

---

CIÊNCIAS BIOLÓGICAS

---

GABRIEL AMARAL

**Análise evolutiva das isoformas  $\alpha$  e  $\beta$  da proteína DACT1,  
importante moduladora da via de sinalização Wnt**



Rio Claro  
2019

GABRIEL AMARAL

ANÁLISE EVOLUTIVA DAS ISOFORMAS  $\alpha$  E  $\beta$  DA PROTEÍNA DACT1,  
IMPORTANTE MODULADORA DA VIA DE SINALIZAÇÃO WNT

Orientadora: Profa. Dra. LÚCIA ELVIRA ALVARES

Supervisora: Profa. Dra. PATRÍCIA PASQUALI PARISE MALTEMPI

Trabalho de Conclusão de Curso apresentado  
ao Instituto de Biociências da Universidade  
Estadual Paulista “Júlio de Mesquita Filho” -  
Câmpus de Rio Claro, para obtenção do grau  
de Bacharel em Ciências Biológicas.

Rio Claro

2019

A485a	<p>Amaral, Gabriel</p> <p>Análise evolutiva das isoformas alfa e beta da proteína DACT1, importante moduladora da via de sinalização Wnt / Gabriel Amaral. -- Rio Claro, 2019</p> <p>50 p. : il., tabs. + 1 CD-ROM</p> <p>Trabalho de conclusão de curso ( - ) - Universidade Estadual Paulista (Unesp), Instituto de Biociências, Rio Claro</p> <p>Orientadora: Lúcia Elvira Alvares</p> <p>1. Evolução do Desenvolvimento. 2. Splicing Alternativo. 3. Bioinformática. I. Título.</p>
-------	---

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

## AGRADECIMENTOS

Não poderia começar meus agradecimentos de outra forma que não mencionando meus pais, Patricia e Alexandre, que sempre me apoiaram em todos os sentidos possíveis, acreditando na minha capacidade e me permitindo ser quem eu sou.

Também quero agradecer à Manuela de Oliveira Ramalho, que me ensinou muito do que eu sei sobre “fazer ciência”. Ela me acompanhou em meu primeiro projeto de pesquisa, compartilhando seus conhecimentos sempre com muita paciência e carinho. Ela me fez entender que momentos de frustração fazem parte do processo e nos ajudam a crescer, porque ciência é tentativa e erro, é repetir experimentos, é começar tudo de novo, é repensar hipóteses, discutir ideias e às vezes chegar a lugar nenhum. E tudo bem.

Agradeço aos amigos que eu fiz ao longo da graduação, em especial à Débora, à Caroline e aos moradores e agregados da república Makuta, onde vivi muitos bons momentos. Falando em amigos, não poderia esquecer da Letícia, do Igor, da Bárbara, da Isabella, da Ana Paula, do Leandro, do Felipe e, principalmente, da Carla, porque ela colocou o meu nome nos agradecimentos do TCC dela também.

Agradeço à Atlética da Biologia, da qual fui membro por quase cinco anos e na qual adquiri experiências muito valiosas e fiz amigos muito queridos. E é impossível não citar também o Interbio, um evento de biólogos para biólogos, o Maior Espetáculo da Terra, cuja 40ª edição ocorreu em Rio Claro graças ao esforço de pessoas extremamente competentes com as quais eu tenho muito orgulho de ter trabalhado.

Por fim, agradeço à minha orientadora, Lúcia Elvira Alvares, por todas as conversas que tivemos ao longo do desenvolvimento desse estudo, pela confiança que ela depositou em mim, por toda atenção que me deu e por contribuir de forma tão positiva na minha formação.

*"Estamos cercados por infindáveis formas  
belíssimas e fascinantes, e não é por acidente,  
mas sim uma consequência direta da evolução  
por seleção natural não-aleatória - a única  
resposta, o maior espetáculo da Terra."*

*Richard Dawkins*

## RESUMO

O gene *Dact1* (Dishevelled binding antagonist of beta catenin 1), exclusivo dos genomas de organismos cordados, é capaz de modular as vias de sinalização Wnt canônica e não canônica através da interação de seu produto com a proteína citoplasmática Dishevelled (DVL). Essa modulação se estende também ao núcleo, onde a DACT1 se liga a fatores de transcrição e outras moléculas envolvidas na regulação da expressão gênica. A via Wnt está relacionada a processos de extrema importância para o desenvolvimento embrionário, como a determinação e a migração celular, a renovação de células-tronco e a organogênese, sendo portanto de grande interesse científico e médico. Sabe-se que, em humanos, a DACT1 apresenta duas isoformas geradas por *splicing* alternativo, as quais provavelmente exercem funções distintas. Neste trabalho, foram coletadas as sequências codificadoras do *Dact1* de 95 espécies de vertebrados (juntamente com as sequências de aminoácidos correspondentes) para estudo da história evolutiva do gene em questão. Através de alinhamentos e outras análises bioinformáticas, foi possível inferir aspectos referentes ao surgimento, a evolução e possíveis funções das isoformas  $\alpha$  e  $\beta$ . Constatou-se que a isoforma ancestral é a  $\alpha$  e que o aparecimento do sítio acceptor alternativo responsável pela síntese da isoforma  $\beta$  se deu em um ancestral comum a todos os Tetrapoda. Esse sítio se perdeu secundariamente em algumas espécies ao longo da filogenia, enquanto que em outras tornou-se constitutivo, sugerindo que nenhuma das isoformas é essencial para o desenvolvimento adequado do indivíduo, embora ao menos uma delas deva estar presente. Foram encontradas assinaturas moleculares de grupo e evidências de evolução acelerada na região submetida ao *splicing* alternativo, que pode estar envolvida na interação entre a DACT1 e o fator de transcrição TCF3. Assim, sugere-se que a atividade da isoforma  $\alpha$  tenha sofrido pequenos ajustes ao longo da evolução, os quais podem ter consequências funcionais no que diz respeito ao controle da via Wnt canônica a nível nuclear.

**Palavras-chave:** Dact1; sinalização Wnt; desenvolvimento embrionário; *splicing* alternativo.

## SUMÁRIO

1. INTRODUÇÃO .....	11
1.1. Estrutura dos Genes Eucarióticos .....	11
1.2. Processamento do mRNA .....	12
1.3. <i>Splicing</i> Alternativo .....	16
1.4. Gene DACT1 .....	17
1.4.1. A Via Wnt .....	18
1.4.2. Funções da Dact1 .....	20
2. OBJETIVOS .....	23
2.1. Objetivos Gerais .....	23
2.2. Objetivos Específicos .....	23
3. PROCEDIMENTOS DE PESQUISA .....	23
3.1. Coleta de Dados .....	23
3.2. Identificação das Isoformas na Amostra de Sequências .....	24
3.3. Alinhamento de Sequências .....	25
3.4. Análise de Possíveis Parceiros Moleculares .....	26
4. RESULTADOS .....	26
5. DISCUSSÃO .....	32
6. CONCLUSÃO .....	36
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	37
APÊNDICES	
I - Espécies Analisadas e Códigos de Acesso .....	41
II - Dados sobre as Sequências Coletadas .....	43
III - Espécies Descartadas .....	45
IV - Alinhamento de Todos os Clados .....	46
V - Alinhamento sem Actinopterygii .....	48
VI - Árvore filogenética Chondrichthyes, Actinopterygii, e Actinistia .....	50

## 1. INTRODUÇÃO

Um gene é um trecho de DNA que coordena a síntese de um produto molecular funcional, comumente uma cadeia polipeptídica (KLUG et al, 2012). Nesse processo, apenas uma das cadeias nucleotídicas é lida para a produção do mRNA, sendo denominada fita molde ou antissenso, pois apresenta-se no sentido  $3' \rightarrow 5'$ . A cadeia complementar, por sua vez, tem sequência quase idêntica à do transcrito a ser sintetizado, diferenciando-se apenas pela substituição das bases nitrogenadas timina por uracila, característica das moléculas de RNA. Essa cadeia é chamada de fita codificadora ou senso, pois apresenta-se no sentido  $5' \rightarrow 3'$  (NELSON & COX, 2013).

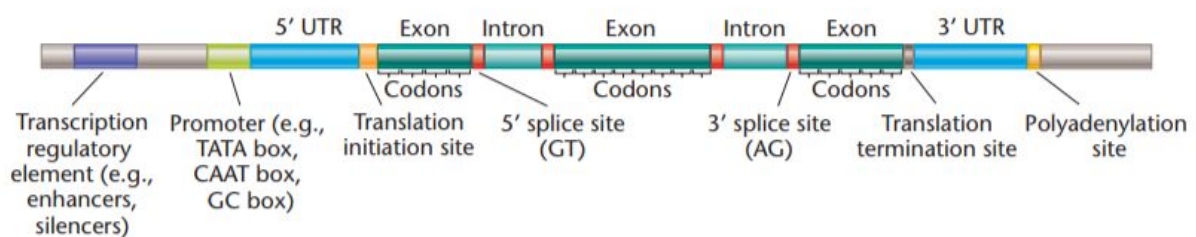
### 1.1. Estrutura dos Genes Eucarióticos

O gene eucariótico típico tem sua sequência codificadora fragmentada em diversos segmentos de DNA denominados éxons, que juntos contém todos os códons a serem transcritos e traduzidos em uma cadeia de aminoácidos (KLUG et al, 2012). Na fita senso, o primeiro éxon geralmente inicia-se com o códon ATG, que atua no mRNA maduro como sítio de iniciação da tradução (na forma AUG). Já o último éxon termina com um dos três possíveis códons de parada: TAG, TAA ou TGA, que atuam como sítio de terminação da tradução (nas formas UAG, UAA ou UGA, respectivamente) (KLUG et al, 2012).

Os éxons são intercalados por segmentos não codificantes denominados íntrons, que são removidos do mRNA após a transcrição (KLUG et al, 2012; NELSON & COX, 2013). Em um gene típico, as sequências intrônicas costumam ser bem mais longas do que as exônicas, de forma que os segmentos codificadores, em geral, correspondem a uma pequena porcentagem do tamanho total dos genes em que estão inseridos (KLUG et al, 2012; NELSON & COX, 2013). Essa característica é reforçada pela presença de diversas regiões regulatórias, também não-codificantes, que são responsáveis pelo controle da expressão gênica e o processamento pós-transcricional do mRNA (KLUG et al, 2012; GILBERT, 2010), conforme apresentado na Figura 1.

O principal elemento regulatório de um gene é o seu promotor, que localiza-se na região  $5'$  da fita codificadora, *upstream* em relação aos éxons e íntrons (KLUG et al, 2012; NELSON & COX, 2013). Ele apresenta sequências consenso que variam de gene para gene e são reconhecidas por fatores de transcrição. Alguns deles, denominados fatores gerais de

transcrição, são responsáveis por recrutar e ativar a enzima RNA polimerase, formando o chamado complexo de pré-iniciação, que é indispensável para a síntese do transcrito primário. Por esse motivo, a região onde se encontram as sequências consenso às quais esses fatores se ligam é conhecida como promotor essencial (KLUG et al, 2012; GILBERT, 2010). Outros fatores atuam na modulação da eficiência e das taxas de transcrição gênica basais, ligando-se a sequências presentes em uma região denominada promotor proximal (KLUG et al, 2012; GILBERT, 2010).



**FIGURA 1 - Estrutura de um gene eucariótico típico, fita senso**

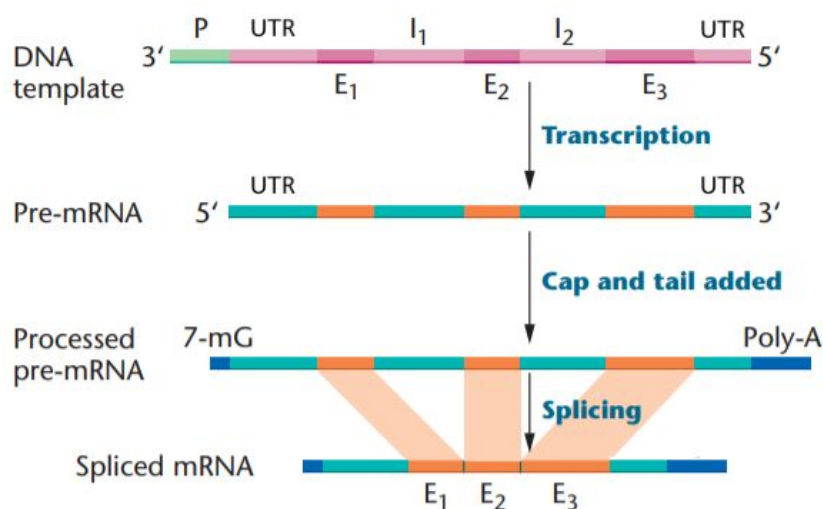
Fonte: KLUG et al, 2011.

É também assim que funcionam os *enhancers* e *silencers*, outros importantes componentes regulatórios dos genes eucarióticos (KLUG et al, 2012). Diferentemente dos promotores, no entanto, esses segmentos podem estar na região 5' da fita senso (*upstream* em relação ao restante do gene, como representado na Figura 1), na região 3' (*downstream*) ou até mesmo entre os trechos codificadores (posicionando-se de maneira a fazerem parte de sequências intrônicas). Como seus próprios nomes sugerem, os *enhancers* e *silencers* são responsáveis por amplificar e minimizar os níveis de transcrição, respectivamente (KLUG et al, 2012). Esses componentes regulatórios são capazes de gerar variações bem mais significativas e refinadas no que diz respeito à expressão gênica, estando envolvidos no controle temporal e tecido-específico da mesma através da ligação de diferentes conjuntos de fatores de transcrição (KLUG et al, 2012; GILBERT, 2010).

## 1.2. Processamento do mRNA

Transcritos primários de mRNA eucariótico sofrem diversas alterações antes de serem finalmente translocados para o citoplasma e utilizados como moldes para a síntese proteica (KLUG et al, 2012; NELSON & COX, 2013), conforme esquematizado na Figura 2. A

primeira dessas alterações ocorre na extremidade 5' da molécula nascente, que é acrescida de um resíduo de 7-metilguanossina, um nucleosídeo modificado conhecido como *cap*. Esse resíduo é adicionado antes mesmo da síntese do transcrito primário ser concluída (NELSON & COX, 2013; CARVALHO & RECCO-PIMENTEL, 2013). Da mesma forma, a extremidade 3' do mRNA também recebe modificações, as quais são catalisadas por um complexo proteico capaz de reconhecer e ligar-se a sequências consenso presentes na 3' UTR. Esse complexo então cliva a molécula e insere, a partir do ponto de clivagem (também chamado de sítio de poliadenilação), até 250 resíduos consecutivos de adenosina monofosfato, produzindo uma estrutura denominada cauda poli-A (NELSON & COX, 2013). Tanto o *cap-5'* quanto a cauda poli-A atuam no sentido de estabilizar o mRNA e protegê-lo contra degradação. Além disso, desempenham papéis essenciais nos processos de translocação núcleo-citoplasmática e tradução gênica (KLUG et al, 2012; CARVALHO & RECCO-PIMENTEL, 2013).



**FIGURA 2 - Esquema geral do processamento pós-transcricional**

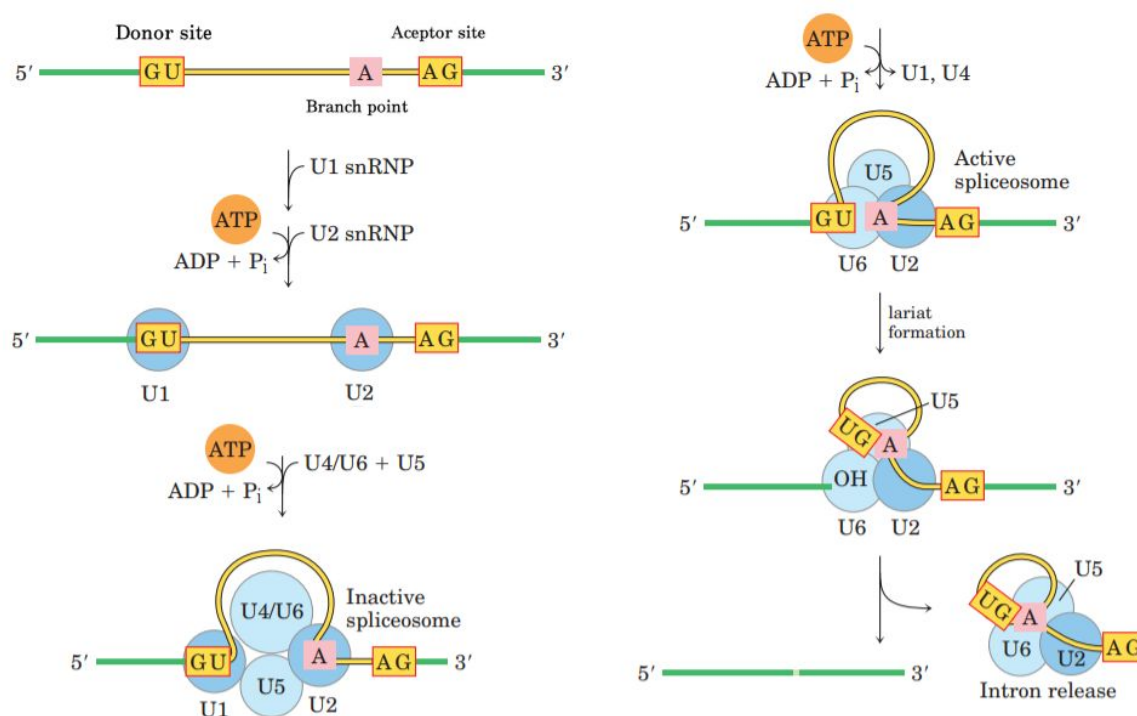
Fonte: KLUG et al, 2011. Adaptado. P = Promoter; E = Exon; I = Intron.

Conforme anteriormente mencionado, o processamento pós-transcricional dos mRNAs nascentes envolve também a remoção de seus segmentos intrônicos, em um processo conhecido como *splicing* (CARVALHO & RECCO-PIMENTEL, 2013). Embora alguns íntrons tenham capacidade de auto-excisão, a maioria conta com o auxílio de um grande complexo denominado spliceossomo, formado por 5 ribonucleoproteínas (snRNPs) altamente



conservada conhecida como sítio de ramificação<sup>1</sup>, ataca a ligação entre a guanina do sítio doador e o nucleotídeo *upstream* adjacente (o último do éxon), promovendo a clivagem da molécula nesse ponto e o estabelecimento de uma nova ligação entre os nucleotídeos G e A, como resultado de uma reação orgânica denominada transesterificação (CLANCY, 2008; KLUG et al, 2012; CARVALHO & RECCO-PIMENTEL, 2013).

Essa mesma reação ocorre então uma segunda vez, mas agora entre o último nucleotídeo do éxon que antecedia o íntron a guanina do sítio aceptor, a qual rompe sua ligação com o nucleotídeo *downstream* adjacente (o primeiro do éxon seguinte). Assim, o íntron separa-se completamente do restante do mRNA e seus dois éxons contíguos ligam-se entre si, encerrando o processo de *splicing* (CLANCY, 2008; CARVALHO & RECCO-PIMENTEL, 2013). Mais detalhes, como o papel dos outros snRNPs e os custos energéticos envolvidos, são apresentados na Figura 4.



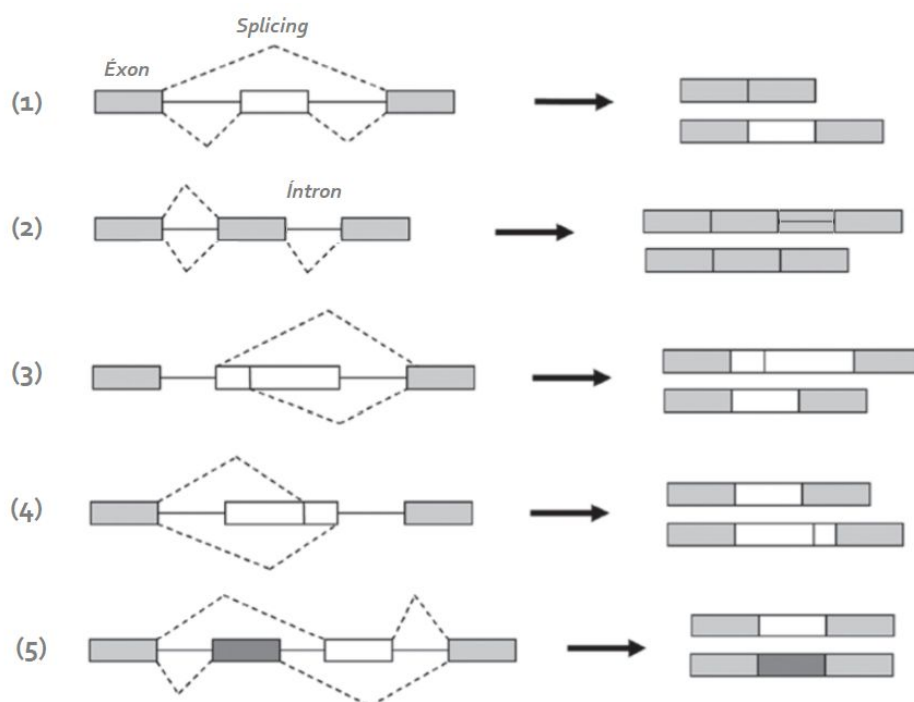
**FIGURA 4 - Esquema do processo de *splicing* e da formação do spliceossomo**

Fonte: NELSON & COX, 2013. Adaptado.

<sup>1</sup> Consiste no consenso 5'-YNYRAY-3', onde Y significa qualquer pirimidina (C ou U), R qualquer purina (A ou G) e N qualquer nucleotídeo (CLANCY, 2008).

### 1.3. *Splicing* Alternativo

Variações na forma como ocorre o *splicing* de um mRNA nascente constituem o princípio de um importante mecanismo eucariótico de regulação da expressão gênica: o *splicing* alternativo (BARALLE & GIUDICE, 2017). Esse mecanismo baseia-se no uso diferencial dos sítios doadores e/ou aceptores presentes em um transcrito primário e permite às células gerar uma vasta diversidade de mRNAs maduros a partir de uma quantidade muito menor de genes (BARALLE & GIUDICE, 2017; NILSEN & GRAVELEY, 2010). Diversos eventos podem caracterizar um *splicing* alternativo, nomeadamente: (1) a remoção de determinado éxon; (2) a retenção de determinado íntron; (3) a utilização de um sítio doador alternativo; (4) a utilização de um sítio aceitador alternativo e (5) a opção entre dois ou mais éxons mutuamente excludentes (WANG et al, 2014), como esquematizado na Figura 5.



**FIGURA 5 - Tipos de *splicing* alternativo**

Fonte: WANG et al, 2014. Adaptado.

Por meio desses processos, um mesmo gene é capaz de produzir uma imensa gama de transcritos maduros únicos que têm potencial para atuar de formas diferentes dentro das

células, contribuindo grandemente para o aumento da diversidade de proteínas (BARALLE & GIUDICE, 2017). Sabe-se que, no genoma humano, aproximadamente 95% dos genes são submetidos a algum tipo de *splicing* alternativo e 37% apresentam como resultado final desse evento a geração de múltiplas isoformas proteicas (KIM et al, 2014; BARALLE & GIUDICE, 2017; WANG et al, 2014). A enorme discrepância entre a quantidade de genes presentes nos genomas da maioria da espécies de metazoários e o número de proteínas sintetizadas pelos mesmos pode ser atribuída, em parte, justamente por esse tipo de processamento pós-transcricional (WANG et al, 2014).

Eventos de *splicing* alternativo dependem de regiões regulatórias do mRNA conhecidas como *enhancers* e *silencers* de *splicing* (ES e SS, respectivamente), que podem estar contidas tanto em segmentos exônicos quanto intrônicos (WANG et al, 2014). Essas regiões apresentam sequências consenso que são reconhecidas e ligadas por proteínas específicas, capazes de modular a funcionalidade dos sítios de *splicing* (NILSEN & GRAVELEY, 2010; WANG et al, 2014). As proteínas da família SR<sup>2</sup>, por exemplo, reconhecem os ES e recrutam componentes do spliceossomo, o que as caracteriza como fatores de ação positiva, pois contribuem para a ativação do processo de *splicing* a partir do sítio que está sob regulação. As ribonucleoproteínas hnRNPs, por outro lado, ligam-se aos SS e prejudicam as interações entre os componentes do spliceossomo, sendo caracterizadas como fatores de ação negativa (NILSEN & GRAVELEY, 2010; WANG et al, 2014).

#### 1.4. Gene *Dact1*

O *Dact1* (Dishevelled binding antagonist of beta catenin 1) é um típico gene eucariótico com sequência codificadora fragmentada, contendo quatro éxons e três longos íntrons (YAU et al, 2005). Acredita-se que tenha surgido tardiamente na linhagem dos deuterostômios, uma vez que não pode ser encontrado nos genomas de organismos invertebrados (SCHUBERT et al, 2014). Seu produto consiste em uma proteína modular homônima que possui uma gama de parceiros moleculares e diversos domínios conservados, notavelmente um zíper de leucina próximo à extremidade N-terminal e um motivo C-terminal de ligação a PDZ (SCHUBERT et al, 2014). Apresenta também sinais de localização e

---

<sup>2</sup> É assim denominada pois suas integrantes apresentam domínios compostos por extensas repetições dos aminoácidos serina e arginina (NILSEN & GRAVELEY, 2010).

exportação nuclear (NLS e NES), indicando que pode deslocar-se entre núcleo e citoplasma e, potencialmente, exercer funções em ambos os compartimentos celulares (GAO et al, 2008).

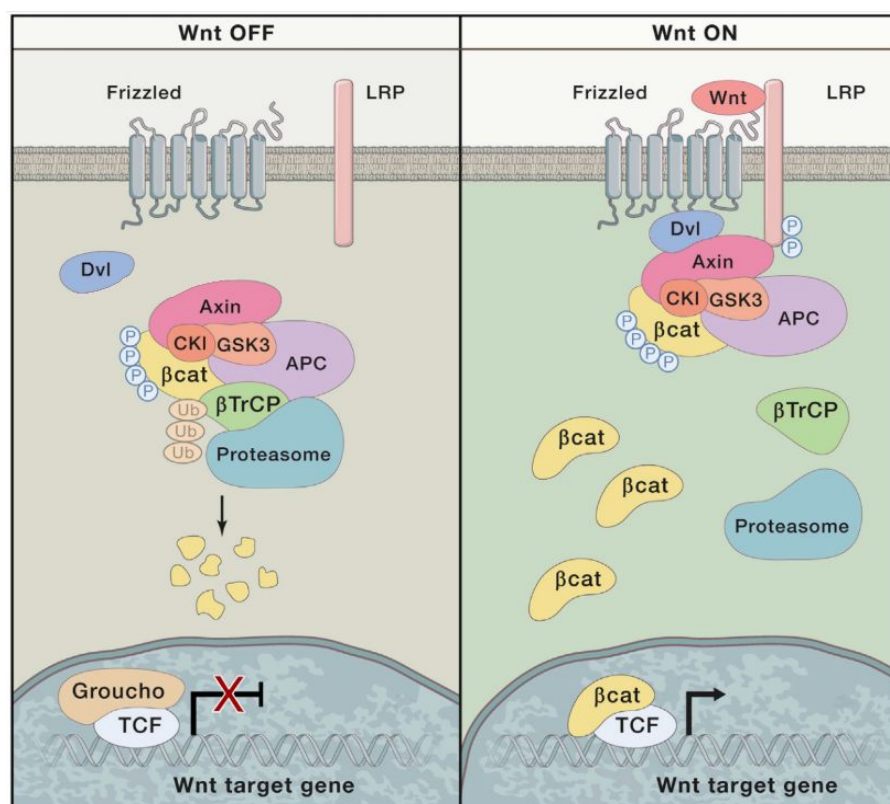
A DACT1 foi descoberta independentemente por dois grupos de pesquisa que, ao conduzirem sistemas duplo-híbrido em levedura, detectaram sua interação com a proteína citoplasmática Dishevelled (DVL), integrante central da via de sinalização Wnt (CHEYETTE et al, 2002 & GLOY et al, 2002). Ambos os grupos coordenaram ensaios com o anuro *Xenopus laevis*, espécie que sabe-se hoje possuir atipicamente dois parálogos muito semelhantes do *Dact1*, sendo um deles identificado por Cheyette et al, 2002 (e nomeado *Dapper*) e o outro por Gloy et al, 2002 (e nomeado *Frodo*). Atualmente esses parálogos são denominados *Dact1a* e *Dact1b*, e entende-se que o segundo está ausente nas demais espécies de vertebrados (SCHUBERT et al, 2014). Além de observarem parálogos distintos, cada um dos grupos também constatou uma forma de atuação diferente para a proteína: enquanto Cheyette et al, 2002 verificou que a DACT1 inibia via de sinalização Wnt, Gloy et al, 2002 concluiu justamente o oposto.

#### 1.4.1. A Via Wnt

A via Wnt é altamente conservada nos metazoários e está envolvida em processos essenciais ao desenvolvimento, como determinação e polaridade celular, migração, organogênese e até mesmo renovação de células-tronco (KOMIYA & HABAS, 2008; NUSSE & CLEVERS, 2017). As Wnt constituem uma família de glicoproteínas que, ao serem secretadas, ligam-se a receptores de membrana denominados Frizzled (FZ). Dependendo de qual é a Wnt responsável pela sinalização, qual FZ está sendo ligado e quais co-receptores estão presentes, a via pode estimular diversas cascatas bioquímicas intracelulares distintas. Todas elas, no entanto, apresentam como intermediário citoplasmático a proteína Dishevelled, que é recrutada e ativada pela FZ (KOMIYA & HABAS, 2008; DAULAT & BORG, 2017).

Em condições normais, sem sinalização externa, as células apresentam um grande complexo proteico que é responsável pela fosforilação da proteína  $\beta$ -catenina, marcando-a para ubiquitinação e consequente degradação proteossomal (NUSSE & CLEVERS, 2017). Na chamada via Wnt canônica (Figura 6), esse complexo é recrutado para a membrana plasmática juntamente com a Dvl, onde ele é inativado por um co-receptor do Wnt denominado LRP5/6 (KOMIYA & HABAS, 2008; MOON et al., 2004). A ausência da

atividade de degradação leva ao acúmulo da  $\beta$ -catenina no citoplasma, o que permite que a mesma seja então translocada para o núcleo, onde atua como co-ativadora de fatores de transcrição da família TCF/LEF, estimulando a expressão controlada de diversos genes (KOMIYA & HABAS, 2008; MOON et al., 2004; NUSSE & CLEVERS, 2017).



**FIGURA 6 - Via de sinalização Wnt canônica**

Fonte: NUSSE & CLEVERS, 2017. Adaptado.

Todas as vias Wnt independentes de  $\beta$ -catenina são ditas não canônicas. Na principal delas, denominada Polaridade Celular Planar (PCP), a DVL é recrutada e ativa pequenas GTPases da família Rho (nomeadamente a RAC1, a RHOA e a CDC42) (KOMIYA & HABAS, 2008; DAULAT & BORG, 2017). Essas proteínas estão envolvidas na remodelação dos componentes de actina do citoesqueleto, contribuindo para a polaridade celular e a migração direcionada de células durante eventos-chave do desenvolvimento embrionário, como gastrulação e neurulação. Além disso, também participam da via JNK, relacionada a proliferação celular e processos apoptóticos e inflamatórios (DAULAT & BORG, 2017).

#### 1.4.2. Funções da DACT1

Alguns outros estudos seguiram-se após as afirmações iniciais contraditórias de Cheyette et al, 2002 e Gloy et al, 2002 sobre a função exercida pela DACT1 na via de sinalização Wnt canônica, mas os mesmos eram igualmente divergentes (WAXMAN et al, 2004; ZHANG et al, 2006). A explicação veio apenas com o trabalho de Teran et al, 2009, no qual demonstrou-se que a atividade da proteína é dependente de seu estado de fosforilação.

Quando não-fosforilada, a DACT1 favorece interações entre a DVL e enzimas do tipo E3 ubiquitina ligase, como a MARCH2 e a pVHL, que são responsáveis por marcar proteínas para degradação (LEE et al, 2018; MA et al, 2015). A via autofágico-lisossômica atua nesse processo tanto quanto a via proteossomal, uma vez que as moléculas de DVL ubiquitinadas tendem a formar agregados citoplasmáticos, tornando-se eventualmente inacessíveis ao proteassoma (GAO et al, 2010). Além disso, a DACT1 também é capaz de estimular o processo de autofagia de maneira global ao interagir com as proteínas Beclin1 e ATG14L e aumentar o estabelecimento de complexos ATG14L-Beclin1-VPS34, os quais são essenciais para a constituição dos autofagossomos (MA et al, 2014). Assim, entende-se que a interrupção da via Wnt promovida pela DACT1 ocorre através de dois mecanismos independentes: ao mesmo tempo em que facilita a ubiquitinação da DVL, também colabora para a síntese da maquinaria que virá a destruí-la (MA et al, 2015).

Um terceiro mecanismo foi observado no trabalho de Gao et al, 2008, onde verificou-se que, no núcleo, a DACT1 é capaz de ligar-se a uma proteína denominada HDAC1 e mediar a interação da mesma com o fator de transcrição LEF1. Dessa forma, o LEF1 torna-se indisponível e não pode ser co-ativado pela  $\beta$ -catenina para estimular a expressão gênica direcionada. Embora esse processo possa contribuir para a inibição da via Wnt canônica, sugere-se que se dê também na ausência de sinalização e atue no controle da atividade basal da  $\beta$ -catenina, uma vez que nem todas as suas moléculas são capturadas pelo complexo de degradação (GAO et al, 2008). Sugere-se também que a DACT1 possa ligar-se diretamente a LEF1 e TCF3, embora os dados de Kivimae et al, 2011 curiosamente não confirmem essas informações (HIKASA & SOKOL, 2004; GAO et al, 2008).

A DACT1 não-fosforilada demonstrou-se essencial para a ocorrência de diversos eventos embrionários. Em mamíferos, por exemplo, observou-se que ela está presente na mesoderme pré-somática dos embriões (FISHER et al, 2006), onde a ativação alternada das

vias de sinalização Wnt e Notch é responsável por promover a formação de estruturas segmentares de mesoderme paraxial denominadas somitos (AULEHLA et al, 2003). Nesse contexto, a DACT1 é produzida ciclicamente e atua, em conjunto com a proteína Axin1, na inibição periódica da via Wnt (SURIBEN, FISHER & CHEYETTE, 2006). Outro exemplo é o processo de delaminação da crista neural, comum a todos os metazoários craniados, que também é dependente dessa mesma inibição (RABADÁN et al, 2016).

Além disso, diversos trabalhos exploraram os efeitos da DACT1 não-fosforilada em células cancerosas, nas quais a interrupção da via canônica mostrou-se capaz de impedir a proliferação celular e induzir a apoptose, apresentando atividade de supressão tumoral (YAU et al, 2005; ZHU et al, 2017; GAO et al, 2018). O mesmo resultado foi observado para células-tronco (JIAO et al, 2018) e células mesangiais renais (JARDIM et al, 2017).

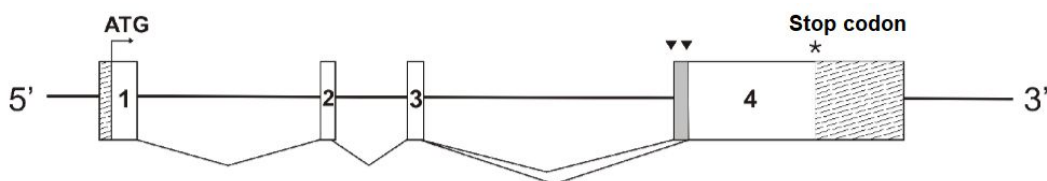
Ao ser fosforilada pela Casein Kinase I $\delta/\epsilon$ , a DACT1 passa a apresentar função oposta àquela anteriormente descrita, estimulando a via Wnt canônica (TERAN, BRANSCOMB & SEELING, 2009). Os processos de estabelecimento de sinapses e formação de dendritos em interneurônios corticais, hipertrofia cardíaca e desenvolvimento placentário são exemplos de eventos cujas ocorrências são dependentes dos efeitos da DACT1 em estado fosforilado (ARGUELO et al, 2013; HAGENMUELLER et al, 2013; HOU et al, 2014).

A DACT1 parece ser capaz de promover e regular também a via PCP, uma vez que interage com a proteína de membrana VANGL2, importante componente dessa cascata bioquímica (SURIBEN et al, 2009; HAGENMUELLER et al, 2014). Isso ocorre, por exemplo, em células da linha primitiva, nas quais a PCP é ativada durante o processo de transição epitélio-mesenquimal. Ratos mutantes para perda de função do gene *Dact1* apresentam desregulação da via nessa fase do desenvolvimento e os indivíduos nascem com diversas má-formações nos sistemas urogenital e digestório (SURIBEN et al, 2009; WEN et al, 2010). Alinhado a essas descobertas está o fato de que, em mamíferos, o *Dact1* é expresso no ducto Mülleriano do embrião e nos ovidutos e útero do indivíduo adulto, sendo essencial ao desenvolvimento e manutenção do sistema reprodutor feminino (SUZUKI et al, 2014; XING et al, 2016).

Diversos outros processos biológicos têm sido relacionados à atividade da DACT1, mas não há consenso sobre a exata forma de atuação da mesma no que diz respeito a muitos deles. Sabe-se, por exemplo, que está envolvida no metabolismo hepático de glicose e lipídeos em ratos acometidos com diabetes tipo II (KUANG et al., 2017); que pode regular

processos de crescimento e orientação axônica no núcleo olfatório anterior, em neurônios receptores olfatórios, em neurônios dos gânglios espinais e em células ganglionares da retina (SUZUKI et al, 2014); que possui papel no processo de odontogênese (KETTUNEN et al, 2010) e que está presente no coração tanto durante o período embrionário como no órgão do indivíduo adulto (SUZUKI et al, 2014). Também está envolvida, através de uma via metabólica semelhante à PCP, em rearranjos do citoesqueleto de actina necessários para a formação de dendritos e espinhas dendríticas e o estabelecimento de sinapses excitatórias durante a diferenciação de neurônios piramidais (OKERLUND et al, 2010).

O estudo de Yau et al, 2005 identificou a existência de duas isoformas da DACT1 em humanos, geradas através de *splicing* alternativo (Figura 7). A isoforma  $\alpha$  é a maior e corresponde àquela originalmente descrita por Cheyette et al, 2002, sendo 37 aminoácidos mais longa do que a isoforma  $\beta$ . Essa variação se dá pois, durante o processamento pós-transcricional, o limite 5' do éxon 4 é estabelecido de acordo com a opção da maquinaria de *splicing* entre dois sítios aceptores alternativos, e sugere-se que a diferença possa refletir em alterações na atividade da proteína final. Através de ensaios com culturas de células de hepatocarcinoma, observou-se que diminuições na síntese da isoforma  $\alpha$  geram acúmulo de  $\beta$ -catenina no citoplasma, levando à promoção da Wnt canônica e o desenvolvimento tumoral.



**FIGURA 7 - *Splicing* alternativo responsável pela síntese das isoformas  $\alpha$  e  $\beta$  da DACT1.**

Fonte: SOBREIRA et al. Não publicado. Adaptado. As setas indicam os dois sítios aceptores alternativos.

Dados do nosso grupo de pesquisa indicam que, além do *Homo sapiens*, outras espécies também produzem as isoformas  $\alpha$  e  $\beta$  da DACT1, e supõe-se que a origem e a diversificação estrutural dessas isoformas podem ter contribuído para o advento de um mecanismo extra de regulação da via Wnt nos vertebrados. Sabendo-se que há evidências de que regiões que sofrem *splicing* alternativo são passíveis de evolução acelerada (XING & LEE, 2005; ERMAKOVA, NURTDINOV & GELFAND, 2006), é razoável presumir que

modificações de sequência na região equivalente à adição relativa à isoforma  $\alpha$  teriam possibilitado ajustes clado-específicos ou até mesmo espécie-específicos no que se refere à função da isoforma maior.

## 2. OBJETIVOS

### 2.1. Objetivos Gerais

Tem-se como objetivo geral desse trabalho proporcionar maior compreensão acerca da história evolutiva das isoformas  $\alpha$  e  $\beta$  da *DACT1* através de análises bioinformáticas. Pretende-se investigar o surgimento das mesmas, a diversificação da região que sofre *splicing* alternativo e possíveis aspectos funcionais advindos desse processo.

### 2.2. Objetivos Específicos

- a) Determinar em que ponto da árvore filogenética dos vertebrados surgem as isoformas  $\alpha$  e  $\beta$  através da análise dos potenciais sítios aceptores de *splicing*;
- b) Alinhar e comparar as sequências codificadoras do *Dact1* a fim de se identificar possíveis assinaturas moleculares de grupo e evidências sugestivas de evolução acelerada na região que sofre *splicing* alternativo;
- c) Buscar na literatura parceiros moleculares capazes de interagir com a *DACT1* na região exclusiva à isoforma  $\alpha$ , possibilitando inferências funcionais.

## 3. PROCEDIMENTOS DE PESQUISA

### 3.1. Coleta de Dados

Buscaram-se as sequências codificadoras (CDS) do gene *DACT1* de 123 espécies representativas dos principais clados de gnatostomados nos bancos de dados Ensembl Genome Browser (<https://www.ensembl.org/index.html>) e National Center for Biotechnology Information - NCBI (<https://www.ncbi.nlm.nih.gov>), juntamente com suas respectivas

sequências de aminoácidos. As pesquisas realizadas foram textuais e o elemento de busca empregado foi o próprio nome do gene em estudo. A escolha por utilizar duas plataformas distintas se deu no intuito de facilitar o encontro das sequências e aumentar a confiabilidade dos dados através da comparação das informações disponíveis em cada uma das fontes.

Notou-se que o NCBI apresenta um portfólio maior do que o do Ensembl, e por isso diversas espécies foram encontradas apenas através dessa ferramenta. Já quando ambos os bancos ofereciam dados relacionados a uma mesma espécie, dois cenários eram possíveis: as sequências eram idênticas e, portanto, bastante confiáveis ou (2) as sequências eram diferentes, e por isso foi necessária uma análise qualitativa das mesmas para que fosse eleita a mais adequada. O critério para essas decisões consistia principalmente em verificar a montagem genômica utilizada como referência por cada uma das ferramentas de busca, levando-se em consideração se eram recentes, se estavam completas (pois muitas sequências apresentam gaps e nucleotídeos indeterminados) e quais eram os trabalhos de origem das mesmas. Também observava-se a qualidade da anotação automática fornecida pela plataforma e a semelhança entre as sequências apresentadas e aquelas sabidas típicas do gene em estudo.

A listagem das espécies analisadas e os códigos de acesso correspondentes ao *DACTI* de cada uma delas estão relacionados no Apêndice 1. O número de éxons e os tamanhos de cada CDS e de cada proteína (em pares de base e em quantidade de aminoácidos, respectivamente) podem ser consultados no Apêndice 2. Espécies descartadas do estudo (por motivos melhor descritos na seção “Resultados”) constam no Apêndice 3.

### **3.2. Identificação das Isoformas na Amostra de Sequências**

Quando os bancos de dados já indicavam a existência das isoformas  $\alpha$  e  $\beta$ , coletava-se a sequência maior e destacava-se os aminoácidos adicionais em relação à menor, sendo feito o mesmo para os nucleotídeos correspondentes. Outras isoformas que não as de interesse foram desconsideradas. Em seguida, com auxílio do software Gene Runner (<http://www.genrunner.net/>) e da ferramenta de alinhamento online CLUSTALW (<https://www.genome.jp/tools-bin/clustalw>), buscou-se e destacou-se manualmente nas sequências das demais espécies a região exclusiva à isoforma  $\alpha$ . Como as anotações automáticas elaboradas pelas plataformas poderiam ter levado em consideração qualquer dos dois possíveis sítios aceptores de *splicing*, a ausência da região em determinada espécie não significava necessariamente que

a mesma não existia, sendo preciso ainda verificar se estaria no íntron antecessor ao éxon 4. Caso se mostrasse presente e potencialmente funcional (sem códons de parada), ela era acrescida às sequências originais (tanto codificadoras quanto proteicas) e, então, também destacada. A situação de cada espécie nesse contexto pode ser conferida no Apêndice 1, com legendas de acordo.

É importante salientar, no entanto, que independentemente da adição relativa à isoforma  $\alpha$  existir, a isoforma  $\beta$  foi contabilizada apenas em espécies nas quais o sítio acceptor (AG) responsável pela ocorrência do *splicing* alternativo que dá origem a essa variante também foi encontrado na extremidade 3' da região em estudo. No caso de espécies que apresentaram códons de parada em suas possíveis regiões adicionais, a isoforma não contabilizada foi a  $\alpha$ , uma vez que o sítio acceptor responsável pela síntese da mesma está necessariamente ausente nesse contexto.

### 3.3. Alinhamento das Sequências

Tanto as sequências nucleotídicas quanto proteicas foram alinhadas através do software MEGA X (<https://www.megasoftware.net/>) e da ferramenta online ClustalW 2.1 (<https://www.genome.jp/tools-bin/clustalw>). Foram produzidos alinhamentos gerais e clado-específicos, tanto de sequências completas quanto apenas da adição relativa à isoforma  $\alpha$ . A partir dos dados obtidos, foi possível a construção de árvores filogenéticas (também pelo MEGA X) e a investigação de assinaturas moleculares de grupo na região submetida a *splicing* alternativo, a qual foi realizada apenas em clados para os quais possuía-se sequências de pelo menos 3 espécies.

Além disso, utilizou-se a ferramenta online SNAP v2.1.1 (<https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html>) para estudo de substituições nucleotídicas sinônimas e não-sinônimas, que pode ou não apresentar resultados sugestivos de evolução acelerada. Alinhamentos especiais foram produzidos para essa análise, abrangendo apenas os clados que já apresentam sítio acceptor alternativo funcional (ou seja, são capazes de produzir ambas as isoformas). Espécies que sofreram perda secundária de um dos sítios foram mantidas. Ademais, como as sequências apresentam éxon 1 mais longo ou mais curto de acordo com o códon utilizado para início da tradução, a extremidade 5' das CDSs de diversas espécies foi cortada de forma que começassem todas exatamente no mesmo ponto.

### 3.4. Investigação de Parceiros Moleculares

Foi realizada uma pesquisa minuciosa na literatura, com auxílio de ferramentas de busca como o PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>) e o Google Scholar (<https://scholar.google.com/>), a respeito das moléculas capazes de interagir com a DACT1. O objetivo foi encontrar parceiros moleculares relacionados à região submetida a *splicing* alternativo.

## 4. RESULTADOS

Em muitos casos, nenhuma das plataformas de busca apresentou informações coerentes a respeito de determinada espécie, seja porque ainda não há montagem genômica disponível para a mesma, porque a montagem está incompleta (presença de *gaps* ou de nucleotídeos não determinados) e não contempla ou contempla apenas parcialmente a região onde o *Dact1* está localizado ou porque a montagem ainda não foi anotada. Em uma situação bastante curiosa, verificou-se que o gene está ausente dos genomas de dois actinopterígios da família Tetraodontidae: *Tetraodon nigroviridis* e *Takifugu rubripes*, conhecidos popularmente como baiacus. Essa informação corrobora os dados de Schubert et al, 2014, que já havia constatado a perda secundária do *Dact1* nesse grupo. Tanto essas duas espécies quanto as que encaixaram-se em qualquer uma das circunstâncias anteriormente descritas foram descartadas (Apêndice 3), restando apenas 95 ao final da triagem.

O tamanho das CDSs do *Dact1* não variou de forma muito considerável (Apêndice 2), sendo a média de pares de base igual a 2476,32 e o desvio padrão aproximadamente 120. As proteínas geradas por essas CDSs, por sua vez, apresentaram média de aminoácidos igual a 823,45 e desvio padrão de aproximadamente 40.

As sequências equivalentes à região que sofre potencial *splicing* alternativo foram encontradas em todas as 95 espécies, sendo que em 36 delas a distinção era apresentada pelas anotações automáticas das próprias ferramentas de busca. Já nas demais 59 essas sequências foram identificadas manualmente, e em 10 delas foi necessária a consulta do íntron antecessor ao éxon 4. Dessas dez, 7 não tiveram suas sequências acrescidas à CDS original (e nem à sequência polipeptídica) pois possuíam códons de parada nessa região, caracterizando-as como não funcionais (Figura 8). A localização da sequência de interesse nas CDSs dos actinopterígios foi definida somente após alinhamento, pois a região é bastante divergente

nessas espécies, tanto entre elas mesmas quanto ao serem comparadas às dos outros grupos. Mais informações estão disponíveis no Apêndice 1.

Com relação à presença do sítio acceptor alternativo (AG) envolvido na coordenação do evento de *splicing* responsável pela síntese da isoforma  $\beta$ , verificou-se que ele aparece pela primeira vez nos anfíbios (estando presente em ambos os homólogos do *Xenopus laevis*, *DAPPER* e *FRODO*) e se mantém conservado ao longo de quase toda a filogenia dos vertebrados, com perdas pontuais no anuro *Nanorana parkeri*, no lagarto *Anolis carolinensis*, na tuatara *Sphenodon punctatus* e no monotremado *Ornithorhynchus anatinus*. Em sete espécies, sendo a maioria delas representante do clado Cetartiodactyla, mutações pontuais do tipo “sem sentido” ocorreram na região que sofria *splicing* alternativo e geraram códons de parada que impossibilitam a síntese da isoforma  $\alpha$  (Figura 8). Dessa forma, o sítio acceptor encarregado pela produção da isoforma  $\beta$  tornou-se constitutivo.

> <i>Elephantulus edwardii</i>			* Cetartiodactyla
MFLCSHSRVGL#SRPLWAPASVFPFCTPQFNSLDFVA			
> <i>Capra hircus</i>			> <i>Tursiops truncatus</i>
DLIG#LEYKEGHCEDQASRAVDCSLSTPQSNFLAVIA		L#HVFRIS#DGWNIKKATVAVGRSPSTSQSNLGVIA	
	*		
> <i>Ovis aries</i>		> <i>Bubalus bubalis</i>	
DLIG#LEYKEGHCEDQASRSVDCSLSTPQSNFLAVIA		DLIGWLEYKEGHCEDQASGAVDCSPST#SNLAVIA	
> <i>Lagenorhynchus obliquidens</i>		> <i>Cavia porcellus</i>	
L#NVFQIS#DGWNIKKATVAVGRSPSTSQSNLGVIA		KSHRMREYKKGCCEDQGSGVMCCSPSTL#FILLDLIA	

**FIGURA 8 - Códons de parada encontrados na região sob potencial *splicing* alternativo**

Fonte: O AUTOR. Códons de parada destacados em vermelho.

Alinhamentos gerais (Apêndices 4 e 5) e clado-específicos das sequências proteicas das espécies envolvidas no estudo permitiram a observação de assinaturas moleculares de grupo e sítios conservados em maior ou menor grau na região do potencial *splicing* alternativo. Devido à discrepância das sequências pertencentes aos peixes actinoptérios quando comparadas às das demais espécies, dois alinhamentos gerais distintos foram produzidos: um contendo o clado em questão (Apêndice 4) e outro excluindo-o (Apêndice 5). Quando os Actinopterygii são desconsiderados, nota-se que três aminoácidos apresentam elevado nível de conservação ao longo de toda a filogenia dos vertebrados. São eles um ácido aspártico (em geral na primeira ou segunda posição), uma alanina (sempre na última posição)

e uma serina (em geral na posição 23). Esses aminoácidos estão presentes em 100%, 95% e 90% das espécies analisadas, respectivamente. Outros 14 sítios se mantêm razoavelmente constantes (mínimo de 80%) e podem ser consultados no Apêndice 5. Em contraste, se os actinopterígios são levados em conta e os mesmos parâmetros são aplicados, apenas 7 sítios além dos três principais se destacam como conservados (Apêndice 4).

Uma parcela considerável das alterações que geraram as discrepâncias observadas nas sequências dos peixes de nadadeiras raiadas provavelmente ocorreram ao longo da história evolutiva do próprio grupo, pois um alinhamento comparativo entre Chondrichthyes, Actinopterygii, e Actinistia (Figura 9) demonstra que as sequências da região exclusiva à isoforma  $\alpha$  dos peixes cartilagosos são mais parecidas com a do celacanto do que com as pertencentes aos actinopterígios. Essa observação é corroborada pela árvore filogenética produzida a partir desse alinhamento (Apêndice 6) e pelas pontuações fornecidas pelo ClustalW, as quais são mais altas para comparações das sequências A e B (*Callorhinchus milii* e *Rhincodon typus*, respectivamente) com C (*Latimeria chalumnae*) do que com qualquer outra (média 44). Mais acerca do tema será abordado na seção “Discussão”.

```

CLUSTAL 2.1 multiple sequence alignment

a. Callorhinchus      EDLIGWRECGEGLL----DYHYVGAVRRSFSAPYSSSLDVVT
b. Rhincodon         EDFIGWRDHGEGLL----DCQYPGIVRRSFSAPYSSSVDAV
c. Latimeria         DDLVSWTEYEGRS----EESNSSTVRRSFSGPYSNSADVVA
d. Lepisosteus      DELVGCADC-EGRC----EDQSSVLVRRSFSAPYSPVDGTP
e. Electrophorus    DELGVCAECEGHCE----DSN-SGTVRRSLSAPYSPSPDGSY
f. Ictalurus        DELNVCAECEEQCE----DSC-SGTVRRSLSVPYSPSPDGNC
g. Danio            DDPGSCAECDCHE----DSS-SGTVRRSLSASYSPPDSSC
h. Carassius       DELGSCAECDCHE----DSS-SGTVRRSLSASYSPPDGS
i. Salmo            DELGSCAECENQCA----EQQTIGTVRKSLSLTPYSPSTDASS
j. Amphiprion      DELAGCLQCDRLGGLCDDSSSTTVRRSLSAPQPAMDADT-
k. Poecilia        EELASCLEYDVLVGGFYDDLGS SGVRRSLSAPHQASLEATS
l. Oreochromis     DELASCLECDVLVGGLCDESSSGTVRCSLSAPNPPTVDAAS
      ::      :      :      ** *:* .

```

**FIGURA 9 - Alinhamento comparativo da região de interesse em Chondrichthyes, Actinopterygii, e Actinistia**

Fonte: O AUTOR. A e B = Chondrichthyes; C = Actinistia; D-L = Actinopterygii. Destaques em amarelo são aminoácidos conservados entre Chondrichthyes e Actinistia, mas não entre Chondrichthyes e Actinopterygii.

Como mencionado, foi possível identificar assinaturas moleculares na região sob potencial *splicing* alternativo em alguns clados, ou seja, aminoácidos (ou sequências de aminoácidos) comuns às espécies de determinado grupo e que não ocorrem nos demais. Essas







canônica de maneira bastante semelhante ao TCF3 (GAO et al, 2008). Embora esse sequestro ocorra através da ligação da DACT1 à HDAC1, os estudos de Gao et al, 2008 e Hikasa & Sokol, 2004 encontraram evidências de que a DACT1 também é capaz de interagir diretamente com o LEF1 e com o TCF3, respectivamente. É importante salientar que Hikasa & Sokol trabalharam com *Xenopus laevis*, e a proteína que demonstrou ligar-se ao TCF3 foi a FRODO, sintetizada a partir do homólogo *Dact1b*, que é exclusivo dessa espécie. A enorme semelhança entre as duas versões do gene, no entanto, permite que as funções atribuídas a um sejam extrapoladas ao outro, como observado em Schubert et al, 2014.

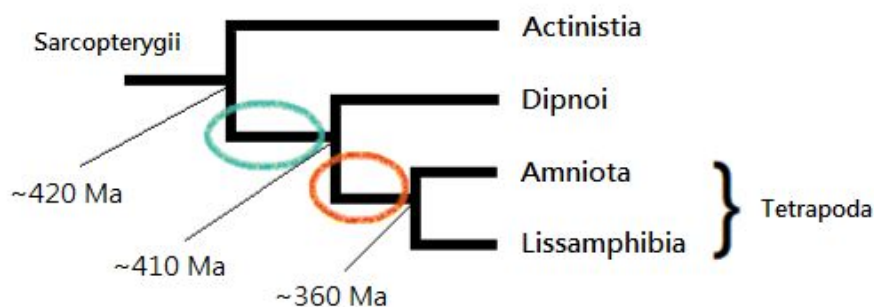
Todas as consequências e inferências evolutivas e funcionais a respeito dos temas aqui levantados serão tratados na seção “Discussão”, a seguir.

## 5. DISCUSSÃO

A presença da região exclusiva à isoforma  $\alpha$  e, concomitantemente, a ausência de sítio acceptor alternativo de *splicing* nas CDSs das espécies *Callorhinchus milii* e *Rhincodon typus*, representantes do clado Chondrichthyes, indica que a isoforma em questão era a única variante produzida pelo gene *Dact1* ancestral. O surgimento da isoforma  $\beta$  provavelmente se deu no início da linhagem dos Tetrapoda, antes do evento de cladogênese responsável pela divergência entre amniotos e lissanfíbios, que estima-se ter ocorrido há aproximadamente 360 milhões de anos (KUMAR & HEDGES, 1998). Esse fato é corroborado quando se verifica que o sítio acceptor alternativo permanece ausente em sarcopterígio primitivos como o celacanto (*Latimeria chalumnae*) e então aparece em anuros, cecílias, lagartos e cobras, sendo mantido ao longo de todo o restante da árvore filogenética dos amniotos. É igualmente possível que o surgimento do sítio alternativo tenha ocorrido ainda mais cedo, há até cerca de 420 milhões de anos, antes da divergência que gerou as linhagens Tetrapoda e Dipnoi (BENTON et al, 2015). Inferências mais específicas a esse respeito são inviáveis, pois não existem montagens genômicas de peixes dipnóicos disponíveis nos bancos de dados consultados. A Figura 16 ilustra ambas as hipóteses levantadas.

É interessante notar que, ao longo da filogenia, ocorreram em diversos clados mutações pontuais que inviabilizaram a síntese de uma das duas isoformas, seja através de alteração do sítio acceptor responsável pela isoforma  $\beta$  (como visto no anuro *Nanorana parkeri*, por exemplo) ou através do aparecimento de códons de parada na região exclusiva à

isoforma  $\alpha$  (como visto nas espécies apontadas pela Figura 8). Essa observação sugere que nenhuma das duas variantes da proteína DACT1 é absolutamente essencial para o desenvolvimento embrionário e a manutenção de processos biológicos fundamentais (embora ao menos uma delas deva estar presente) e, portanto, possíveis funções atribuídas à região em estudo devem estar relacionadas a ajustes bioquímicos bastante refinados.



**FIGURA 16 - Surgimento da isoforma  $\beta$  da DACT1 na história evolutiva de Vertebrata**

Fonte: O AUTOR, a partir dos dados de KUMAR & HEDGES, 1998 e BENTON et al, 2015. As regiões circuladas correspondem ao trechos da filogenia nos quais o sítio acceptor alternativo pode ter surgido.

Embora as perdas pontuais mencionadas pareçam muitas vezes ocorrer ao acaso, é curioso o surgimento de códons de parada em 5 espécies representantes do clado Cetartiodactyla. Ao que tudo indica, essas alterações foram recentes na história evolutiva do grupo e ocorreram diversas vezes, de forma independente e em linhagens distintas. Dentro da família Delphinidae, por exemplo, observa-se que os golfinhos *Lagenorhynchus obliquidens* e *Tursiops truncatus* possuem os códons ocre e âmbar em posições idênticas, enquanto que a mesma região na espécie *Ornicus orca* apresenta-se completamente funcional. A partir disso, supõe-se que as mutações que deram origem a esses códons ocorreram em um ancestral comum aos gêneros das duas primeiras espécies, mas não ao da terceira. Já na família Bovidae é provável que a perda de funcionalidade da isoforma  $\alpha$  tenha se dado pelo menos duas vezes: em um ancestral comum à subfamília Caprinae (pois um códon âmbar foi encontrado na mesma posição em *Ovis aries* e *Capra hircus*) e em um ancestral exclusivo ao gênero *Bubalus* ou à espécie *Bubalus bubalis*. Essa última afirmação pode ser explicada pelo fato de que não há códon de parada em *Bos taurus* e pelo entendimento de que os gêneros *Bos*, *Bison* e *Bubalus* formam um clado monofilético (ZURANO et al, 2019; BIBI, 2013).

Inferências mais detalhadas e consistentes a respeito da história evolutiva das isoformas  $\alpha$  e  $\beta$  em Cetartiodactyla são difíceis devido ao reduzido número de espécies do grupo que apresentam montagens genômicas completas e anotadas. Ainda assim, os dados coletados nesse trabalho indicam que pode ter havido uma pressão de seleção positiva ao longo da evolução do clado no sentido de fixar o aparecimento de códons de parada na região exclusiva à isoforma  $\alpha$ , impossibilitando a síntese dessa variante da DACT1. Mais estudos são necessários para que seja avaliado o significado evolutivo da característica mencionada.

Alguns outros grupos também se destacam pela peculiaridade das sequências referentes à região sob potencial *splicing* alternativo, mas um caso bastante curioso é o de Actinopterygii, no qual as sequências divergem substancialmente também entre si. Uma possível explicação para essa diversidade é a duplicação genômica adicional que ocorreu no início da história evolutiva dos teleósteos, há cerca de 320 milhões de ano, além de outras mais recentes que foram detectadas em algumas linhagens específicas dentro do mesmo clado. Acredita-se que esses eventos de duplicação genômica completa tenham sido motor da rápida especiação e radiação adaptativa sofridas pelos teleósteos, que resultou em quase 30 mil espécies viventes atualmente (RAVI & VENKATESH, 2018). No alinhamento comparativo produzido entre Chondrichthyes e Actinopterygii e Actinistia (Figura 9), contudo, a sequência do holósteo *Lepisosteus oculatus* apresenta mais semelhança àquelas referentes aos teleósteos do que às referentes aos peixes cartilagosos (pontuações médias 29 e 55, respectivamente, fornecidas pelo ClustalW). Esses dados sugerem que uma parcela das mudanças que tornaram a região exclusiva à isoforma  $\alpha$  tão diferente em actinopterígios já haviam surgido antes do evento de cladogênese que separou as linhagens Holostei e Teleostei. Mais do que isso: é provável que tenham aparecido em um ancestral comum entre Actinopterygii e Sarcopterygii, uma vez que o alinhamento entre *Lepisosteus oculatus* e *Latimeria chalumnae* (Actinistia) apresenta pontuação 54 (quase o dobro de 29).

A partir de Tetrapoda, as sequências começam a ficar mais consistentes e a quantidade de sítios conservados dentro de cada grupo é, de forma geral, bastante elevada. No que diz respeito a assinaturas moleculares, no entanto, o único que se destaca é Serpentes (Figura 11, molduras retangulares). Isso ocorre porque a maioria dos sítios se mantém conservados de maneira consistente ao longo de toda a árvore filogenética dos vertebrados, não sendo exclusivos de nenhum clado específico (Apêndice 5). Existem também posições que são representadas por 2 ou 3 aminoácidos que se alternam durante a história evolutiva de tal

maneira que, sem uma análise bastante aprofundada, parece incoerente. Talvez assinaturas de grupo seriam mais facilmente encontradas em táxons menos amplos do que aqueles estabelecidos para esse estudo (como é o caso de Serpentes e Anura, nos quais de fato foram identificados padrões interessantes), mas a reduzida quantidade de espécies a que tivemos acesso limita muito esse tipo de análise, sendo um impedimento à busca de assinaturas moleculares até ao lidar com clados grandes como Perissodactyla e Lagomorpha.

Independentemente disso, a análise de substituições sinônimas e não-sinônimas mostrou que a região exclusiva à isoforma  $\alpha$  sofre, em média, mais mutações do segundo tipo em comparação ao restante da CDS do gene (o dobro), que por sua vez apresenta uma proporção de substituições sinônimas bem maior. Esses dados estão de acordo com a ideia de que regiões submetidas a *splicing* alternativo tendem a evoluir mais rapidamente do que regiões constitutivas do genoma, uma vez que alterações de aminoácidos causam alterações tanto na estrutura proteica quanto, possivelmente, em sua conformação tridimensional e atividade. A maior ocorrência de substituições sinônimas na CDS constitutiva pode ser reflexo de uma necessidade de conservação mais intensa nessas regiões, devido a suas importâncias estrutural e funcional. Contudo, embora esse padrão também tenha sido relatado no trabalho de Xing & Lee, 2005, ele não parece ser o mais comum, pois ao estudar sequências subordinadas a *splicing* alternativo de mais de 3 mil genes, Ermakova et al, 2006 verificou aumentos nas proporções tanto de mutações sinônimas quanto não-sinônimas. Sob esse ponto de vista, portanto, o *Dact1* apresenta-se como uma exceção.

O possível envolvimento da região exclusiva à isoforma  $\alpha$  na interação entre a DACT1 e o fator de transcrição TCF3 é bastante interessante. No estudo de Hikasa & Sokol, 2004, sugere-se que essa ligação estaria relacionada à ativação da expressão gênica direcionada dependente de TCF3, o que contradiz as observações de Gao et al, 2008 e Yau et al, 2005, onde verificou-se o papel da DACT1 no núcleo é impedir a via Wnt canônica. Nesse último trabalho, especificamente, é demonstrado que a redução na síntese da isoforma  $\alpha$  gera acúmulo de  $\beta$ -catenina citoplasmática em células cancerosas. Inferir funções sem ensaios bioquímicos específicos é muito difícil para uma proteína como a DACT1, cuja atividade depende do estado seu estado de fosforilação e pode ser contexto/tecido-específica. Além disso, sabe-se que há uma certa redundância funcional entre os parálogos da família DACT (KIVIMAE, YANG & CHEYETTE, 2011), o que pode levar a conclusões incorretas.

De qualquer forma, é muito provável que a interação entre a DACT1 e o TCF3 esteja envolvida em ajustes da via de sinalização Wnt canônica a nível nuclear, assim como é o caso de ligações com proteínas como a HDAC1 e a LEF1. O sítio exato através do qual essa interação ocorre ainda não é conhecido, mas a região que o abrange também contém a sequência submetida a *splicing* alternativo, que apresenta evidências de evolução acelerada. Assim, sugere-se que as substituições de aminoácidos que ocorreram nessa sequência ao longo da história evolutiva dos vertebrados podem ter gerado alterações nas propriedades da ligação entre a DACT1 e o TCF3, conseqüentemente ajustando-a de maneira refinada e específica. A ausência da isoforma  $\alpha$  pode ser tanto um impedimento para a interação como pode também apenas ocasionar ajustes, uma vez que não se sabe se o sítio ativo localiza-se de fato na região sob *splicing* alternativo ou somente próxima a ela. Mais estudos são necessários para que seja compreendida a relevância desses ajustes para a cascata bioquímica em curso.

## 6. CONCLUSÃO

Conclui-se que a isoforma ancestral da DACT1 é a  $\alpha$  (maior), tendo o sítio acceptor alternativo responsável pela síntese da isoforma  $\beta$  surgido em um ancestral comum a todos os Tetrapoda. Em algumas espécies, esse sítio foi secundariamente perdido e, em outras, tornou-se constitutivo (como é o caso de diversos representantes do clado Cetartiodactyla). Também verificou-se que a região que sofre *splicing* alternativo apresenta evidências de evolução acelerada, a qual pode estar envolvida em ajustes funcionais no que diz respeito à interação entre a DACT1 e fatores de transcrição como o TCF3 e outras proteínas nucleares responsáveis pela regulação da via Wnt canônica.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ARGUELO, Annie et al. **Dapper antagonist of catenin-1 cooperates with Dishevelled-1 during postsynaptic development in mouse forebrain GABAergic interneurons.** PLoS One, v.8, n.6, 2013.

AULEHLA, Alexander et al. **Wnt3a plays a major role in the segmentation clock controlling somitogenesis.** Developmental Cell, v.4, p.395-406, 2003.

BARALLE, Francisco E.; GIUDICE, Gimena. **Alternative splicing as a regulator of development and tissue identity.** Nature Reviews - Molecular Cell Biology, v. 18, n. 7, p. 437-451, 2017.

BENTON, Michael J. et al. **Constraints on the timescale of animal evolutionary history.** Paleontologia Electronica, v.18, n.1, p.1-116, 2015.

BIBI, Faysal. **A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics.** BMC Evolutionary Biology, v.13, n.166, 2013.

CARVALHO, Hernandes F.; RECCO-PIMENTEL, Shirley Maria. **A Célula.** 3.ed. Barueri: Manole, 2013.

CHEYETTE, Benjamin N. R. et al. **Dapper, a Dishevelled-associated antagonist of  $\beta$ -catenin and JNK signaling, is required for notochord formation.** Developmental Cell, v.2, p.449-461, 2002.

CLANCY, Suzanne. **RNA splicing: introns, exons and spliceosome.** Nature Education, v.1, n.10, 2008.

DAULAT, Avais M.; BORG, Jean-Paul. **Wnt/planar cell polarity signaling: new opportunities for cancer treatment.** Trends in Cancer, v.3, n.2, p.113-125, 2017.

ERMAKOVA, NURTDINOV & GELFAND, 2006

FISHER, Daniel A. et al. **Three Dact gene family members are expressed during embryonic development and in the adult brains of mice.** Developmental Dynamics, v.235, p.2620-2630, 2006.

GAO, Xia et al. **Dapper1 is a nucleocytoplasmic shuttling protein that negatively modulates Wnt signaling in the nucleus.** The Journal of Biological Chemistry, v.283, n.51, p.35679-35688, 2008.

GAO, Jinlan et al. **Cyclin G2 suppresses Wnt/ $\beta$ -catenin signaling and inhibits gastric cancer cell growth and migration through Dapper1.** Journal of Experimental & Clinical Cancer Research, v.37, n.317, 2018.

GILBERT, Scott F. **Developmental Biology**. 9.ed. Sunderland: Sinauer Associates, Inc., 2010.

GLOY, Joachim; HIKASA, Hiroki; SOKOL, Sergei Y. **Frodo interacts with Dishevelled to transduce Wnt signals**. *Nature Cell Biology*, v.4, p.351-357, 2002.

HAGENMUELLER, Marco et al. **Dapper-1 induces myocardial remodeling through activation of canonical Wnt signaling in cardiomyocytes**. *Hypertension*, v.61, n.6, p.1177-1186, 2013.

HAGENMUELLER, Marco et al. **Dapper-1 is essential for Wnt5a induced cardiomyocyte hypertrophy by regulating the Wnt/PCP pathway**. *FEBS Letters*, v.588, n.14, p.2230-2237, 2014.

HIKASA, Hiroki; SOKOL, Sergei Y. **The involvement of Frodo in TCF-dependent signaling and neural tissue development**. *Development*, v.131, n.19, p.4725-4734, 2004.

HOU, Jian et al. **DACT1 is involved in human placenta development by promoting Wnt signaling**. *Archives of Gynecology and Obstetrics*, v.291, n.6, p.1289-1296, 2014.

JARDIM, Daniele P.; POÇO, Paula C.E.; CAMPOS, Alexandre H. **Dact1, a Wnt-pathway inhibitor, mediates human mesangial cell TGF- $\beta$ 1-induced apoptosis**. *Journal of Cellular Physiology*, v.232, n.8, p.2104-2111, 2017.

JIAO, Shujie et al. **miR-124 promotes proliferation and neural differentiation of neural stem cells through targeting DACT1 and activating Wnt/ $\beta$ -catenin pathways**. *Molecular and Cellular Biochemistry*, v.449, n.1-2, p.305-314, 2018.

KETTUNEN, Paivi et al. **Dact1-3 mRNAs exhibit distinct expression domains during tooth development**. *Gene Expression Patterns*, v.10, n.2-3, p.140-143, 2010.

KIVIMAE, Saul; YANG, Xio Y.; CHEYETTE, Benjamin N. R. **All Dact (Dapper/Frodo) scaffold proteins dimerize and exhibit conserved interactions with Vangl, Dvl, and serine/threonine kinases**. *BMC Biochemistry*, v.12, n.33, 2011.

KLUG, William S. **Concepts of Genetics**. 10.ed. San Francisco: Pearson Benjamin Cummings, 2012.

KOMIYA, Yuko; HABAS, Raymond. **Wnt signal transduction pathways**. *Organogenesis*, v.4, n.2, p.68-75, 2008.

KUANG, Jian-Ren et al. **Dapper1 attenuates hepatic gluconeogenesis and lipogenesis by activating PI3K/Akt signaling**. *Molecular and Cellular Endocrinology*, v.447, p.106-115, 2017.

KUMAR, Sudhir; HEDGES, Blair. **A molecular timescale for vertebrate evolution**. *Nature*, v.392, n.6679, p.917-920, 1998.

LEVINE, Aaron. **Bioinformatics Approaches to RNA Splicing**. Tese (Master of Philosophy) - The Sanger Centre, University of Cambridge. Hinxton, 2001.

MOON, Randall T. et al. **Wnt and  $\beta$ -catenin signalling: diseases and therapies**. Nature Reviews Genetics, v.5, n.9, p.691-701, 2004.

NELSON, David L.; COX, Michael M. **Lehninger Principles of Biochemistry**. 6.ed. W. H. Freeman, 2013.

NILSEN, Timothy W.; GRAVELEY, Brenton R. **Expansion of the eukaryotic proteome by alternative splicing**. Nature, v.463, n.7280, p.457-63, 2010.

NUSSE, Roel; CLEVERS, Hans. **Wnt/b-catenin signaling, disease, and emerging therapeutic modalities**. Cell, v.169, n.6, p.985-999, 2017.

OKERLUND, Nathan D. et al. **Dact1 is a postsynaptic protein required for dendrite, spine, and excitatory synapse development in the mouse forebrain**. The Journal of Neuroscience, v.30, n.12, p.4362-4368, 2010.

PYRON, R. Alexander; BURBRINK, Frank T.; WIENS, John J. **A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes**. BMC Evolutionary Biology, v.13, n.93, 2013.

PYRON, R. Alexander; WIENS, John J. **A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians**. Molecular Phylogenetics and Evolution, v.61, n.2, p.543-583, 2011.

RABADÁN, M. Angeles et al. **Delamination of neural crest cells requires transient and reversible Wnt inhibition mediated by DACT1/2**. Development, v.143, n.12, p.2194-2205, 2016.

RAVI, Vydianathan; VENKATESH, Byrappa. **The divergent genomes of teleosts**. Annual Review of Animal Biosciences, v.6, p.47-68, 2018.

SCHUBERT, Frank R. et al. **Dact genes are chordate specific regulators at the intersection of Wnt and Tgf- $\beta$  signaling pathways**. BMC Evolutionary Biology, v.14, n.157, 2014.

SURIBEN, Rowena; FISHER, Daniel A.; CHEYETTE, Benjamin N.R. **Dact1 presomitic mesoderm expression oscillates in phase with Axin2 in the somitogenesis clock of mice**. Developmental Dynamics, v.235, n.11, p.3177-3183, 2006.

SURIBEN, Rowena et al. **Posterior malformations in Dact1 mutant mice arise through misregulated Vangl2 at the primitive streak**. Nature Genetics, v.41, n.9, p.977-985, 2009.

SUZUKI, Daisuke et al. **Expression analysis of Dact1 in mice using a LacZ reporter**. Gene Expression Patterns, v.15, n.1, p.21-30, 2014.

TERAN, Evelyn; BRANSCOMB, Aron D.; SEELING, Joni M. **Dpr acts as a molecular switch, inhibiting Wnt signaling when unphosphorylated, but promoting Wnt signaling when phosphorylated by Casein Kinase I $\delta$ / $\epsilon$ .** PLoS One, v.4, n.5, 2009.

WANG, Yan et al. **Mechanism of alternative splicing and its regulation (Review).** Biomedical Reports, v.3, n.2, p.150-158, 2014.

WAXMAN, Joshua S. et al. **Zebrafish Dapper1 and Dapper2 play distinct roles in Wnt-mediated developmental processes.** Development, v.131, n.23, p.5909-5921, 2004.

WEN, Jun et al. **Loss of Dact1 disrupts planar cell polarity signaling by altering dishevelled activity and leads to posterior malformation in mice.** The Journal of Biological Chemistry, v.285, n.14, p.11023-11030, 2010.

XING, Qiong et al. **Genetic analysis of DACT1 in 100 Chinese Han women with Müllerian duct anomalies.** Reproductive Biomedicine Online, v.32, n.4, p.420-426, 2016.

YAU, Tai-On et al. **HDPRI, a novel inhibitor of the WNT/b-catenin signaling, is frequently downregulated in hepatocellular carcinoma: involvement of methylation-mediated gene silencing.** Oncogene, v.24, n.9, p.1607-1614, 2005.

ZHANG, Long et al. **Dapper 1 antagonizes Wnt signaling by promoting Dishevelled degradation.** The Journal of Biological Chemistry, v.281, n.13, p.8607-8612, 2006.

ZHU, Ke et al. **DACT1 overexpression inhibits proliferation, enhances apoptosis, and increases daunorubicin chemosensitivity in KG-1 $\alpha$  cells.** Tumor Biology, v.39, n.10, 2017.

ZURANO, Juan P. et al. **Cetartiodactyla: Updating a time-calibrated molecular phylogeny.** Molecular Phylogenetics and Evolution, v.133, p.256-262, 2019.

## APÊNDICE I - Espécies Analisadas e Códigos de Acesso

CLADOS		ESPÉCIES	Dact1		
			NCBI	Ensembl (releases 94/95/96)	
Códigos de Acesso					
CHONDRICHTHYES		<i>Callorhynchus milii</i> *	103178514	-----	
		<i>Rhincodon typus</i> *	109928507	-----	
ACTINOPTERYGII	HOLOSTEI	<i>Lepisosteus oculatus</i> *	102691086	ENSLOC00000009312	
	TELEOSTEI	<i>Danio rerio</i> *	405799	ENSDARG00000101635	
		<i>Oreochromis niloticus</i> *	100700415	ENSONIG00000011734	
		<i>Electrophorus electricus</i> *	113579357	-----	
		<i>Salmo salar</i> *	106602221	-----	
		<i>Amphiprion ocellaris</i> *	111585187	ENSAOCG00000004652	
		<i>Ictalurus punctatus</i> *	108270448	ENSIPUG00000001860	
		<i>Carassius auratus</i> *	113117655	-----	
		<i>Poecilia formosa</i> *	103156184	ENSPFOG00000023754	
ACTINISTIA		<i>Latimeria chalumnae</i> *	2353319	ENSLACG00000018807	
SARCOPTERYGII	LISSAMPHIBIA	<i>Rhinatrema bivittatum</i>	115089854	-----	
		<i>Nanorana parkeri</i> *	108792045	-----	
		<i>Xenopus laevis (homolog A)</i>	398315	-----	
		<i>Xenopus laevis (homolog B)</i>	399186	-----	
		<i>Xenopus tropicalis</i>	493325	ENSXETG000000011962	
	SAUROPSIDA	RHYNCHOCEPHALIA	<i>Sphenodon punctatus</i> *	-----	ENSSPUG000000018322
		SQUAMATA	<i>Gekko japonicus</i>	107120639	-----
			<i>Anolis carolinensis</i> *	100555885	ENSACAG000000014812
			<i>Python bivittatus</i>	103057879	-----
			<i>Pseudonaja textilis</i>	113434940	-----
			<i>Protobothrops mucrosquamatus</i>	107292182	-----
			<i>Thamnophis sirtalis</i>	106550598	-----
		TESTUDINES	<i>Pelodiscus sinensis</i>	102444102	ENSPSIG000000000545
			<i>Terrapene mexicana triunguis</i>	112125543	-----
			<i>Chrysemys picta</i>	101945567	ENSCPBG000000024848
		CROCODYLIA	<i>Alligator sinensis</i>	102374409	-----
			<i>Alligator mississippiensis</i>	102561856	-----
			<i>Crocodylus porosus</i>	109320887	ENSCPRG00000005017955
		AVES	<i>Anas platyrhynchos</i>	101795068	ENSAPLG000000007362
			<i>Numida meleagris</i>	110401514	ENSNMEG000000016998
	<i>Gallus gallus</i>		723789	ENSGALG000000012023	
	<i>Pygoscelis adeliae</i>		103923078	-----	
	<i>Sturnus vulgaris</i>		106850230	-----	
	<i>Empidonax traillii</i>		114067303	-----	
	<i>Lonchura striata domestica</i>		110479635	ENLSLDG000000007102	
	<i>Taeniopygia guttata</i>		100221895	ENSTGUG000000013023	
	MONOTREMATA		<i>Ornithorhynchus anatinus</i> *	100086796	-----
	MARSUPIALIA	<i>Phascolarctos cinereus</i>	110208422	ENSPCIG000000017476	
		<i>Sarcophilus harrisii</i>	100926607	ENSSHAG000000018459	
		<i>Monodelphis domestica</i>	100021268	ENSMODG000000008148	
XENARTHRA	<i>Dasyus novemcinctus</i>	101426347	ENSDNOG000000000222		
AFROTHERIA	<i>Loxodonta africana</i>	100669653	ENSLAFG000000007484		
	<i>Elephantulus edwardii</i> #	102867458	-----		
EULIPOTYPHLA	<i>Condylura cristata</i>	101619054	-----		
	<i>Erinaceus europaeus</i>	103110097	ENSEEUG000000009193		
PERISSODACTYLA	<i>Equus caballus</i>	100059241	ENSECAG000000009167		
	<i>Equus asinus</i>	106845848	ENSEASG00000005016185		

(Continua)

SARCOPTERYGII	TETRAPODA	MAMMALIA	CARNIVORA	<i>Acinonyx jubatus</i>	106974570	-----
				<i>Felis catus</i>	101098146	ENSCFAG00000045905
				<i>Vulpes vulpes</i>	112921298	-----
				<i>Canis lupus familiaris</i>	100687551	ENSCAFG00000015431
				<i>Ursus arctos horribilis</i>	113242317	-----
				<i>Mustela putorius furo</i>	101687549	ENSMPUG00000006171
				<i>Odobenus rosmarus</i>	101373515	-----
			CETARTIODACTYLA	<i>Delphinapterus leucas</i>	111173108	-----
				<i>Physeter catodon</i>	102979852	-----
				<i>Orcinus orca</i>	101289587	-----
				<i>Lagenorhynchus obliquidens #</i>	113622812	-----
				<i>Tursiops truncatus #</i>	101315540	ENSTRTRG00000008825
				<i>Sus scrofa</i>	100621416	ENSSSCG00000031856
				<i>Ovis aries #</i>	101116320	ENSOARG00000021114
				<i>Capra hircus #</i>	102175028	ENSCHIG00000020832
				<i>Bubalus bubalis #</i>	102391933	-----
				<i>Bos taurus</i>	538778	ENSBTAG00000019421
			CHIROPTERA	<i>Rousettus aegyptiacus</i>	107502131	-----
				<i>Pteropus vampyrus</i>	105308851	ENSPVAT00000004564
				<i>Phyllostomus discolor</i>	114492869	-----
				<i>Hipposideros armiger</i>	109383892	-----
				<i>Myotis lucifugus</i>	102418277	ENSMLUG00000007848
			LAGOMORPHA	<i>Ochotona princeps</i>	101519084	-----
				<i>Oryctolagus cuniculus</i>	100339172	ENSOCUG00000015140
			RODENTIA	<i>Cavia porcellus #</i>	100728819	ENSCPOG00000039410
				<i>Ictidomys tridecemlineatus</i>	101963894	ENSSTOG00000028057
				<i>Marmota marmota</i>	107138321	-----
				<i>Jaculus jaculus</i>	101608927	ENSJJAG00000023135
				<i>Rattus norvegicus</i>	500666	ENSRNOG00000008445
				<i>Urocitellus parryii</i>	113197631	ENSUPAG00010002490
				<i>Mus musculus</i>	59036	ENSMUSG00000044548
				<i>Mesocricetus auratus</i>	101836710	ENSMAUG00000008693
				<i>Cricetulus griseus</i>	100768384	ENSCGRG00001013082
			PRIMATES	<i>Otolemur garnettii</i>	100962280	ENSOGAG00000027814
<i>Microcebus murinus</i>	105882914	ENSMICG00000004813				
<i>Callithrix jacchus</i>	100400795	ENSCJAG00000018902				
<i>Rhinopithecus bieti</i>	108524793	ENSRBIG00000029351				
<i>Aotus nancymae</i>	105716364	ENSANAG00000035326				
<i>Chlorocebus sabaues</i>	103229090	ENSCSAG00000014335				
<i>Papio anubis</i>	101017740	ENSPANG00000004888				
<i>Macaca mulatta</i>	701218	ENSMMUG00000001090				
<i>Nomascus leucogenys</i>	100606126	ENSNLEG00000013400				
<i>Pongo abelii</i>	100437942	ENSPPYG00000005861				
<i>Gorilla gorilla</i>	101147080	ENSGGOG00000015389				
<i>Pan troglodytes</i>	452938	ENSPTRG00000006397				
<i>Homo sapiens</i>	51339	ENSG00000165617				

Código	Significado
	Ambas as isoformas são apresentadas pelos bancos de dados
	Adição relativa à isoforma $\alpha$ identificada manualmente.
	Adição relativa à isoforma $\alpha$ encontrada no íntron que precede o último éxon.
	Adição relativa à isoforma $\alpha$ encontrada no íntron que precede o último éxon, mas apresenta códon(s) de parada.
	A informação obtida dessa fonte foi considerada imprecisa, sendo então complementada ou simplesmente descartada.
#	O sítio acceptor alternativo <i>upstream</i> está ausente, impossibilitando a síntese da isoforma $\alpha$ .
*	O sítio acceptor alternativo <i>downstream</i> está ausente, impossibilitando a síntese da isoforma $\beta$ .

## APÊNDICE II - Dados sobre as Sequências Coletadas

Espécie	Dact1				Isoformas produzidas	
	Fita	Éxons	CDS (pb)	Proteína (aa)	Qtd	Tamanhos (aa)
<i>Callorhynchus milii</i> *	(+)	4	2479	825	1	825
<i>Rhincodon typus</i> *	(+)	4	2497	831	1	831
<i>Lepisosteus oculatus</i> *	(-)	4	2476	824	1	824
<i>Danio rerio</i> *	(-)	4	2467	821	1	821
<i>Oreochromis niloticus</i> *	(-)	4	2503	833	1	833
<i>Electrophorus electricus</i> *	(+)	4	2491	829	1	829
<i>Salmo salar</i> *	(+)	4	2659	885	1	885
<i>Amphiprion ocellaris</i> *	(+)	4	2404	800	1	800
<i>Ictalurus punctatus</i> *	(-)	4	2515	837	1	837
<i>Carassius auratus</i> *	(+)	4	2464	820	1	820
<i>Poecilia formosa</i> *	(+)	4	2560	852	1	852
<i>Latimeria chalumnae</i> *	(-)	4	2449	815	1	815
<i>Rhinatrema bivittatum</i>	(-)	4	2422	806	2	769/806
<i>Nanorana parkeri</i> *	(+)	4	2446	814	1	814
<i>Xenopus laevis (homolog A)</i>	(+)	4	2458	818	2	783/818
<i>Xenopus laevis (homolog B)</i>	(-)	4	2476	824	2	789/824
<i>Xenopus tropicalis</i>	(-)	4	2476	824	2	789/824
<i>Sphenodon punctatus</i> *	(-)	4	2503	833	2	795/883
<i>Gekko japonicus</i>	(+)	5	2263	753	2	715/753
<i>Anolis carolinensis</i> *	(+)	4	2389	795	1	795
<i>Python bivittatus</i>	(-)	5	2497	831	2	794/831
<i>Pseudonaja textilis</i>	(-)	4	2512	836	2	801/836
<i>Protobothrops mucrosquamatus</i>	(-)	4	2401	799	2	764/799
<i>Thamnophis sirtalis</i>	(+)	5	2443	813	2	778/813
<i>Pelodiscus sinensis</i>	(+)	3	2203	733	2	695/733
<i>Terrapene mexicana triunguis</i>	(-)	4	2494	830	2	796/830
<i>Chrysemys picta</i>	(+)	4	2488	828	2	794/828
<i>Alligator sinensis</i>	(-)	4	2194	730	2	692/730
<i>Alligator mississippiensis</i>	(-)	5	2338	778	2	741/778
<i>Crocodylus porosus</i>	(+)	5	2401	799	2	764/799
<i>Anas platyrhynchos</i>	(+)	4	2512	835	2	797/835
<i>Numida meleagris</i>	(-)	4	2464	820	2	782/820
<i>Gallus gallus</i>	(-)	4	2470	822	2	784/822
<i>Pygoscelis adeliae</i>	(-)	6	2209	735	2	697/735
<i>Sturnus vulgaris</i>	(-)	4	2509	835	2	797/835
<i>Empidonax traillii</i>	(-)	5	2491	829	2	791/829
<i>Lonchura striata domestica</i>	(-)	4	2506	834	2	796/834
<i>Taeniopygia guttata</i>	(-)	4	2470	822	2	784/822
<i>Ornithorhynchus anatinus</i> *	(-)	4	2470	822	1	822
<i>Phascolarctos cinereus</i>	(-)	4	2563	853	2	821/853
<i>Sarcophilus harrisi</i>	(+)	5	2551	849	2	817/849
<i>Monodelphis domestica</i>	(+)	4	2557	851	2	814/851
<i>Dasyurus novemcinctus</i>	(+)	4	2536	844	2	807/844
<i>Loxodonta africana</i>	(+)	4	2515	837	2	800/837
<i>Elephantulus edwardii</i> #	(+)	4	2413	803	1	803
<i>Condylura cristata</i>	(-)	4	2533	843	2	806/843
<i>Erinaceus europaeus</i>	(+)	4	2539	845	2	808/845
<i>Equus caballus</i>	(+)	5	2848	948	2	911/948
<i>Equus asinus</i>	(-)	4	2494	830	2	793/830
<i>Acinonyx jubatus</i>	(+)	4	2524	840	2	803/840
<i>Felis catus</i>	(+)	4	2524	840	2	803/840
<i>Vulpes vulpes</i>	(-)	4	2074	690	2	653/690
<i>Canis lupus familiaris</i>	(+)	4	2584	860	2	823/860

(Continua)

<i>Ursus arctos horribilis</i>	(-)	4	2428	808	2	771/808
<i>Mustela putorius furo</i>	(+)	5	2281	759	2	722/759
<i>Odobenus rosmarus</i>	(-)	4	2539	845	2	808/845
<i>Delphinapterus leucas</i>	(-)	4	2536	844	2	807/844
<i>Physeter catodon</i>	(+)	4	2419	805	2	768/805
<i>Orcinus orca</i>	(-)	4	2526	841	2	809/841
<i>Lagenorhynchus obliquidens #</i>	(+)	4	2425	807	1	807
<i>Tursiops truncatus #</i>	(-)	4	2425	807	1	807
<i>Sus scrofa</i>	(+)	4	2425	807	2	770/807
<i>Ovis aries #</i>	(+)	4	2416	804	1	804
<i>Capra hircus #</i>	(-)	4	2415	804	1	804
<i>Bubalus bubalis #</i>	(-)	4	2416	804	1	804
<i>Bos taurus</i>	(+)	4	2527	841	2	804/841
<i>Rousettus aegyptiacus</i>	(+)	4	2539	845	2	808/845
<i>Pteropus vampyrus</i>	(-)	4	2539	845	2	808/845
<i>Phyllostomus discolor</i>	(-)	4	2545	847	2	810/847
<i>Hipposideros armiger</i>	(-)	4	2542	846	2	809/846
<i>Myotis lucifugus</i>	(-)	4	2206	734	2	697/734
<i>Ochotona princeps</i>	(-)	4	2575	857	2	820/857
<i>Oryctolagus cuniculus</i>	(+)	4	2536	844	2	807/844
<i>Cavia porcellus #</i>	(-)	4	2086	694	1	694
<i>Ictidomys tridecemlineatus</i>	(+)	4	2518	838	2	801/838
<i>Marmota marmota</i>	(+)	2	2507	838	2	801/838
<i>Jaculus jaculus</i>	(+)	4	2533	843	2	806/843
<i>Rattus norvegicus</i>	(+)	4	2449	815	2	778/815
<i>Urocyon parryi</i>	(-)	4	2512	836	2	800/836
<i>Mus musculus</i>	(+)	4	2449	815	2	778/815
<i>Mesocricetus auratus</i>	(-)	4	2446	814	2	777/814
<i>Cricetulus griseus</i>	(+)	4	2452	816	2	780/816
<i>Otolemur garnettii</i>	(+)	4/5	2476	824	2	788/824
<i>Microcebus murinus</i>	(+)	4	2506	834	2	797/834
<i>Callithrix jacchus</i>	(+)	4	2515	837	2	800/837
<i>Rhinopithecus bieti</i>	(+)	4	2515	837	2	800/837
<i>Aotus nancymae</i>	(+)	4	2512	836	2	799/836
<i>Chlorocebus sabaeus</i>	(+)	4	2998	998	2	961/998
<i>Papio anubis</i>	(+)	4	2515	837	2	800/837
<i>Macaca mulatta</i>	(+)	4	2515	837	2	800/837
<i>Nomascus leucogenys</i>	(+)	4	2515	837	2	800/837
<i>Pongo abelii</i>	(+)	4	2515	837	2	800/837
<i>Gorilla gorilla</i>	(+)	4	2512	836	2	799/836
<i>Pan troglodytes</i>	(+)	4	2512	836	2	799/836
<i>Homo sapiens</i>	(+)	4	2512	836	2	799/836

### APÊNDICE III - Espécies Descartadas

CLADO	ESPÉCIE	Dact1			
		NCBI	Ensembl (release 94/95)		
		Gene ID			
ACTINOPTERYGII	<i>Gadus morhua</i>	?	?		
	<i>Gasterosteus aculeatus</i>	?	?		
	<i>Tetraodon nigroviridis</i>	---	-----		
	<i>Takifugu rubripes</i>	---	-----		
SARCOPTERYGII	LUNGFISH	<i>Protopterus annectens</i>	?	?	
	TESTUDINES	<i>Chelonia mydas</i>	102944538	-----	
		SAUROPSIDA	AVES	<i>Struthio camelus</i>	104141427
	<i>Calypte anna</i>			103535069	-----
	<i>Chaetura pelagica</i>			104398472	-----
	<i>Columba livia</i>			102095338	-----
	<i>Aptenodytes forsteri</i>			103907653	-----
	<i>Pelecanus crispus</i>			104028957	-----
	<i>Colius striatus</i>			104561636	-----
	<i>Ficedula albicollis</i>			101810362	ENSFALG00000003535
	AFROTHERIA	<i>Chrysochloris asiatica</i>	102820649	-----	
	EULIPOTYPHILA	<i>Sorex araneus</i>	101558294	-----	
	MAMMALIA	CARNIVORA	<i>Ursus maritimus</i>	103674666	ENSUMAG00000000074
			<i>Ailuropoda melanoleuca</i>	100477023	ENSAMEG00000002655
		CETARTIODACTYLA	<i>Leptonychotes weddellii</i>	102735702	-----
			<i>Vicugna pacos</i>	102529541	-----
			<i>Camelus ferus</i>	102515786	-----
			<i>Bison bison</i>	104994255	-----
		CHIROPTERA	<i>Balaenoptera acutorostrata</i>	102999695	-----
			<i>Pteropus alecto</i>	102887779	-----
			<i>Rhinolophus sinicus</i>	?	?
			<i>Eptesicus fuscus</i>	103283921	-----
	PRIMATES	<i>Myotis davidii</i>	102751902	-----	
<i>Carlito syricta</i>	103259037	-----			

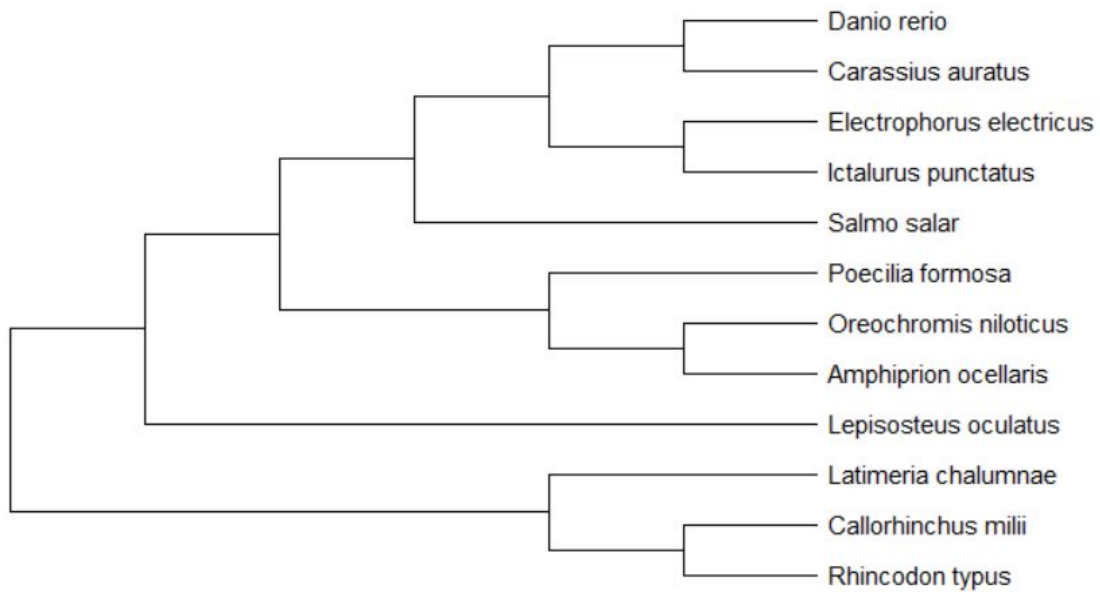
Código	Significado
	CDS é parcial, pois a montagem genômica e/ou a anotação está(ão) incompleta(s).
	Dact1 não foi encontrado porque a montagem genômica e/ou a anotação não está(ão) completa(s).
	O gene Dact1 está ausente do genoma dessa espécie.
	Não há montagem genômica disponível para essa espécie.



60. <i>Bos taurus</i>	- - -	D L I G W L E Y	- - -	K E G H C	E D Q A S	- R A V D C	S P S T P Q S N S	L A V I A
61. <i>Rousettus aegyptiacus</i>	- - -	D L I G W L D Y	- - -	K E G R C	E D Q A S	- E T V C C	S P S T S Q F N P	L D V I A
62. <i>Pteropus vampyrus</i>	- - -	D L I G W L D Y	- - -	K E G H C	E D Q A S	- E T V C C	S P S T S Q F N P	L D V I A
63. <i>Phyllostomus discolor</i>	- - -	D L I G W L D Y	- - -	T E G H C	E D Q A S	- G A M C C	F P S T P Q F N P	L G V I A
64. <i>Hipposideros armiger</i>	- - -	D L I G W L D H	- - -	K E G H C	E D Q A S	- E A V C R	S P S T S K F N P	L G V I A
65. <i>Myotis lucifugus</i>	- - -	D L I G W L D C	- - -	T E G R C	E E Q A S	- E A V G C	F P S T S Q F N P	L G V I A
66. <i>Ochotona princeps</i>	- - -	D L L R R L D C	- - -	K E A Q R	D D Q A S	- G A G G R	P P C T L H F N V	L D V I A
67. <i>Oryctolagus cuniculus</i>	- - -	D L L R R L E F	- - -	R E G H C	E D Q A S	- G A V G R	S P S A A P C H A	L D V I A
68. <i>Ictidomys tridecemlineatus</i>	- - -	D L I G W L E Y	- - -	K E G H C	E D Q A S	- G A I C C	S P S T P Q F N S	L D V I A
69. <i>Marmota marmota</i>	- - -	D L I G W L E Y	- - -	K E G H C	E D Q A S	- E A I C C	S P S T P Q F N S	L D V I A
70. <i>Jaculus jaculus</i>	- - -	D L I G W L E C	- - -	K G S H C	E D Q A S	- G A I C S	S H S T P Q F N S	L D V I T
71. <i>Rattus norvegicus</i>	- - -	D L I G W L E C	- - -	K G G H C	E D Q A S	- G T M C S	S P S T P Q F N P	L D V I A
72. <i>Urocyon parryi</i>	- - -	D L I G W L E Y	- - -	K E G H C	E D Q A S	- G X I C C	S P S T P Q F N S	L D V I A
73. <i>Mus musculus</i>	- - -	D L I G W L E C	- - -	K G G P C	E D Q A S	- G T V C S	S P S T P Q F N S	L D V I A
74. <i>Mesocricetus auratus</i>	- - -	D L I G W L E C	- - -	K G G H C	E D Q A S	- G A I C S	P P S T P Q F N S	L D V I A
75. <i>Cricetulus griseus</i>	- - -	D L I G W L E C	- - -	K G S R	- E D Q A S	- G A I C S	S P S T P Q F N S	L D V I A
76. <i>Otolemur garnettii</i>	E D D	D Y L R W L E Y	- - -	K E G H C	L G	- - - -	V V C P F I C T P Q F H A	L D V I A
77. <i>Microcebus murinus</i>	- - -	D F I G W L E Y	- - -	K E G H C	E D H V S	- G A V C R	S F S T P Q F N S	L D V I T
78. <i>Callithrix jacchus</i>	- - -	D L I G W L E Y	- - -	K E G H C	E D Q A S	- G A V C H	S L S T P Q F N S	L D V I A
79. <i>Rhinopithecus bieti</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
80. <i>Aotus nancymae</i>	- - -	D L I G W L E Y	- - -	K E G H C	E D Q A S	- G A V C H	S L S T P Q F N S	L D V I A
81. <i>Chlorocebus sabaeus</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
82. <i>Papio anubis</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
83. <i>Macaca mulatta</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
84. <i>Nomascus leucogenys</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
85. <i>Pongo abelii</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
86. <i>Gorilla gorilla</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
87. <i>Pan troglodytes</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A
88. <i>Homo sapiens</i>	- - -	D L I G L L E Y	- - -	K E G H C	E D Q A S	- G A V C R	S L S T P Q F N S	L D V I A



57. Ochotona princeps	- - -	D L L R R L D C -	K E A Q R D D Q A S G A G G R P P C T L H F N V L D V I A
58. Oryctolagus cuniculus	- - -	D L L R R L E F -	R E G H C E D Q A S G A V G R S P S A A P C H A L D V I A
59. Ictidomys tridecemlineatus	- - -	D L I G W L E Y -	K E G H C E D Q A S G A I C C S P S T P Q F N S L D V I A
60. Marmota marmota	- - -	D L I G W L E Y -	K E G H C E D Q A S E A I C C S P S T P Q F N S L D V I A
61. Jaculus jaculus	- - -	D L I G W L E C -	K G S H C E D Q A S G A I C S S H S T P Q F N S L D V I T
62. Rattus norvegicus	- - -	D L I G W L E C -	K G G H C E D Q A S G T M C S S P S T P Q F N P L D V I A
63. Urocyon parryi	- - -	D L I G W L E Y -	K E G H C E D Q A S G X I C C S P S T P Q F N S L D V I A
64. Mus musculus	- - -	D L I G W L E C -	K G G P C E D Q A S G T V C S S P S T P Q F N S L D V I A
65. Mesocricetus auratus	- - -	D L I G W L E C -	K G G H C E D Q A S G A I C S P P S T P Q F N S L D V I A
66. Cricetulus griseus	- - -	D L I G W L E C -	K G S R - E D Q A S G A I C S S P S T P Q F N S L D V I A
67. Ootomys garnettii	E D D	D Y L R W L E Y -	K E G H C L G - - - V C P F I C T P Q F H A L D V I A
68. Microcebus murinus	- - -	D F I G W L E Y -	K E G H C E D H V S G A V C R S F S T P Q F N S L D V I T
69. Callithrix jacchus	- - -	D L I G W L E Y -	K E G H C E D Q A S G A V C H S L S T P Q F N S L D V I A
70. Rhinopithecus bieti	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
71. Aotus nancymaae	- - -	D L I G W L E Y -	K E G H C E D Q A S G A V C H S L S T P Q F N S L D V I A
72. Chlorocebus sabaeus	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
73. Papio anubis	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
74. Macaca mulatta	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
75. Nomascus leucogenys	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
76. Pongo abelii	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
77. Gorilla gorilla	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
78. Pan troglodytes	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A
79. Homo sapiens	- - -	D L I G L L E Y -	K E G H C E D Q A S G A V C R S L S T P Q F N S L D V I A

**APÊNDICE VI - Árvore filogenética Chondrichthyes, Actinopterygii, e Actinistia**

Fonte: O AUTOR. Inferências obtidas através do método de Máxima Verossimilhança e o modelo de matrizes JTT. A análise foi conduzida no software MEGA X.