.122-123. 2006.
- SRIDHAR, G. R.; NIRMALA, G.; RAO, A. A.; MADHAVI, A. S.; SREELATHA, S.; RANI, J. S. Serum butyrylcholinesterase in type 2 diabetes mellitus: a biochemical and bioinformatics approach. **Lipids in Health and Disease**, v.4, n.18. 2005.
- SUTHERLAND, D.; MCCLELLAN, J. S.; MILNER, D.; SOONG, W.; AXON, N., SANDERS, M. Two cholinesterase activities and genes are present in amphioxus. **J.Exp.Zool**, v.277, n.3, p.213-229. 1997.
- TREE OF LIFE. Disponível em: <http://tolweb.org>. Acesso em:2013.

APÊNDICES

APÊNDICE 1 - MATERIAL E MÉTODOS

O objeto do presente trabalho é o gene RAPH1, também chamado informalmente de gene da lamelipodina e anteriormente nominado através de um grande número de termos ainda presentes na maior parte dos trabalhos referentes a ele. O HUGO Gene Nomenclature Committee fixou o nome do gene em Ras Association (RalGDS/AF-6) and Pleckstrin Homology Domains 1 (RAPH1), deixando como sinonímia os seguintes termos (NCBI) (GeneBank):

amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 18
(ALS2CR18)

amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 9
(ALS2CR9)

KIAA1681

Lamellipodin (LPD)

OTTHUMP00000163779

OTTHUMP00000163780

OTTHUMP00000206258

OTTHUMP00000208081

OTTHUMP00000208082

OTTHUMP00000208083

OTTHUMP00000208084

OTTHUMP00000208085

OTTHUMP00000208086

OTTHUMP00000208087

OTTHUMP00000208088

Proline-rich EVH1 ligand 2 (PREL-2)

Protein RMO13

RaIGDS/AF-6

RMO1

O gene, no genoma humano, está posicionado em 2q33 e tem como código identificador, perante o HUGO Gene Nomenclature Committee, o termo HGNC 14436.

Os principais materiais utilizados no trabalho foram os conjuntos de sequências nucleotídicas e aminoacídicas obtidas dos bancos de DNA disponibilizados através da rede internacional de computadores (Internet). As principais páginas através das quais foram obtidas as sequências foram a Ensembl, do European Bioinformatics Institute e do Wellcome Trust Sanger Institute, acessada através do Universal Resource Locator (URL) <<http://www.ensembl.org/index.html>> e a do National Center for Biotechnology Information (NCBI), acessada através do URL <<http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>>. (NCBI) (GeneBank) (DNAsp)

As sequências nucleotídicas depositadas nestes bancos tanto correspondem a nucleotídeos de ácido desoxirribonucleico (DNA) do gene como um todo, incluindo íntrons e éxons, quanto a nucleotídeos de ácido ribonucleico (RNA) correspondentes principalmente a ácidos ribonucleicos mensageiros (RNAm), em uma ou mais das composições alternativas utilizadas pelo organismo e seus correspondentes DNAs complementares (DNAc).

Para avaliar a completude e a correspondência destas sequências, foi feito o alinhamento delas através do programa alinhador de múltiplas sequências ClustalW, contido no programa Bioedit, da North Carolina State University, obtido de sua página na rede, acessada através do URL <<http://www.mbio.ncsu.edu/BioEdit/BioEdit>>. (BioEdit)

Foram utilizadas no trabalho também sequências de aminoácidos correspondentes às sequências nucleotídicas referidas. Estas correspondem aos RNAm identificados e sequenciados que tiveram suas sequências nucleotídicas convertidas em aminoácidos pelos próprios bancos fornecedores de sequências.

Como a cada gene correspondem muitas vezes mais de um RNAm, em suas várias composições alternativas, muitas vezes foi necessário descobrir a que sequência de aminoácidos corresponderia um RNAm que incluísse todos os éxons do gene, utilizando para isto a informação contida nas sequências de aminoácidos correspondentes a todos os RNAm de um determinado gene. (Lewin, 2008) (NCBI)

Isto foi, também, realizado através de alinhamentos diversos com o programa ClustalW. (Clustalw)

A busca de semelhanças entre o gene RAPH1 e outros trechos no genoma de quaisquer espécies foi realizada através do programa BLAST do NCBI, acessado através do URL <<http://www.ncbi.nlm.nih.gov/blast/blast.cgi>>. (Blast N)

As espécies apresentadas na Tabela 2 foram aquelas das quais se obtiveram sequências nucleotídicas de DNA, de bancos de sequências disponibilizados na rede, anotadas sob a denominação de RAPH1 ou sob todas as demais denominações equivalentes. (NCBI) (Tree of Life) (GeneBank)

Tabela 2 - Espécies com sequências nucleotídicas de DNA do gene RAPH1

* +	Ailuropoda melanoleuca	panda	Ursidae
* #	Anolis carolinensis	lagarto americano	Polychrotidae
* N	Bos taurus	boi	Bovidae
* #	Caenorhabditis elegans	nematoda de clima temperado	Rhabditidae
* #	Callithrix jacchus	sagui comum	Callithricidae
* #	Canis familiaris	cão	Canidae
* #	Cavia porcellus	cobaia	Caviidae
* N	Choloepus hoffmanni	preguiça	Megalonychidae
* #	Danio rerio (2)	paulistinha	Cyprinidae
* N	Dipodomys ordii	rato-canguru	Heteromyidae
* # +	Equus caballus	cavalo	Equidae
* N	Erinaceus europaeus	ouriço europeu	Erinaceidae
* N	Felis catus	gato doméstico	Felidae
* # +	Gallus gallus	galinha	Phasianidae
* #	Gasterosteus aculeatus (2)	esgana-gato	Gasterosteidae
* #	Gorilla gorilla	gorila	Hominidae
* # +	Homo sapiens	homem	Hominidae
* #	Loxodonta africana	elefante africano	Elephantidae
* N	Macropus eugenii	Canguru	Macropodidae
* #	Meleagris gallopavo	peru	Phasianidae
* N	Microcebus murinus	lêmure-camundongo	Cheirogaleidae
* #	Monodelphis domestica	guaiquica	Didelphidae
* # +	Mus musculus	camundongo	Muridae
* N	Myotis lucifugus	morcego norte americano	Vespertilionidae
* N	Nomascus leucogenys	gibão da bochecha branca	Hylobatidae
* N	Ochotona princeps	lebre assobiadora	Ochotonidae
* N	Oryctolagus cuniculus	coelho	Leporidae
* #	Oryzias latipes	medaca	Adrianichthyidae
* N	Otolemur garnettii	gálago	Galagidae
* # +	Pan troglodytes	chimpanzé	Hominidae
* #	Pongo abelli	orangotango de Sumatra	Hominidae
* N	Procavia capensis	damão do cabo	Procaviidae
* #	Pteropus vampyrus	morcego asiático	Pteropodidae
*	Rattus norvegicus	ratazana	Muridae
* N	Sorex araneus	musaranho	Soricidae
* # +	Sus scrofa	porco doméstico	Suidae

* #	Taeniopygia guttata	mandarim	Passeridae
* #	Takifugu rubripes (2)	baiacu japonês	Tetraodontidae
* N	Tarsius syrichta	társio das Filipinas	Tarsiidae
* #	Tetraodon nigroviridis (2)	baiacu do sul da Ásia	Tetraodontidae
* N	Tupaia belangeri	musaranho arborícola	Tupaiidae
* N	Tursiops truncatus	golfinho comum	Delphinidae
* N	Vicugna pacos	Alpaca	Camelidae
* # +	Xenopus laevis	rã africana	Pipidae

* Espécies das quais foi capturada sequência de todo o gene.

+ Espécies das quais foi capturada sequência de éxons a partir de RNA.

Espécies das quais foi capturada sequência de éxons, a partir da sequência de todo o gene.

N Espécies com falhas em qualquer ponto da sequência do gene.

() Espécies com genes RAPH1 parálogos

Dentre as muitas sequências disponíveis, foram utilizadas, no que diz respeito aos polipeptídios, as seguintes:

Tabela 3 - Sequências de aminoácidos utilizadas

Ailuropoda melanoleuca	ENSAMET00000003248
Anolis carolinensis	ENSACAT00000004500
Bos taurus	ENSBTAT00000014638
Caenorhabditis elegans	F10E9.6c
Callithrix jacchus	ENSCJAT00000010390 modif
Canis familiaris	ENSCAFT00000020408
Cavia porcellus	ENSCPOT00000022088
Choloepus hoffmanni	ENSCHOT00000000246
Danio rerio gene1	ENSDART00000061156 modif
Danio rerio gene2	ENSDART00000088290
Dipodomys ordii	ENSDORT00000007527
Equus caballus	ENSECAT00000023862
Erinaceus europaeus	ENSEEUT00000012046
Felis catus	ENSFCAT00000019231
Gadus morhua gene1	ENSGMOT00000007333
Gadus morhua gene2	ENSGMOT00000000606 modif
Gallus gallus	ENSGALT00000014120
Gasterosteus aculeatus gene1	ENSGACT00000011348
Gasterosteus aculeatus gene2	ENSGACT00000020022
Gorilla gorilla	ENSGGOT00000024255 modif
Homo sapiens	ENST00000374493
Latimeria chalumnae	ENSLACT00000013196
Loxodonta africana	ENSLAFT00000015138 modif
Macropus eugenii	ENSMEUT00000002528
Meleagris gallopavo	ENSMGAT00000009866
Microcebus murinus	ENSMICT00000001086
Monodelphis domestica	ENSMODT00000016314 modif
Mus musculus	ENSMUST00000127573
Myotis lucifugus	ENSMLUT00000008182
Nomascus leucogenys	ENSNLET00000009003
Ochotona princeps	ENSOPRT00000005690

Ornithorhynchus anatinus	ENSOANT00000010805
Oryctolagus cuniculus	ENSOCUT00000016729
Oryzias latipes	ENSORLT00000006623
Otolemur garnettii	ENSOGAT00000015105
Pan troglodytes	ENSPTRT00000056178
Petromyzon marinus	ENSPMAT00000006307
Pongo abelli	ENSPPYT00000015226
Procavia capensis	ENSPCAT00000006276
Pteropus vampyrus	ENSPVAT00000014164
Sarcophilus harrisii	ENSSHAT00000008942
Sorex araneus	ENSSART00000000357
Sus scrofa	ENSSSCT00000017550
Taeniopygia guttata	ENSTGUT00000010165
Takifugu rubripes gene1	ENSTRUT00000020869
Takifugu rubripes gene2	ENSTRUT00000040469
Tarsius syrichta	ENSTSYT00000014267
Tetraodon nigroviridis gene1	ENSTNIT00000017769
Tetraodon nigroviridis gene2	ENSTNIT00000014686
Tupaia belangeri	ENSTBET00000001734
Tursiops truncatus	ENSTTRT00000005203
Vicugna pacos	ENSVPAT00000009507
Xenopus tropicalis	ENSXETT00000061421

De acordo com a disponibilidade das sequências nos bancos de dados, foram obtidas sequências de diferentes graus de completude, mas de modo geral abrangendo todos os éxons do gene, tanto no que diz respeito às sequências aminoacídicas quanto às nucleotídicas, como pode-se observar abaixo:

Tabela 4 - Tamanho das sequências nucleotídicas - reunião de éxons

Ailuropoda melanoleuca	Panda	1789
Anolis carolinensis	lagarto americano Polychrotidae	3726
Bos taurus	Boi	4714
Caenorhabditis elegans	nematoda de clima temperado	1847
Callithrix jacchus	sagui comum	6989
Canis familiaris	Cão	3135
Cavia porcellus	Cobaia	1962
Choloepus hoffmanni	Preguiça	
Ciona intestinalis	ascídia Urochordata	2152
Ciona savignyi	ascídia Urochordata	718
Danio rerio (2)	Paulistinha	4592/6484
Dasyopus novemcinctus	tatu comum	
Dipodomys ordii	rato-canguru	
Echinops telfairi	ouriço de Madagascar	
Equus caballus	Cavalo	2079
Erinaceus europaeus	ouriço europeu	
Felis catus	Gato	
Gallus gallus	Galinha	2267

Gasterosteus aculeatus (2)	esgana-gato	2205/3819
Gorilla gorilla	Gorila	3990
Homo sapiens sapiens	Homem	17466
Homo sapiens neanderthalensis	homem neandertal	10243
Loxodonta africana	elefante africano	2480
Macaca mulatta	macaco resus	5284
Macropus eugenii	Canguru	
Meleagris gallopavo	Peru	3615
Microcebus murinus	lêmure-camundongo	
Monodelphis domestica	guaiquica	3647
Mus musculus	Camundongo	3689
Myotis lucifugus	morcego norte americano	
Nomascus leucogenys	gibão da bochecha branca	2289
Ochotona princeps	lebre assobiadora	
Ornithorhynchus anatinus	Ornitorrinco	2025
Oryctolagus cuniculus	Coelho	3165
Oryzias latipes	Medaca	3498
Otolemur garnettii	Gálago	
Pan troglodytes	Chimpanzé	10243
Pongo abelli	orangotango de Sumatra	3900
Procavia capensis	damão do cabo	
Pteropus vampyrus	morcego asiático	
Rattus norvegicus	Ratazana	
Sorex araneus	Musaranho	
Spermophilus tridecemlineatus	esquilo listrado	
Sus scrofa	porco doméstico	1572
Taeniopygia guttata	Mandarim	1799
Takifugu rubripes (2)	baiacu japonês	4155/4734
Tarsius syrichta	társio das Filipinas	
Tetraodon nigroviridis (2)	baiacu sul-asiático	2853/2708
Tupaia belangeri	musaranho arborícola	
Tursiops truncatus	golfinho comum	
Vicugna pacos	Alpaca	

Tabela 5 - Tamanho das sequências de aminoácidos

Ailuropoda melanoleuca	Panda	598
Anolis carolinensis	lagarto americano Polychrotidae	1241
Bos taurus	Boi	1145
Caenorhabditis elegans	nematoda de clima temperado	779
Callithrix jacchus	sagui comum	1255
Canis familiaris	Cão	1045
Cavia porcellus	Cobaia	654
Choloepus hoffmanni	Preguiça	895
Ciona intestinalis	ascídia Urochordata	
Ciona savignyi	ascídia Urochordata	
Danio rerio (gene 1)	Paulistinha	1367
Danio rerio (gene 2)	Paulistinha	1511
Dasyopus novemcinctus	tatu comum	
Dipodomys ordii	rato-canguru	1288
Echinops telfairi	ouriço de Madagascar	
Equus caballus	Cavalo	118

<i>Erinaceus europaeus</i>	ouriço europeu	1284
<i>Felis catus</i>	Gato	1029
<i>Gadus morhua</i> (gene 1)	Bacalhau	639
<i>Gadus morhua</i> (gene 2)	Bacalhau	318
<i>Gallus gallus</i>	Galinha	601
<i>Gasterosteus aculeatus</i> (gene1)	esgana-gato	735
<i>Gasterosteus aculeatus</i> (gene2)	esgana-gato	1272
<i>Gorilla gorilla</i>	Gorila	1308
<i>Homo sapiens sapiens</i>	Homem	1302
<i>Homo sapiens neanderthalensis</i>	homem neandertal	1302
<i>Loxodonta africana</i>	elefante africano	821
<i>Macaca mulatta</i>	macaco resus	
<i>Macropus eugenii</i>	Canguru	1303
<i>Meleagris gallopavo</i>	Peru	1204
<i>Microcebus murinus</i>	lêmure-camundongo	656
<i>Monodelphis domestica</i>	guaiquica	1182
<i>Mus musculus</i>	Camundongo	716
<i>Myotis lucifugus</i>	morcego norte americano	691
<i>Nomascus leucogenys</i>	gibão da bochecha branca	733
<i>Ochotona princeps</i>	lebre assobiadora	1301
<i>Ornithorhynchus anatinus</i>	Ornitorrinco	213
<i>Oryctolagus cuniculus</i>	Coelho	1054
<i>Oryzias latipes</i>	Medaca	1166
<i>Otolemur garnettii</i>	Gálago	1308
<i>Pan troglodytes</i>	Chimpanzé	1302
<i>Petromyzon marinus</i>	Lampreia	892
<i>Pongo abelli</i>	orangotango de Sumatra	1298
<i>Procavia capensis</i>	damão do cabo	1241
<i>Pteropus vampyrus</i>	morcego asiático	1309
<i>Rattus norvegicus</i>	Ratazana	
<i>Sarcophilus harrisii</i>	diabo da Tasmânia	1160
<i>Sorex araneus</i>	Musaranho	1303
<i>Spermophilus tridecemlineatus</i>	esquilo listrado	
<i>Sus scrofa</i>	porco doméstico	524
<i>Taeniopygia guttata</i>	Mandarim	600
<i>Takifugu rubripes</i> (gene 1)	baiacu japonês	1329
<i>Takifugu rubripes</i> (gene 2)	baiacu japonês	1346
<i>Tarsius syrichta</i>	társio das Filipinas	1218
<i>Tetraodon nigroviridis</i> (gene 1)	baiacu sul-asiático	1290
<i>Tetraodon nigroviridis</i> (gene 2)	baiacu sul-asiático	1305
<i>Tupaia belangeri</i>	musaranho arborícola	1297
<i>Tursiops truncatus</i>	golfinho comum	1296
<i>Vicugna pacos</i>	Alpaca	1018
<i>Xenopus laevis</i>	rã sul-africana	1235

APÊNDICE 2 – ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO GENE RAPH1

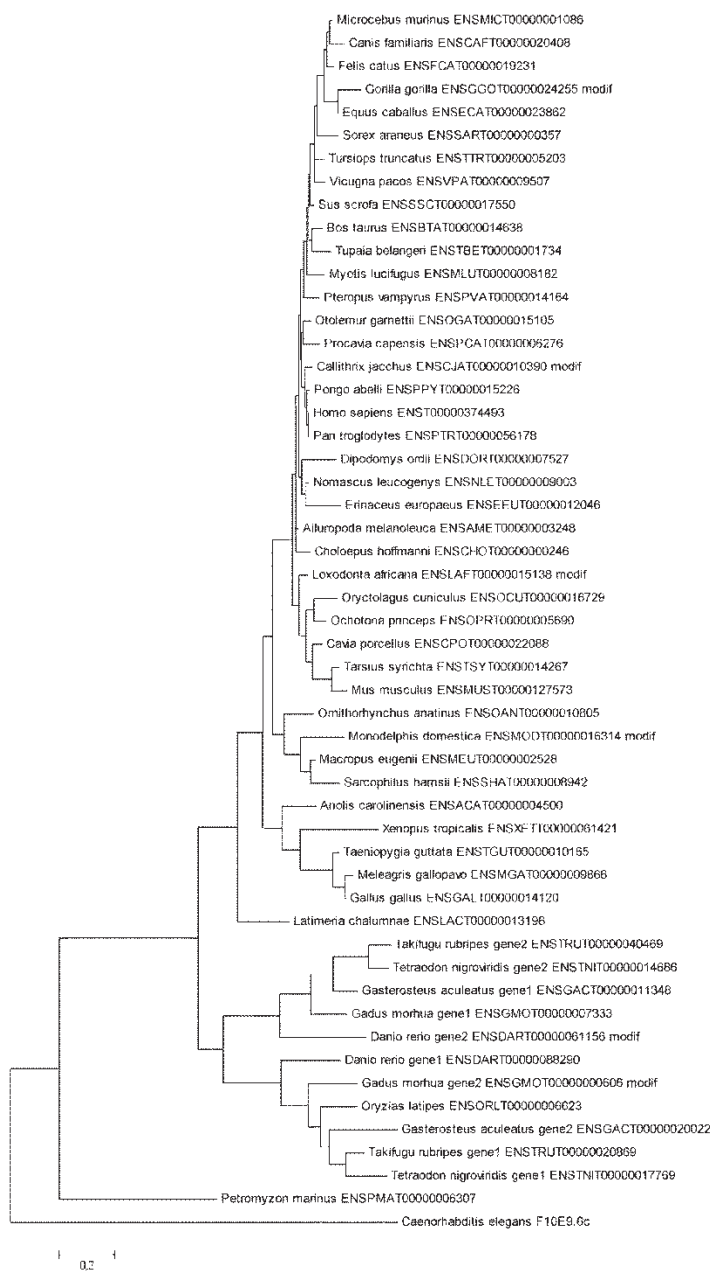


Figura 11 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos de 53 sequências, utilizando o método de Máxima Verossimilhança (ML).

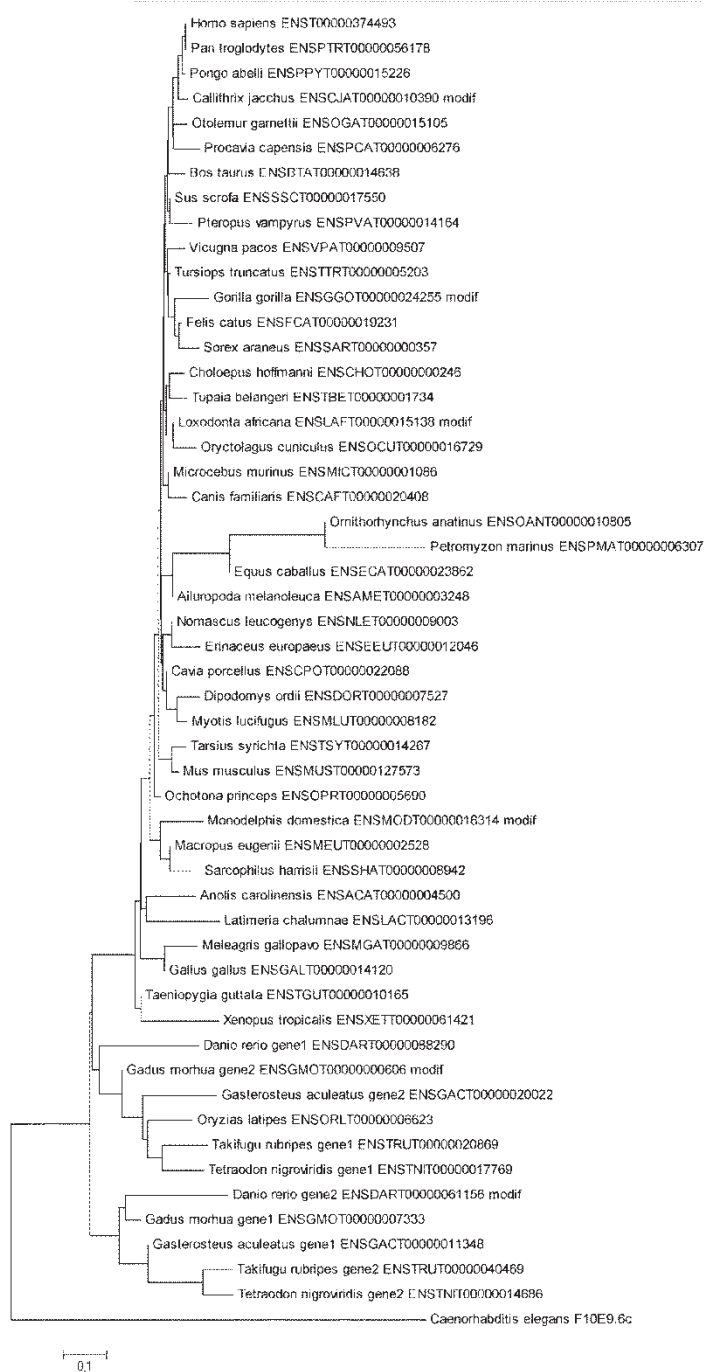


Figura 12 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos de 53 sequências, utilizando o método de Neighbor Joining (NJ)

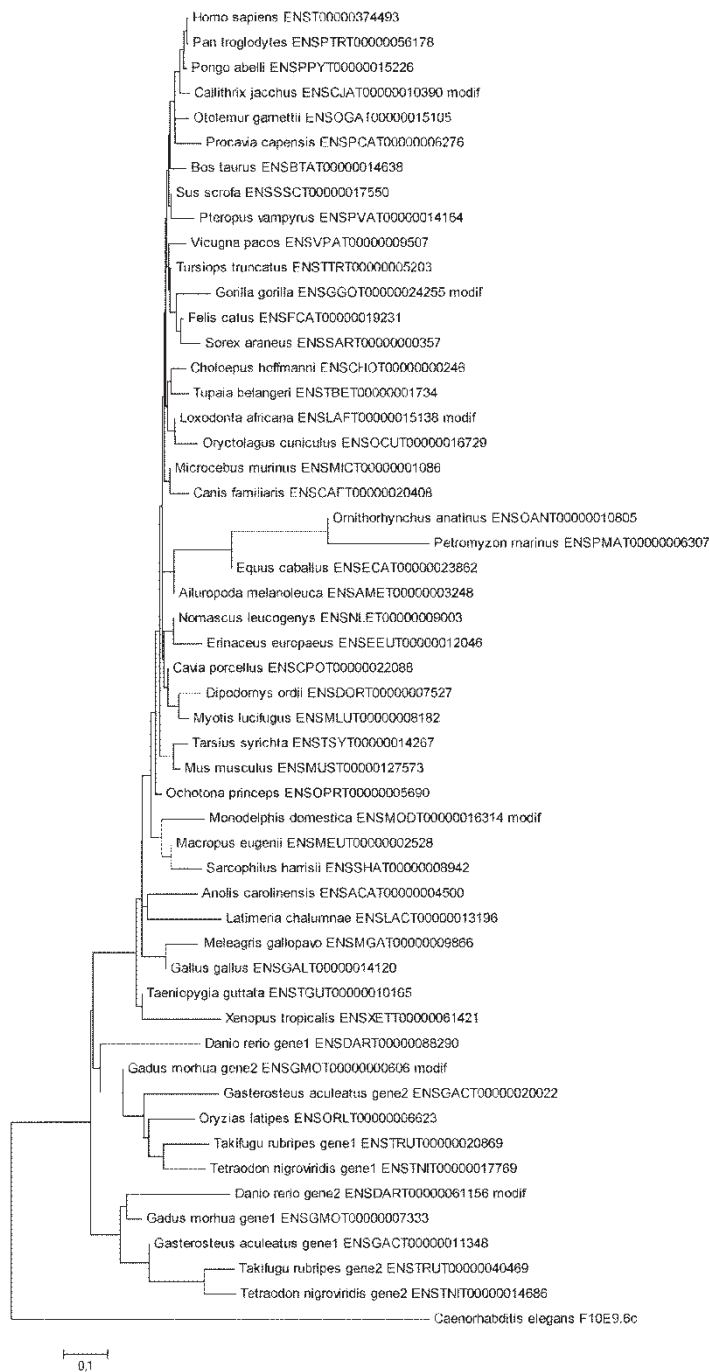


Figura 13 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos de 53 sequências, utilizando o método de Evolução Mínima (ME).

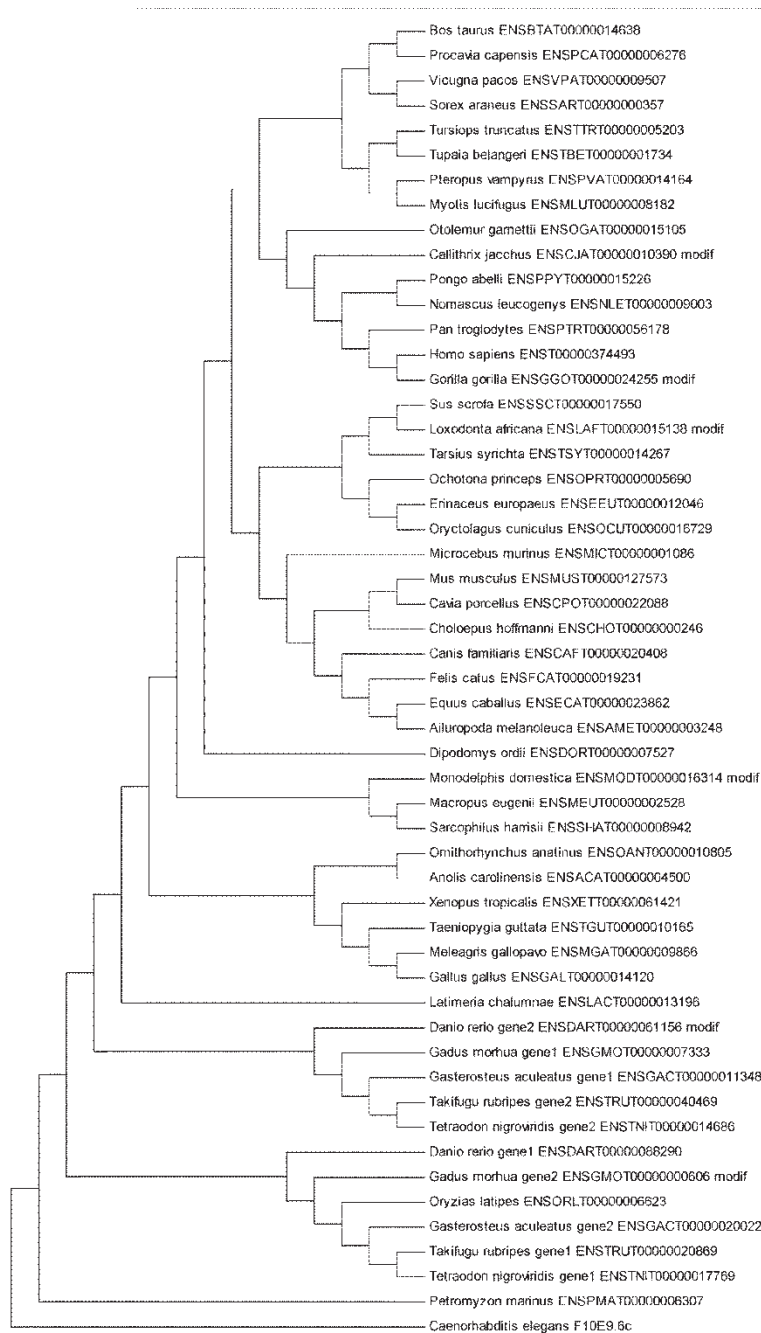


Figura 14 - Árvore filogenética do gene RAPH1 elaborada com seqüências de aminoácidos de 53 seqüências, utilizando o método de Máxima Parcimônia (MP).

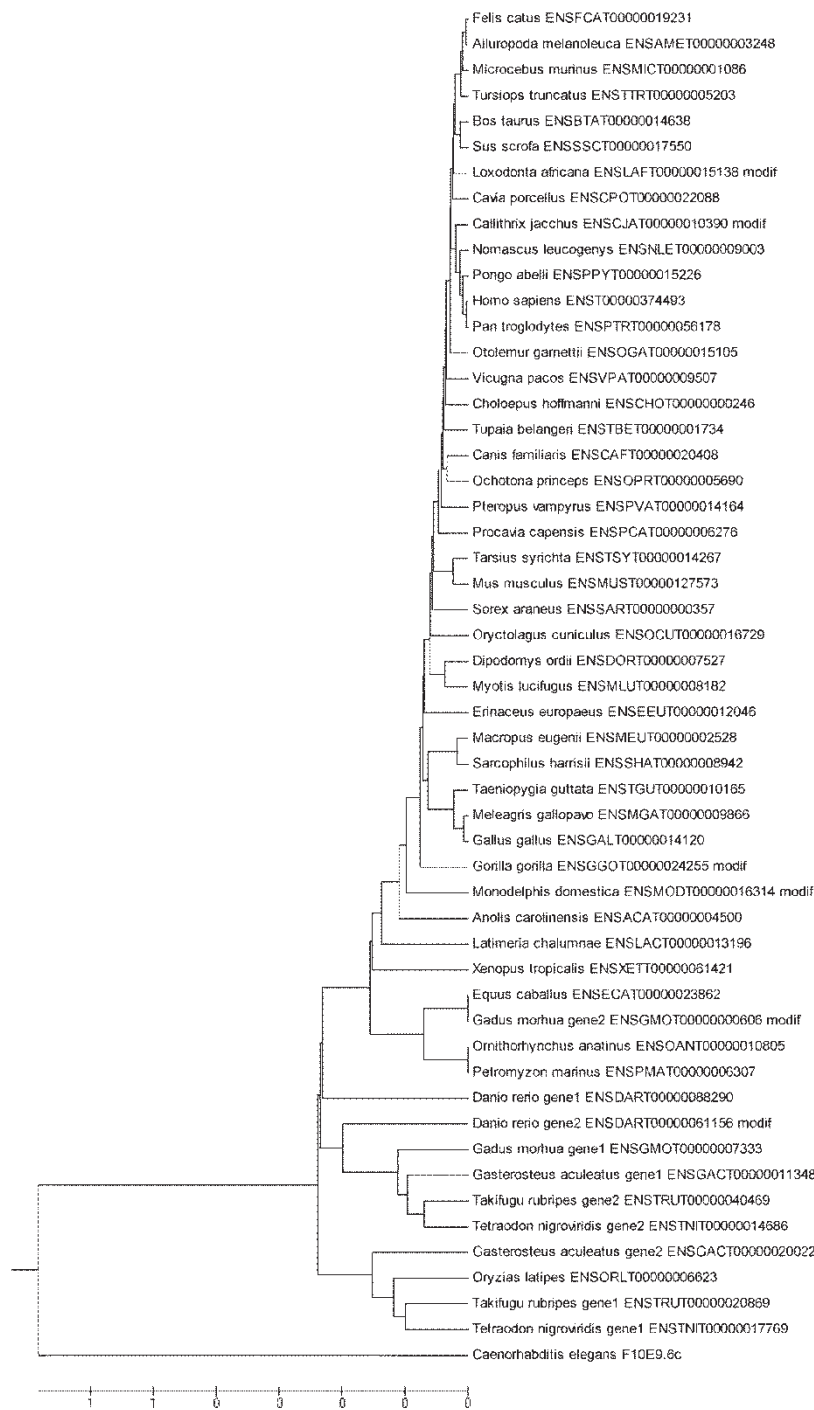


Figura 15 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos de 53 sequências, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

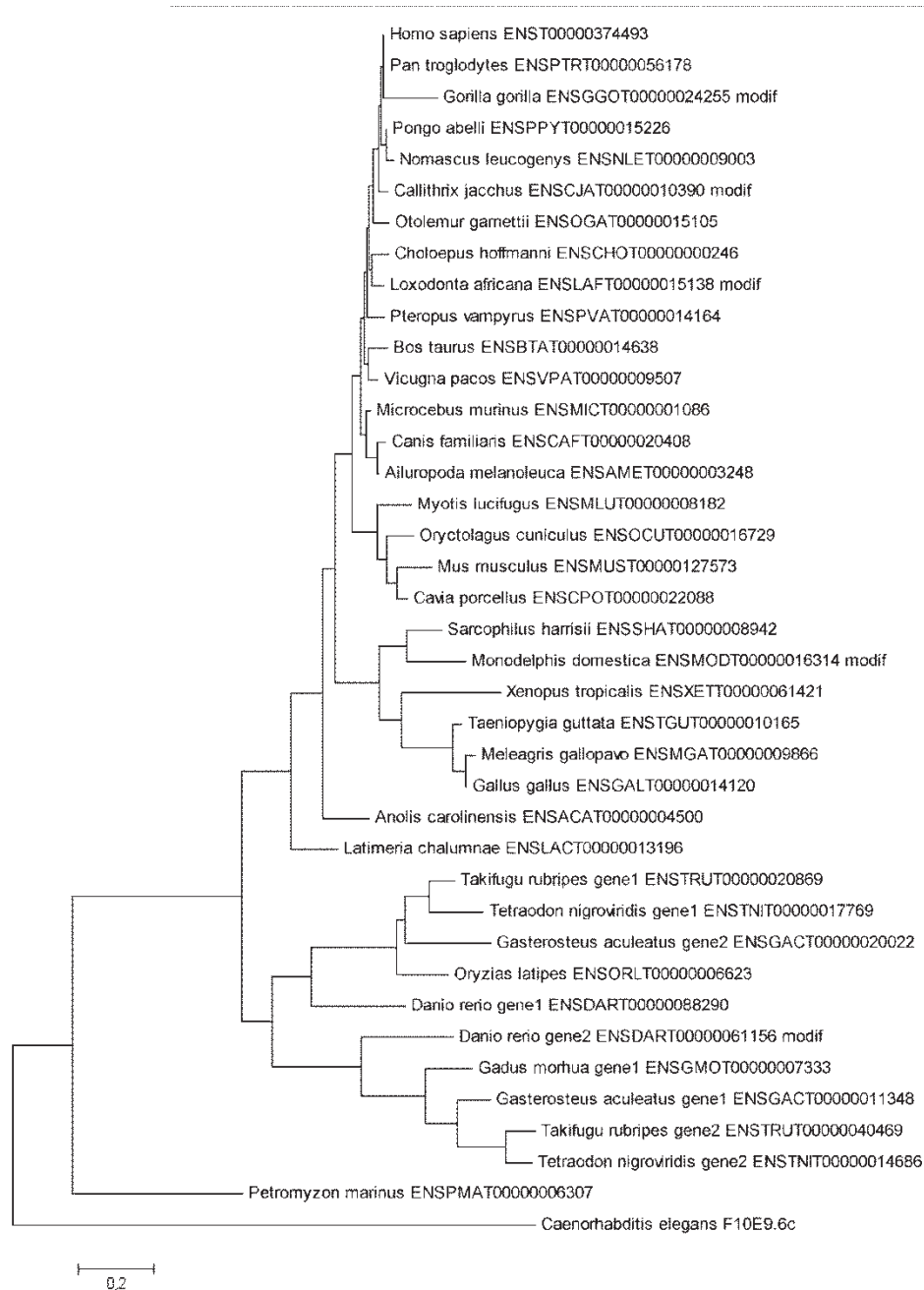


Figura 16 - Árvore filogenética do gene RAPH1 elaborada com seqüências de aminoácidos de 39 seqüências, utilizando o método de Máxima Verossimilhança (ML).

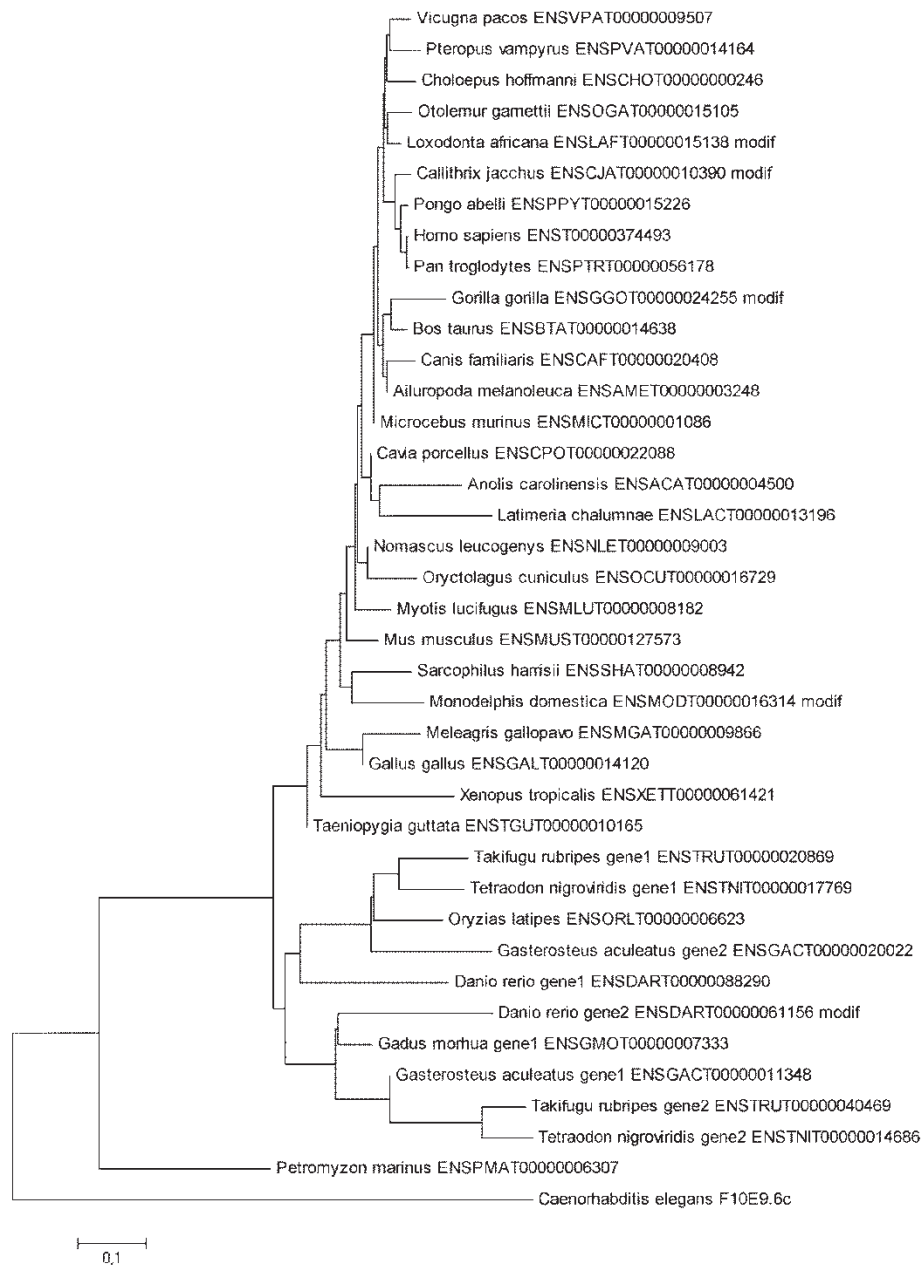


Figura 17 - Árvore filogenética do gene RAPH1 elaborada com seqüências de aminoácidos de 39 seqüências, utilizando o método de Neighbor Joining (NJ)

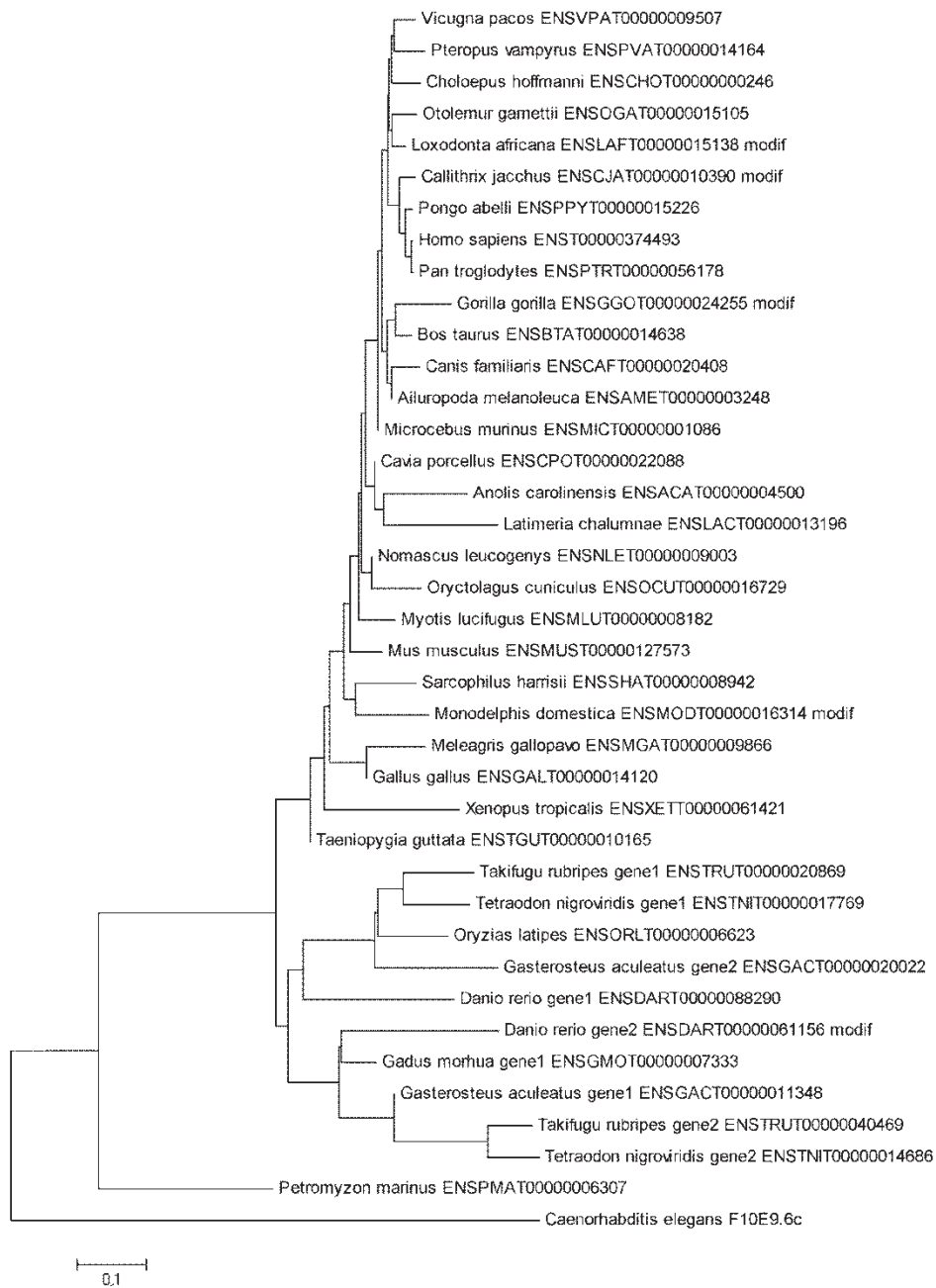


Figura 18 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos de 39 sequências, utilizando o método de Evolução Mínima (ME)

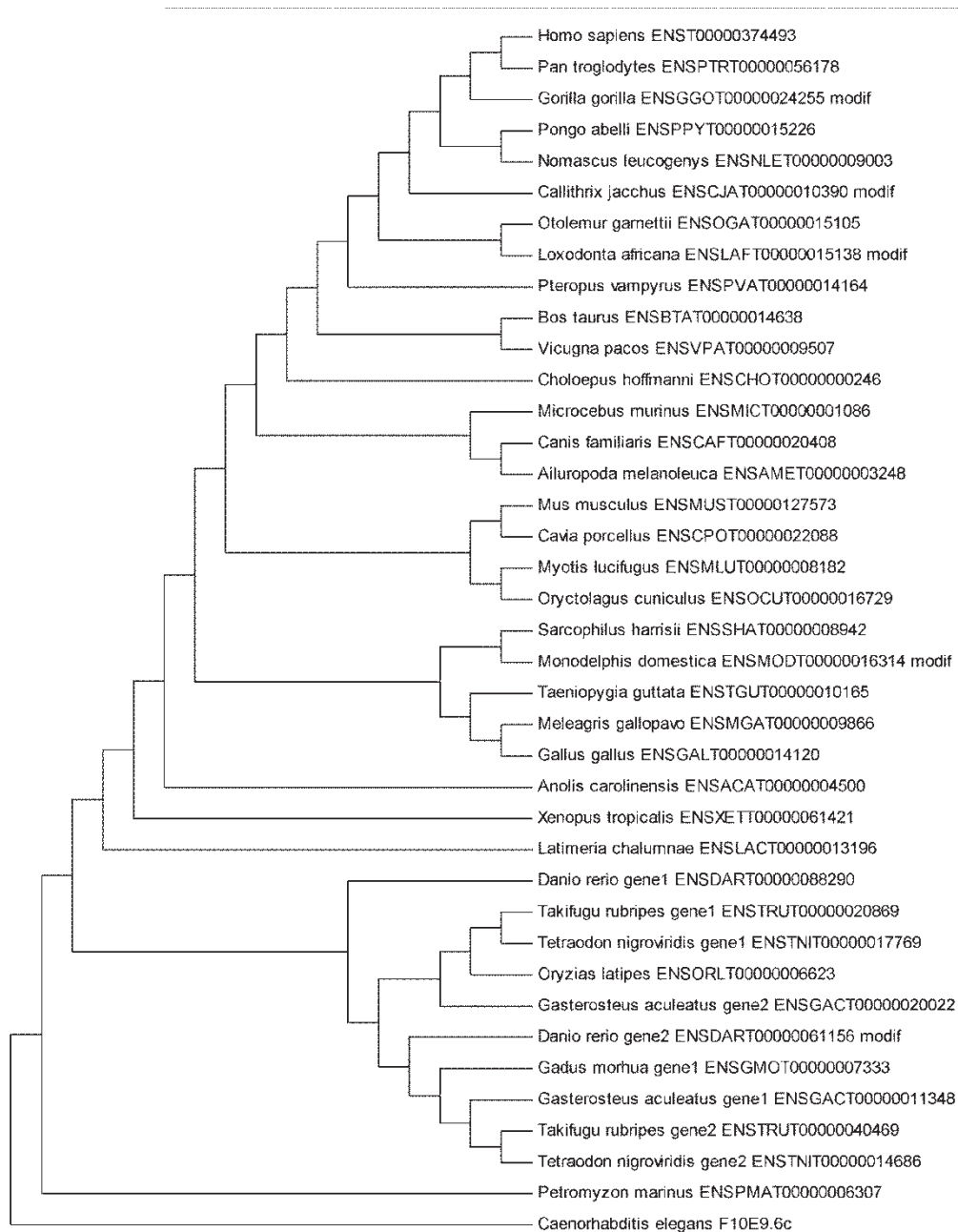


Figura 19 - Árvore filogenética do gene RAPH1 elaborada com seqüências de aminoácidos de 39 seqüências, utilizando o método de Máxima Parcimônia (MP)

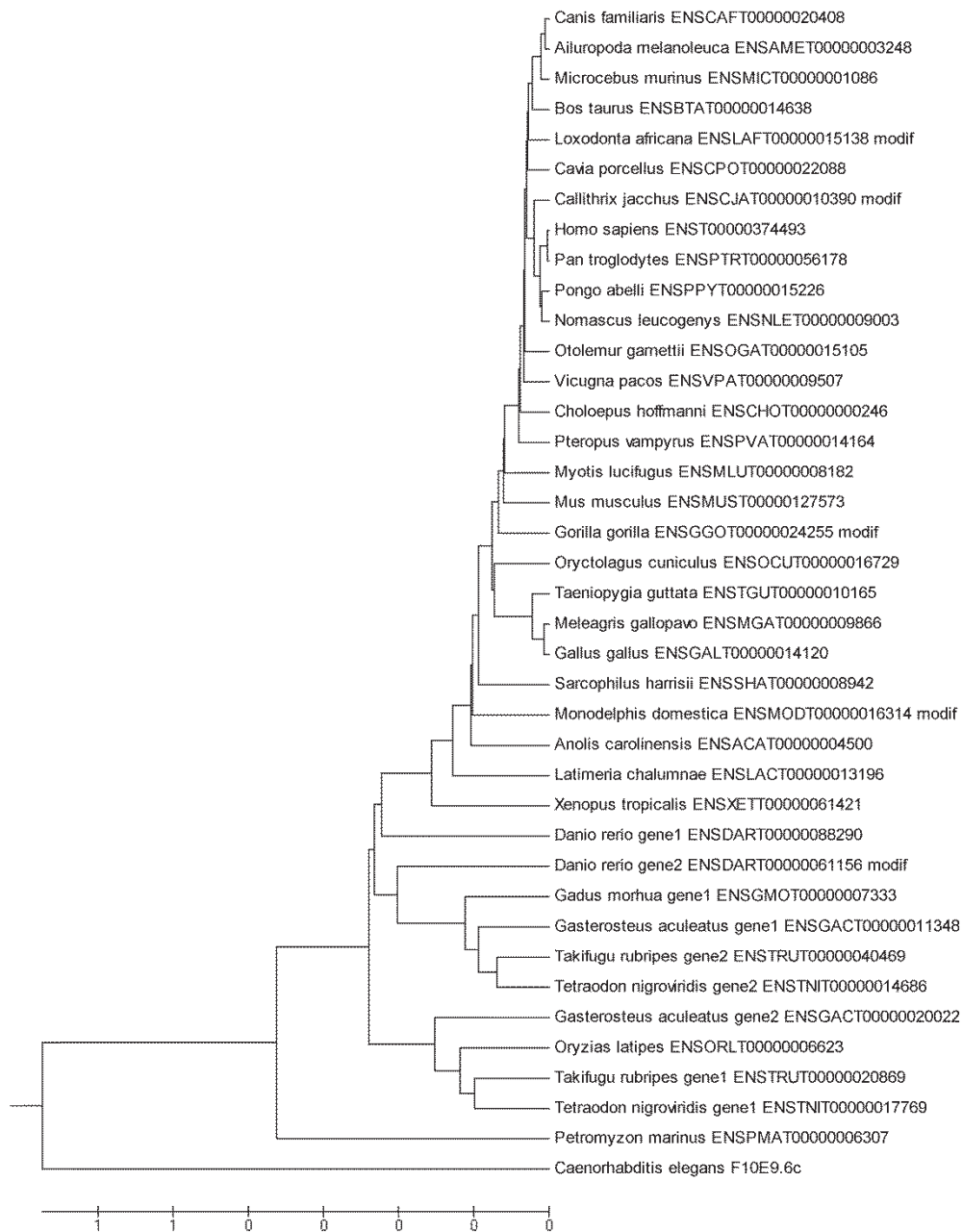


Figura 20 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos de 39 sequências, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

APÊNDICE 3 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE cDNA DO GENE RAPH1

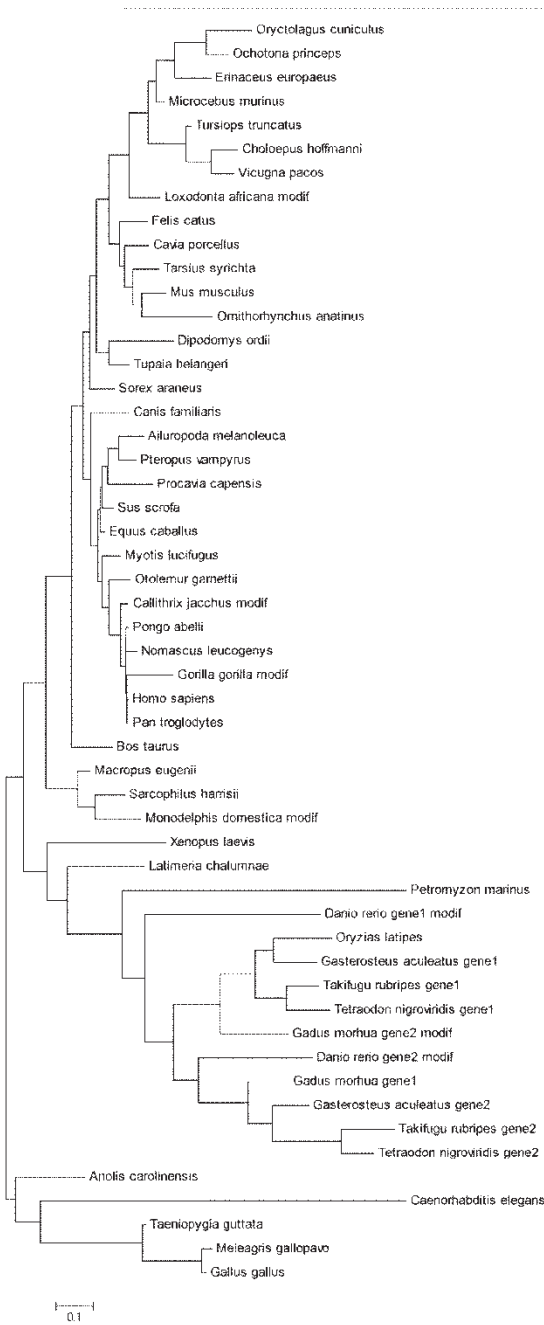


Figura 21 - Árvore filogenética do gene RAPH1 elaborada com sequências de cDNA de 53 sequências, utilizando o método de Máxima Verossimilhança (ML)

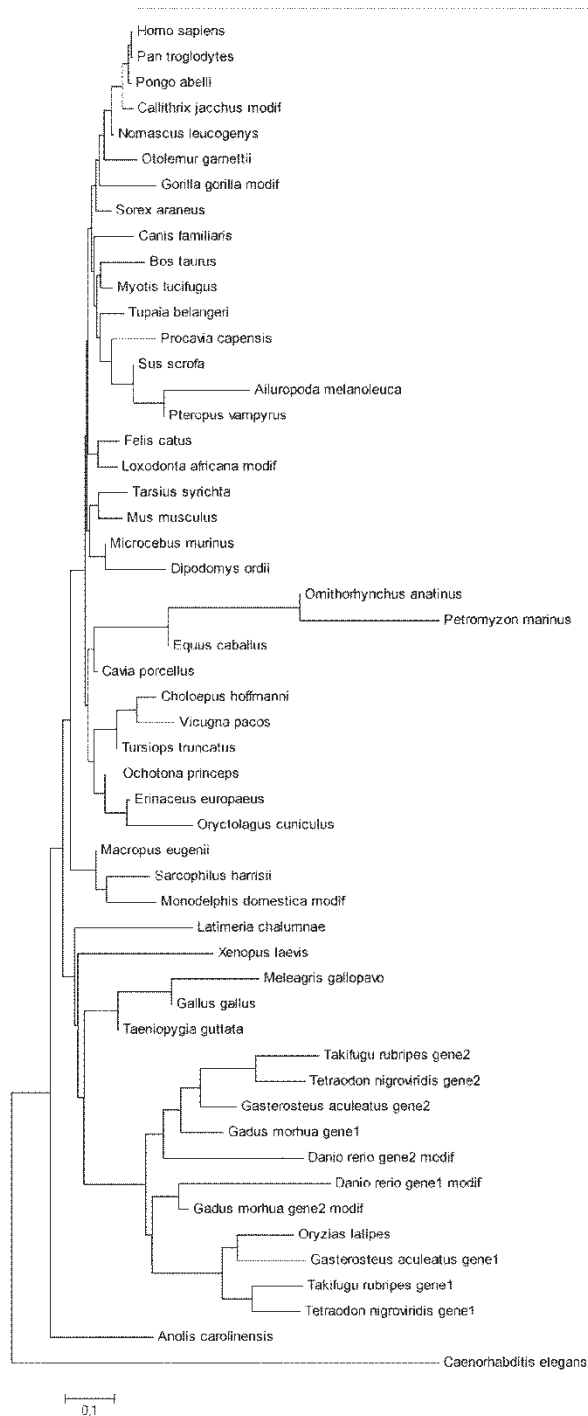


Figura 22 - Árvore filogenética do gene RAPH1 elaborada com sequências de cDNA de 53 sequências, utilizando o método de Neighbor Joining (NJ)

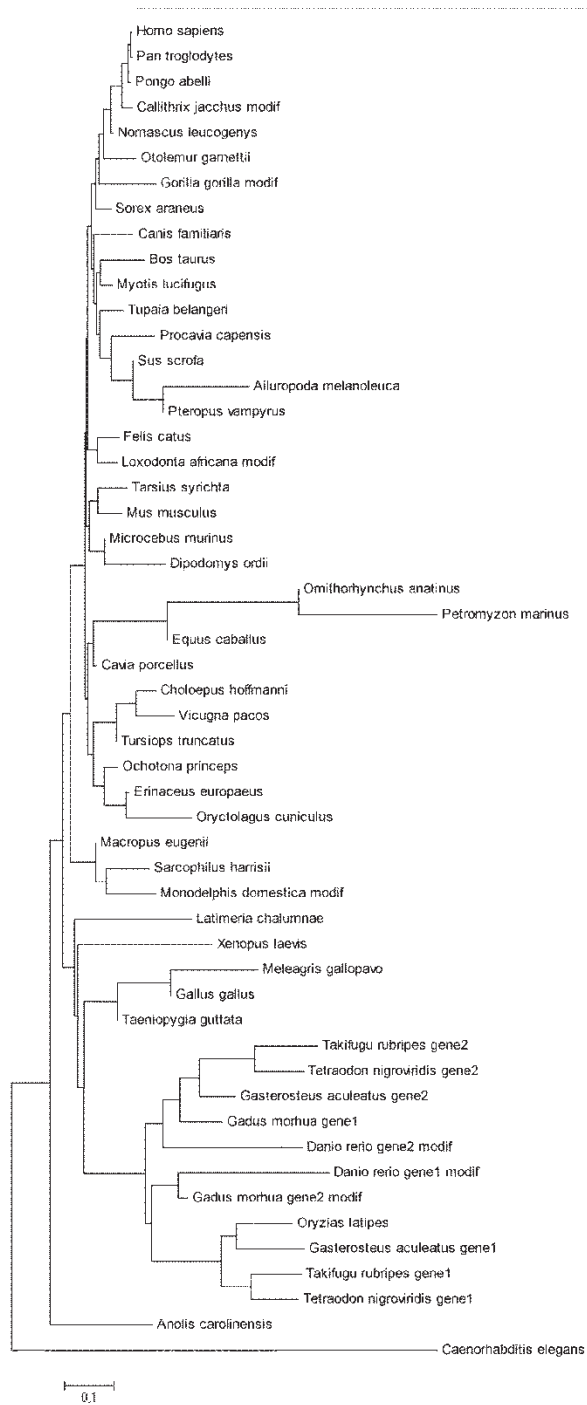


Figura 23 - Árvore filogenética do gene RAPH1 elaborada com seqüências de cDNA de 53 seqüências, utilizando o método de Evolução Mínima (ME)



Figura 24 - Árvore filogenética do gene RAPH1 elaborada com sequências de cDNA de 53 sequências, utilizando o método de Máxima Parcimônia (MP).



Figura 25 - Árvore filogenética do gene RAPH1 elaborada com seqüências de cDNA de 53 seqüências, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

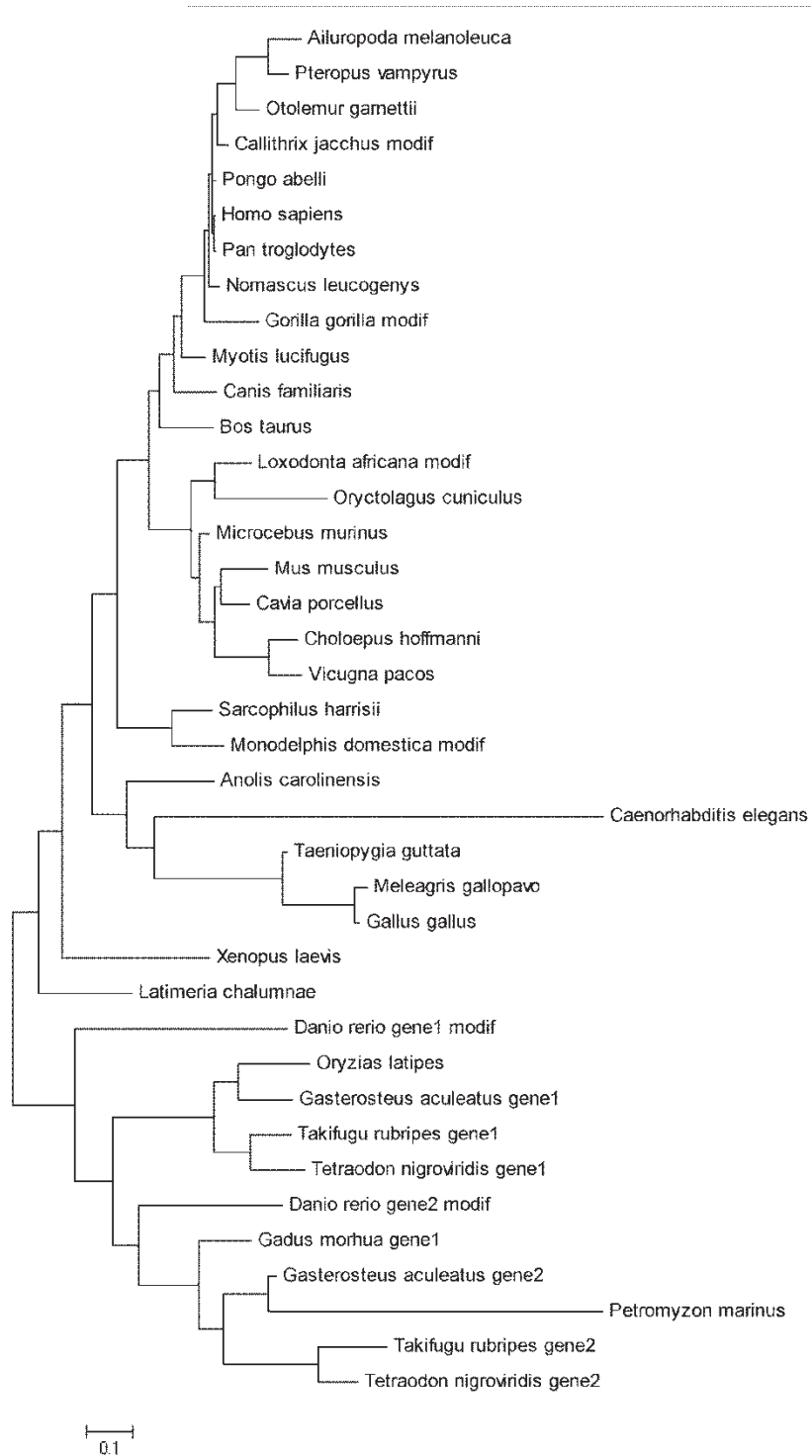


Figura 26 - Árvore filogenética do gene RAPH1 elaborada com seqüências de cDNA de 39 seqüências, utilizando o método de Máxima Verossimilhança (ML)

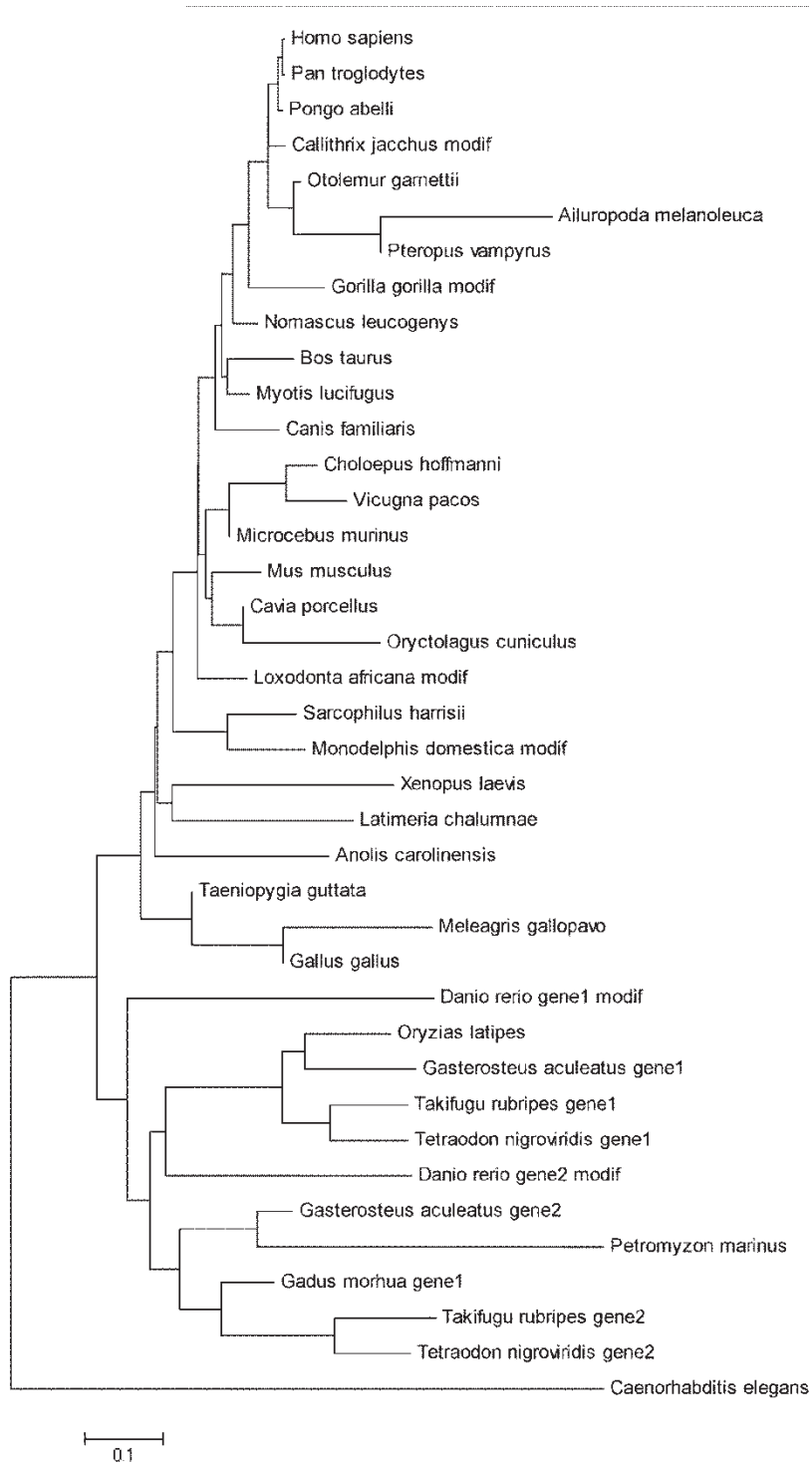


Figura 27 - Árvore filogenética do gene RAPH1 elaborada com sequências de cDNA de 39 sequências, utilizando o método de Neighbor Joining (NJ)

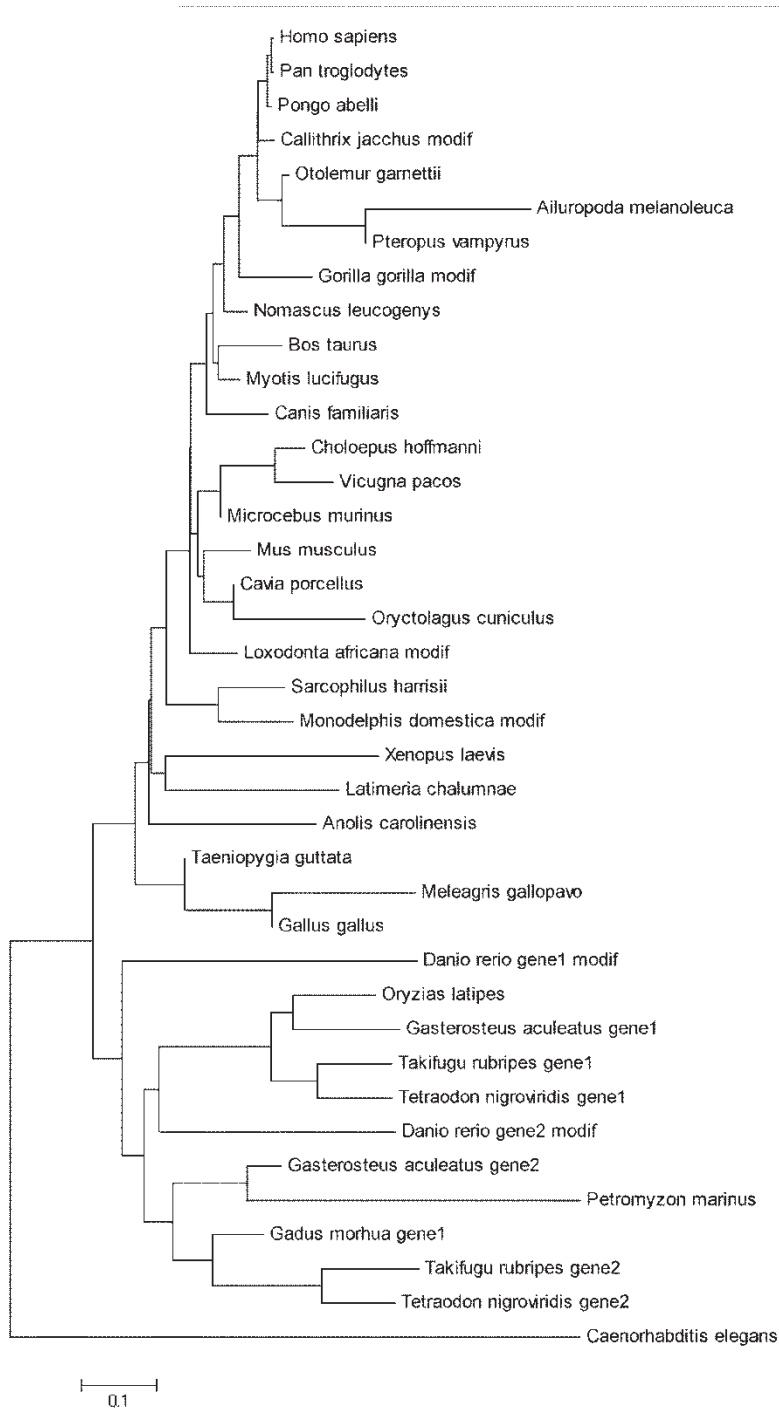


Figura 28 - Árvore filogenética do gene RAPH1 elaborada com seqüências de cDNA de 39 seqüências, utilizando o método de Evolução Mínima (ME)

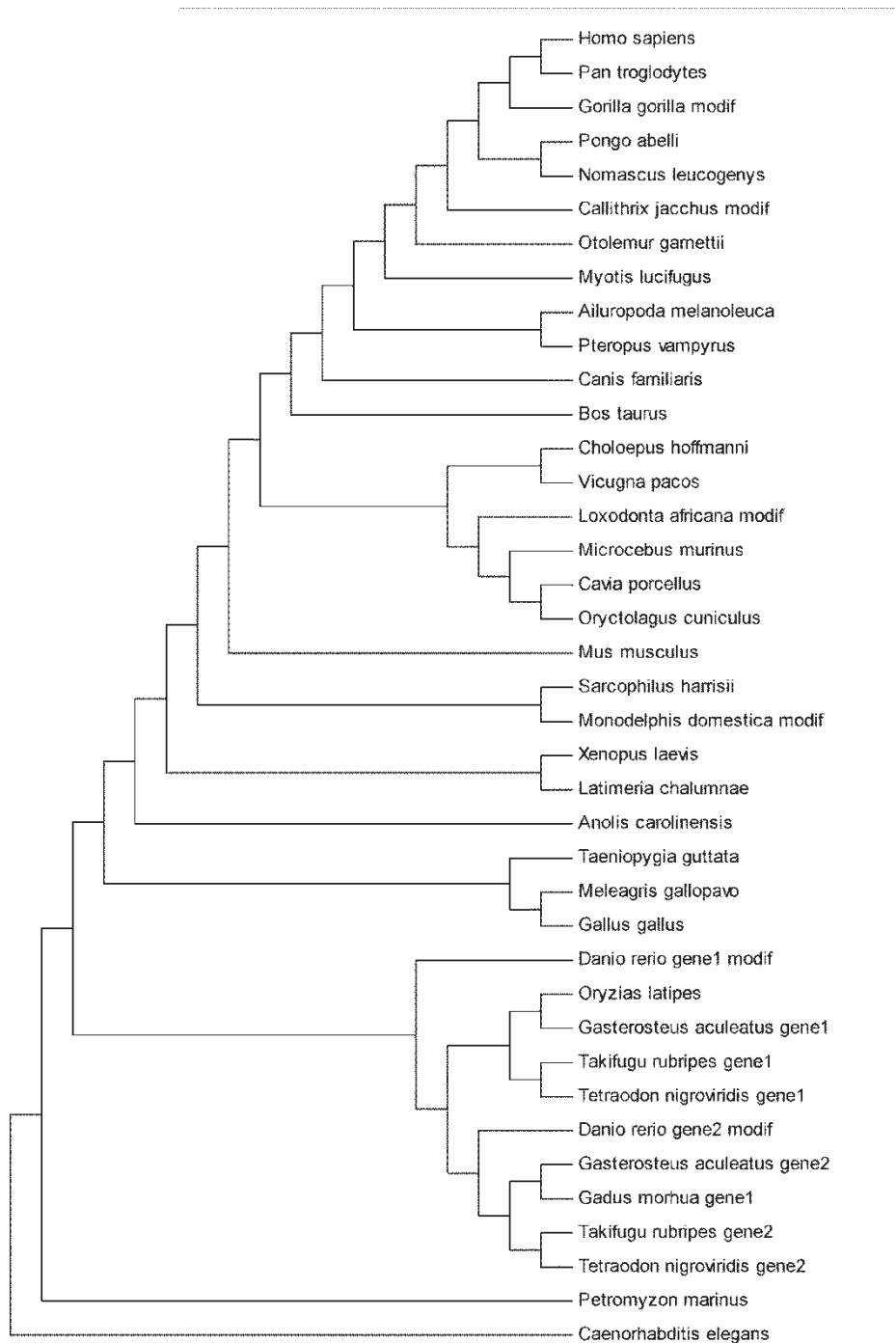


Figura 29 - Árvore filogenética do gene RAPH1 elaborada com sequências de cDNA de 39 sequências, utilizando o método de Máxima Parcimônia (MP)

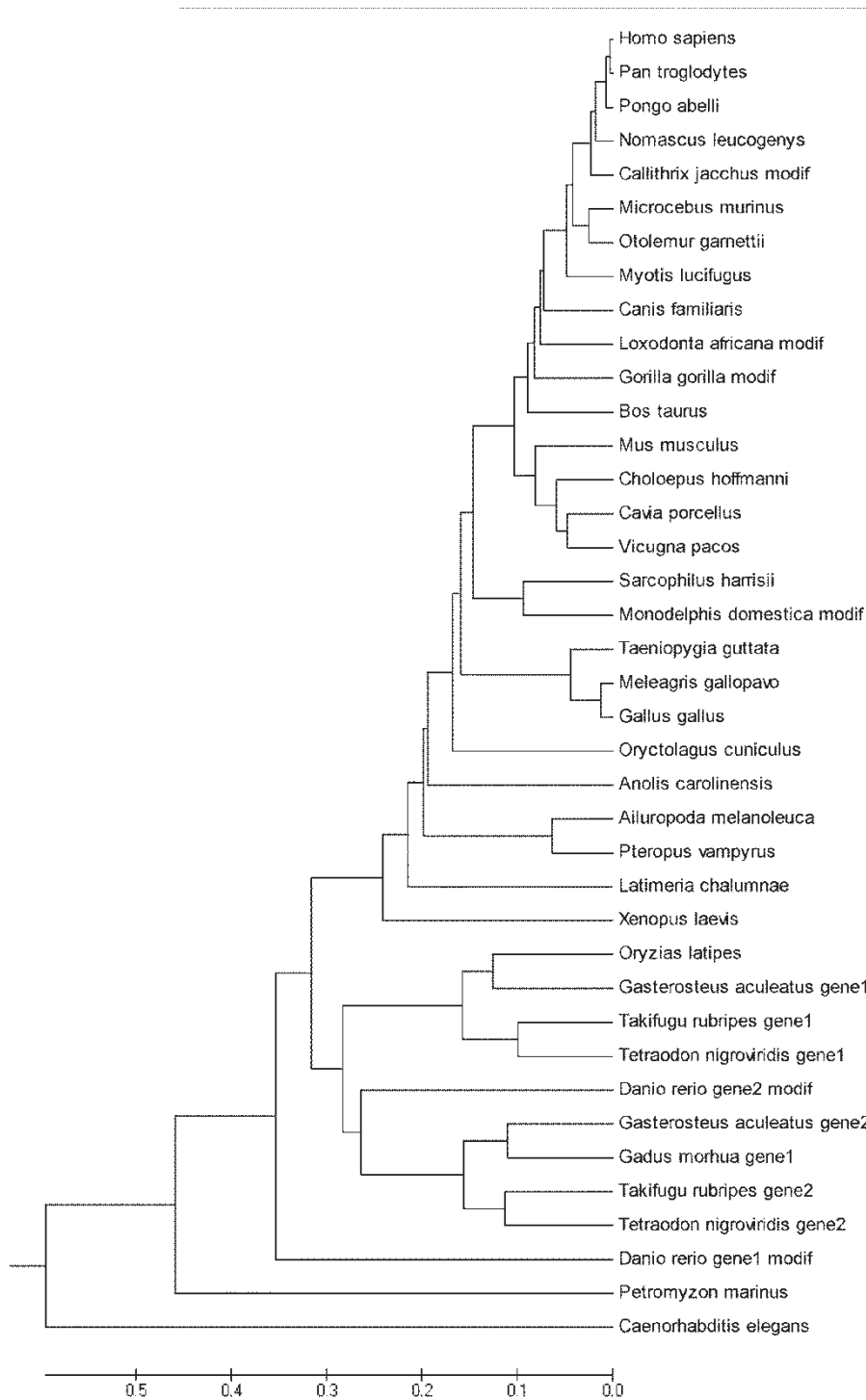


Figura 30 - Árvore filogenética do gene RAPH1 elaborada com sequências de cDNA de 39 sequências, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

APÊNDICE 4 – ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO RA DO GENE RAPH1



Figura 31 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio RA, utilizando o método de Máxima Verossimilhança (ML)

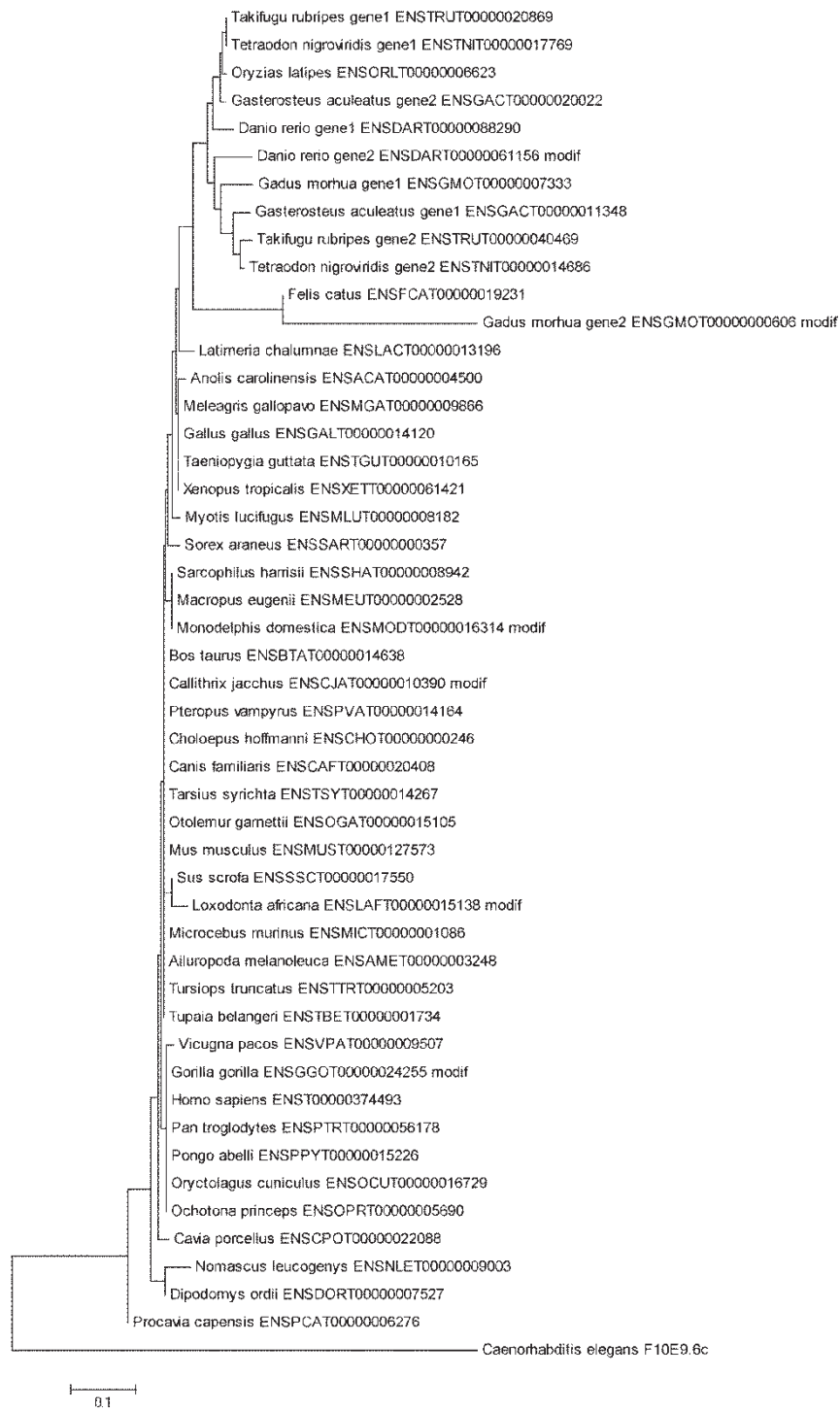


Figura 32 - Árvore filogenética do gene RAPH1 elaborada com seqüências de aminoácidos da região do domínio RA, utilizando o método de Neighbor Joining (NJ)

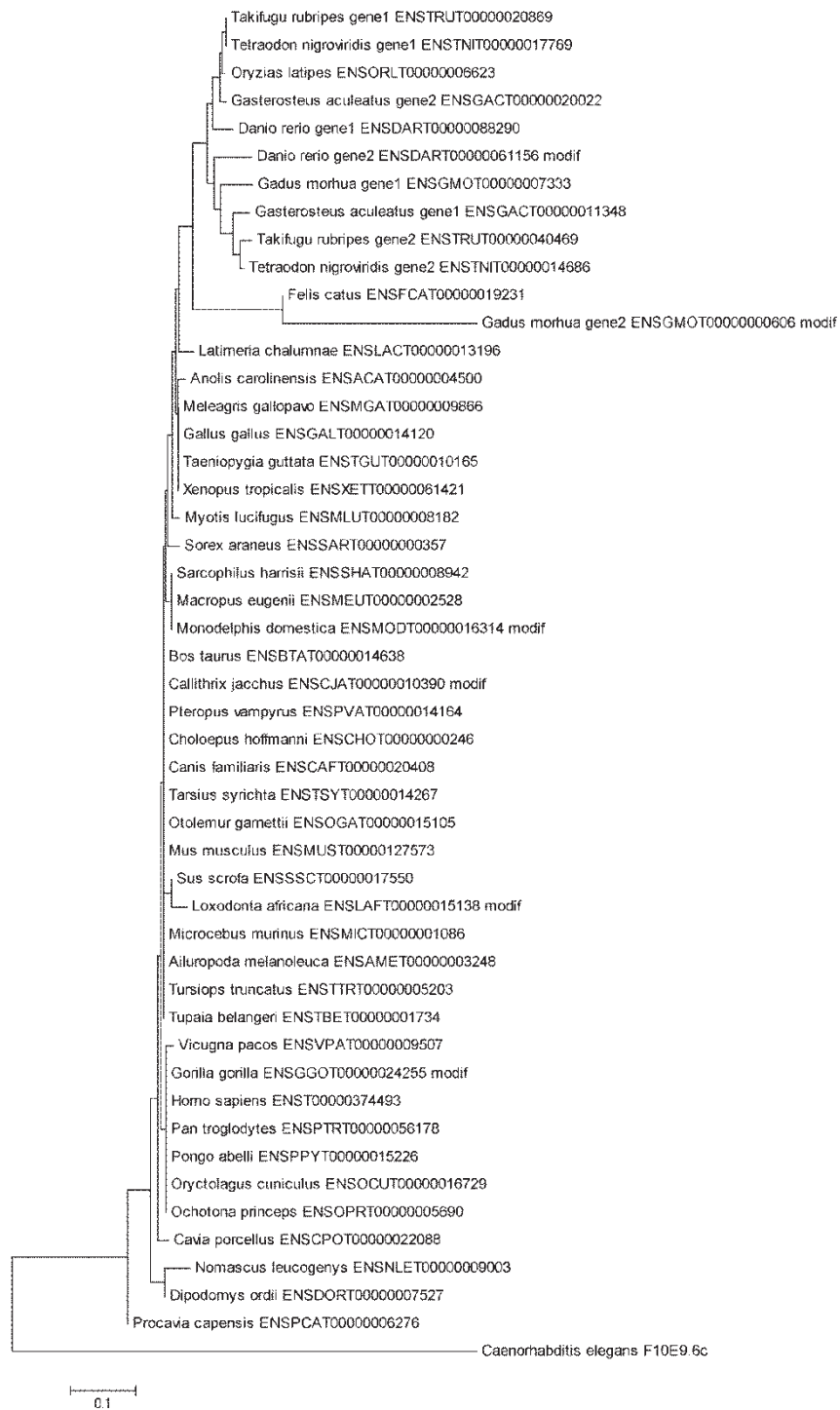


Figura 33 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio RA, utilizando o método de Evolução Mínima (ME)

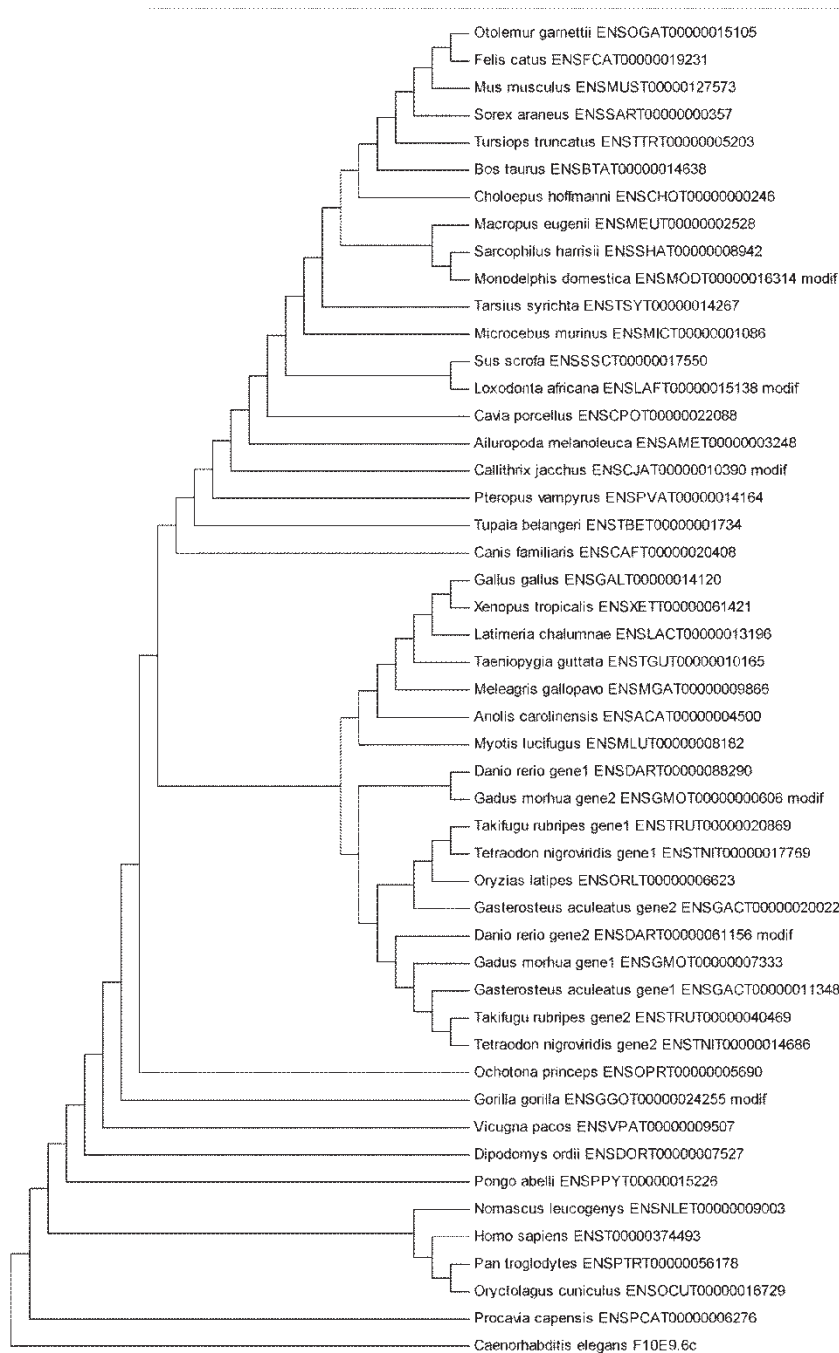


Figura 34 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio RA, utilizando o método de Máxima Parcimônia (MP)

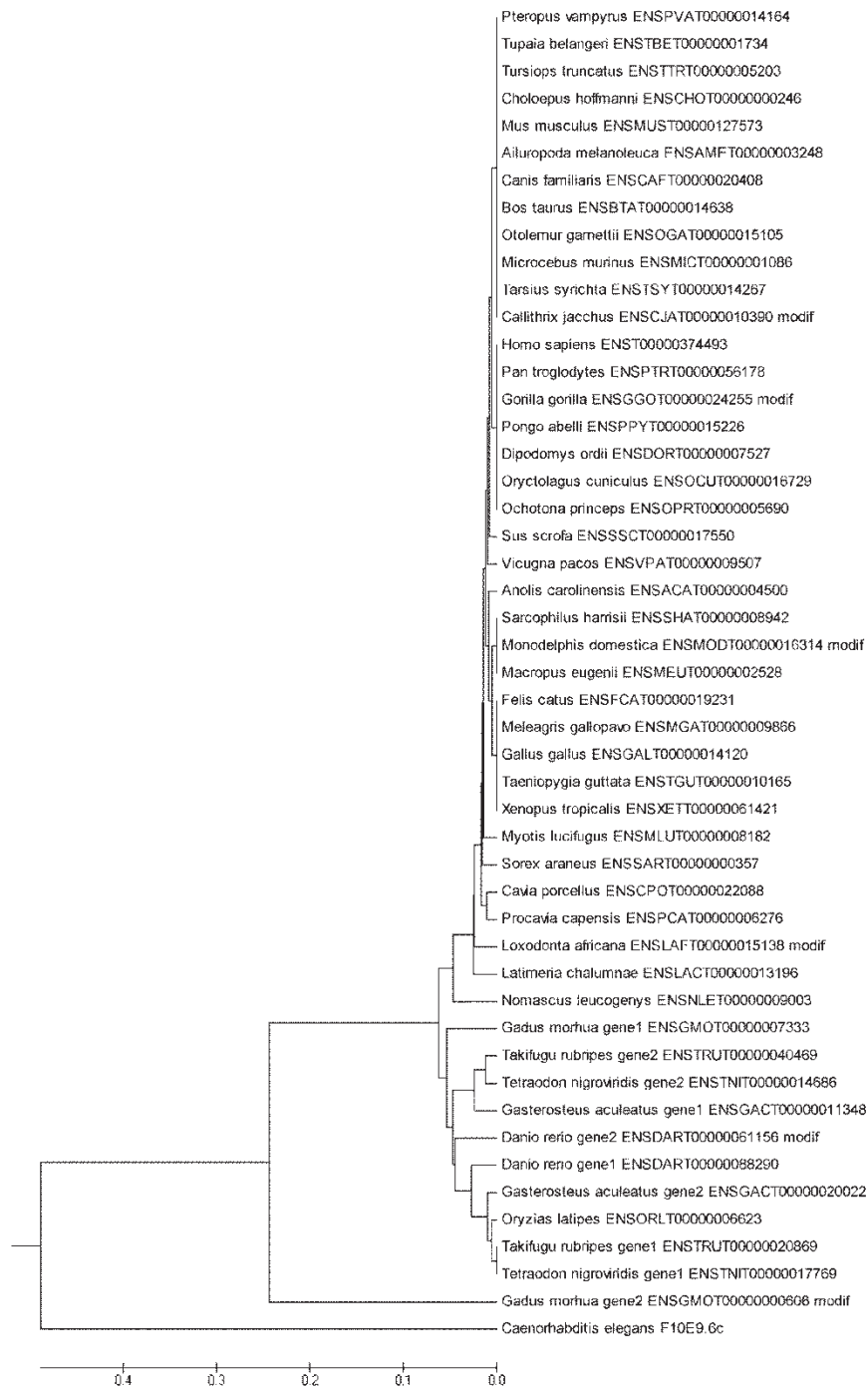


Figura 35 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio RA, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

APÊNDICE 5 – ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO RA DO GENE RAPH1



Figura 36 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PH, utilizando o método de Máxima Verossimilhança (ML)



Figura 37 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PH, utilizando o método de Neighbor Joining (NJ)



Figura 38 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PH, utilizando o método de Evolução Mínima (ME)

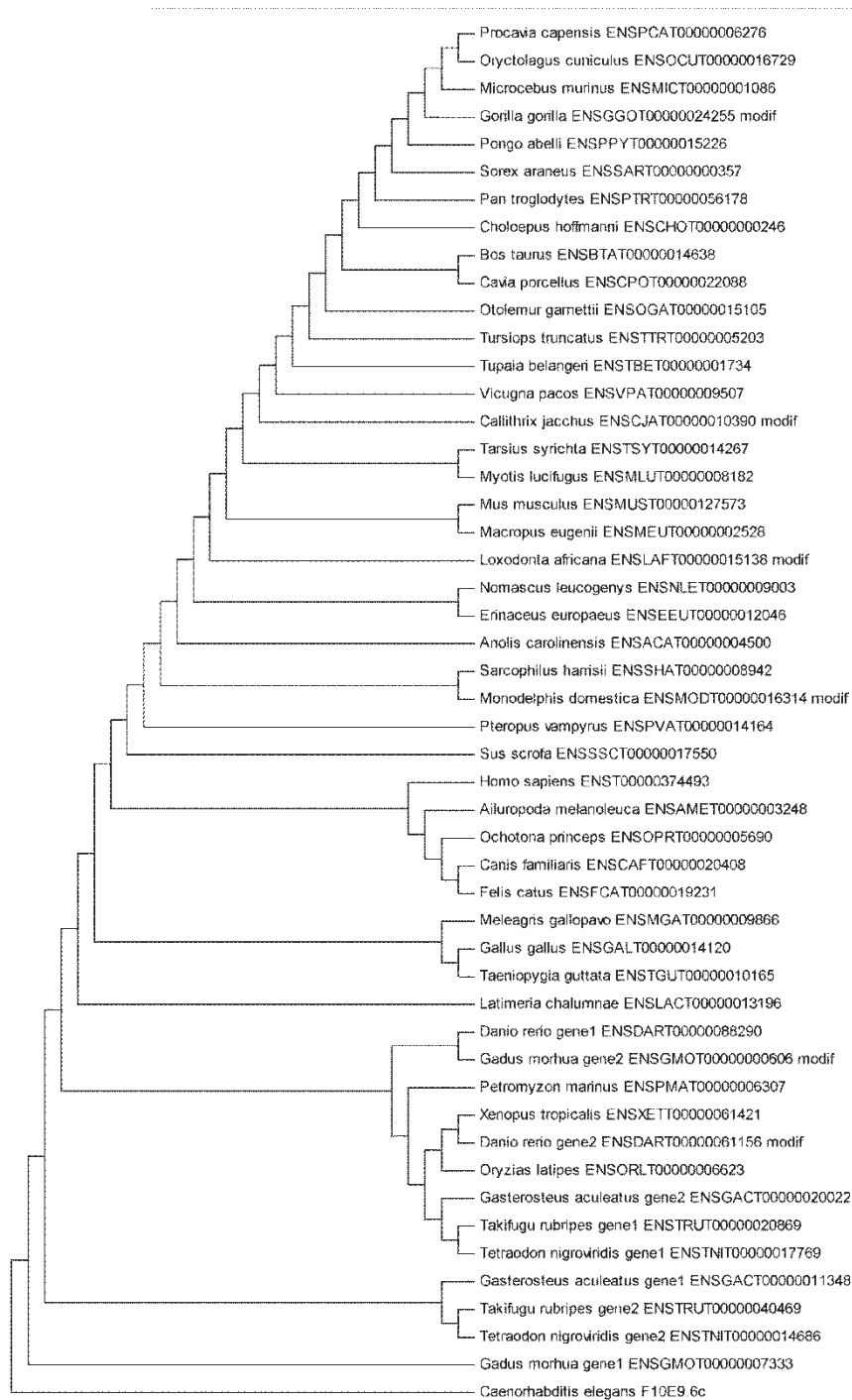


Figura 39 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PH, utilizando o método de Máxima Parcimônia (MP)

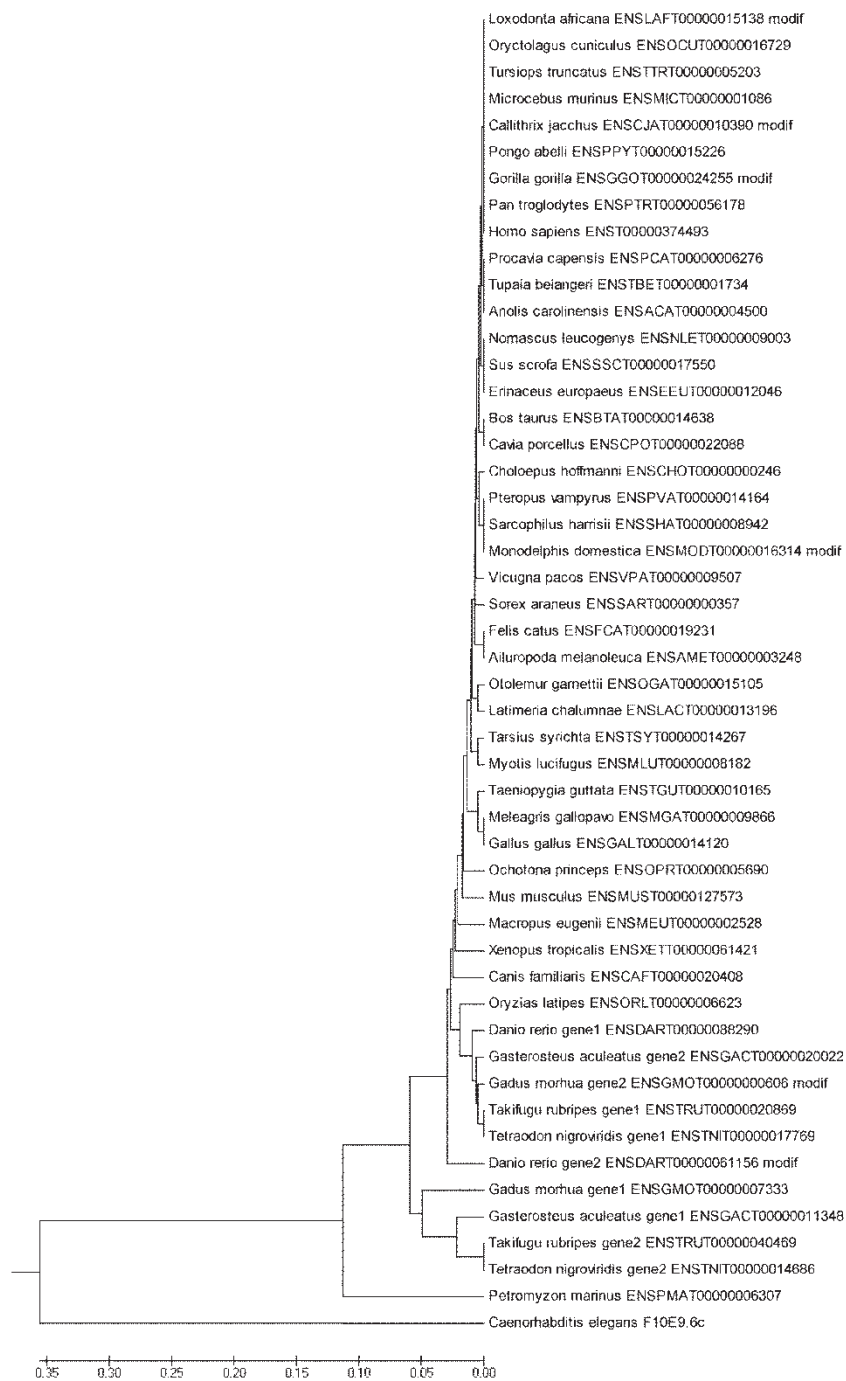


Figura 40 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PH, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

APÊNDICE 6 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO GM DO GENE RAPH1

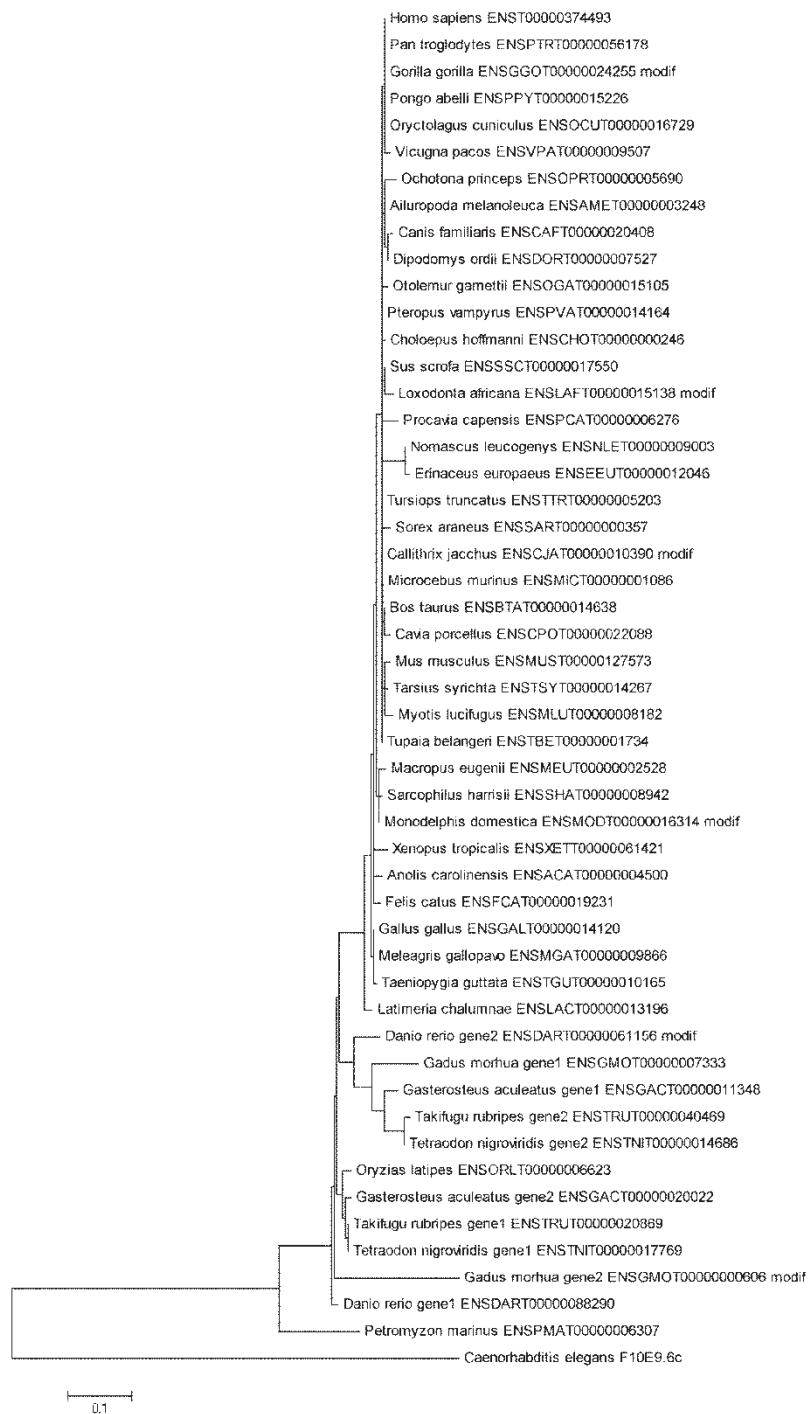


Figura 41 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio GM, utilizando o método de Máxima Verossimilhança (ML)



Figura 42 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio GM, utilizando o método de Neighbor Joining (NJ)



Figura 43 - Árvore filogenética do gene RAPH1 elaborada com seqüências de aminoácidos da região do domínio GM, utilizando o método de Evolução Mínima (ME)

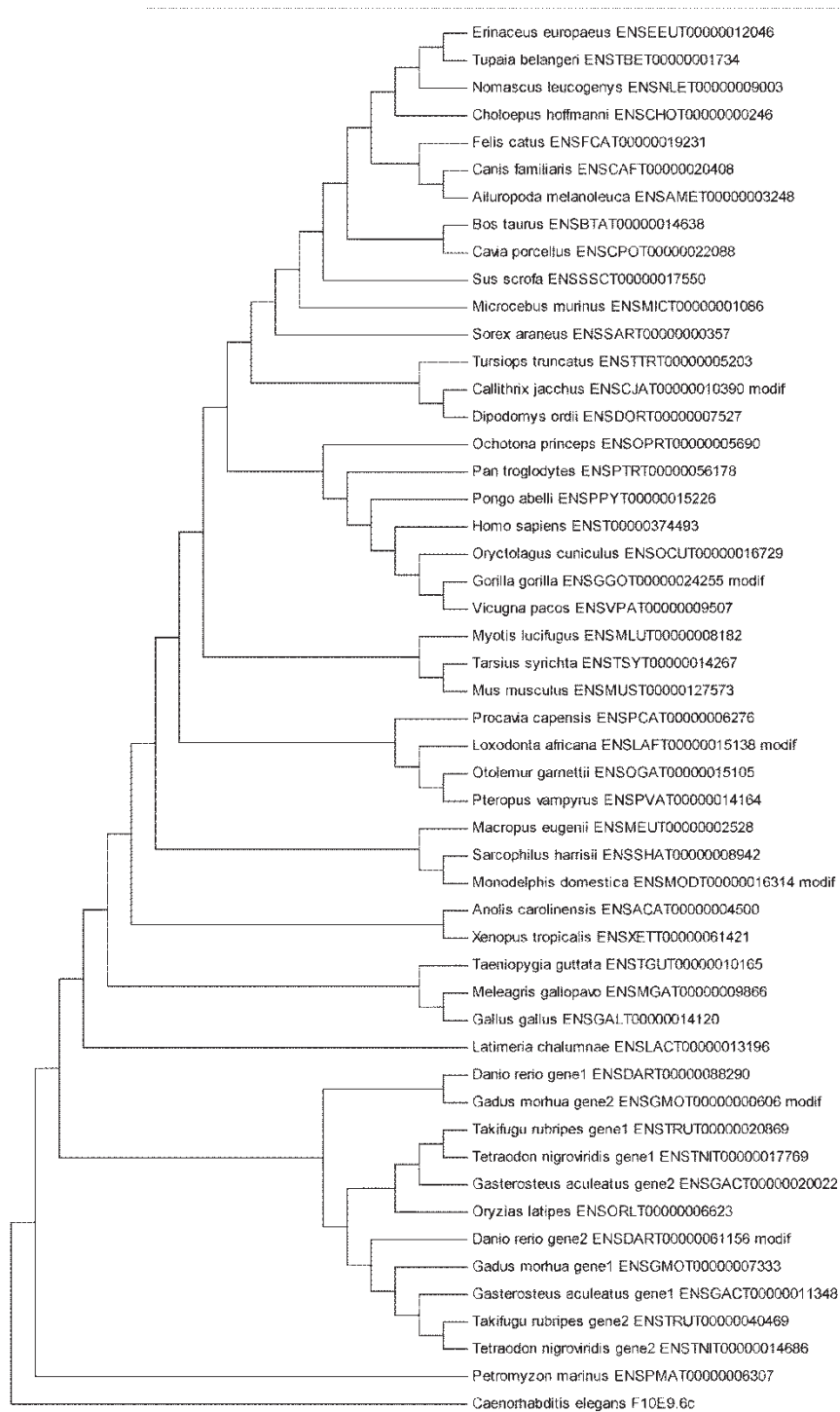


Figura 44 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio GM, utilizando o método de Máxima Parcimônia (MP)

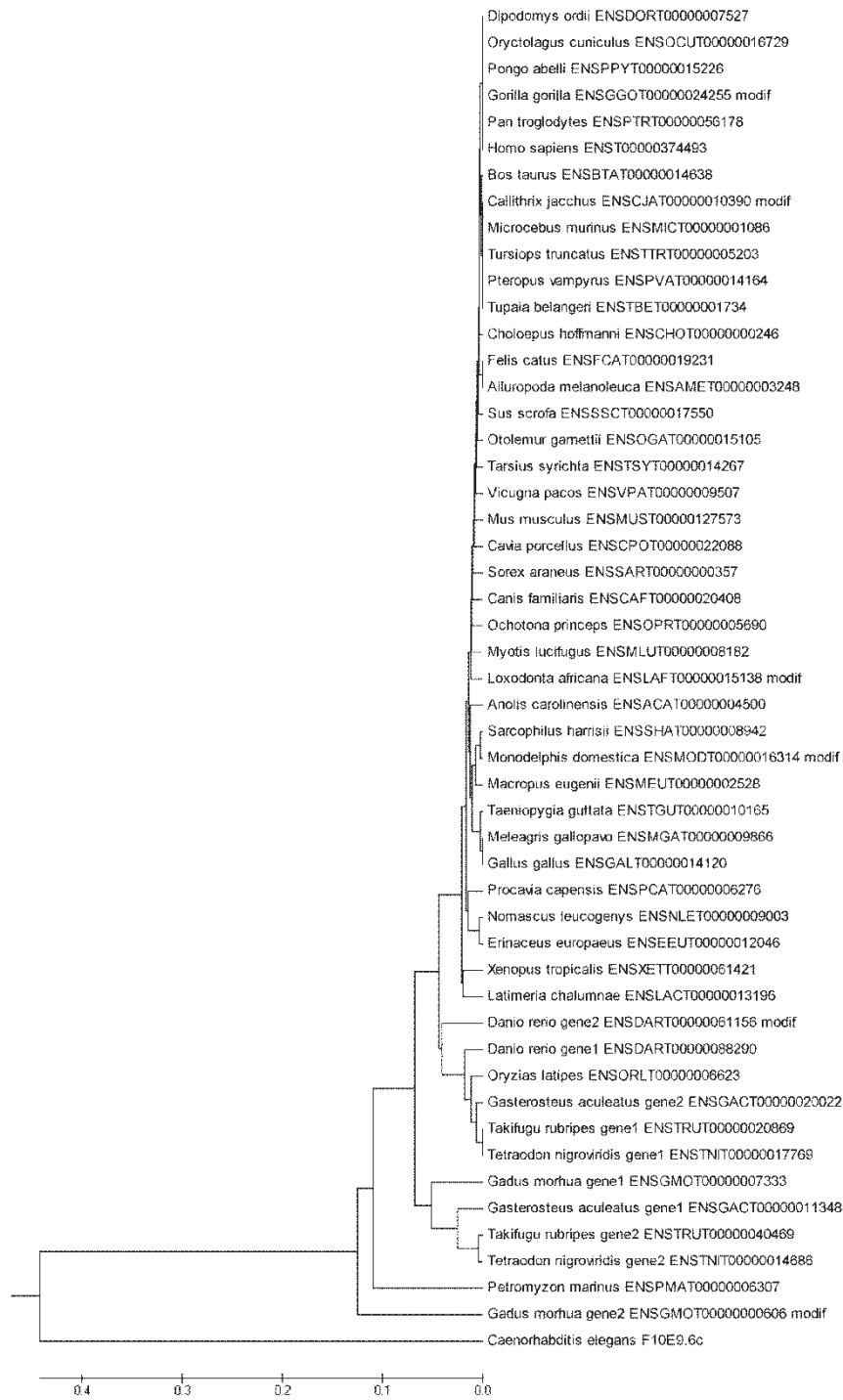


Figura 45 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio GM, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

APÊNDICE 7 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE
AMINOÁCIDOS DO DOMÍNIO PBS DO GENE RAPH1

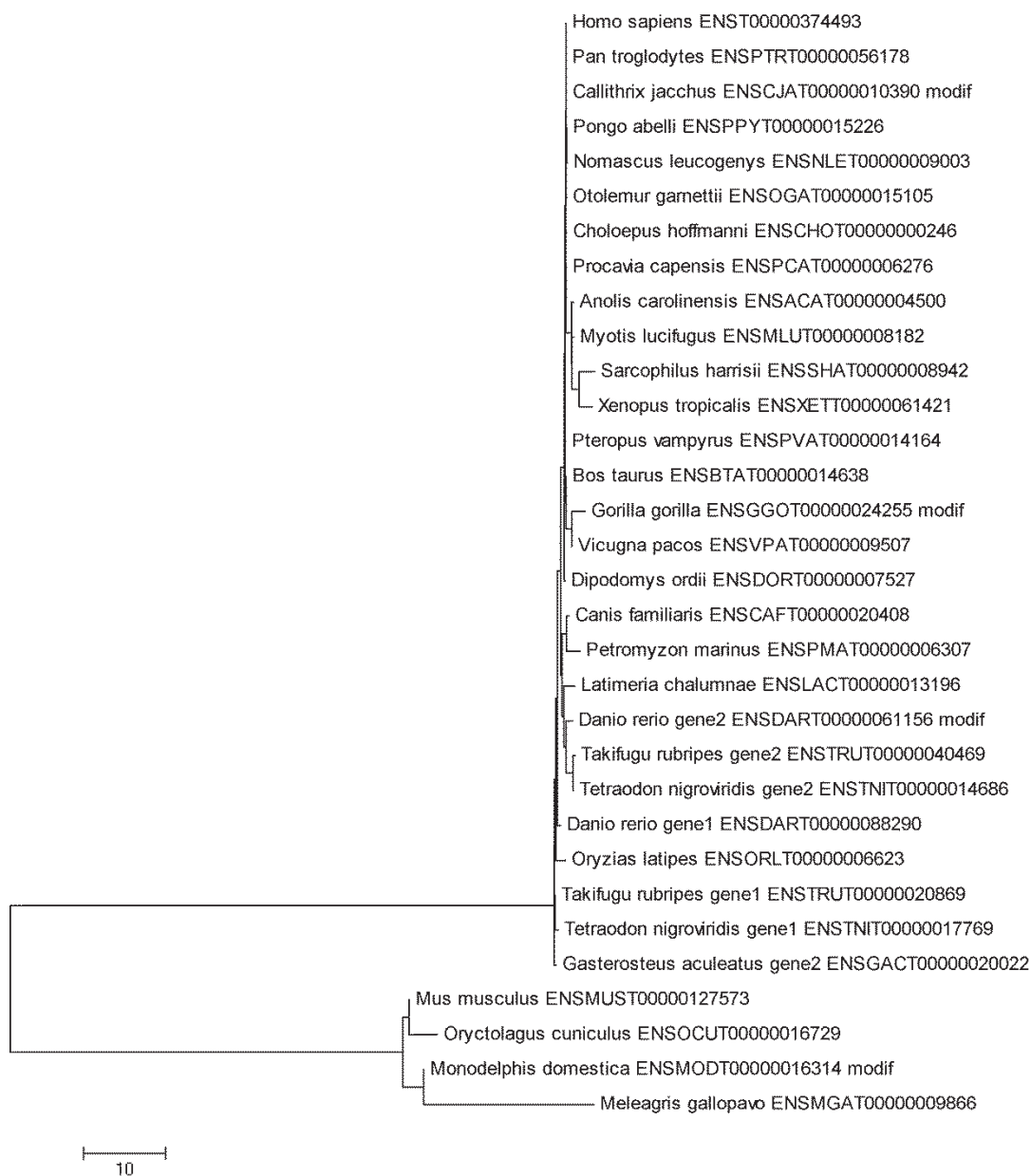


Figura 46 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PBS, utilizando o método de Máxima Verossimilhança (ML)

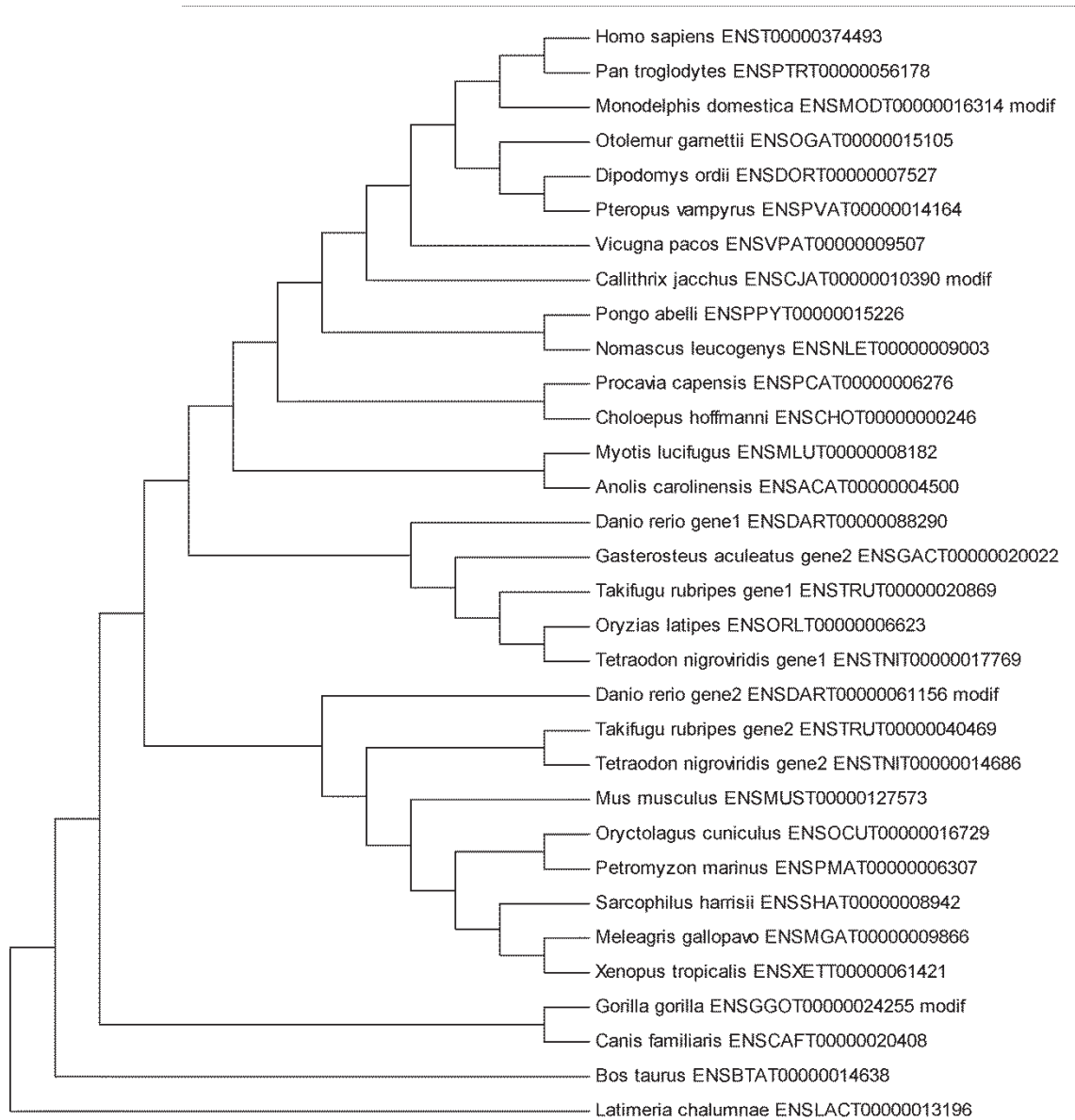


Figura 47 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio PBS, utilizando o método de Máxima Parcimônia (MP)

APÊNDICE 8 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE
AMINOÁCIDOS DO DOMÍNIO EVH1 1 DO GENE RAPH1

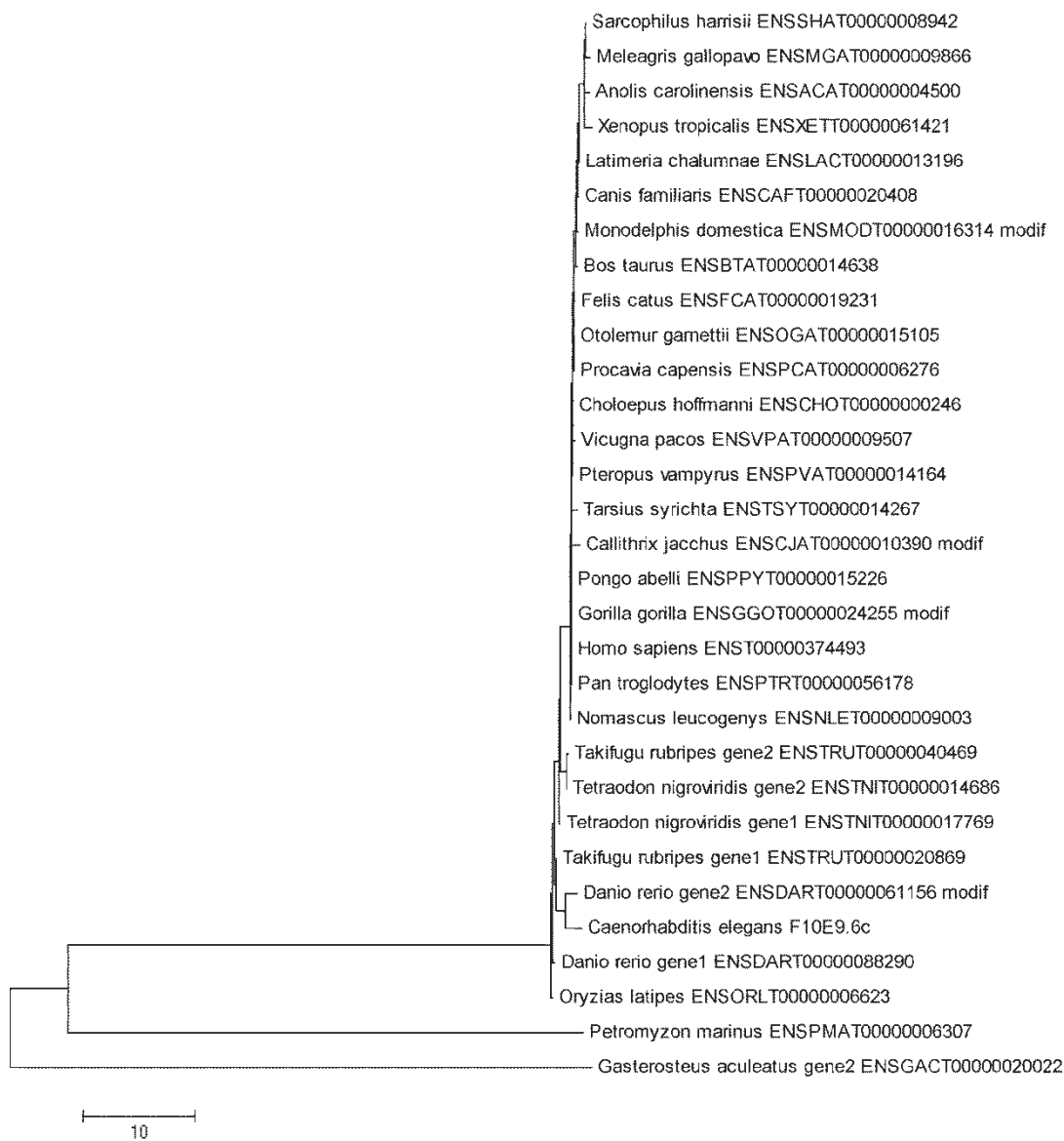


Figura 48 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 1, utilizando o método de Máxima Verossimilhança (ML)

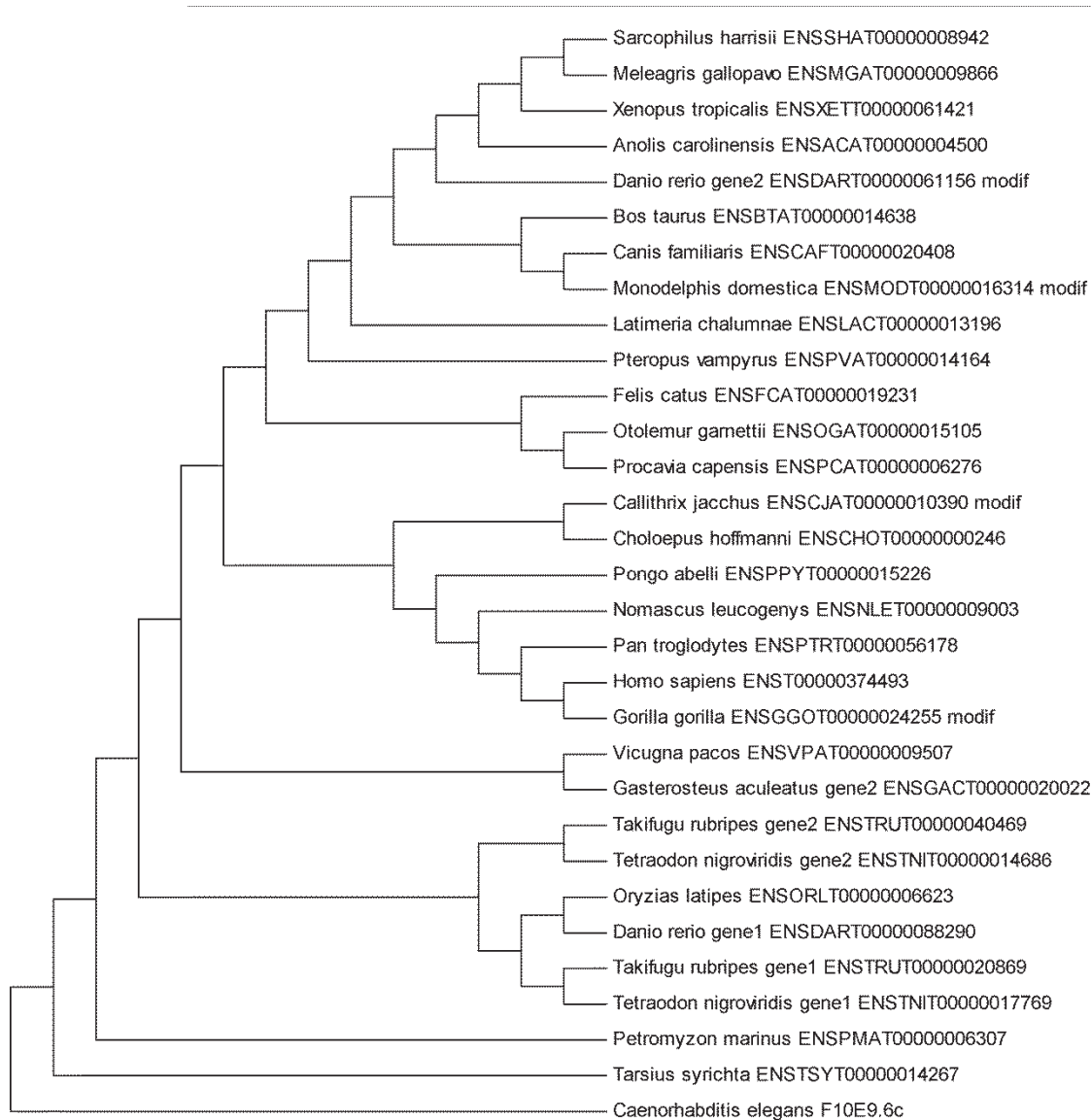


Figura 49 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 1, utilizando o método de Máxima Parcimônia (MP)

APÊNDICE 9 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE
AMINOÁCIDOS DO DOMÍNIO EVH1 2 DO GENE RAPH1

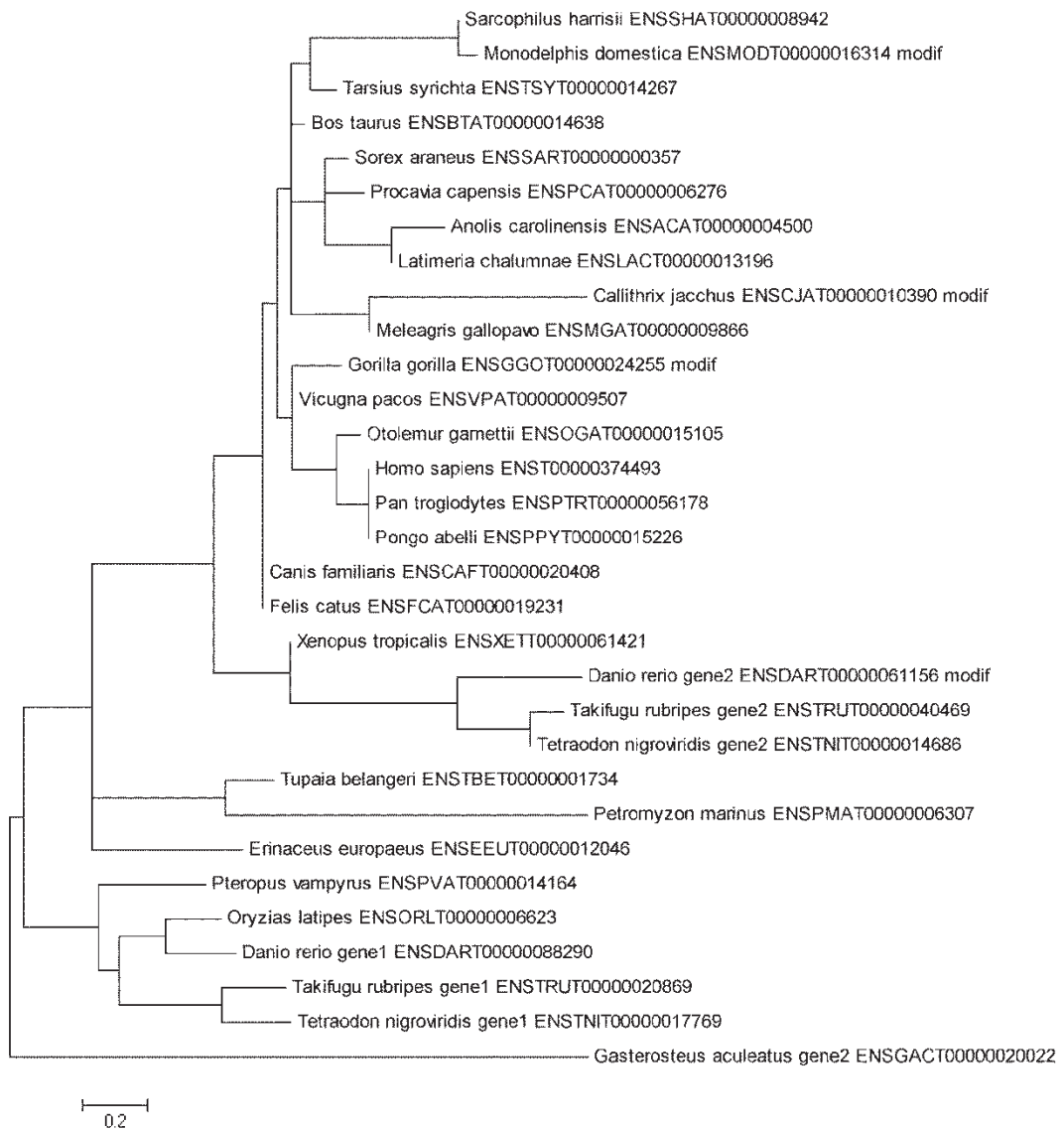


Figura 50 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 2, utilizando o método de Máxima Verossimilhança (ML)

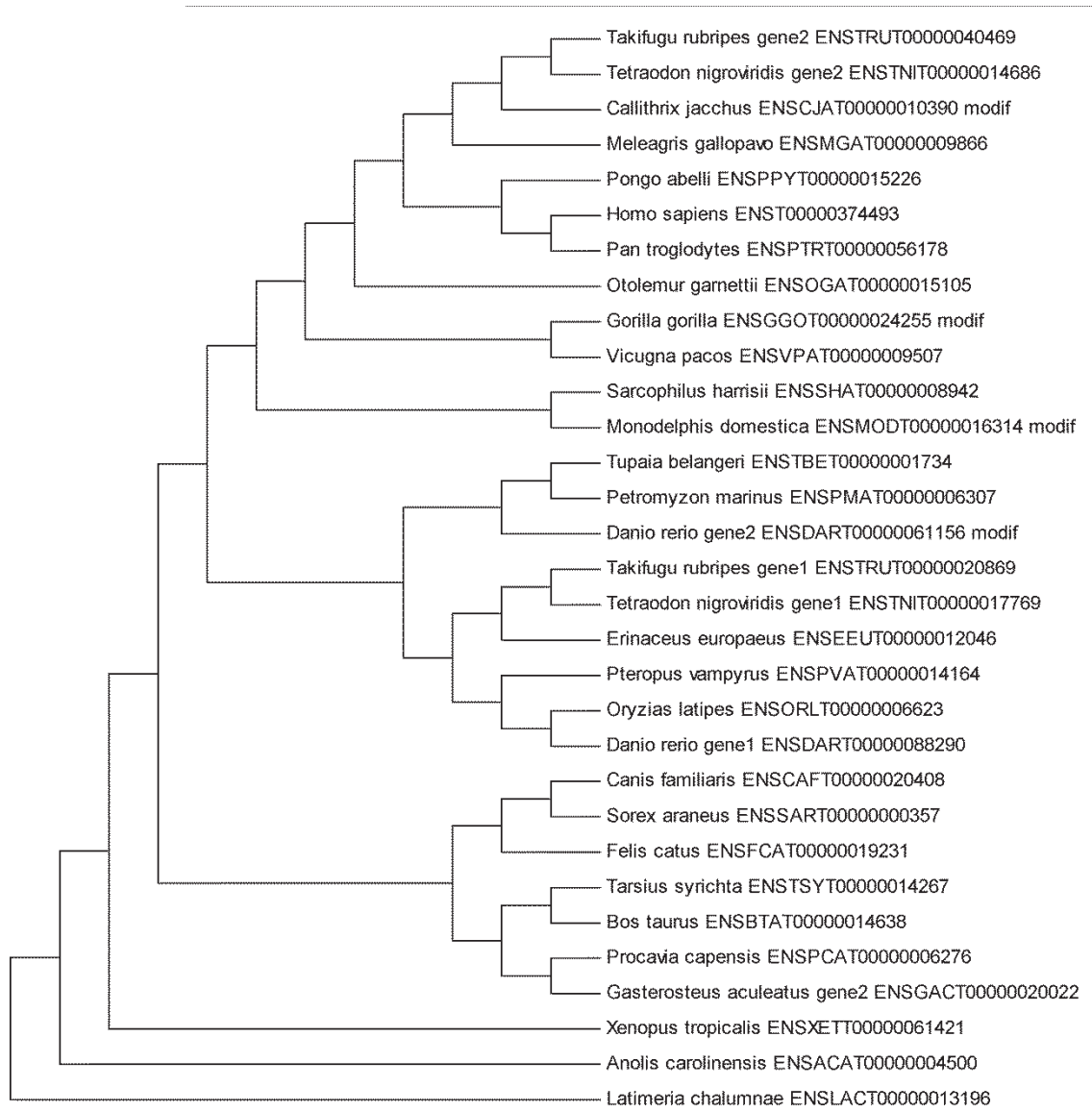


Figura 51 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 2, utilizando o método de Máxima Parcimônia (MP)

APÊNDICE 10 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO EVH1 3 DO GENE RAPH1

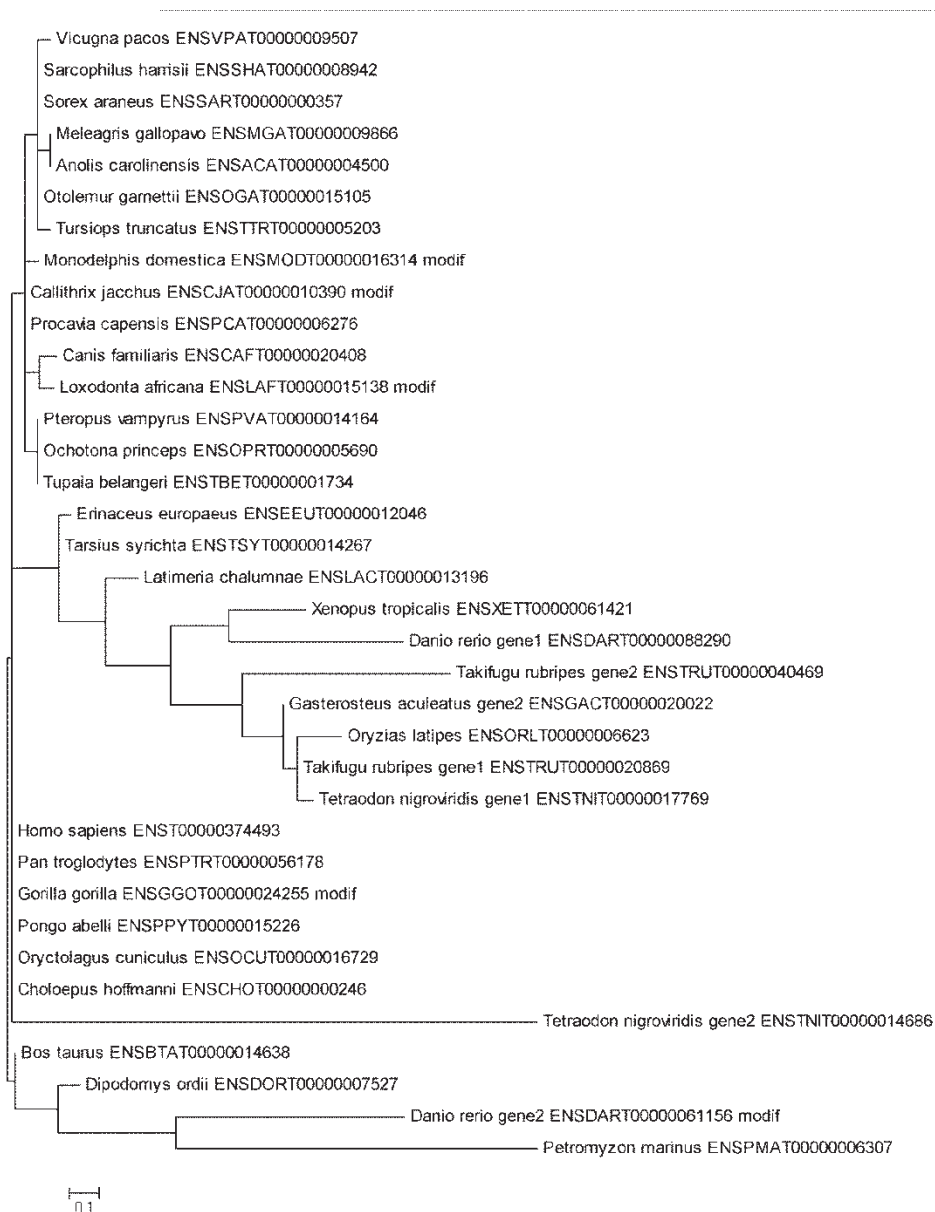


Figura 52 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 3, utilizando o método de Máxima Verossimilhança (ML)

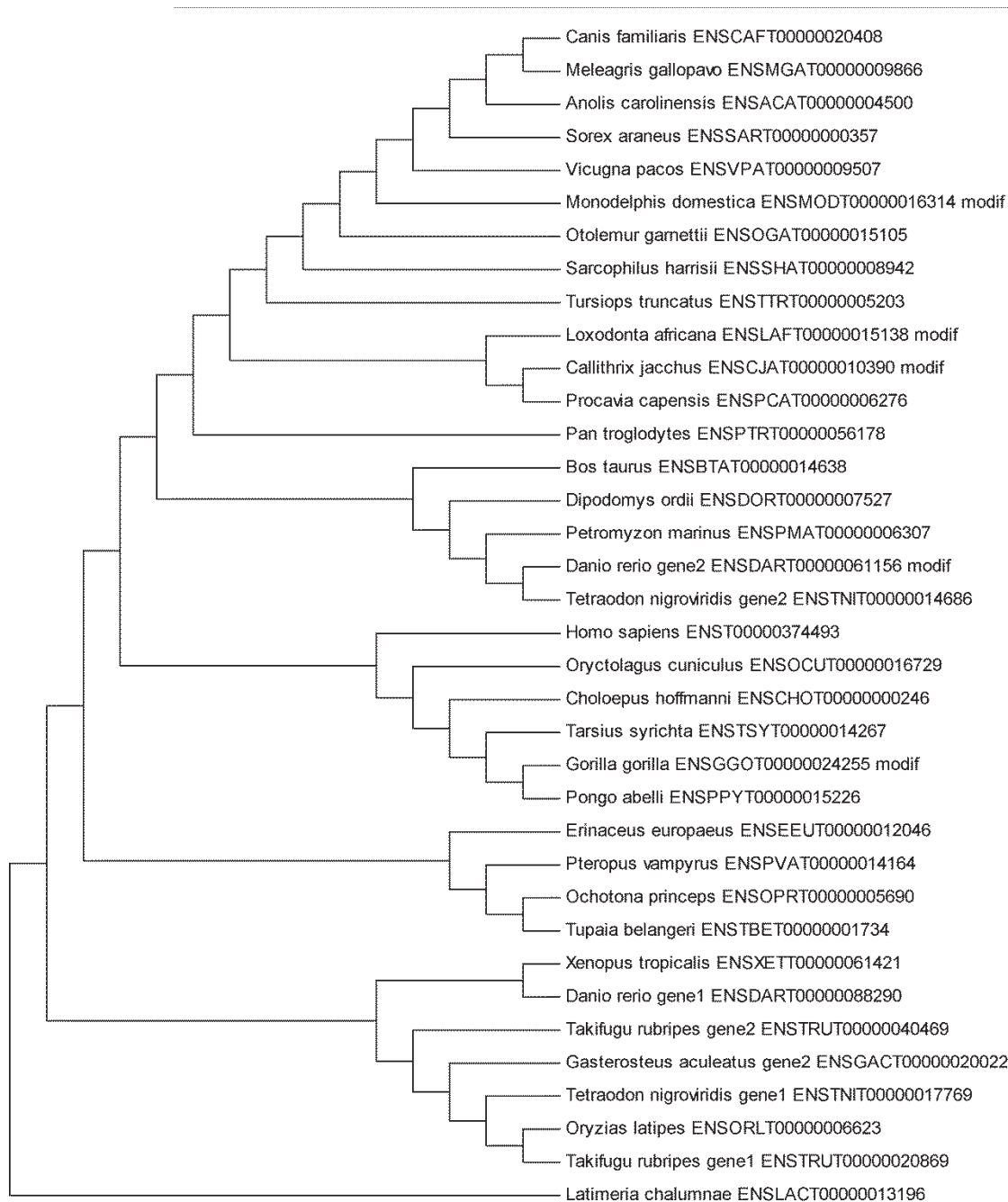


Figura 53 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 3, utilizando o método de Máxima Parcimônia (MP)

APÊNDICE 11 - ÁRVORES ELABORADAS COM SEQUÊNCIAS DE AMINOÁCIDOS DO DOMÍNIO EVH1 4 DO GENE RAPH1

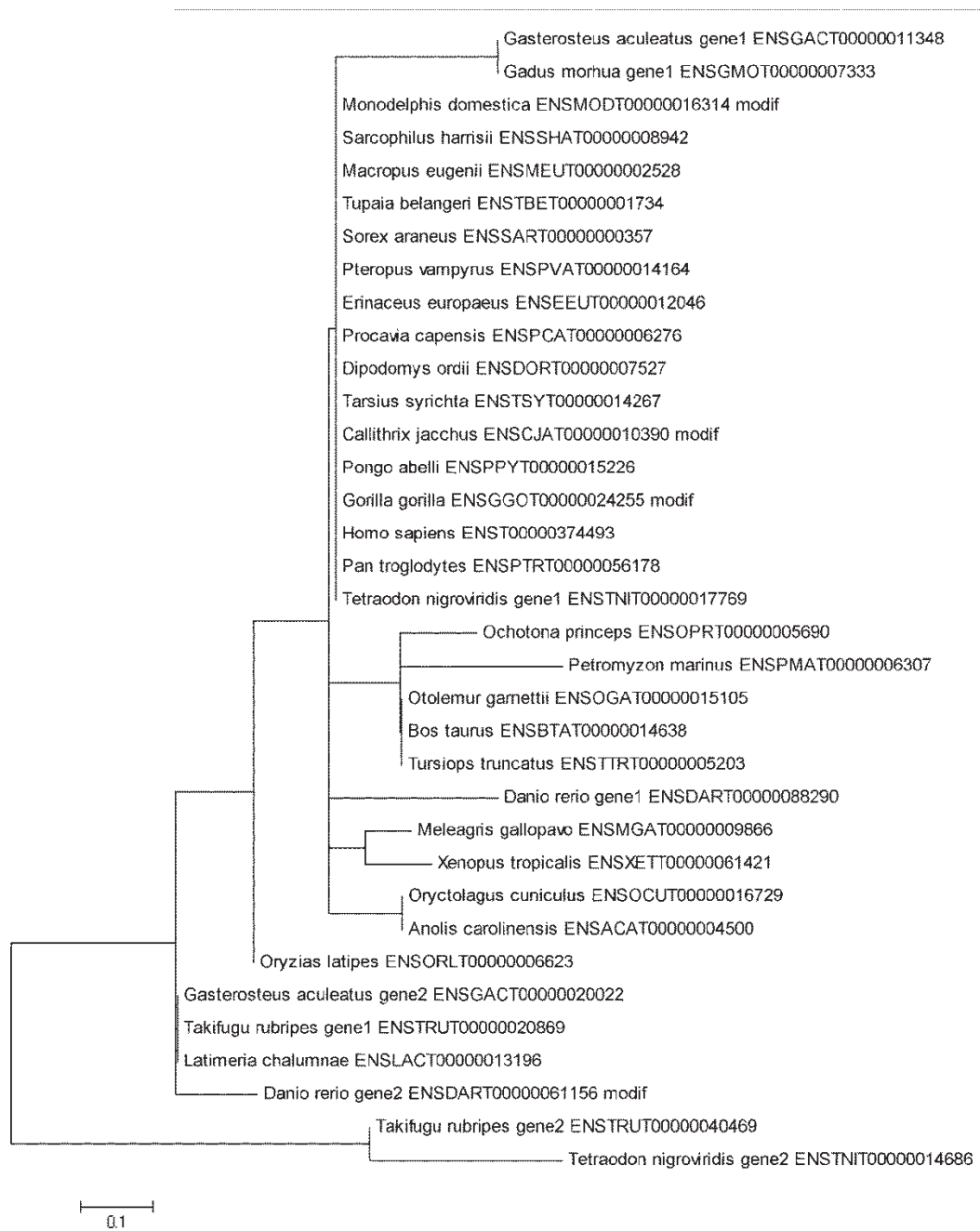


Figura 54 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 4, utilizando o método de Máxima Verossimilhança (ML)

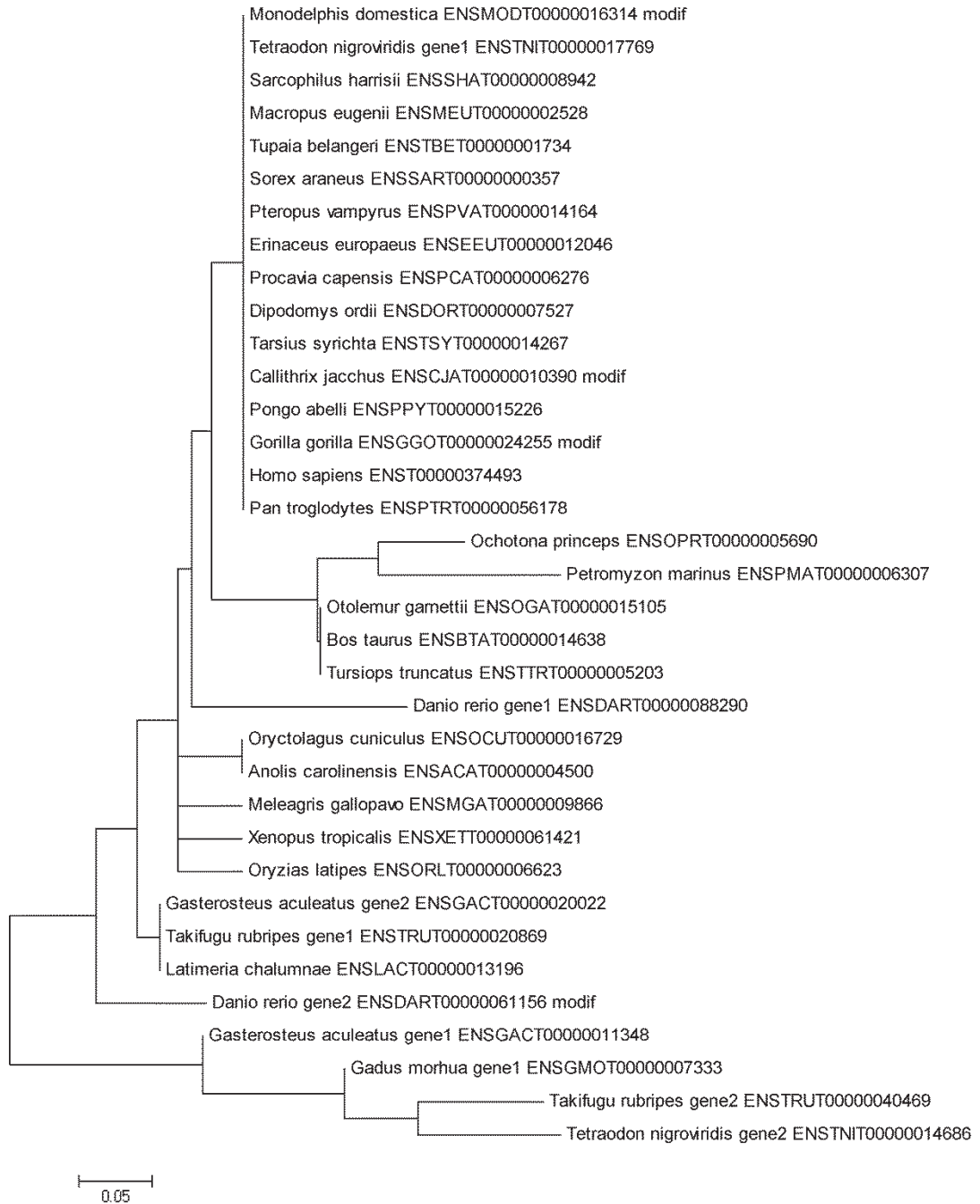


Figura 55 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 4, utilizando o método de Neighbor Joining (NJ)

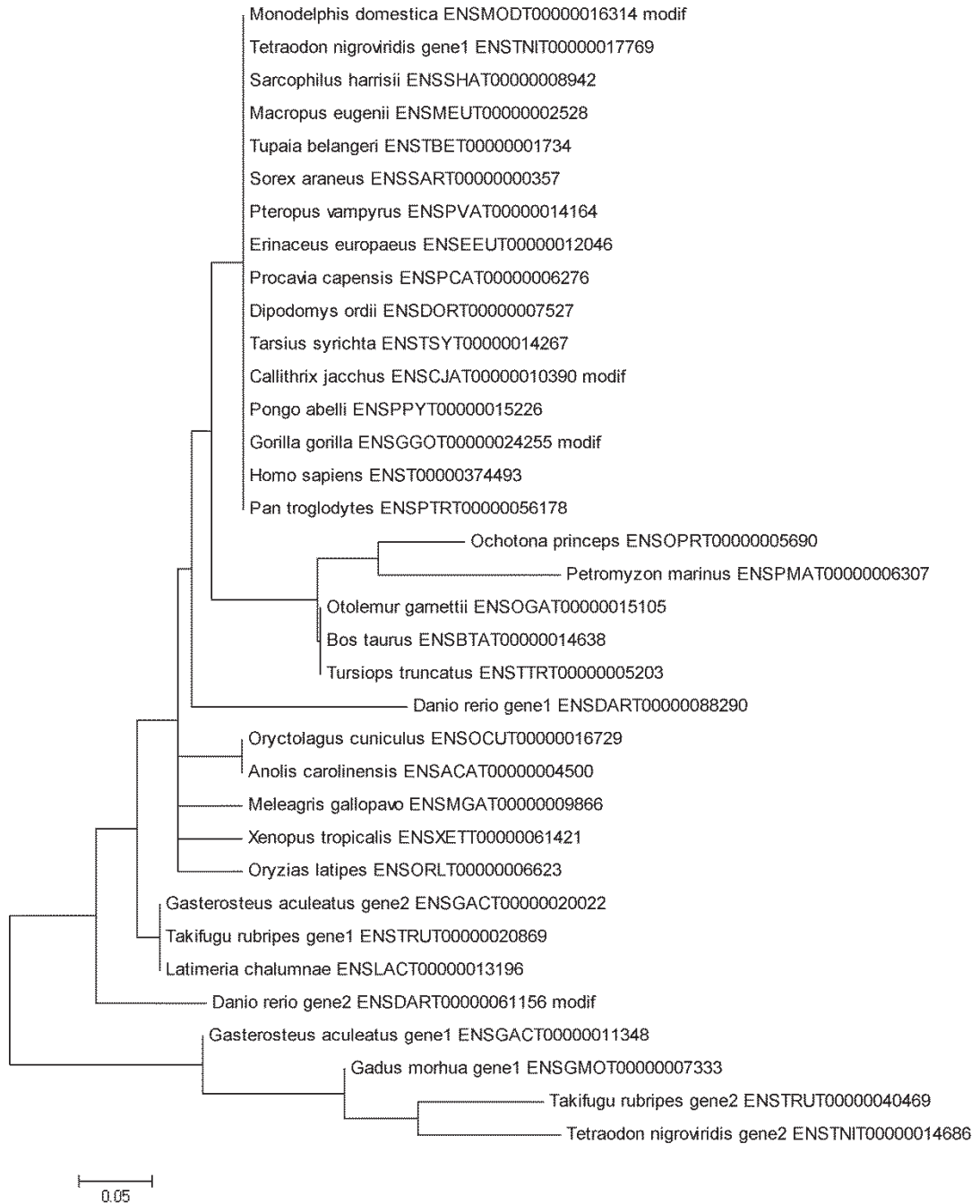


Figura 56 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 4, utilizando o método de Evolução Mínima (ME)

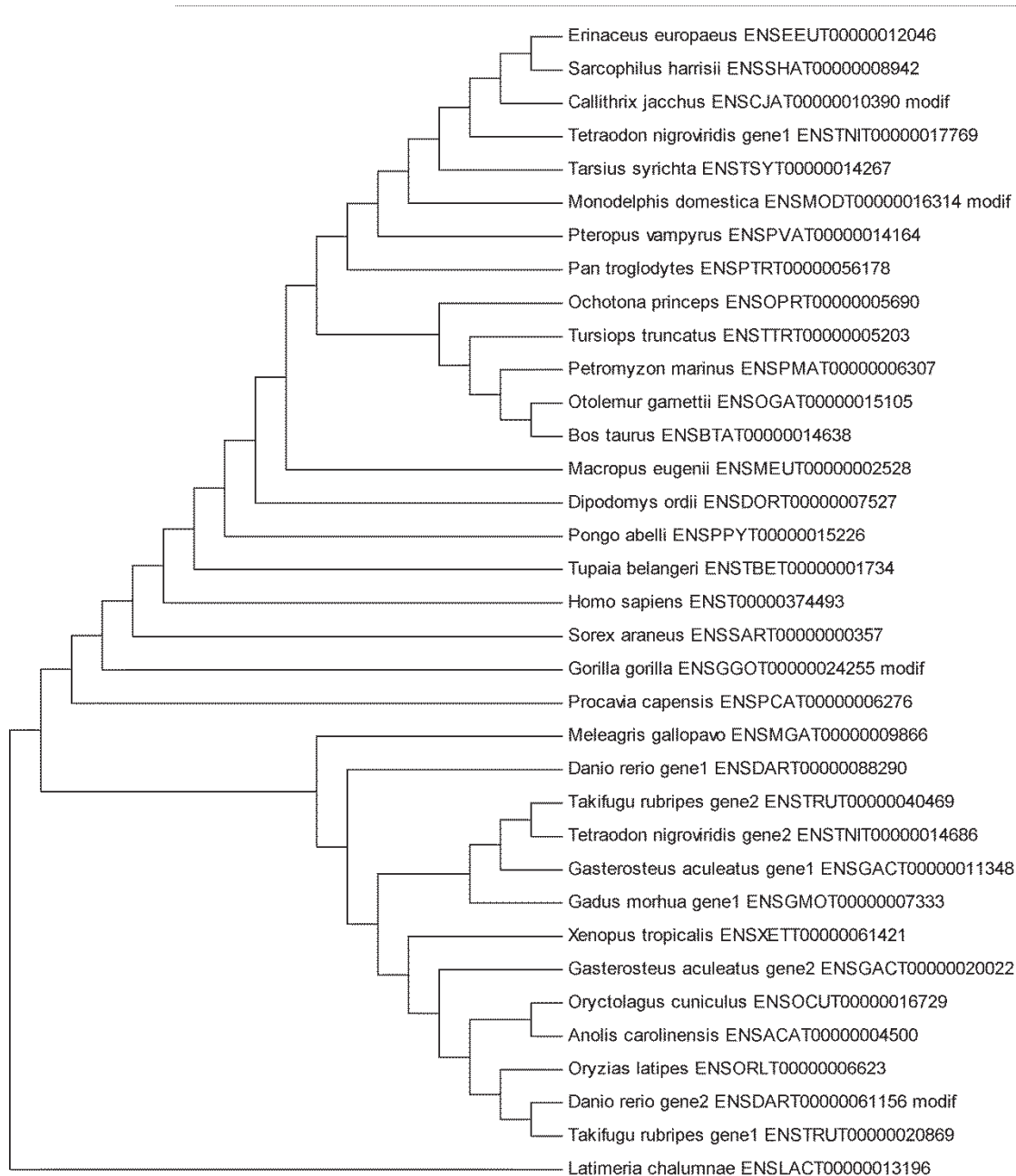


Figura 57 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 4, utilizando o método de Máxima Parcimônia (MP)

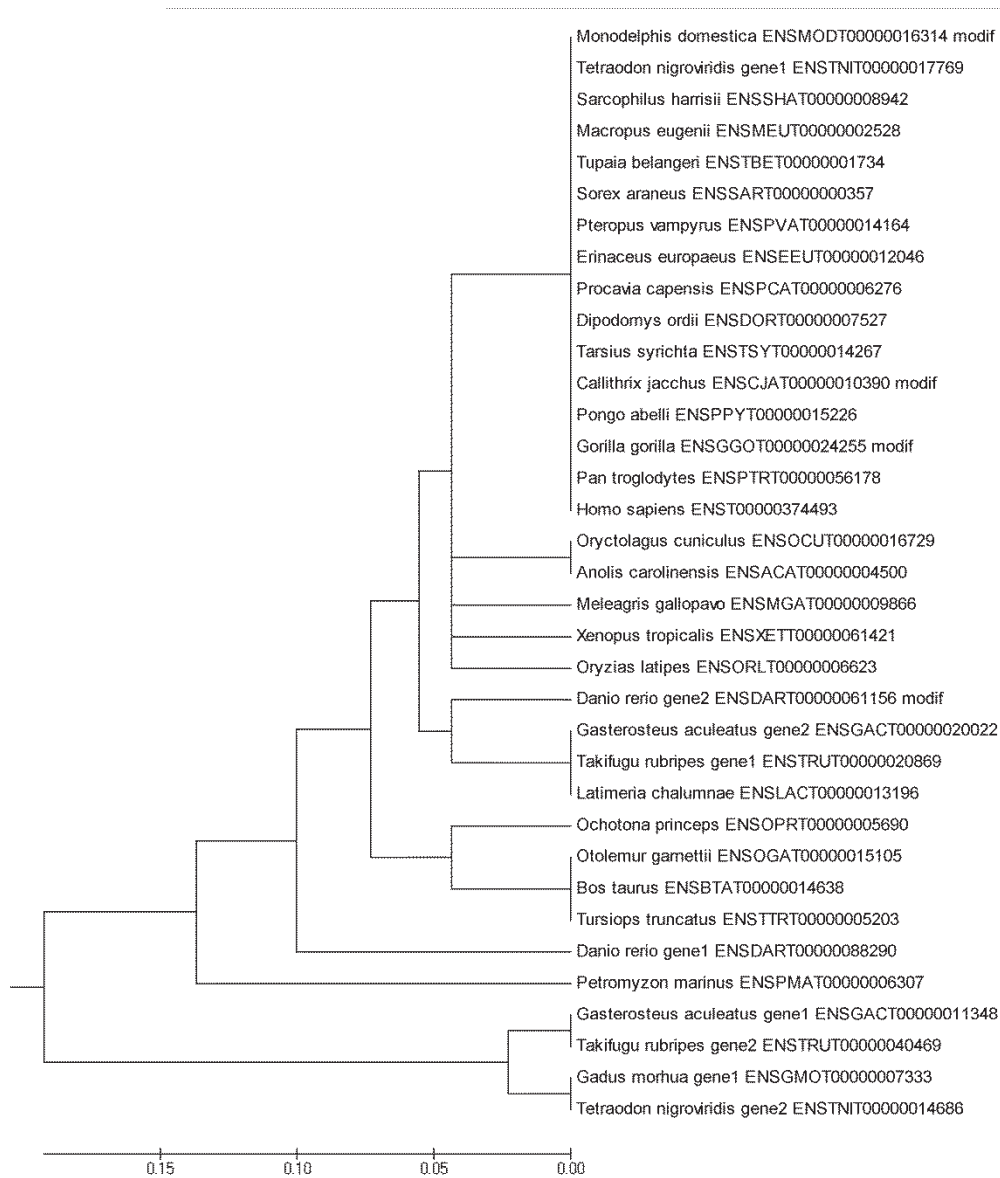


Figura 58 - Árvore filogenética do gene RAPH1 elaborada com sequências de aminoácidos da região do domínio EVH1 4, utilizando o método de Unweighted Pair-Group Method Using Arithmetic Averages (UPGMA)

APÊNDICE 12 – ÁRVORES AMINOACÍDICAS DE GENES PARÁLOGOS E
ORTÓLOGOS RAPH1

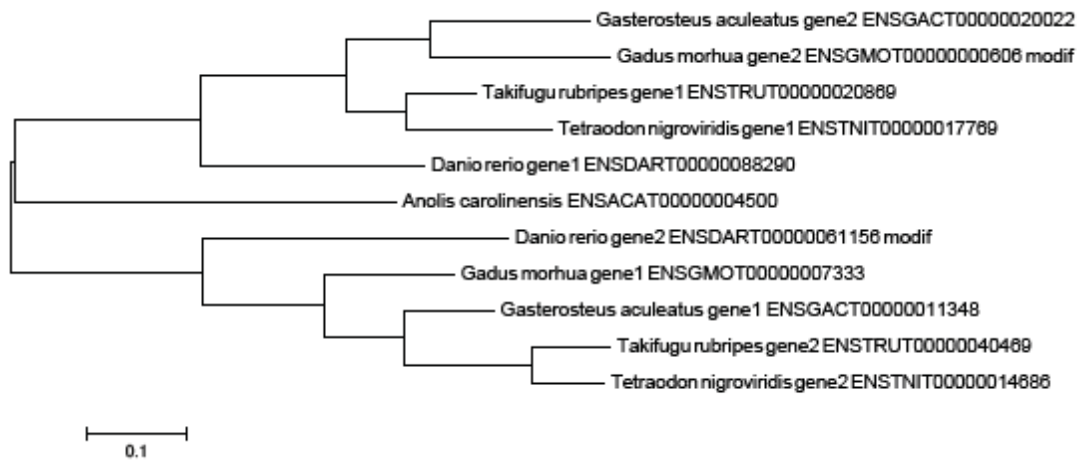


Figura 59 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Anolis carolinensis*

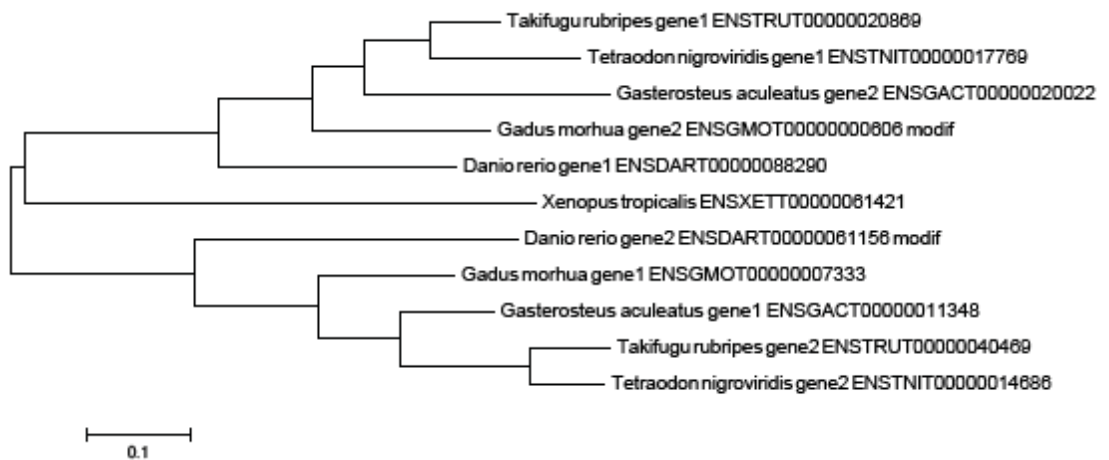


Figura 60 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Xenopus laevis*

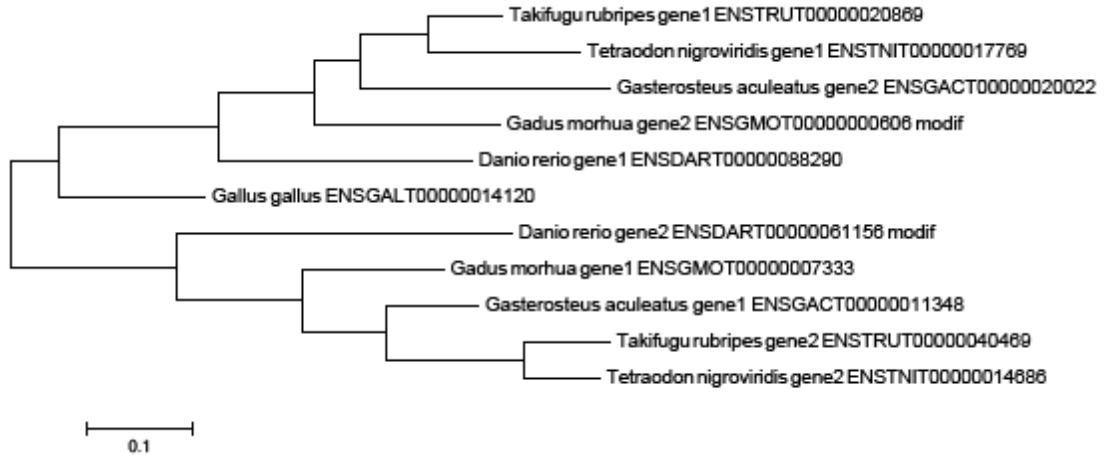


Figura 61 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Gallus gallus*

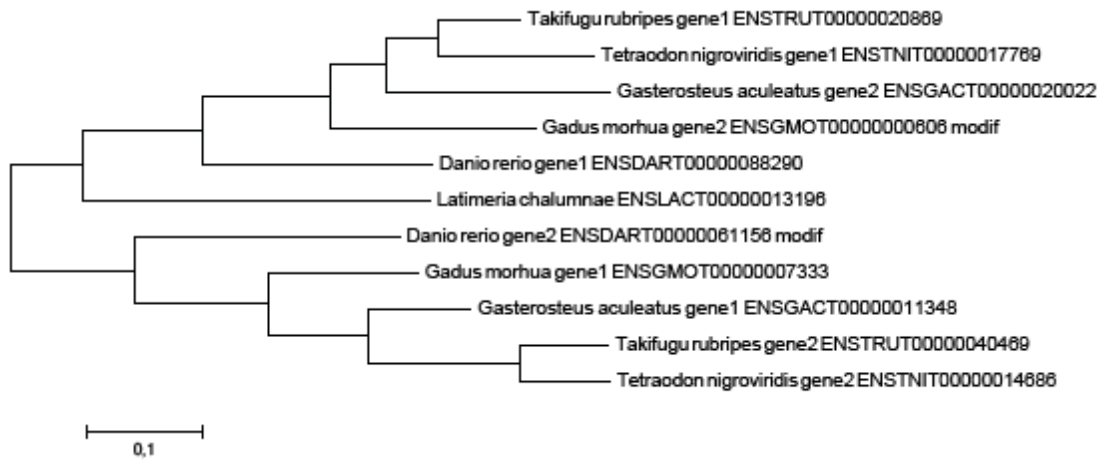


Figura 62 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Latimeria chalumnae*

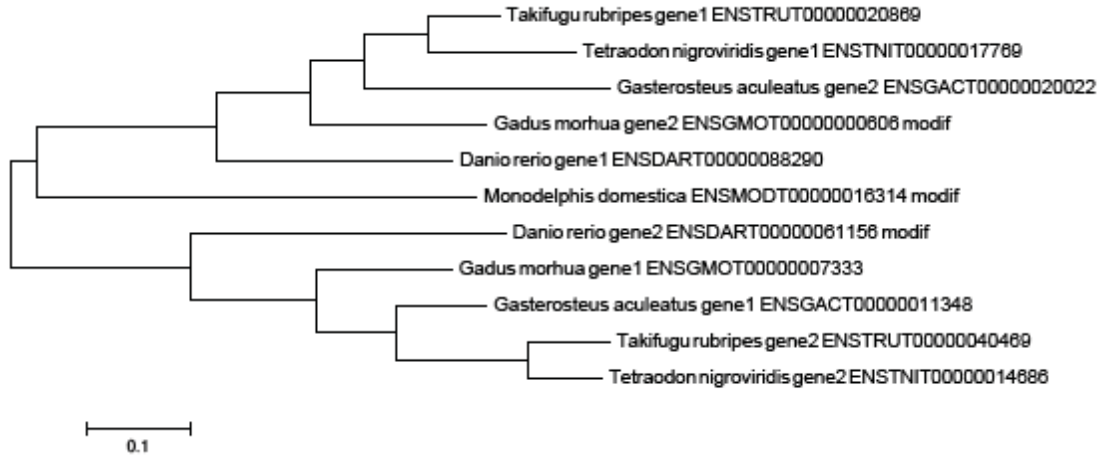


Figura 63 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Monodelphis domestica*

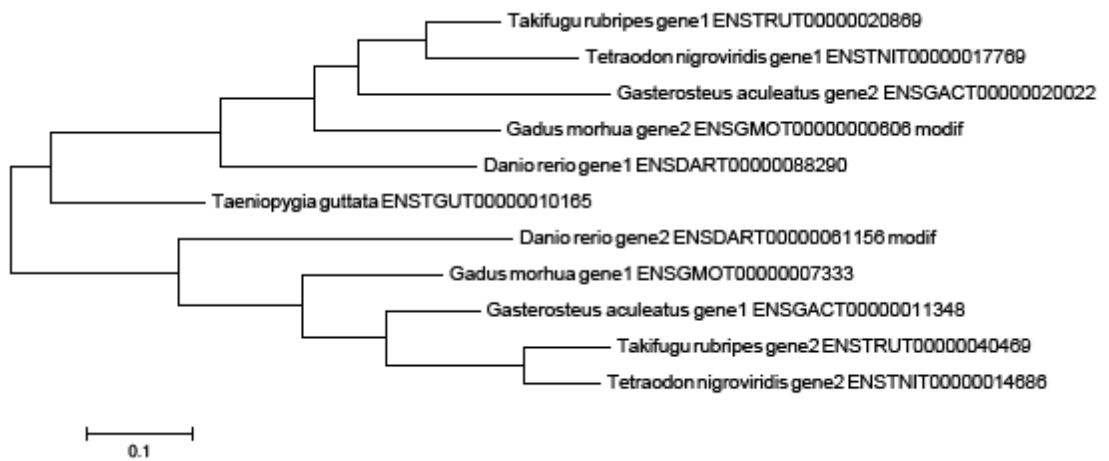


Figura 64 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Taeniopygia guttata*

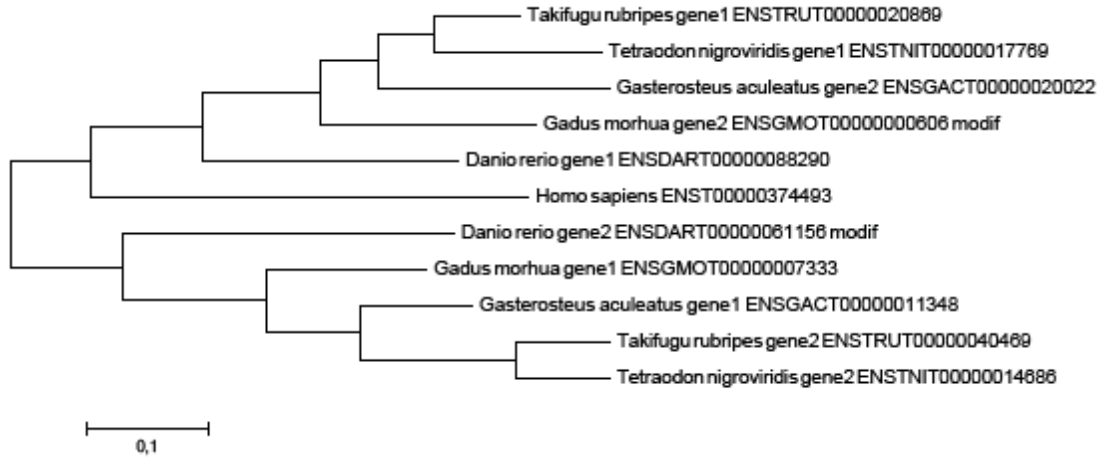


Figura 65 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos e *Homo sapiens*

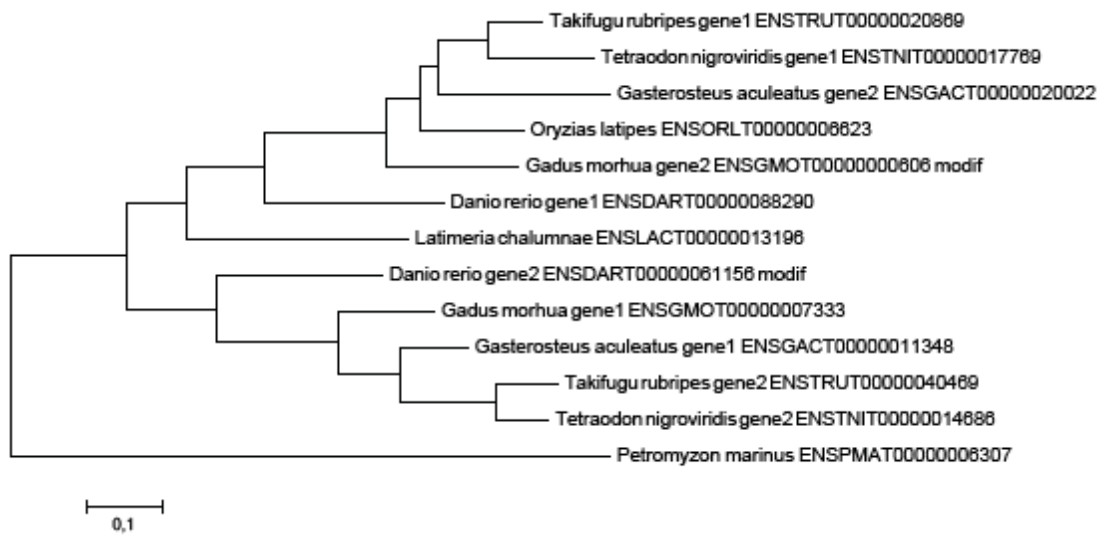


Figura 66 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos, *Latimeria chalumnae*, *Petromyzon marinus* e *Oryzias latipes*

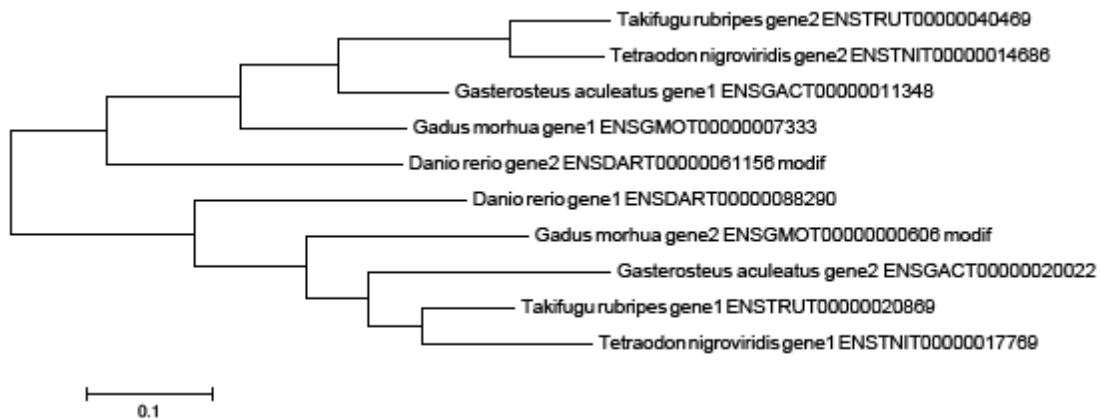


Figura 67 – Árvore Aminoacídica de Máxima Verossimilhança – Genes Duplos

APÊNDICE 13 – TABELAS DE COMPOSIÇÃO AMINOACÍDICA DO GENE

RAPH1

Tabela 6 – Composição Aminoacídica (parte 1 de 2)

	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu
Homo sapiens	88	12	57	66	24	49	26	44	93	101
Pan troglodytes	87	12	57	66	24	49	26	44	93	102
Gorilla gorilla	77	15	59	70	28	49	24	48	101	106
Pongo abelli	84	12	57	66	24	49	24	45	92	100
Nomascus leucogenys	43	8	42	48	14	26	15	28	54	64
Callithrix jacchus	79	12	55	66	23	50	27	45	91	99
Tarsius syrichta	70	12	46	44	16	48	20	31	70	84
Microcebus murinus	39	7	42	51	11	23	16	28	48	61
Otolemur garnettii	86	11	57	66	26	53	25	45	92	101
Equus caballus	8	1	13	14	3	6	1	5	7	13
Bos taurus	82	11	55	62	19	43	19	37	84	90
Sus scrofa	33	7	39	45	10	19	14	23	43	56
Canis familiaris	67	13	50	63	21	32	25	40	76	79
Felis catus	40	9	23	30	12	22	12	21	56	42
Ailuropoda melanoleuca	36	9	42	47	10	23	17	29	50	63
Mus musculus	41	11	44	52	17	27	23	31	54	75
Dipodomys ordii	87	5	44	44	14	34	20	32	54	76
Cavia porcellus	46	9	43	51	11	28	17	27	51	62
Procavia capensis	75	7	41	47	22	44	16	37	75	76
Choloepus hoffmanni	59	10	42	41	15	22	22	38	58	74
Tursiops truncatus	59	9	51	58	15	39	21	33	65	79
Loxodonta africana	51	9	46	55	17	32	21	38	61	70
Vicugna pacos	68	13	47	58	21	35	24	33	76	85
Erinaceus europaeus	67	8	46	47	17	48	18	23	54	69
Pteropus vampyrus	90	10	55	64	23	47	27	45	79	95
Myotis lucifugus	43	10	44	48	12	27	14	28	50	63
Oryctolagus cuniculus	86	9	58	63	18	53	20	29	75	87
Ochotona princeps	55	11	50	55	14	29	17	31	60	71
Sorex araneus	59	11	50	57	18	39	16	33	72	76
Tupaia belangeri	68	8	52	55	16	41	17	33	65	73
Macropus eugenii	41	8	48	48	12	29	17	32	45	63
Sarcophilus harrisii	69	13	57	66	22	54	22	41	84	97
Monodelphis domestica	71	15	60	71	31	52	19	49	96	109
Ornithorhynchus anatinus	27	0	18	18	3	9	3	6	13	19

Meleagris gallopavo	78	13	60	65	32	50	21	42	99	101
Gallus gallus	38	8	42	46	10	24	14	23	50	57
Taeniopygia guttata	43	8	41	47	10	22	13	20	50	59
Anolis carolinensis	93	8	57	60	29	49	21	39	89	96
Xenopus tropicalis	64	13	57	61	34	43	28	41	86	119
Oryzias latipes	81	11	59	59	21	86	28	30	73	88
Gasterosteus aculeatus g1	57	12	41	52	16	35	22	28	54	71
Gasterosteus aculeatus g2	100	11	55	63	31	86	28	34	68	107
Danio rerio g1	80	15	62	66	30	66	33	39	88	112
Danio rerio g2	101	11	54	59	31	62	40	52	85	108
Gadus morhua g1	42	9	33	33	13	28	16	16	38	62
Gadus morhua g2	16	7	17	15	9	17	9	11	21	37
Takifugu rubripes g1	95	14	48	62	25	80	37	34	83	99
Takifugu rubripes g2	85	13	60	71	35	58	31	47	82	117
Tetraodon nigroviridis g1	108	17	51	65	25	81	32	32	74	108
Tetraodon nigroviridis g2	96	10	60	69	28	62	32	34	73	118
Petromyzon marinus	68	17	35	36	26	41	19	29	49	70
Caenorhabditis elegans	50	15	42	72	19	20	19	34	49	67
Médias	64,92	10,37	47,38	53,90	19,37	41,15	20,92	33,02	66,31	80,31

Tabela 7 – Composição Aminoacídica (parte 2 de 2)

	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Totais
Homo sapiens	26	36	212	75	48	156	73	78	11	27	1302
Pan troglodytes	27	36	212	74	49	156	73	77	11	27	1302
Gorilla gorilla	28	44	176	73	53	163	75	75	15	29	1308
Pongo abelli	26	35	214	72	50	158	74	78	11	27	1298
Nomascus leucogenys	19	25	63	46	33	89	39	43	8	24	731
Callithrix jacchus	27	34	188	69	50	156	78	69	10	27	1255
Tarsius syrichta	15	28	129	49	47	113	52	43	8	22	947
Microcebus murinus	15	27	17	35	33	78	35	28	7	20	621
Otolemur garnettii	22	37	209	75	51	150	78	80	11	26	1301
Equus caballus	5	6	3	7	4	11	6	2	1	2	118
Bos taurus	19	31	180	62	46	139	66	65	11	24	1145
Sus scrofa	14	21	17	27	27	54	29	26	5	15	524
Canis familiaris	18	33	167	54	41	127	56	51	10	22	1045
Felis catus	10	16	106	30	21	78	29	42	8	19	626
Ailuropoda melanoleuca	15	23	16	33	29	67	33	28	7	21	598
Mus musculus	14	25	29	39	33	91	42	36	9	23	716
Dipodomys ordii	19	24	147	51	42	114	44	32	8	16	907
Cavia porcellus	17	23	20	39	33	81	32	34	9	21	654
Procavia capensis	22	30	202	56	41	133	60	68	7	18	1077
Choloepus hoffmanni	15	26	113	51	28	84	41	48	6	17	810
Tursiops truncatus	16	25	62	47	44	105	47	46	10	21	852
Loxodonta africana	14	26	81	50	32	99	43	45	8	23	821
Vicugna pacos	15	26	145	48	32	108	49	59	8	20	970
Erinaceus europaeus	10	27	92	43	49	98	39	33	8	21	817
Pteropus vampyrus	29	37	210	73	53	154	71	67	9	23	1261
Myotis lucifugus	21	24	47	38	35	88	37	34	8	20	691
Oryctolagus cuniculus	18	30	104	59	54	138	54	61	11	26	1053
Ochotona princeps	16	26	58	46	42	87	46	41	9	18	782
Sorex araneus	18	28	97	40	41	101	45	47	10	21	879
Tupaia belangeri	18	25	100	46	41	110	44	44	9	20	885
Macropus eugenii	17	29	42	45	37	91	47	31	8	19	709
Sarcophilus harrisii	29	37	143	60	57	155	56	56	14	28	1160
Monodelphis domestica	22	42	130	59	48	138	58	63	15	34	1182
Ornithorhynchus anatinus	7	7	6	15	9	32	12	5	1	3	213
Meleagris gallopavo	25	43	132	61	50	151	64	72	13	31	1203
Gallus gallus	16	24	18	34	28	77	28	37	8	19	601

Taeniopygia guttata	15	22	16	33	29	76	28	41	8	19	600
Anolis carolinensis	32	35	200	70	46	158	62	64	11	22	1241
Xenopus tropicalis	38	41	147	71	51	171	58	69	12	31	1235
Oryzias latipes	20	50	154	68	57	138	60	49	11	23	1166
Gasterosteus aculeatus g1	19	31	25	38	44	95	32	37	8	18	735
Gasterosteus aculeatus g2	26	49	186	75	72	132	56	56	9	28	1272
Danio rerio g1	30	44	189	84	55	203	68	59	11	33	1367
Danio rerio g2	37	55	304	109	47	187	72	63	10	24	1511
Gadus morhua g1	13	21	31	30	37	75	21	36	8	17	579
Gadus morhua g2	7	7	15	16	16	42	13	22	6	15	318
Takifugu rubripes g1	24	52	218	83	52	149	78	59	10	27	1329
Takifugu rubripes g2	29	46	169	90	50	212	62	53	12	24	1346
Tetraodon nigroviridis g1	20	40	179	82	71	143	72	54	10	26	1290
Tetraodon nigroviridis g2	29	41	169	68	60	205	61	53	12	25	1305
Petromyzon marinus	20	27	156	37	38	102	43	50	5	24	892
Caenorhabditis elegans	20	35	70	36	35	87	42	38	8	21	779
Médias	20,06	31,00	117,60	53,29	41,75	117,40	49,67	48,98	9,10	22,13	948,63

Tabela 8 – Composição Aminoacídica Percentual (parte 1 de 2)

	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu
Homo sapiens	6,76	0,92	4,38	5,07	1,84	3,76	2,00	3,38	7,14	7,76
Pan troglodytes	6,68	0,92	4,38	5,07	1,84	3,76	2,00	3,38	7,14	7,83
Gorilla gorilla	5,89	1,15	4,51	5,35	2,14	3,75	1,83	3,67	7,72	8,10
Pongo abelli	6,47	0,92	4,39	5,08	1,85	3,78	1,85	3,47	7,09	7,70
Nomascus leucogenys	5,88	1,09	5,75	6,57	1,92	3,56	2,05	3,83	7,39	8,76
Callithrix jacchus	6,29	0,96	4,38	5,26	1,83	3,98	2,15	3,59	7,25	7,89
Tarsius syrichta	7,39	1,27	4,86	4,65	1,69	5,07	2,11	3,27	7,39	8,87
Microcebus murinus	6,28	1,13	6,76	8,21	1,77	3,70	2,58	4,51	7,73	9,82
Otolemur garnettii	6,61	0,85	4,38	5,07	2,00	4,07	1,92	3,46	7,07	7,76
Equus caballus	6,78	0,85	11,02	11,86	2,54	5,08	0,85	4,24	5,93	11,02
Bos taurus	7,16	0,96	4,80	5,41	1,66	3,76	1,66	3,23	7,34	7,86
Sus scrofa	6,30	1,34	7,44	8,59	1,91	3,63	2,67	4,39	8,21	10,69
Canis familiaris	6,41	1,24	4,78	6,03	2,01	3,06	2,39	3,83	7,27	7,56
Felis catus	6,39	1,44	3,67	4,79	1,92	3,51	1,92	3,35	8,95	6,71
Ailuropoda melanoleuca	6,02	1,51	7,02	7,86	1,67	3,85	2,84	4,85	8,36	10,54
Mus musculus	5,73	1,54	6,15	7,26	2,37	3,77	3,21	4,33	7,54	10,47
Dipodomys ordii	9,59	0,55	4,85	4,85	1,54	3,75	2,21	3,53	5,95	8,38
Cavia porcellus	7,03	1,38	6,57	7,80	1,68	4,28	2,60	4,13	7,80	9,48
Procavia capensis	6,96	0,65	3,81	4,36	2,04	4,09	1,49	3,44	6,96	7,06
Choloepus hoffmanni	7,28	1,23	5,19	5,06	1,85	2,72	2,72	4,69	7,16	9,14
Tursiops truncatus	6,92	1,06	5,99	6,81	1,76	4,58	2,46	3,87	7,63	9,27
Loxodonta africana	6,21	1,10	5,60	6,70	2,07	3,90	2,56	4,63	7,43	8,53
Vicugna pacos	7,01	1,34	4,85	5,98	2,16	3,61	2,47	3,40	7,84	8,76
Erinaceus europaeus	8,20	0,98	5,63	5,75	2,08	5,88	2,20	2,82	6,61	8,45
Pteropus vampyrus	7,14	0,79	4,36	5,08	1,82	3,73	2,14	3,57	6,26	7,53
Myotis lucifugus	6,22	1,45	6,37	6,95	1,74	3,91	2,03	4,05	7,24	9,12
Oryctolagus cuniculus	8,17	0,85	5,51	5,98	1,71	5,03	1,90	2,75	7,12	8,26
Ochotona princeps	7,03	1,41	6,39	7,03	1,79	3,71	2,17	3,96	7,67	9,08
Sorex araneus	6,71	1,25	5,69	6,48	2,05	4,44	1,82	3,75	8,19	8,65
Tupaia	7,68	0,90	5,88	6,21	1,81	4,63	1,92	3,73	7,34	8,25

belangeri											
Macropus eugenii	5,78	1,13	6,77	6,77	1,69	4,09	2,40	4,51	6,35	8,89	
Sarcophilus harrisii	5,95	1,12	4,91	5,69	1,90	4,66	1,90	3,53	7,24	8,36	
Monodelphis domestica	6,01	1,27	5,08	6,01	2,62	4,40	1,61	4,15	8,12	9,22	
Ornithorhynchus anatinus	12,68	0,00	8,45	8,45	1,41	4,23	1,41	2,82	6,10	8,92	
Meleagris gallopavo	6,48	1,08	4,99	5,40	2,66	4,16	1,75	3,49	8,23	8,40	
Gallus gallus	6,32	1,33	6,99	7,65	1,66	3,99	2,33	3,83	8,32	9,48	
Taeniopygia guttata	7,17	1,33	6,83	7,83	1,67	3,67	2,17	3,33	8,33	9,83	
Anolis carolinensis	7,49	0,64	4,59	4,83	2,34	3,95	1,69	3,14	7,17	7,74	
Xenopus tropicalis	5,18	1,05	4,62	4,94	2,75	3,48	2,27	3,32	6,96	9,64	
Oryzias latipes	6,95	0,94	5,06	5,06	1,80	7,38	2,40	2,57	6,26	7,55	
Gasterosteus aculeatus g1	7,76	1,63	5,58	7,07	2,18	4,76	2,99	3,81	7,35	9,66	
Gasterosteus aculeatus g2	7,86	0,86	4,32	4,95	2,44	6,76	2,20	2,67	5,35	8,41	
Danio rerio g1	5,85	1,10	4,54	4,83	2,19	4,83	2,41	2,85	6,44	8,19	
Danio rerio g2	6,68	0,73	3,57	3,90	2,05	4,10	2,65	3,44	5,63	7,15	
Gadus morhua g1	7,25	1,55	5,70	5,70	2,25	4,84	2,76	2,76	6,56	10,71	
Gadus morhua g2	5,03	2,20	5,35	4,72	2,83	5,35	2,83	3,46	6,60	11,64	
Takifugu rubripes g1	7,15	1,05	3,61	4,67	1,88	6,02	2,78	2,56	6,25	7,45	
Takifugu rubripes g2	6,32	0,97	4,46	5,27	2,60	4,31	2,30	3,49	6,09	8,69	
Tetraodon nigroviridis g1	8,37	1,32	3,95	5,04	1,94	6,28	2,48	2,48	5,74	8,37	
Tetraodon nigroviridis g2	7,36	0,77	4,60	5,29	2,15	4,75	2,45	2,61	5,59	9,04	
Petromyzon marinus	7,62	1,91	3,92	4,04	2,91	4,60	2,13	3,25	5,49	7,85	
Caenorhabditis elegans	6,42	1,93	5,39	9,24	2,44	2,57	2,44	4,36	6,29	8,60	
Médias	6,84	1,09	5,00	5,68	2,04	4,34	2,21	3,48	6,99	8,47	

Tabela 9 – Composição Aminoacídica Percentual (parte 2 de 2)

	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Totais
Homo sapiens	2,00	2,76	16,28	5,76	3,69	11,98	5,61	5,99	0,84	2,07	1302
Pan troglodytes	2,07	2,76	16,28	5,68	3,76	11,98	5,61	5,91	0,84	2,07	1302
Gorilla gorilla	2,14	3,36	13,46	5,58	4,05	12,46	5,73	5,73	1,15	2,22	1308
Pongo abelli	2,00	2,70	16,49	5,55	3,85	12,17	5,70	6,01	0,85	2,08	1298
Nomascus leucogenys	2,60	3,42	8,62	6,29	4,51	12,18	5,34	5,88	1,09	3,28	731
Callithrix jacchus	2,15	2,71	14,98	5,50	3,98	12,43	6,22	5,50	0,80	2,15	1255
Tarsius syrichta	1,58	2,96	13,62	5,17	4,96	11,93	5,49	4,54	0,84	2,32	947
Microcebus murinus	2,42	4,35	2,74	5,64	5,31	12,56	5,64	4,51	1,13	3,22	621
Otolemur garnettii	1,69	2,84	16,06	5,76	3,92	11,53	6,00	6,15	0,85	2,00	1301
Equus caballus	4,24	5,08	2,54	5,93	3,39	9,32	5,08	1,69	0,85	1,69	118
Bos taurus	1,66	2,71	15,72	5,41	4,02	12,14	5,76	5,68	0,96	2,10	1145
Sus scrofa	2,67	4,01	3,24	5,15	5,15	10,31	5,53	4,96	0,95	2,86	524
Canis familiaris	1,72	3,16	15,98	5,17	3,92	12,15	5,36	4,88	0,96	2,11	1045
Felis catus	1,60	2,56	16,93	4,79	3,35	12,46	4,63	6,71	1,28	3,04	626
Ailuropoda melanoleuca	2,51	3,85	2,68	5,52	4,85	11,20	5,52	4,68	1,17	3,51	598
Mus musculus	1,96	3,49	4,05	5,45	4,61	12,71	5,87	5,03	1,26	3,21	716
Dipodomys ordii	2,09	2,65	16,21	5,62	4,63	12,57	4,85	3,53	0,88	1,76	907
Cavia porcellus	2,60	3,52	3,06	5,96	5,05	12,39	4,89	5,20	1,38	3,21	654
Procavia capensis	2,04	2,79	18,76	5,20	3,81	12,35	5,57	6,31	0,65	1,67	1077
Choloepus hoffmanni	1,85	3,21	13,95	6,30	3,46	10,37	5,06	5,93	0,74	2,10	810
Tursiops truncatus	1,88	2,93	7,28	5,52	5,16	12,32	5,52	5,40	1,17	2,46	852
Loxodonta africana	1,71	3,17	9,87	6,09	3,90	12,06	5,24	5,48	0,97	2,80	821
Vicugna pacos	1,55	2,68	14,95	4,95	3,30	11,13	5,05	6,08	0,82	2,06	970
Erinaceus europaeus	1,22	3,30	11,26	5,26	6,00	12,00	4,77	4,04	0,98	2,57	817
Pteropus vampyrus	2,30	2,93	16,65	5,79	4,20	12,21	5,63	5,31	0,71	1,82	1261
Myotis lucifugus	3,04	3,47	6,80	5,50	5,07	12,74	5,35	4,92	1,16	2,89	691
Oryctolagus cuniculus	1,71	2,85	9,88	5,60	5,13	13,11	5,13	5,79	1,04	2,47	1053
Ochotona princeps	2,05	3,32	7,42	5,88	5,37	11,13	5,88	5,24	1,15	2,30	782
Sorex araneus	2,05	3,19	11,04	4,55	4,66	11,49	5,12	5,35	1,14	2,39	879
Tupaia belangeri	2,03	2,82	11,30	5,20	4,63	12,43	4,97	4,97	1,02	2,26	885
Macropus eugenii	2,40	4,09	5,92	6,35	5,22	12,83	6,63	4,37	1,13	2,68	709
Sarcophilus	2,50	3,19	12,33	5,17	4,91	13,36	4,83	4,83	1,21	2,41	1160

harrisii												
Monodelphis domestica	1,86	3,55	11,00	4,99	4,06	11,68	4,91	5,33	1,27	2,88	1182	
Ornithorhynchus anatinus	3,29	3,29	2,82	7,04	4,23	15,02	5,63	2,35	0,47	1,41	213	
Meleagris gallopavo	2,08	3,57	10,97	5,07	4,16	12,55	5,32	5,99	1,08	2,58	1203	
Gallus gallus	2,66	3,99	3,00	5,66	4,66	12,81	4,66	6,16	1,33	3,16	601	
Taeniopygia guttata	2,50	3,67	2,67	5,50	4,83	12,67	4,67	6,83	1,33	3,17	600	
Anolis carolinensis	2,58	2,82	16,12	5,64	3,71	12,73	5,00	5,16	0,89	1,77	1241	
Xenopus tropicalis	3,08	3,32	11,90	5,75	4,13	13,85	4,70	5,59	0,97	2,51	1235	
Oryzias latipes	1,72	4,29	13,21	5,83	4,89	11,84	5,15	4,20	0,94	1,97	1166	
Gasterosteus aculeatus g1	2,59	4,22	3,40	5,17	5,99	12,93	4,35	5,03	1,09	2,45	735	
Gasterosteus aculeatus g2	2,04	3,85	14,62	5,90	5,66	10,38	4,40	4,40	0,71	2,20	1272	
Danio rerio g1	2,19	3,22	13,83	6,14	4,02	14,85	4,97	4,32	0,80	2,41	1367	
Danio rerio g2	2,45	3,64	20,12	7,21	3,11	12,38	4,77	4,17	0,66	1,59	1511	
Gadus morhua g1	2,25	3,63	5,35	5,18	6,39	12,95	3,63	6,22	1,38	2,94	579	
Gadus morhua g2	2,20	2,20	4,72	5,03	5,03	13,21	4,09	6,92	1,89	4,72	318	
Takifugu rubripes g1	1,81	3,91	16,40	6,25	3,91	11,21	5,87	4,44	0,75	2,03	1329	
Takifugu rubripes g2	2,15	3,42	12,56	6,69	3,71	15,75	4,61	3,94	0,89	1,78	1346	
Tetraodon nigroviridis g1	1,55	3,10	13,88	6,36	5,50	11,09	5,58	4,19	0,78	2,02	1290	
Tetraodon nigroviridis g2	2,22	3,14	12,95	5,21	4,60	15,71	4,67	4,06	0,92	1,92	1305	
Petromyzon marinus	2,24	3,03	17,49	4,15	4,26	11,43	4,82	5,61	0,56	2,69	892	
Caenorhabditis elegans	2,57	4,49	8,99	4,62	4,49	11,17	5,39	4,88	1,03	2,70	779	
Médias	2,11	3,27	12,40	5,62	4,40	12,38	5,24	5,16	0,96	2,33	948,6	