

Universidade Federal de Minas Gerais
Instituto Ciências Biológicas

Departamento de Genética, Ecologia e Evolução
Programa de Pós-Graduação em Genética

Sarah Aparecida Soares de Oliveira Henz

**Impacto da Existência de Múltiplas Isoformas da Proteína RAD51D na Avaliação da
Conservação Evolutiva dos Resíduos de Aminoácidos e, conseqüentemente, na
Classificação de suas Variantes Genéticas quanto à Patogenicidade**

Belo Horizonte, 2025

Sarah Aparecida Soares de Oliveira Henz

Impacto da Existência de Múltiplas Isoformas da Proteína RAD51D na Avaliação da Conservação Evolutiva dos Resíduos de Aminoácidos e, conseqüentemente, na Classificação de suas Variantes Genéticas quanto à Patogenicidade

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação em Genética da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de mestre.

Orientadora: Profa. Dra. Maria Raquel Santos Carvalho

Belo Horizonte, 2025

043

Henz, Sarah Aparecida Soares de Oliveira.

Impacto da Existência de Múltiplas Isoformas da Proteína RAD51D na Avaliação da Conservação Evolutiva dos Resíduos de Aminoácidos e conseqüentemente, na Classificação de suas Variantes Genéticas quanto à Patogenicidade / Sarah Aparecida Soares de Oliveira Henz. – 2025.

126 f. : il. ; 29,5 cm.

Orientadora: Profa. Dra. Maria Raquel Santos Carvalho.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética. 2. Mutações genéticas. 3. Síndrome Hereditária de Câncer de Mama e Ovário. I. Carvalho, Maria Raquel Santos. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 575



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ICB - COLEGIADO DE PÓS-GRADUAÇÃO EM GENÉTICA - SECRETARIA

ATA DE DEFESA DE DISSERTAÇÃO

ATA DA DEFESA DE DISSERTAÇÃO	372/2025
	entrada
	1º/2023
SARAH APARECIDA SOARES DE OLIVEIRA HENZ	CPF: 130.660.706-03

Às quinze horas do dia **04 de julho de 2025**, reuniu-se a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: "**Impacto da Existência de Múltiplas Isoformas da Proteína Rad51d na Avaliação da Conservação Evolutiva dos Resíduos de Aminoácidos E, Consequentemente, na Classificação de Suas Variantes Genéticas Quanto À Patogenicidade**", requisito para obtenção do grau de Mestre em **Genética**. Abrindo a sessão, a Presidente da Comissão, **Dra. Maria Raquel Santos Carvalho**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	Instituição	Indicação
Dra. Maria Raquel Santos Carvalho	UFMG	APROVADA
Dr. Vasco Ariston Carvalho de Azevedo	UFMG	APROVADA
Dra. Célia Regina Moreira Lanza	UFMG	APROVADA

Pelas indicações, a candidata foi considerada: **APROVADA**

O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 04 de julho de 2025.

Dra. Maria Raquel Santos Carvalho - Orientadora

Dr. Vasco Ariston Carvalho de Azevedo

Dra. Célia Regina Moreira Lanza

Assinatura dos membros da banca examinadora:



Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo, Professor do Magistério Superior**, em 08/07/2025, às 09:56, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Célia Regina Moreira Lanza, Professora do Magistério Superior**, em 09/07/2025, às 14:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria Raquel Santos Carvalho, Membro**, em 11/07/2025, às 13:12, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador 4344300 e o código CRC E7DD9385.



UNIVERSIDADE FEDERAL DE MINAS GERAIS

"Impacto da Existência de Múltiplas Isoformas da Proteína Rad51d na Avaliação da Conservação Evolutiva dos Resíduos de Aminoácidos E, Conseqüentemente, na Classificação de Suas Variantes Genéticas Quanto À Patogenicidade"

Sarah Aparecida Soares de Oliveira Henz

Dissertação aprovada pela banca examinadora constituída pelos Professores:

Profa Maria Raquel Santos Carvalho - Orientadora
UFMG

Prof. Vasco Ariston Carvalho de Azevedo
UFMG

Profa Célia Regina Moreira Lanza
UFMG

Belo Horizonte, 04 de julho de 2025.



Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo, Professor do Magistério Superior**, em 08/07/2025, às 09:57, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

Documento assinado eletronicamente por **Celia Regina Moreira Lanza, Professora do**



Magistério Superior, em 09/07/2025, às 14:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria Raquel Santos Carvalho, Membro**, em 11/07/2025, às 13:12, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4344316** e o código CRC **52C5C54D**.

“Dedico este trabalho aos meus pais, que sempre me apoiaram e me deram a melhor educação possível”.

Agradecimentos

Agradeço primeiramente a Deus, que me permitiu chegar até aqui. A caminhada foi longa e difícil, mas ele me fortaleceu todos dias para continuar firme.

Gostaria de agradecer minha mãe, que sempre me apoiou na minha jornada acadêmica e me apoiou incondicionalmente para a realização desse sonho. Ao meu pai, agradeço também pelo apoio que tive durante tantas dificuldades, principalmente no último ano. A maior herança que um pai e uma mãe podem deixar para um filho é o conhecimento. Obrigado pai e mãe por terem investido durante toda minha vida na melhor educação possível, além de apoio e amor.

Ao meu companheiro, melhor amigo, meu marido, Alexandre meu agradecimento por estar sempre ao meu lado nos momentos de alegria e dificuldade, na saúde e na doença. Trazendo luz aos meus dias mais difíceis, sendo meu apoio incondicional em tudo que precisei até aqui. Essa vitória é nossa, você é minha base, que trouxe equilíbrio nos dias mais difíceis. Te amo.

Aos meus irmãos que tanto amo. Matheus, que você tenha orgulho da irmã e que siga crescendo cada dia mais lindo e inteligente, sei que terei cada vez mais orgulho de você também. Gabi, obrigada por toda paciência e por estar ao meu lado sempre. A vocês, Jhonathan, Kimberly e agora Helena o meu muito obrigada. A minha tia Bel e a minha vó, o meu muito obrigado por todo amor e apoio que sempre me deram, me fazendo ser quem sou hoje.

Aos meus tios, primos, Brenda e Hugo, obrigada por tornarem momentos difíceis em mais leves, por estarem aqui sempre que eu precisei.

Aos meus amigos Alicia, Ana, Giselle, Nayara, obrigada por todo apoio. Em especial quero dedicar um agradecimento ao Matheus Vieira e a NADINY que nos momentos mais difíceis estiveram comigo e me ajudaram muito a chegar até aqui também.

A Carol Trevelin e a Gabrielle Mapa, minhas amigas que o mestrado me deu, agradeço por compartilhar tantos momentos difíceis, me ajudarem nos momentos de aflição e preocupação que sabemos que é o peso que o mestrado traz. Vocês foram muito importantes nesta caminhada.

Agradeço ao Laboratório de Genética Humana e Médica da UFMG, onde fui muito bem recebida. Agradeço à Marlene, técnica do laboratório, por toda ajuda. Agradeço também ao Arthur e Tamires pela ajuda com Bioinformática.

Gostaria de agradecer também à minha mentora acadêmica de vida, Maria Leticia, que desde a faculdade me ajuda em tudo que preciso e me dá muito apoio também. Obrigada.

À minha orientadora, Maria Raquel, quero agradecer pelo voto de confiança em me orientar, por toda ajuda e paciência que teve comigo em todo este período que não foi fácil. Mas conseguimos, aprendi muito com você e levarei comigo todo este aprendizado para vida.

Também agradeço aos professores Frederico e Marcelo por me auxiliarem durante todo o mestrado em tudo que precisei.

Ao programa de Pós Graduação em Genética e a UFMG, meu muito obrigada.

Por fim, agradeço a bolsa CAPES pelo o investimento neste projeto, tornando-o possível.

Resumo

O gene *RAD51D* está envolvido nas vias de reparo das quebras da dupla fita de DNA durante a replicação, fazendo parte das vias dos genes *BRCA1* e *BRCA2*. Variantes genéticas patogênicas em *RAD51D* têm sido associadas aos cânceres de mama e de ovário e apresentam alta penetrância. A classificação das variantes genéticas quanto à patogenicidade é fundamental, pois permite nortear as condutas clínicas. Variantes genéticas são classificadas em benignas (B), provavelmente benignas (PB), variantes de significado incerto (VUS), provavelmente patogênicas (PP) ou patogênicas (P). Os principais critérios para classificação de uma variante genética são: 1. envolvimento do gene com o fenótipo; 2. frequência alélica da variante na população em geral e em amostras de pacientes; 3. Conservação evolutiva do resíduo (aminoácido) afetado pela variante; 4. Impacto funcional predito nos níveis de RNA e proteína; 5. segregação da variante em famílias e estudos funcionais avaliando o impacto da variante. A existência de isoformas dificulta a avaliação da conservação evolutiva dos resíduos, uma vez que os programas de alinhamento múltiplo (por exemplo, BLAST e CLUSTAL Omega) não separam as isoformas *antes* de fazerem o alinhamento. O *RAD51D* apresenta pelo menos 8 isoforma proteicas e a presença destas isoformas, ou qual é a isoforma canônica, varia conforme a espécie/classe. O objetivo do presente estudo é avaliar a conservação evolutiva de cada resíduo da isoforma canônica humana da proteína *RAD51D*, e outras isoformas relevantes clínica ou evolutivamente.

Palavras-chave: *RAD51D*; câncer; conservação evolutiva.

Abstract

The *RAD51D* gene is involved in the repair pathways of double-strand breaks in DNA during replication and is part of the *BRCA1* and *BRCA2* gene pathways. Pathogenic genetic variants in *RAD51D* have been associated with breast and ovarian cancers and have high penetrance. Classification of genetic variants according to pathogenicity is essential, as it allows guiding clinical conduct. Genetic variants are classified as benign (B), probably benign (PB), variants of uncertain significance (VUS), probably pathogenic (PP), or pathogenic (P). The main criteria for classifying a genetic variant are: 1. involvement of the gene with the phenotype; 2. allele frequency of the variant in the general population and patient samples; 3. evolutionary conservation of the residue (amino acid) affected by the variant; 4. predicted functional impact on RNA and protein levels; 5. Variant segregation in families and functional studies evaluating the impacts of the variant. The existence of isoforms makes it difficult to assess the evolutionary conservation of residues, since multiple alignment programs (e.g., BLAST and CLUSTAL Omega) do not separate isoforms before performing the alignment. *RAD51D* has at least 8 protein isoforms and the presence of these isoforms, or which one is the canonical isoform, varies according to the species/class. The objective of the present study is to evaluate the evolutionary conservation of each residue of the human canonical isoform of the *RAD51D* protein and other clinically or evolutionarily relevant isoforms.

Keywords: *RAD51D*; cancer; evolutionary conservation.

Lista de imagens

FIGURA 1 Marcas do Câncer	22
FIGURA 2 Genes envolvidos no processo de reversão da forquilha	27
FIGURA 3. Posição cromossômica e estrutura do RAD51D	29
FIGURA 4 Contexto genômico no entorno do RAD51D	30
FIGURA 5 - Genes na região 17q12	31
FIGURA 6 Splicing e splicing alternativo no gene RAD51D	32
A) No nível do mRNA	32
B) Estrutura da isoforma canônica	32
C) No nível da proteína, na isoforma canônica, destacando-se o primeiro aminoácido de cada exon.	32
FIGURA 7 Domínios RAD51D	33
FIGURA 8 Alinhamento isoforma canônica X isoforma 2	42
FIGURA 9 Alinhamento isoforma canônica X isoforma 3	42
FIGURA 10 Alinhamento isoforma canônica X isoforma 4	43
FIGURA 11 Alinhamento isoforma canônica X isoforma 5	43
FIGURA 12 Alinhamento isoforma canônica X isoforma 6	44
FIGURA 13 Alinhamento isoforma canônica X isoforma 7	44
FIGURA 14 Alinhamento isoforma canônica X isoforma 8	45
FIGURA 15 Comparação da Isoforma canônica em humana com a isoforma 329	46
FIGURA 16 Alinhamento 328	52
FIGURA 17 Grau de profundidade da isoforma canônica	54
FIGURA 18 - Parte inicial do alinhamento do grupo de aves	57
FIGURA 19 Parte do alinhamento Gallus gallus	58
FIGURA 20 Parte do alinhamento Gallus gallus.	58
FIGURA 21 - Alinhamento múltiplo, usando-se a isoforma 329 do gene RAD51D	60
FIGURA 22 Localização do éxon número 9	114
FIGURA 22A) Ponto de diminuição de conservação evolutiva em primatas	114
FIGURA 23 - Ponto de diminuição de conservação evolutiva em roedores	115
FIGURA 24 Localização no éxon número 5	115
FIGURA 24A Ponto de início da diminuição da conservação evolutiva em aves, répteis, anfíbios e peixes	115
FIGURA 25-Linha do tempo de acordo com a isoforma canônica	116
FIGURA 26- Isoforma 297	117
Figura 27 - Alinhamento múltiplo de todas as isoformas presentes em seres humanos	118

Lista de Tabelas

Tabela 1 Programas e bancos de dados utilizado nas análises	38
Tabela 2 - Isoformas de RAD51D em humanos	41
Tabela 3	55
Tabela 4 - Conservação e profundidade em mamíferos por aminoácido	62
Tabela 5-Conservação e profundidade em aves, répteis, anfíbios e peixes por aminoácido	79
Tabela 6-Conservação e profundidade em primatas por aminoácido	99

Lista de abreviaturas, siglas e unidades de medida

Termo	Significado
NCBI	<i>National Center for Biotechnology Information</i>
DNA	ácido desoxirribonucleico
ssDNA	DNA de fita simples
INCA	<i>Instituto nacional de câncer</i>
<i>RAD51</i>	<i>Radiation Sensitive 51</i>
<i>RAD51C</i>	<i>Radiation Sensitive 51 C</i>
<i>RAD51D</i>	<i>Radiation Sensitive 51 D</i>
<i>RB</i>	gene supressor de tumor
<i>TP53</i>	gene supressor de tumor
<i>MITF</i>	Fator de Transcrição Associado à Microftalmia
<i>RAS</i>	<i>Rat Sarcoma</i>
<i>NF2</i>	neurofibromatose tipo 2
<i>BRCA1</i>	<i>Breast Cancer gene 1</i>
<i>BRCA2</i>	<i>Breast Cancer gene 2</i>
<i>CHEK2</i>	<i>checkpoint kinase 2</i>
<i>ATM</i>	<i>Ataxia-Telangiectasia Mutated</i>
<i>PALB2</i>	<i>Partner and Localizer of BRCA2</i>
ACMG	<i>American College of Medical Genetics and Genomics</i>
APA	<i>American Assiciation of Pathology</i>

GTPases	Enzimas guanosina trifosfatases
CYP17	Enzima citocromo P450 17A1
CYP19	Enzima citocromo P450 19
<i>GSTM1</i>	gene que codifica a enzima glutationa S-transferase mu 1
<i>NQO2</i>	<i>N-ribosildihyronicotinamide</i>
FT	Fator de transcrição
<i>RecA</i>	Recombinação A
ATP	Adenosina Trifosfato
<i>DMC1</i>	<i>DNA Meiotic Recombinase 1</i>
<i>RFFL</i>	<i>ring finger and FYVE like domain containing E3 ubiquitin protein ligase</i>
<i>FNDC8</i>	<i>fibronectin type III domain containing 8</i>
<i>Pb</i>	<i>Pares de base</i>
<i>HOXA5</i>	gene <i>homeobox</i>
SMAD4	componente essencial da via de sinalização TGF- β
<i>TMEM132E</i>	Proteína transmembrana
<i>UNC45</i>	Proteína chaperona
<i>CCL</i>	Ligante da quimiocina
<i>PLM</i>	Fosfolambana
<i>RARA</i>	Receptor de Ácido Retinoico Alfa
<i>PLZF</i>	<i>Promyelocytic Leukemia Zinc Finger</i>

<i>BRAF</i>	<i>B-Raf proto-oncogene, serine/threonine kinase</i>
VUS	Variante de significado incerto
<i>SMARCA11</i>	Helicase de Annealing de Reestrutura de Cromatina Relacionada à SNF2
<i>ZRANB3</i>	<i>Zinc finger RANBP2-type containing 3</i>
<i>HLTF</i>	<i>Helicase-like Transcription Factor</i>
FBH1	Gene membro da família F-BOX
<i>XRCC2</i>	<i>X-ray repair cross complementing 2</i>
HR	<i>Recombinação homóloga</i>

Sumário

1. Introdução	20
1.1 O câncer	20
1.2 Os cânceres de mama e de ovário	24
1.2.1 Reparo por recombinação homóloga	26
1.3 Origem e evolução das famílias RAD e RAD51	27
1.4 O gene RAD51D	29
1.5 A proteína RAD51D	33
1.6 Variantes genéticas e câncer	34
1.7 Classificação de variantes	34
1.8 Importância da conservação evolutiva na classificação de cada resíduo quanto a patogenicidade	35
1.9 Impacto das isoformas na avaliação da conservação evolutiva	36
1.10 Objetivos	36
1.10.1 Objetivo geral	36
1.10.2 Objetivos específicos	37
2. Materiais e Métodos	37
2.1 Como as sequências foram obtidas?	37
2.2 Como as sequências foram analisadas?	38
2.3 Análise dos resultados	39
3.0 Resultados e Discussão	40
3.1 Quais as isoformas da proteína RAD51D estão presentes em humanos? Quais as diferenças e qual o mecanismo de origem de cada uma delas?	40
3.2 Qual a diferença entre as isoformas 328 e 329?	46
3.3 Como a isoforma presente em humanos aparece nas outras classes?	47
3.3.1 Mamíferos	47
3.3.2 Aves	48
3.3.3 Répteis	49
3.3.4 Anfíbios	50
3.3.5 Peixes	51
3.4 Como é o alinhamento das sequências da isoforma canônica humana (328 aminoácidos) com suas correspondentes em outras espécies?	51
3.5 Qual a profundidade de conservação evolutiva em mamíferos na isoforma canônica?	52
3.6 Qual o grau de profundidade da conservação evolutiva isoforma canônica de humanos em outras classes?	53
3.7 Qual a classe tem a isoforma semelhante canônica de humanos mais bem conservada?	56
3.8 Como é o alinhamento da isoforma 329 de roedores e qual o grau de profundidade de conservação evolutiva?	59

3.9 Qual o grau de conservação entre os roedores?	60
3.10 Quais as principais posições suscetíveis a trocas de aminoácidos na RAD51D?	62
3.11 Qual plataforma podemos utilizar para discriminar as isoformas?	116
3.12 Conservação E Classificação	118
5. Referências eletrônicas e bibliográficas	120
6.0 Sequências	123

1. Introdução

O desenvolvimento dos métodos de Biologia Molecular tem permitido grandes avanços na compreensão dos mecanismos subjacentes às doenças e das suas aplicações. Isto é particularmente verdadeiro em relação ao câncer. Os primeiros genes relacionados ao câncer identificados foram Rb e TP53, que permitiram identificar as primeiras síndromes de cânceres familiares. Atualmente, são centenas de genes e vias metabólicas conhecidas, abrindo caminho para o desenvolvimento de novos métodos para diagnóstico e terapia. A contribuição da Genética varia de acordo com o subtipo de câncer e com diversos fatores epidemiológicos, como idade de início e recorrência familiar. A seguir, serão discutidos aspectos gerais do câncer, seus mecanismos celulares e genético-moleculares. O foco do trabalho será nos cânceres de mama e ovário e a contribuição do gene *RAD51D*, com enfoque na dificuldade de classificação de variantes genéticas quanto à patogenicidade, em função da existência de isoformas.

1.1 O câncer

O câncer é definido como o crescimento descontrolado e progressivo de células que, ao longo do tempo, perdem suas características de diferenciação e adquirem a capacidade de invadir tecidos adjacentes e metastatizar para órgãos distantes. De forma simplificada, esse processo corresponde à formação tumoral e, conseqüentemente, ao desenvolvimento da neoplasia maligna (INCA).

Historicamente, o câncer era considerado uma doença predominante em indivíduos idosos. Contudo, nas últimas décadas, observou-se um aumento significativo da incidência em todas as faixas etárias, incluindo crianças, adolescentes e adultos jovens. Tal fenômeno está associado, em parte, às mudanças nos hábitos de vida e à maior exposição a fatores ambientais e comportamentais que favorecem alterações genéticas, especialmente em proto-oncogenes e genes supressores tumorais. Esses eventos moleculares desencadeiam uma cascata de mutações que afetam múltiplas vias de regulação celular.

Durante a tumorigênese, as células neoplásicas adquirem propriedades biológicas que favorecem sua sobrevivência e expansão. Hanahan (2022) descreve um conjunto de características fundamentais — denominadas “marcas do câncer” — que transformam uma célula normal em uma célula maligna. Entre essas características estão: (i) capacidade de sustentar sinais proliferativos; (ii) evasão de mecanismos supressores de crescimento; (iii) resistência à morte celular; (iv) potencial replicativo ilimitado; (v) indução de angiogênese; (vi) ativação da invasão e metástase; (vii) reprogramação do metabolismo celular; e (viii) evasão da detecção e destruição pelo sistema imunológico (HANAHAN, 2022). A Figura 1 ilustra essas marcas do câncer.

No mesmo artigo, Hanahan (2022) descreve o conceito de “características habilitadoras”, que correspondem às condições neoplásicas que permitem aos tumores adquirir e manter suas características funcionais essenciais. Entre essas condições estão: desbloqueio da plasticidade fenotípica, reprogramação epigenética não mutacional, microbiomas polimórficos e presença de células senescentes.

A plasticidade é a capacidade de adaptação a diferentes mudanças. No contexto oncológico, a plasticidade fenotípica é um componente crítico da patogênese do câncer, pois envolve a perda da capacidade de diferenciação celular ou até mesmo a desdiferenciação. Hanahan (2022) cita como exemplo de desdiferenciação a carcinogênese do câncer de cólon, em que a diferenciação celular é interrompida pela ação de dois fatores de transcrição do desenvolvimento: *HOXA5* (ENSG00000106004) e *SMAD4* (ENSG00000141646). Outro exemplo é a desdiferenciação de melanócitos, ocasionada pela supressão da atividade do regulador mestre *MITF*.

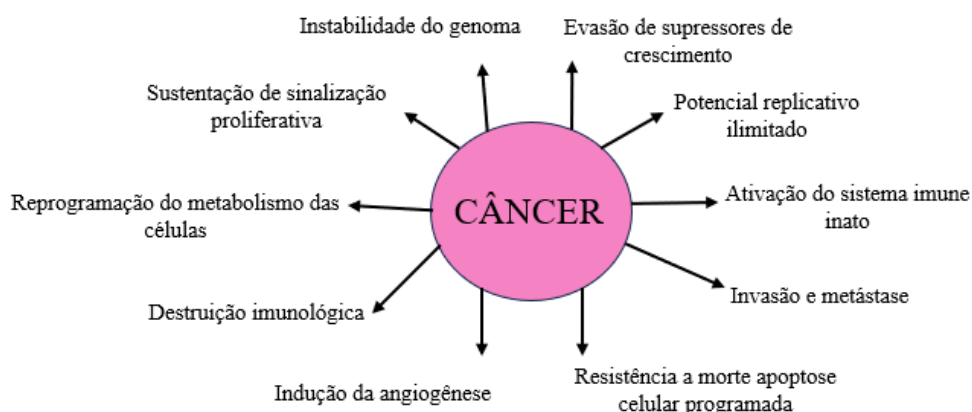
Ainda no contexto da diferenciação celular, destaca-se o fenômeno da transdiferenciação, ou metaplasia tecidual, que ocorre quando células adquirem um fenótipo diferente daquele que lhes é característico. Um exemplo clássico é o esôfago de Barrett.

A regulação epigenética durante a fase embrionária já é amplamente reconhecida. No câncer, há evidências que sustentam a hipótese de que alterações epigenéticas contribuem para a aquisição de propriedades celulares que favorecem a progressão maligna. Além disso, o microambiente tumoral, formado de maneira heterogênea por células cancerígenas e outros componentes, desempenha um papel crucial no desenvolvimento do câncer, pois

diferentes características físicas e bioquímicas presentes nesse microambiente podem induzir alterações no epigenoma.

A senescência celular, ou envelhecimento celular, corresponde à parada irreversível do ciclo celular, ou seja, à interrupção permanente da proliferação celular. Trata-se de um mecanismo de defesa contra o surgimento de neoplasias. Essa condição é acompanhada por alterações morfológicas, mudanças no metabolismo e ativação de um fenótipo secretor associado à liberação de diversas proteínas bioativas. Diversas modalidades terapêuticas, como quimioterapia e radioterapia, utilizam a indução da senescência como estratégia para suprimir a progressão tumoral (HANAHAN, 2022).

FIGURA 1 Marcas do Câncer



Características que as células malignas devem adquirir para sobreviver.

FONTE: Imagem adaptada pela autora de HANAHAN, Douglas. Hallmarks of cancer: new dimensions. *Cancer discovery*, v. 12, n. 1, p. 31-46, 2022. Doi: <https://doi.org/10.1158/2159-8290.CD-21-1059>.

Em um tecido normal, a progressão do ciclo celular é regulada por sinalizadores que controlam o crescimento e a divisão celular. Quando se torna indiferenciada, a célula cancerígena passa a manter o controle dessa função com o objetivo de sustentar sua própria proliferação.

As células neoplásicas podem garantir essa capacidade proliferativa por meio da produção de ligantes de fatores de crescimento, o que possibilita a estimulação proliferativa autócrina. Além disso, a sinalização pode ser corrompida pelo aumento da expressão de proteínas receptoras na superfície da célula tumoral, intensificando a resposta a sinais de crescimento (HANAHAN, 2011).

Um dos mecanismos de defesa contra a sinalização desregulada é o *feedback* negativo. Um gene relevante nesse processo é a oncoproteína **RAS**, cuja função, quando comprometida, afeta sua atividade como GTPase. Essa atividade está associada ao mecanismo intrínseco de *feedback* negativo que, em condições normais, assegura que a transmissão ativa do sinal seja transitória (HANAHAN, 2011).

Outro fator que permite às células cancerígenas controlar sua própria sinalização é a capacidade de contornar a ação de genes reguladores negativos da proliferação celular — os genes supressores de tumor. O mais importante entre eles é o *TP53* (*ENSG00000141510*), que desempenha papel central na tomada de decisões relacionadas à proliferação celular, ativação da senescência e/ou indução da apoptose (HANAHAN, 2011).

Outro gene supressor relevante é o *NF2* (*ENSG00000186575*), cujo mecanismo de ação envolve a inibição do ciclo celular por meio do contato célula-célula. A perda desse gene está associada ao desenvolvimento da neurofibromatose.

Além disso, é importante destacar os proto-oncogenes — genes que, em condições fisiológicas, estimulam a divisão celular e atuam como fatores de crescimento, transdutores de sinais e reguladores da transcrição. Quando sofrem mutações, esses genes tornam-se oncogenes, levando à proliferação celular exacerbada. Nos genomas de aves e mamíferos, além dessas funções usuais, os oncogenes também desempenham papel no controle da diferenciação celular (KONTOMANOLIS et al., 2020).

De acordo com o Instituto Nacional de Câncer (INCA), estima-se que, em 2025, sejam diagnosticados mais de 700 mil casos de câncer no Brasil, sendo o câncer de pele não melanoma o mais incidente, seguido pelo câncer de mama feminino, que representa 10,5% dos casos.

1.2 Os cânceres de mama e de ovário

O câncer de mama é a neoplasia maligna mais prevalente entre as mulheres em todo o mundo. Aproximadamente 1 em cada 10 casos de câncer em mulheres corresponde ao câncer de mama (MENON et al., 2024). Em 2018, mais de 630 mil mulheres morreram em decorrência dessa doença, e, no mesmo período, foram diagnosticados mais de 2 milhões de novos casos.

A frequência do câncer de mama ao longo da vida é estimada em 12,5% na população feminina e 0,1% na população masculina. Em cerca de 80% dos casos, a doença acomete principalmente mulheres com mais de 50 anos. Já o câncer de ovário apresenta uma frequência estimada de 1 a 2% ao longo da vida das mulheres (CASAUBON et al., 2023; OBEAGU; OBEAGU, 2024; KATSURA; OGUNMWONYI; KANKAM; SAHA, 2022).

Quase 85% dos casos de câncer de mama apresentam-se na forma de adenocarcinoma, acometendo ductos mamários e epitélio lobular (KATSURA; OGUNMWONYI; KANKAM; SAHA, 2022).

Para o diagnóstico do câncer de mama, é fundamental a realização de exames radiológicos, bem como análises histológicas, como biópsia e citologia, visando um tratamento mais eficaz e direcionado. As opções terapêuticas incluem cirurgia, radioterapia, quimioterapia e terapia endócrina, cuja indicação depende dos achados histológicos e genéticos (KATSURA; OGUNMWONYI; KANKAM; SAHA, 2022).

O fator hereditário exerce papel significativo na ocorrência do câncer de mama, estando presente em aproximadamente 10 a 15% dos casos de câncer relacionados a mutações genéticas germinativas que se segregam em famílias. Em mulheres jovens, com até 30 anos, essa proporção pode chegar a até 25% dos casos (MENON et al., 2024).

Os principais genes associados ao câncer de mama são os genes supressores de tumor *BRCA1* (*ENSG0000012048*) e *BRCA2* (*ENSG00000139618*). Variantes germinativas patogênicas nesses genes conferem um risco aproximado de 85% para o desenvolvimento da doença nas mulheres portadoras. O gene *BRCA1* está associado a cerca de 35% dos casos hereditários de câncer de mama, enquanto o *BRCA2* corresponde a aproximadamente 25% desses casos (CASAUBON et al., 2023).

O processo inicia-se na linhagem germinativa, envolvendo genes como *BRCA1/2*, *CHEK2* (*ENSG00000183765*), *PALB2* (*ENSG00000083093*) e *TP53*, que apresentam susceptibilidade moderada a alta para a doença. Em seguida, ocorre a inativação de genes supressores tumorais, incluindo *ATM*, *BRCA1/2*, *CHEK2* e *PALB2* (XIONG et al., 2025).

Genes ativos na via de síntese de estrogênio, como *CYP17*, *CYP19*, *GSTM1* (*ENSG00000134184*) e *NQO2* (*ENSG00000124588*), também são alvo de estudos para identificar possíveis polimorfismos associados ao câncer de mama.

O câncer de mama pode ser classificado em carcinoma invasivo e carcinoma lobular, além de mais de 18 subtipos histológicos (XIONG et al., 2025).

Situação semelhante ocorre no câncer de ovário (OBEAGU; OBEAGU, 2024), onde variantes germinativas patogênicas nos mesmos genes aumentam o risco para tipos específicos de malignidades epiteliais (CASAUBON et al., 2023).

Sabe-se que a incidência do câncer de ovário varia estatisticamente entre diferentes etnias, embora a causa dessa diferença ainda seja desconhecida. Algumas hipóteses indicam que fatores como o uso de contraceptivos orais, laqueadura tubária e ooforectomia reduzem o risco de desenvolvimento do câncer de ovário, enquanto o uso de terapia hormonal na menopausa pode aumentar esse risco em cerca de 20% (TORRE et al., 2018).

O câncer epitelial de ovário é o tipo mais comum entre as mulheres e pode ser subdividido histologicamente em carcinomas serosos, endometrióides, mucinosos e de células claras. Esses tumores também podem ser classificados como tipo I ou tipo II, conforme sua instabilidade genética. Os cânceres epiteliais tipo I são geralmente de baixo grau, enquanto os tipo II apresentam comportamento agressivo e alto grau, apresentando baixa taxa de sobrevivência quando diagnosticados tardiamente. Os cânceres de ovário não epiteliais tendem a ser menos agressivos.

Os tumores de alto grau estão frequentemente associados a mutações nos genes *BRCA1* e *BRCA2*, presentes em aproximadamente 40% dos casos, assim como nos genes *RAD51C* (*ENSG00000108384*) e *RAD51D* (*ENSG00000185379*). Esses genes estão relacionados a mutações raras de penetrância moderada, envolvendo a interação entre os genes *BRCA1*, *BRCA2* e a helicase C-terminal 1.

O diagnóstico do câncer de ovário depende significativamente do subtipo tumoral, bem como da etnia do paciente. O diagnóstico precoce é fundamental para melhorar o prognóstico; contudo, essa etapa é dificultada pela ausência de sintomas evidentes nos estágios iniciais da doença (TORRE et al., 2018).

1.2.1 Reparo por recombinação homóloga

As vias de reparo por recombinação homóloga são mediadas por diversos genes, incluindo o *RAD51D*. As proteínas *BRCA1* e *BRCA2* formam um complexo que recruta a proteína *RAD51* aos locais de ruptura da dupla fita do DNA. A proteína *BRCA2* é responsável pela montagem do filamento de nucleoproteína da *RAD51* na extremidade ressecada do DNA (WASSING; ESASHI, 2021). Esse processo também é mediado por proteínas da família *RAD52* (*ENSG0000002016*) (LI et al., 2019).

Dessa forma, a função principal da família *RAD51* é promover a proteção contra o estresse da replicação celular, desacelerando e estabilizando a abertura da dupla fita e a formação das forquilhas de replicação do DNA. Essas ações são essenciais para garantir a integridade do reparo por recombinação homóloga ao longo da vida celular (BONILLA et al., 2020; WASSING; ESASHI, 2021).

O reparo por recombinação homóloga pode ser subdividido em quatro etapas principais (BONILLA et al., 2020):

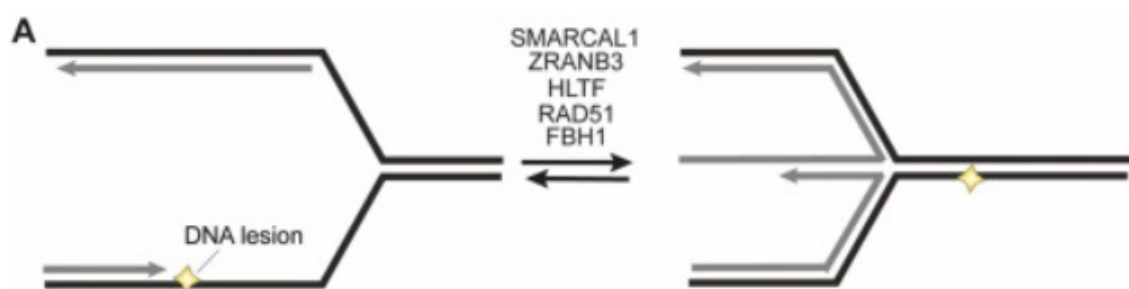
1. Reversão das forquilhas;
2. Proteção das forquilhas invertidas;
3. Reparo e reinicialização das forquilhas de replicação;
4. Preenchimento das lacunas pós-replicação.

A reversão das forquilhas de replicação auxilia no enfrentamento do estresse replicativo, por meio do reanelamento das fitas molde e da geração de duplex de DNA. Qualquer desregulação desse processo pode desencadear centenas de doenças humanas. Além disso,

a paralisação prolongada das forquilhas pode levar à morte celular (ADOLPH; CORTEZ, 2024).

Essa reversão faz com que o número de vias de junção aumente de três para quatro, como, por exemplo, as junções de Holliday. Esse processo envolve diversos genes, entre eles a família *RAD51*, conforme ilustrado na Figura 2 (ADOLPH; CORTEZ, 2024).

FIGURA 2 Genes envolvidos no processo de reversão da forquilha



A reversão da forquilha de replicação insere uma lesão de DNA molde de volta no DNA duplex. Esse processo requer *RAD51* e é mediado por translocases dependentes de ATP.

Fonte: ADOLPH, Madison B.; CORTEZ, David. Mechanisms and regulation of replication fork reversal. *Dna Repair*, [S.L.], v. 141, p. 103731, set. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.dnarep.2024.103731>. Adaptado pela autora.

O principal gatilho para a reversão da forquilha de replicação é o desacoplamento das atividades da helicase e da polimerase, o que resulta na exposição de DNA de fita simples (ssDNA) (ADOLPH; CORTEZ, 2024).

A proteção da forquilha replicativa ocorre por meio da estabilização do filamento de *RAD51*, mediada pelo gene *BRCA2*, que previne a degradação dos nucleotídeos causada pela abertura e pela formação das forquilhas paralisadas (BONILLA et al., 2020).

1.3 Origem e evolução das famílias RAD e RAD51

Os genes que compõem a família RAD pertencem a uma nova classe de proteínas ligadoras de GTP. Esses genes exibem uma dependência única de magnésio para se ligarem à guanina, processo este regulado por uma GTPase específica para o RAD. A

expressão do gene RAD é particularmente elevada em tecidos como coração, pulmão e músculo esquelético (MOYERS; BILAN; ZHU; KAHN, 1997).

A superfamília *RAD51* originou-se de um ancestral comum aos domínios Archaea e Eukarya, tendo sido identificada a partir do estudo do gene *RecA* em *Escherichia coli*. O gene ***RAD51*** é o homólogo eucariótico do gene bacteriano *RecA*. Essas proteínas apresentam domínios altamente conservados, como os motivos Walker A e Walker B, relacionados à ligação e hidrólise de ATP (LIN et al., 2006).

A partir da análise de conservação evolutiva, a superfamília foi dividida em três grandes grupos: *recA*, *RAD α* e *RAD β* (BONILLA et al., 2020).

O grupo *recA* está presente em genes de diversos reinos, como Monera, Protista, Fungi e Plantae. O grupo *RAD α* inclui as recombinases primárias dos eucariotos (*RAD51* e *DMC1*) e Archaea. Por fim, o grupo *RAD β* compreende os parálogos eucarióticos canônicos do RAD51 — entre eles, *RAD51C*, *RAD51D*, *XRCC2* (*ENSG00000196584*), *XRCC3* (*ENSG00000126215*) — além do gene *radB* de Archaea.

A hipótese proposta para explicar a diversidade dentro dessa família é que, entre Archaea e eucariotos, ocorreram duplicações nos genes ancestrais do grupo *RAD β* , resultando nas diferenças observadas nas proteínas *RAD51* entre os reinos Animalia, Plantae e Fungi. Também se sugere a ocorrência de transferência horizontal por meio de eventos endossimbióticos (LIN et al., 2006; BONILLA et al., 2020).

A família *RAD α* apresenta maior conservação evolutiva em comparação à família *RAD β* . Alguns genes do grupo *RAD β* sofreram perdas evolutivas em fungos e em alguns invertebrados (LIN et al., 2006).

Focando no grupo de interesse, *RAD β* , observa-se uma subdivisão adicional dentro da família RAD51, que inclui os genes *RAD51* (também denominado *RAD51A*), *RAD51B*, *RAD51C* e *RAD51D*. O gene *RAD51D* aparentemente surgiu mais recentemente nos organismos eucariotos e é o membro mais divergente dessa família (LIN et al., 2006; BONILLA et al., 2020).

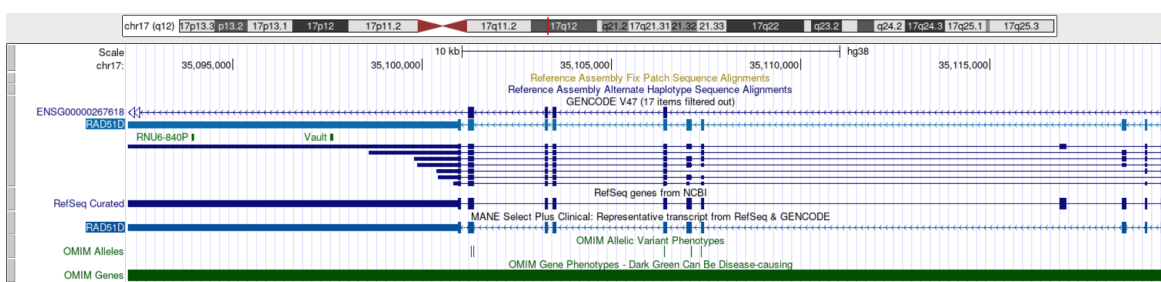
Segundo Lin e colaboradores (2006), os genes da família RAD51 compartilham pouca homologia sequencial, o que resulta em proteínas distintas entre diferentes grupos de organismos. Em vertebrados, a proteína RAD51D conserva cerca de 74% de identidade com as sequências encontradas em organismos como leveduras e plantas. A sequência canônica humana da RAD51D, composta por 328 aminoácidos, pode apresentar até 99% de similaridade com a de camundongos.

1.4 O gene *RAD51D*

O gene *RAD51D* é o homólogo 4 da família gênica *RAD51* e foi descrito na levedura *Saccharomyces cerevisiae*, juntamente com outro gene, o *DMC1*. Estes dois genes são estruturalmente semelhantes aos de *E. coli* (JIANG *et al.*, 2018).

O gene *RAD51D* mapeia em 17q12, nas coordenadas chr17:35,092,221-35,119,860, ocupa 27.640 bp (GRCh38/hg38), estando localizado na fita *minus* (UCSC Genome Browser, acesso em 25/2/2025). Esta é uma região rica em genes (Figura 3). Upstream ao *RAD51D*, está o gene *RFFL* (ENSG00000092871), para o qual está descrito um produto *readthrough RAD51D-RFFL*.

FIGURA 3. Posição cromossômica e estrutura do *RAD51D*



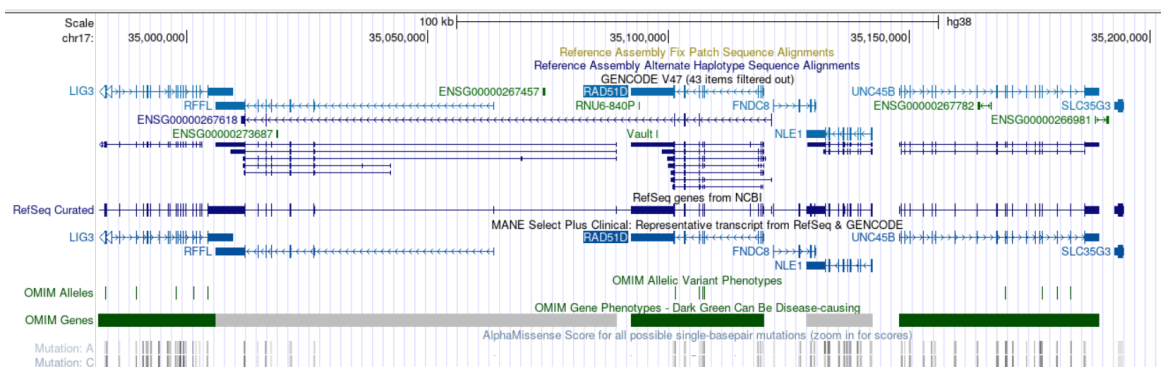
Localização do gene *RAD51D*

FONTE: Criada pela autora com base na imagem do UCSC Genome Browser (acesso em 27/2/2025)

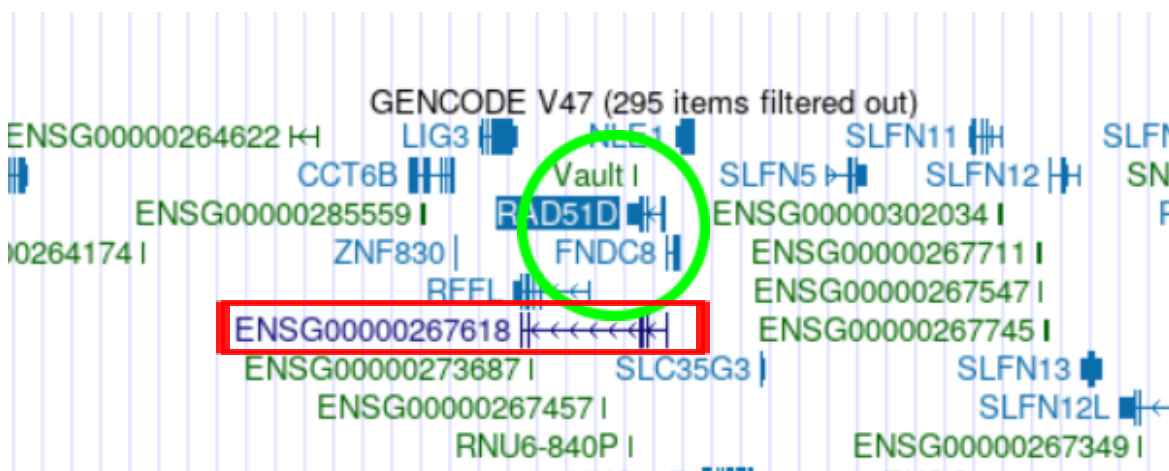
Outro gene próximo da região é o *FNDC8* (ENSG00000073598) (hg38, chr17:35.121.615-35.130.732), localizado na fita *plus*, está 1755 pb do *RAD51D*, com potencial para formar *overlapping gene* com *RAD51D* (*RAD51D-FNDC8*), do tipo cabeça-cabeça, com compartilhamento de elementos regulatórios à 5'-UTR (Figura 4).

FIGURA 4 Contexto genômico no entorno do *RAD51D*

A)



B)



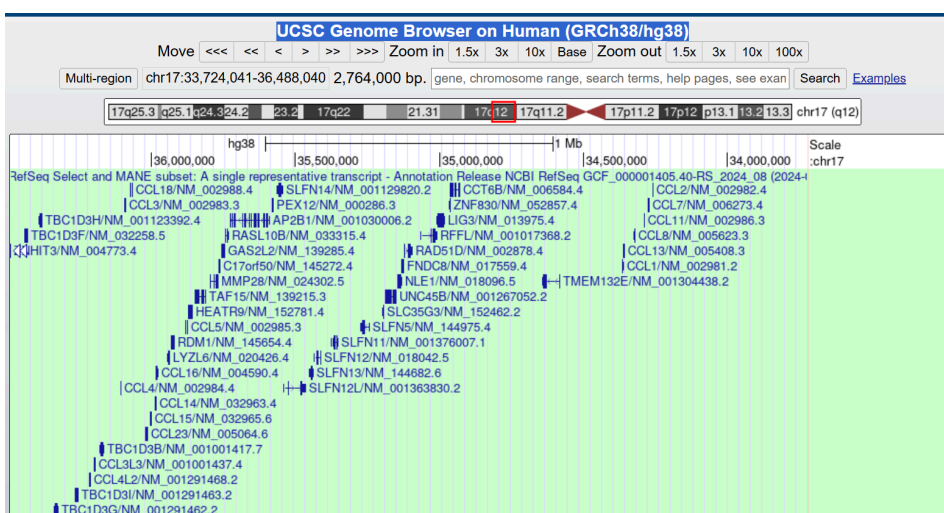
Capturas de imagem da região no entorno do gene *RAD51D* no UCSC Genome Browser, com a montagem hg38. Em A) região gênica onde está localizada o *RAD51D* Em B), essa região é apresentada em maior resolução. o transcrito do *readthrough* composto pelos genes *RAD51D-RFFL* está identificado por um retângulo vermelho (ENSG00000267618). A região dos *overlapping genes* *RAD51D-FND C8* está identificada por um círculo verde. São overlapping genes na orientação cabeça-cabeça.

FONTE: Criada pela autora com base na imagem do UCSC Genome Browser (acesso em 27/2/2025)

O *RAD51D* se localiza em uma região muito rica em genes, apesar da região 17q12 ser heterocromática, no cariótipo com bandamento G. Nesta região, se encontram, por exemplo, os genes, *TMEM132E* (ENSG00000181291) proteína transmembrana altamente conservada em mamíferos e peixes expressão cóclea de camundongos cérebro, rim, pulmão e fígado, está altamente ligada alguns tipos de surdez. *UNC45* (ENSG00000140553) trata-se de uma proteína chaperona ligada a proliferação celular e fusão de mioblastos. Além disso, nesta região estão os clusters dos genes CCL, TBC.

Esta região é importante em outras neoplasias como a leucemia, cuja iniciação é associada aos produtos de fusão *RARA-PLZF*, *RARA-PML* e *TIF1A* com *RARA* (ENSG00000131759) e *BRAF* (ENSG00000157764) geram produção de oncoproteínas.

FIGURA 5 - Genes na região 17q12



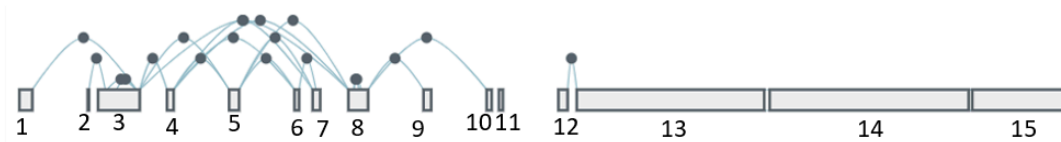
Mapa de genes localizados na região 17q12

Genome Browser 27/02/2025

Segundo as bases de dados UniProt e GTex, o *RAD51D* possui como isoforma canônica uma proteína composta por 328 aminoácidos (O75771-1) e mais sete isoformas em humanos. A isoforma canônica é composta por 11 exons e é bastante expressa em células de linfócitos, testículo, rins, tireóide, útero, tecido vaginal, tecido mamário, ovário e em várias regiões do cérebro, entre outras. Os principais exons expressos nestes tecidos são 10 e 11. Sabendo disso, podemos pressupor que uma alteração nestes exons podem ter mais chances de alterar a função (Figura 6).

FIGURA 6 Splicing e splicing alternativo no gene *RAD51D*

A) No nível do mRNA



Segundo o Gtex esses são os exões presentes no *RAD51D* e as diversas possibilidades de splicing alternativo.

FONTE: captura de imagem do Gtex adaptada pela autora.

B) Estrutura da isoforma canônica



Segundo o Gtex esses são os exões presentes no *RAD51D*, isoforma canônica.

FONTE: captura de imagem do Gtex adaptada pela autora.

C) No nível da proteína, na isoforma canônica, destacando-se o primeiro aminoácido de cada exon.



Sequência de éxons da isoforma canônica do gene *RAD51D* evidenciando o aminoácido inicial para cada éxon.

FONTE: imagem criada pela autora

1.5 A proteína *RAD51D*

O *RAD51* constitui um peptídeo ligante e um domínio da *RecA*. Esses domínios Walker A e Walker B participam da troca de fita de ATP e superfícies de ligação de DNA/proteína (TOH; NGEOW, 2021).

Pouco estudada, o gene *RAD51D* apresenta variantes benignas, patogênicas e VUS (variante de significado incerto). De acordo com o banco de dados Varsome (acesso em 02/04/2025) o *RAD51D* tem 228 variantes patogênicas ou provavelmente patogênicas, 989 VUS e 388 benignas ou provavelmente benigna. (Banco de dados Varsome consultado 02/04/2025)

A proteína *RAD51D*, em sua isoforma canônica, se caracteriza por apresentar os domínios Walker A e Walker B (Figura 7). O Walker A é responsável pela ligação de ATP e está localizado nas posições 107 a 114. O Walker B é responsável pela hidrólise de ATP e está localizado nas posições 219 a 223, que são altamente conservados. A importância destes domínios é que são cruciais para reversão da forquilha de replicação.

FIGURA 7 Domínios *RAD51D*



Localização dos domínios Rad51

Fonte: ADOLPH, Madison B.; CORTEZ, David. Mechanisms and regulation of replication fork reversal. **Dna Repair**, [S.L.], v. 141, p. 103731, set. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.dnarep.2024.103731>. Adaptado pela autora

A proteína *RAD51D* está envolvida nas vias de vários genes relacionados à reversão da forquilha de Holliday, tais como *SMARCAL1*, *ZRANB3*, *HLTF* e *FBH1*.

O *RAD51D* desempenha diversas funções; sua inativação pode bloquear completamente a reversão da forquilha de replicação. Atua como mediadora na inversão da fita e na formação do laço D durante a recombinação homóloga de quebras de fita dupla, formando filamentos de nucleoproteína em substratos. Juntamente com a *BRCA2*, promove a

nucleação do filamento e inibe a atividade ATPase do *RAD51*, estabilizando o filamento (ADOLPH; CORTEZ, 2024).

1.6 Variantes genéticas e câncer

O gene *RAD51D* está envolvido na via de reparo por recombinação homóloga, responsável pela correção de quebras na dupla fita do DNA durante a replicação. Esse gene está associado ao câncer de mama e de ovário, devido à sua atuação nas vias de reparo mediadas pelos genes *BRCA1* e *BRCA2* (YOSHIDA; MIKI, 2004).

Entre as funções da *RAD51D* destaca-se o papel fundamental na recombinação homóloga (HR), essencial para a manutenção da estabilidade do DNA e prevenção do desenvolvimento do câncer. Contudo, embora a inibição da família RAD51 possa causar danos genômicos e aumentar a propensão ao câncer, a regulação positiva e a indução das proteínas dessa família também têm sido associadas à resistência terapêutica e à progressão metastática (GRUNDY; BUCKANOVICH; A BERNSTEIN, 2020).

1.7 Classificação de variantes

Pessoas portadoras de mutações classificadas como patogênicas nos genes *RAD51D* e *RAD51C* apresentam maior risco de desenvolver câncer de mama e ovário. A função do *RAD51D* pode ser afetada por diversos fatores, incluindo metilação do promotor e mutações de diferentes tipos, como alterações no número de cópias, por exemplo, deleções (GRUNDY; BUCKANOVICH; BERNSTEIN, 2020).

Entretanto, variantes genéticas são comuns em todos os genes e nem sempre implicam em repercussões funcionais. Por isso, é fundamental destacar a importância da classificação das variantes do gene *RAD51D*.

As variantes genéticas são classificadas em cinco categorias, conforme critérios estabelecidos pelo *American College of Medical Genetics and Genomics* (ACMG) e pela *American Association of Pathology* (APA):

- Benignas (B);

- Provavelmente benignas (PB);
- Variantes de significado incerto (VUS);
- Provavelmente patogênicas (PP);
- Patogênicas (P).

Essa classificação baseia-se nos seguintes critérios (RICHARDS et al., 2015):

1. Evidência da frequência populacional obtida em diversos bancos de dados;
2. Estudos funcionais realizados em cultivo celular ou modelos animais;
3. Análise bioinformática, que inclui conservação evolutiva do resíduo, tipos de variantes previamente descritas para determinada posição e predição de impacto funcional por meio de diversos algoritmos;
4. Estudos de segregação, demonstrando co-segregação do alelo com o fenótipo em famílias;
5. Relatos de caso descrevendo a variante em indivíduos afetados;
6. Impacto no splicing alternativo, entre outros.

1.8 Importância da conservação evolutiva na classificação de cada resíduo quanto a patogenicidade

O *RAD51D* foi recentemente incluído em painéis multigênicos devido à sua associação com o aumento do risco para câncer de mama e ovário. Regiões ou aminoácidos altamente conservados indicam que uma mutação pontual nessas posições pode gerar variantes potencialmente patogênicas.

A importância desse estudo reside em identificar quais resíduos de aminoácidos são mais ou menos conservados, inclusive até espécies distantes como peixes, para melhor compreender a classificação das variantes relacionadas a esse gene.

1.9 Impacto das isoformas na avaliação da conservação evolutiva

O gene *RAD51D* apresenta múltiplas isoformas em humanos e, ao longo da evolução, possui diferentes isoformas canônicas distribuídas entre diversas classes de organismos. A isoforma humana possui 328 aminoácidos.

Este trabalho tem como objetivo classificar as variantes do gene com base na conservação evolutiva. Contudo, um desafio importante é realizar essa análise considerando a grande diversidade de isoformas existentes.

Por exemplo, nos mamíferos, há uma isoforma com 329 aminoácidos presente em roedores, a qual também pode ser encontrada em outras classes, como aves, répteis e peixes. Outra isoforma provável é a isoforma 8, com 348 aminoácidos, observada em alguns primatas. A isoforma 4, com 213 aminoácidos, é encontrada em várias classes. Além disso, existe uma isoforma não descrita no ser humano, com sequência de 312 e 313 aminoácidos, principalmente em aves.

Quando a conservação evolutiva é avaliada considerando diferentes isoformas simultaneamente, há um risco significativo de erro na determinação da profundidade da conservação de resíduos específicos de aminoácidos.

1.10 Objetivos

1.10.1 Objetivo geral

Investigar o impacto das isoformas na avaliação da conservação evolutiva dos resíduos de aminoácidos da proteína *RAD51D*, que afetam na aplicação clínica na classificação de variantes genéticas quanto à patogenicidade.

1.10.2 Objetivos específicos

Desenvolver um mapa da profundidade de conservação evolutiva de cada aminoácido considerando separadamente as isoformas ENST00000345365.6 (canônica, 328 aminoácidos), (329 aminoácidos, roedores e morcegos).

1.10.3 Justificativa e relevância

Tipicamente, a avaliação da conservação evolutiva é feita considerando-se todas as sequências disponíveis nas bases de dados, ou seja, mistura as diferentes isoformas conhecidas em humanos e em outros organismos. A conservação evolutiva tem muito peso na classificação de variantes e, sozinha, direciona a classificação de uma substituição de aminoácidos qualquer para benigna. Isto leva a uma subestimativa da potencial patogenicidade das substituições de aminoácidos.

2. Materiais e Métodos

2.1 Como as sequências foram obtidas?

O NCBI (*National Center for Biotechnology Information*) é a principal plataforma utilizada para ampliar a busca e encontrar o maior número possível de isoformas do gene *RAD51D*. Para os alinhamentos, foram utilizados os softwares Clustal Ômega e MEGAX. A metodologia adotada neste trabalho foi dividida em três etapas:

Primeira etapa:

Foi realizada uma busca no NCBI utilizando o BLASTP da isoforma canônica para obter as primeiras sequências. Como critérios iniciais de exclusão, foram descartadas sequências preditas, hipotéticas e aquelas com homologia inferior a 70%. Priorizaram-se para análise as sequências canônicas e aquelas com tamanho próximo, como a isoforma de 329 aminoácidos. Considerando a sequência canônica, foram contabilizadas 16 sequências de mamíferos, 16 de aves, uma de répteis, três de peixes e nenhuma de anfíbios.

Segunda etapa:

Esta etapa consistiu na coleta da lista completa de ortólogos do *RAD51D* disponível no NCBI. Foram contabilizadas ao todo 584 sequências: 218 de mamíferos, 191 de peixes,

123 de aves, 41 de répteis e 11 de anfíbios. Inicialmente, não foi aplicado nenhum critério de exclusão. Após o primeiro alinhamento realizado com Clustal Ômega, estabeleceram-se os seguintes critérios para filtragem das sequências: remoção de gaps centrais únicos; exclusão de sequências com início tardio, que poderiam ser readthrough; e exigência de que as sequências tivessem 328 e 329 aminoácidos.

Terceira etapa:

Esta etapa teve como objetivo identificar a isoforma típica de cada classe taxonômica. Para isso, foi utilizado o modelo de cada classe como referência para realizar buscas via BLASTP no NCBI. Os organismos escolhidos para referência foram: *Danio rerio* para peixes; *Gallus gallus* para aves; *Caretta caretta* para répteis; e *Xenopus tropicalis* para anfíbios. Esta metodologia não foi aplicada para mamíferos, pois seria semelhante à da primeira etapa. Para esta etapa, também foram aplicados os critérios de exclusão: homologia inferior a 70%, sequências preditas e hipotéticas; e a seleção das sequências foi interrompida quando houve mudança de classe na lista fornecida pelo NCBI.

2.2 Como as sequências foram analisadas?

Para a análise das sequências, foram realizados diversos alinhamentos visando a máxima exploração do conteúdo disponível. Para isso, utilizamos as ferramentas BLAST do NCBI, MEGAX e Clustal Ômega. Após a obtenção da lista final, dentro dos critérios estabelecidos, a análise passou a ser realizada por posição, avaliando a profundidade de conservação de cada resíduo dentro de cada classe taxonômica, e, em algumas situações, por espécie e isoforma.

Para essa análise detalhada, utilizamos os softwares MEGAX, Alpha Minsesse e ConSurf.

A seguir, na Tabela 1, são apresentados os bancos de dados utilizados nesta metodologia, bem como sua finalidade dentro do projeto.

Tabela 1 Programas e bancos de dados utilizado nas análises

SOFTWARE/SITES	FINALIDADE	LINK
Mega X	Alinhamento e análise	PROGRAMA

Clustal omega	Alinhamento	https://www.ebi.ac.uk/jdispatcher/msa/clustalo
NCBI	Alinhamento e busca de sequências	https://blast.ncbi.nlm.nih.gov/Blast.cgi
GTex	Busca de exon	https://gtexportal.org/home/gene/RAD51D#genetranscript-browser-block
ENSEMBLE	Busca de isoformas	https://grch37.ensembl.org/Homo_sapiens/Transcript/ProteinSummary?db=core;g=ENSG000001379;r=17:33426811-33448541;t=ENST0000036276
UNIPROT	informações do RAD51	https://www.uniprot.org/uniprotkb/O75771/entry/O75771-5
Varsome	Busca por variantes	https://varsome.com/gene/hg38/Rad51d
AlphaMissense	patogenicidade de acordo com a posição	https://alphamissense.hegelab.org/hotspot
ConSurf	trocas de aminoácidos	https://consurf.tau.ac.il/barak/progress.php?number=1736449221
Tax-on-tree	teste de árvore filogenética	Programa
ClinVar	Busca por variante	https://www.ncbi.nlm.nih.gov/clinvar/

Fonte: Feito pela autora

2.3 Análise dos resultados

A última etapa da análise consistiu na construção de tabelas comparativas utilizando os dados obtidos pelo MEGAX, ferramenta que avalia a conservação evolutiva. Foram elaboradas tabelas específicas para mamíferos, primatas e para as demais classes (aves, répteis, anfíbios e peixes).

Observou-se que determinadas regiões apresentavam maior conservação do que outras. Por esse motivo, foi necessário utilizar uma ferramenta da **Thermo Fisher** (<https://www.thermofisher.com/br/en/home/life-science/genome-editing/invitrogen-truedes-ign-genome-editor.html>) para o estudo dos éxons e íntrons conservados.

Com base nesses dados, construímos uma linha do tempo evolutiva para o estudo.

3.0 Resultados e Discussão

O *RAD51D* possui diversas isoformas, e essa diversidade é ainda maior entre diferentes classes taxonômicas. Em outras classes, a maior homologia nem sempre é com a isoforma canônica humana; algumas sequências apresentam maior semelhança com a isoforma 8 (348 aminoácidos) ou com a isoforma 4 (283 aminoácidos). A maioria dessas isoformas é gerada por splicing alternativo (Figura 13), sendo que, por exemplo, a isoforma 8 apresenta a adição de um éxon. Embora a isoforma 8 não tenha sido o foco principal do nosso estudo, ela é encontrada especialmente em primatas e também em uma sequência de aves, sugerindo que essa variante é de ocorrência natural na evolução.

3.1 Quais as isoformas da proteína RAD51D estão presentes em humanos? Quais as diferenças e qual o mecanismo de origem de cada uma delas?

Na tabela 2, são listadas principais características das isoformas da proteína RAD51D descritas em humanos.

Tabela 2 - Isoformas de RAD51D em humanos

isoformas	Quantidades de resíduos de aminoácidos	Principais diferenças com a canônica
canônica	328	
Isoforma 2	49	49-49: A → S 50-328: AUSENTE
Isoforma 3	216	50-161: AUSENTE
Isoforma 4	283	116-160: AUSENTE
Isoforma 5	328	88-101: SLDKLLDAGLYTGE → RQKLSGGSRWCMHL
Isoforma 6	118	88-118: SLDKLLDAGLYTGEVTEIVGGPGSGKTQVCL → RHGGRTQVGTWEDCSCLRSPQGDRGVGSGML 119-328: AUSENTE
Isoforma 7	212	193-212: VTGSSGTVKVVVDSVTAVV → DGIPEHLNHIPHCLHVLPC 213-328: AUSENTE
Isoforma 8	348	49-88: ALVALRRVLLAQFSAFPVNGADLYEELKTSTALILSTGIG S → TWRAHSSGNLGGQLPQVPAGRSWSGVRNALKKAGLGH GGTGSLSLNAFDERGTAVSTSR

FONTE: Tabela criada pela autora, utilizando dados da UniProt (acesso em março de 2023).

A isoforma canônica é a principal isoforma do gene *RAD51D* encontrada em humanos contendo 328 aminoácidos. Além desta, estão registradas outras sete isoformas encontradas em humanos, com tamanhos e homologias diferentes. No quadro 2, são mostradas as isoformas da proteína RAD51D, de acordo com uniprot, com respectivos tamanhos e diferenças entre cada uma delas é a canônica.

Realizando um alinhamento par-a-par entre a isoforma canônica e cada uma das demais isoformas, conseguimos observar as principais diferenças entre elas. Na maioria, o mecanismo observado é o *splicing* alternativo e/ou presença ou ausência de éxon terminal.

A isoforma número dois, O75771-2, tem 49 aminoácidos e foi gerada por *stop codon* prematuro.

FIGURA 8 Alinhamento isoforma canônica X isoforma 2

sp|O75771-2|RA51D_HUMAN Isoform 2 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
 Sequence ID: **Query_2748211** Length: **49** Number of Matches: **1**

Range 1: 1 to 49 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
100 bits(250)	5e-31	Compositional matrix adjust.	48/49(98%)	49/49(100%)	0/49(0%)

Go to previous match #0 for lcl|Query_2748211

Query	1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKA	49
Sbjct	1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYK+	49

Alinhamento da canônica com a isoforma 2, na qual a isoforma 2 é gerada por um stop códon prematuro
 FONTE: Captura de tela do alinhamento realizado pelo NCBI

A isoforma três, O75711-3 possui 216 aminoácidos é formada por *splicing* alternativo. O alinhamento com a isoforma canônica é dividido, da posição 1° a posição 83°, se alinham primeiro com a canônica e depois voltam a se alinhar na posição 160°, ou seja, cerca de 80 aminoácidos estão fora do alinhamento. Esta falha pode corresponder a um éxon inteiro que está presente na isoforma três e não se alinha com a canônica (Figura 9).

FIGURA 9 Alinhamento isoforma canônica X isoforma 3

sp|O75771-3|RA51D_HUMAN Isoform 3 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
 Sequence ID: **Query_2748212** Length: **216** Number of Matches: **2**

Range 1: 48 to 216 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
338 bits(866)	1e-121	Compositional matrix adjust.	168/169(99%)	169/169(100%)	0/169(0%)

Query	160	QAEALRRIQVHAFDIFQMLDVLQELRGTVAAQQTGSSGTVKVVVDSVTAVVSPLLGGQ	219
Sbjct	48	+AEALRRIQVHAFDIFQMLDVLQELRGTVAAQQTGSSGTVKVVVDSVTAVVSPLLGGQ	107
Query	220	QREGALMMQLARELKTLDLGMVAVVTHITRDRSGRLKPALGRSWSFVPSTRILLD	279
Sbjct	108	QREGALMMQLARELKTLDLGMVAVVTHITRDRSGRLKPALGRSWSFVPSTRILLD	167
Query	280	TIEGAGASGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT	328
Sbjct	168	TIEGAGASGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT	216

Range 2: 1 to 83 [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Score	Expect	Method	Identities	Positives	Gaps
109 bits(272)	2e-32	Compositional matrix adjust.	61/86(71%)	69/86(80%)	4/86(4%)

Query	1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRV-LLA	59
Sbjct	1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKA ALRR+ ++	59
Query	60	QFSAFPVNGADLYEELKTSTAILSTG	85
Sbjct	60	F F + D+ +EL+ + A TG	83

Alinhamento canônica com a isoforma 3, na qual possui a ausência do éxon número 4
 FONTE: Captura de tela do alinhamento realizado pelo NCBI

A isoforma 4 possui 283 aminoácidos (O75711-4) e ao se alinhar com a isoforma canônica é notável um splicing alternativo próximo a posição 110° (Figura 10)

A isoforma 5 também possui 328 aminoácidos (O75711-5) e é a mais parecida com a isoforma canônica. Isto é, não parece sofrer splicing alternativo e sim apenas mudanças pontuais de aminoácidos. (Figura 11)

A isoforma 6 possui 118 de aminoácidos (O75711-6) e se alinha com a isoforma canônica até a posição 87 apenas. (Figura 12).

FIGURA 10 Alinhamento isoforma canônica X isoforma 4

sp|O75771-4|RA51D_HUMAN Isoform 4 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
 Sequence ID: **Query_2748213** Length: **283** Number of Matches: **1**

Range 1: 1 to 283 [Graphics](#) [▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
544 bits(1401)	0.0	Compositional matrix adjust.	283/328(86%)	283/328(86%)	45/328(13%)
Query 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Sbjct 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Query 61		FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM			120
Sbjct 61		FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQ			115
Query 121		AANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVWHAFDIFQMLD			180
Sbjct 116		-----AEALRRIQVWHAFDIFQMLD			135
Query 181		VLQELRGTVAQQTGSSGTVKVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLD			240
Sbjct 136		VLQELRGTVAQQTGSSGTVKVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLD			195
Query 241		LGMVVVNTNHIRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGRRMACLAKSSR			300
Sbjct 196		LGMVVVNTNHIRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGRRMACLAKSSR			255
Query 301		QPTGFQEMVDIGTWTSEQSATLQGDQT	328		
Sbjct 256		QPTGFQEMVDIGTWTSEQSATLQGDQT	283		

Alinhamento canônica e isoforma 4, na qual apresenta um GAP no éxon número 4 e o término da proteína é no éxon número 9

FONTE: Captura de tela do alinhamento realizado pelo NCBI

FIGURA 11 Alinhamento isoforma canônica X isoforma 5

sp|O75771-5|RA51D_HUMAN Isoform 5 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
Sequence ID: Query_2748214 Length: 328 Number of Matches: 1

Range 1: 1 to 328 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
619 bits(1597)	0.0	Compositional matrix adjust.	316/329(96%)	317/329(96%)	2/329(0%)
Query 1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ				60
Sbjct 1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ				60
Query 61	FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGE-VTEIVGGPGSGKTQVCLC				119
Sbjct 61	FSAFPVNGADLYEELKTSTAILSTGIG KL + VTEIVGGPGSGKTQVCLC				119
Query 120	MAANVAHGLQQNVLVYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVHAFDIFQML				179
Sbjct 120	MAANVAHGLQQNVLVYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVHAFDIFQML				179
Query 180	DVLQELRGTVAAQVGTGSSGTVKVVVDSVTAVVSPLLGGQOREGLALMMQLARELKTLLAR				239
Sbjct 180	DVLQELRGTVAAQVGTGSSGTVKVVVDSVTAVVSPLLGGQOREGLALMMQLARELKTLLAR				239
Query 240	DLGMVVVNTNHIRDRDGRGRLKPALGRSWSFVPSTRILLDTIEGAGASGRRMACLAKSS				299
Sbjct 240	DLGMVVVNTNHIRDRDGRGRLKPALGRSWSFVPSTRILLDTIEGAGASGRRMACLAKSS				299
Query 300	RQPTGFQEMVDIGTWTGTEQSATLQGDQT		328		
Sbjct 300	RQPTGFQEMVDIGTWTGTEQSATLQGDQT		328		

Alinhamento isoforma canônica com isoforma 5. Ambas têm 328, não há *splicing* alternativo, apenas mudanças pontuais de aminoácidos.

FONTE: Captura de tela do alinhamento realizado pelo NCBI

FIGURA 12 Alinhamento isoforma canônica X isoforma 6

sp|O75771-6|RA51D_HUMAN Isoform 6 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
Sequence ID: Query_2748215 Length: 118 Number of Matches: 1

Range 1: 1 to 87 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
174 bits(441)	1e-58	Compositional matrix adjust.	87/87(100%)	87/87(100%)	0/87(0%)
Query 1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ				60
Sbjct 1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ				60
Query 61	FSAFPVNGADLYEELKTSTAILSTGIG		87		
Sbjct 61	FSAFPVNGADLYEELKTSTAILSTGIG		87		

Alinhamento isoforma canônica com a isoforma 6, na qual a isoforma 6 só alinha até o aminoácido 87, ou seja, até o éxon número 3.

FONTE: Captura de tela do alinhamento realizado pelo NCBI

A isoforma 7 possui 212 aminoácidos (O75711-7), quando alinhada com a isoforma canônica, assim como a isoforma seis, não se alinha por completo. O alinhamento ocorre da posição 1° até a posição 192°. (Figura 13)

FIGURA 13 Alinhamento isoforma canônica X isoforma 7

sp|O75771-7|RA51D_HUMAN Isoform 7 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
Sequence ID: Query_2748216 Length: 212 Number of Matches: 1

Range 1: 1 to 192 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
385 bits(988)	4e-140	Compositional matrix adjust.	192/192(100%)	192/192(100%)	0/192(0%)
Query 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Sbjct 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Query 61		FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM			120
Sbjct 61		FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM			120
Query 121		AANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVVFHAFDIFQMLD			180
Sbjct 121		AANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVVFHAFDIFQMLD			180
Query 181		VLQELRGTVAAQ 192			
Sbjct 181		VLQELRGTVAAQ 192			

Alinhamento da isoforma canônica com a isoforma 7, na qual o alinhamento vai até o aminoácido 192, ou seja, até o éxon de número 6.

FONTE: Captura de tela do alinhamento realizado pelo NCBI

Tanto a isoforma 6 quanto a isoforma 7 parecem ter um último éxon que não é presente na isoforma canônica.

A isoforma 8 possui 348 aminoácidos (O75711-8) e é comumente vista em primatas. Quando alinhada com a isoforma canônica podemos observar um *splicing* alternativo ocorrendo na isoforma canônica, ou seja a isoforma 8 possui uma região a mais que a isoforma 328 não tem. (Figura 14)

FIGURA 14 Alinhamento isoforma canônica X isoforma 8

sp|O75771-8|RA51D_HUMAN Isoform 8 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D
Sequence ID: Query_2748217 Length: 348 Number of Matches: 1

Range 1: 1 to 348 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
535 bits(1378)	0.0	Compositional matrix adjust.	294/350(84%)	299/350(85%)	24/350(6%)
Query 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Sbjct 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVDLVSADLEEVAQKCGLSYK A L			60
Query 61		FSAFPVNGADLYEELKTSTAILSTGIGS-----LDKLLDAGLY			98
Sbjct 61		LQLPQVPAGRSWSGVR--NALKKAGLGHGGTDGLSLNAFDERGTVAVSTRDKLLDAGLY			118
Query 99		TGEVTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEE			158
Sbjct 119		TGEVTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEE			178
Query 159		EQAEALRRIQVVFHAFDIFQMLDVLQELRGTVAAQVGTGSSGTVKVVVDSVTAVVSPLLGG			218
Sbjct 179		EQAEALRRIQVVFHAFDIFQMLDVLQELRGTVAAQVGTGSSGTVKVVVDSVTAVVSPLLGG			238
Query 219		QOREGLALMMQLARELKT LARDLGMVVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILL			278
Sbjct 239		QOREGLALMMQLARELKT LARDLGMVVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILL			298
Query 279		DTIEGAGASGGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT			328
Sbjct 299		DTIEGAGASGGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT			348

Alinhamento da isoforma canônica com a isoforma 8, na qual podemos observar o *splicing* alternativo sendo iniciado próximo ao aminoácido 48 (éxon 3 e 4) e sofrendo um GAP no éxon número 4.

FONTE: Captura de tela do alinhamento realizado pelo NCBI e adaptado pela autora

3.2 Qual a diferença entre as isoformas 328 e 329?

Neste estudo, o objetivo é focar, principalmente, nas isoformas 328 e 329. Além disso, são vistas mais 56 substituições de aminoácidos na isoforma 329, mas parecem não comprometer a parte funcional da proteína ao ponto de ser visto como uma variante genética.(Figura 15). Estas substituições aparecem distribuídas ao longo da sequência, sugerindo que não se trate de *splicing* alternativo.

Nas nossas buscas pelo NCBI, Uniprot e outras bases de dados, é perceptível que a isoforma canônica (328 aminoácidos) é a principal isoforma presente na classe dos mamíferos, estando pouco representada nas demais classes. Ao contrário, a isoforma 329 se encontra bastante presente nas demais classes e pouco presente em mamíferos. Em mamíferos, a isoforma 329 só foi encontrada em algumas espécies de roedores e em um morcego (*Sturnira hondurensis*). Dentre tantas isoformas do *RAD51D* escolhemos a isoforma 328 por ser a canônica de humanos e a 329 por ser a isoforma comprovadamente existente com mais similaridade com a isoforma canônica, pois mesmo com alterações pontuais de aminoácidos, a diferença mais marcante é adição de uma arginina na posição 288°.

A isoforma 329 se apresenta com mais frequência nas classes de aves, peixes, e répteis e anfíbios, não necessariamente sendo a canônica das espécies e classes. Observamos que em aves a sequência com maior predominância foram sequência que não são alvos deste estudo, as sequências com 312/313 aminoácidos, por exemplo. Em répteis e anfíbios, as sequências 339 aminoácidos se apresentaram bastante presentes também.

FIGURA 15 Comparação da Isoforma canônica em humana com a isoforma 329

```

Query: unnamed protein product Query ID: lcl|Query_44281 Length: 328

>NP_001100499.1:1-329 DNA repair protein RAD51 homolog 4 [Rattus norvegicus]
Sequence ID: Query_44283 Length: 329
Range 1: 1 to 329

Score:555 bits(1429), Expect:0.0,
Method:Compositional matrix adjust.,
Identities:275/329(84%), Positives:299/329(90%), Gaps:1/329(0%)

Query 1   MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Sbjct 1   ..M..A.....IV....GRK....A..AA..... 60

Query 61  FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM 120
Sbjct 61  .....L.....L.....V 120

Query 121 AANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD 180
Sbjct 121 .....S.....M.....R.....K..S..Q.....S..... 180

Query 181 VLQELRGTVAAQQTGSSGTVKVVVDSVTAVVSPLLGGQQREGLALMMQLARELKT LARD 240
Sbjct 181 M..D....M...A.A.....I.....A.....I.... 240

Query 241 LGMVVVNTNHRDRDRSGRLKPALGRSWSFVPSRILLDTIEGAGASG-GRRMACLAKSS 299
Sbjct 241 ..V.....L.....R.F.....EVF....TL.RSQ.TVR.I..P 300

Query 300 RQPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Sbjct 301 .....L..VI....L..E...PE.P.K.. 329

```

Alinhamento pairwise da isoforma canônica com a 329.

FONTE: Captura de tela do alinhamento realizado pelo NCBI

3.3 Como a isoforma presente em humanos aparece nas outras classes?

3.3.1 Mamíferos

A isoforma canônica do RAD51D está amplamente presente em mamíferos e é pouco representada nas demais classes. Em roedores, foi identificada uma isoforma semelhante, contendo 329 aminoácidos, altamente conservada dentro do grupo, com cerca de 99% de homologia entre as diferentes espécies de roedores (BONILLA et al., 2020).

Nossa análise foi realizada por classe taxonômica, conforme descrito a seguir:

Metodologia 1: Baseou-se em um BLASTP utilizando a isoforma canônica (ENST00000345365.6) como sequência de consulta. No primeiro levantamento, foram encontradas 101 sequências. Alinhamentos subsequentes foram realizados para avaliar similaridade e homologia em relação à sequência canônica. Aplicando os critérios de exclusão para sequências preditas, hipotéticas e aquelas com homologia inferior a 70%, o conjunto foi reduzido a 66 sequências. Destas, 16 representaram as isoformas em estudo, sendo 9 isoformas com 328 aminoácidos e 7 com 329 aminoácidos.

Metodologia 2: Consistiu na análise da lista de ortólogos, aplicando critérios rigorosos de exclusão, como remoção de sequências com gaps centrais e/ou únicos, início tardio ou precoce, e sequências readthrough, independentemente da classe. Inicialmente, 104 sequências foram excluídas por apresentarem muitos gaps ou por estarem muito distantes da isoforma canônica. Posteriormente, a análise foi realizada por classe, considerando os mesmos critérios. Para mamíferos, havia inicialmente 218 sequências, das quais 9 foram excluídas por início precoce e 3 por gaps únicos.

Com foco nas isoformas de 328 e 329 aminoácidos, restaram 77 sequências da isoforma 328 e 22 da isoforma 329.

Metodologia 3: Baseada em BLASTP do principal modelo para cada classe, não foi aplicada para mamíferos, pois seria redundante com a metodologia 1.

Assim, ao final da análise, contamos com um total de 86 sequências da isoforma 328 e 29 sequências da isoforma 329 para mamíferos.

3.3.2 Aves

Na metodologia 1, foi realizado um BLASTP filtrando para que a busca incluísse apenas a classe Aves. Foram encontradas 279 sequências. Aplicando os critérios de exclusão previamente estabelecidos e realizando os seguintes alinhamentos:

- 328/329 aves × canônica;
- 329 aves × 329 mamíferos (*Rattus norvegicus*, modelo);
- 312/313/315 aves × canônica;
- Isoforma 283 aves × canônica;
- Isoforma 348 aves × canônica.

Após o refinamento, obtiveram-se 14 sequências com 328 aminoácidos e 2 sequências com 329 aminoácidos.

Na metodologia 2, após a divisão de classes da lista de ortólogos, havia inicialmente 118 sequências. Após o alinhamento e aplicação dos critérios de exclusão, restaram 45 sequências, todas com 329 aminoácidos.

Na metodologia 3, utilizou-se como modelo *Gallus gallus* (galo doméstico). Seguindo os critérios de exclusão, obtiveram-se 298 sequências de aves. Parte dessas sequências corresponde a isoformas distintas, sendo as mais comuns as de 312, 313 e 315 aminoácidos, amplamente presentes em aves. Ao final, restaram 28 sequências de 329 aminoácidos e 2 de 328 aminoácidos.

Total geral para Aves: 16 sequências de 328 aminoácidos e 75 sequências de 329 aminoácidos.

3.3.3 Répteis

Na metodologia 1, o BLASTP foi filtrado para selecionar apenas sequências da classe Reptilia. Foram inicialmente obtidas 11 amostras. Após a aplicação dos critérios de exclusão previamente definidos, o conjunto foi reduzido para 7 sequências. Entre estas, observou-se um problema de homologia: muitas eram mais semelhantes a outros membros da família RAD51 do que ao RAD51D.

Dessa forma, restaram apenas 3 sequências, todas com 329 aminoácidos, que foram utilizadas nos seguintes alinhamentos:

- 329 répteis × canônica;
- 329 répteis × 329 mamíferos;
- 3 sequências × 328 canônica;
- 329 répteis × demais sequências de répteis.

Apenas uma dessas sequências apresentou correspondência com os critérios do estudo, sendo de 329 aminoácidos.

Na metodologia 2, após a separação por classes, foram obtidas 41 amostras de répteis. Aplicando os critérios de exclusão, 5 sequências foram removidas. Entre as 36 sequências restantes, identificou-se apenas 1 com 328 aminoácidos e 18 com 329 aminoácidos.

Na metodologia 3, utilizou-se *Caretta caretta* (tartaruga-comum) como modelo para o BLASTP. Foram encontradas 27 sequências, das quais apenas 2 eram compatíveis com as isoformas de interesse (328 e 329 aminoácidos), sendo ambas de 329 aminoácidos.

Total geral para Répteis: 23 sequências de 329 aminoácidos e 1 sequência de 328 aminoácidos.

3.3.4 Anfíbios

Na metodologia 1, foi realizado um BLASTP filtrando apenas para a classe Amphibia. Foram inicialmente encontradas 59 amostras, das quais 20 foram selecionadas após a aplicação dos critérios de exclusão. No entanto, muitas dessas sequências correspondiam a outras isoformas com baixa similaridade ao RAD51D, chegando a apresentar maior homologia com outros membros da família RAD51. Por essa razão, nenhuma sequência foi aproveitada nesta metodologia.

Na metodologia 2, após a separação por classes, foram identificadas 11 sequências de anfíbios. Entre estas, apenas 1 sequência correspondia à isoforma de 328 aminoácidos.

Na metodologia 3, utilizou-se *Xenopus tropicalis* (rã tropical) como modelo para o BLASTP. Entretanto, os resultados retornados já se iniciavam com sequências da classe Reptilia, não sendo identificada nenhuma sequência de interesse para anfíbios.

Total geral para Anfíbios: 1 sequência de 328 aminoácidos e nenhuma sequência de 329 aminoácidos.

3.3.5 Peixes

Na metodologia 1, o BLASTP foi filtrado para buscar apenas sequências da classe dos peixes, resultando em 120 sequências iniciais. Após a aplicação dos critérios de exclusão previamente definidos, restaram 58 sequências. Ao remover as isoformas que não eram de interesse para este estudo, permaneceram apenas 3 sequências correspondentes à isoforma de 329 aminoácidos.

Para análise, foram realizados os seguintes alinhamentos:

- 329 peixes × canônica;
- 329 peixes × 329 mamíferos (*modelo: Rattus norvegicus*);
- 329 peixes × 348 mamíferos.

Na metodologia 2, após a separação por classes, foram obtidas 191 sequências de peixes. Seguindo os critérios de exclusão, 8 amostras foram descartadas. Entre as sequências restantes, apenas 2 sequências correspondiam à isoforma de 328 aminoácidos e 2 sequências à isoforma de 329 aminoácidos.

Na metodologia 3, utilizou-se *Danio rerio* (peixe-zebra) como modelo para o BLASTP. Após aplicar os critérios de exclusão, obteve-se um conjunto de 140 amostras, das quais apenas 2 sequências correspondiam à isoforma de 329 aminoácidos.

Total geral para Peixes: 2 sequências de 328 aminoácidos e 7 sequências de 329 aminoácidos.

3.4 Como é o alinhamento das sequências da isoforma canônica humana (328 aminoácidos) com suas correspondentes em outras espécies?

Utilizando todas as sequências obtidas nas três metodologias, foi realizado um alinhamento das sequências com 328 aminoácidos, abrangendo 86 mamíferos, 2 peixes, 1 réptil, 1 anfíbio e 16 aves.

Com exceção dos mamíferos, as demais classes são compostas, em sua maioria, por outras isoformas que não serão investigadas neste estudo.

O alinhamento mostrou-se uniforme, sem *gaps* significativos, inícios tardios e precoces, apresentando boa conservação em parte da proteína, com cerca de 132 resíduos conservados até peixes. Observou-se maior divergência no final das sequências, especialmente nos últimos 30 resíduos.

No mesmo alinhamento, comparando-se à isoforma canônica, identificou-se uma alteração característica de *splicing alternativo*, presente em 16 espécies — todas passeriformes — iniciando-se na posição 161^o e estendendo-se até a posição 177^o. O *gap* complementar a essa alteração ocorre nas últimas 16 posições da proteína. Parte das substituições de aminoácidos identificadas nesse alinhamento é considerada benigna e/ou de baixo impacto funcional (Figura 16).

FIGURA 16 Alinhamento 328

XP_033834654.1	PDRSTHSFSSLRSATLTKSSRQPCHRVVEFDLSMHSRSKEVPSAKRHS	XP_033834654.1	CFGVAHVIALHLKQGVIFIDTTGGITATRLHQMLQTESSSSLEEQI-----
XP_043910077.1	KDEARSVRSAKRIATLLKSSRQPIRISKDFLEDGCFMEEHPTDIDAGR	XP_043910077.1	CLGVAVNVSYELKQVNLVYDSSGGFTASRLQLQWKSRTDCIEDQ-----
XP_029468865.1	KEGCGNGTNRQIATLSKSPRQAGLCVALEIGSSGVLEENPVLKSEGL	XP_029468865.1	CLSIAVNVSHSLKQVNLVYDSTGGFNASRLFLQVSKTQEEEEQV-----
XP_020659597.1	SKDSAGKATFRKAVLHKSPPRQAGIQVVLNLEHMDQTEDTLFISSCEGP	XP_020659597.1	CLNIAAHITHSLQQVNLVYDSTGGFTATRLQLLQGRTEEEEQ-----
NWR48469.1	REATHKATTQRAASLAKSPRLPTGIQVELDIG-----	NWR48469.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEQARQILGRL-RLCFVLQ
NXN03091.1	KEATHDKATTQRAATLAKSPRQPTGIQVELDIG-----	NXN03091.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLHQMLQARTDEEEQARQVLRGRL-RLCFVLQ
NWR04180.1	KEATHDKASTQRAATLAKSPRQPTGIQVELDIG-----	NWR04180.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLQARTDEEEQARQVLRGRL-RLCFVLQ
NXR49687.1	KEATHKATTQRAASLAKSPRQPTGIQVELDLG-----	NXR49687.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVLQ
NXQ41541.1	REGTHKGTQRAASLAKSPRQPTGIQVELDIG-----	NXQ41541.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVLQ
NWV65419.1	REGTHKATTQRAASLAKSPRQPTGIQVELDIG-----	NWV65419.1	CLGIAASVSLGLKQVHFLDSTAGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVPO
NWV46519.1	REATHKATTQRAASLAKSPRQPTGIQVELDIG-----	NWV46519.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLQARTDEEEQARQVLRGRL-RLCFVLQ
NXN51848.1	KEATHDKATTQRAATLAKSPRQPTGIQVELDIG-----	NXN51848.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVLQ
NXX33842.1	KEATHKGTQRAASLAKSPRQPTGIQVELDIG-----	NXX33842.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVLQ
NXW76011.1	KEATHKATTQRAASLAKSPRQPTGIQVELDIG-----	NXW76011.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVLQ
NWU23788.1	KEATHKASTRRVASLAKSPRQPTGIQVELDI-----	NWU23788.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQANCLTLMTPPELCKXQ
NXQ45909.1	REGSHKGTTHRAASLAKSPRQPTGIQVELDIG-----	NXQ45909.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLSFLVQ
NXI78041.1	REATHDKATTRRVASLAKSPRQPTGIQVELDIG-----	NXI78041.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVLQ
NXB20062.1	KEATHKATTQRLASLAKSPRQPTGIQVELDIG-----	NXB20062.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARRVLRGRL-RLCFVPO
NXJ29801.1	KEATHKATTQRAASLAKSPRQPTGIQVELDIG-----	NXJ29801.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVPO
NXH32278.1	KEATHKATTQRAASLAKSPRQPTGIQVELDIG-----	NXH32278.1	CLGIAASVSLGLKQVHFLDSTGGFTASRLYQLRAQTEDEEEQARQVLRGRL-RLCFVPO
XP_001373739.2	REGSQVPH-GCRVASLTKSPRQPTGIQEVVEIGSWGTEPELSPSLERNKP	XP_001373739.2	CLCTAANVAHSLGQVNLVYDSTGGFTASRLQLLQARTDEEEEQ-----
XP_020861786.1	REGGAPAG-GCRVASLTKSPRQPTGIQEVVEIGSWGTEPELSPSLERNKP	XP_020861786.1	CLCTAANVAHSLGQVNLVYDSTGGFTASRLQLLQARTDEEEEQ-----
XP_036621911.1	REGGAPAG-GCRVASLTKSPRQPTGIQEVVEIGSWGTEPELSPSLERNKP	XP_036621911.1	CLCTAANVAHSLGQVNLVYDSTGGFTASRLQLLQARTDEEEEQ-----
XP_006890969.1	REGGAPAG-GCRVASLTKSPRQPTGIQEVVEIGSWGTEPELSPSLERNKP	XP_006890969.1	CLCTAANVAHSLGQVNLVYDSTGGFTASRLQLLQARTDEEEEQ-----
XP_003469653.1	GEGAGLLG-SHRVACLTKSSRQATGIQETVYDVGSGWTEPELSPSLERNKP	XP_003469653.1	CLCVAANVAHSLGQVNLVYDSTGGFTASRLQLLQARTDEEEEQ-----
XP_006890969.1	GEGAGLLG-SHRVACLTKSSRQATGIQETVYDVGSGWTEPELSPSLERNKP	XP_006890969.1	CLCVAANVAHSLGQVNLVYDSTGGFTASRLQLLQARTDEEEEQ-----

Alinhamento 328, observa-se 2 *gaps*, um a partir da posição 161^o e o outro no final da proteína. *Gaps* formados por sequências de passeriformes.

FONTE: Captura de tela do alinhamento realizado pelo Clustal Ômega e adaptado pela autora

3.5 Qual a profundidade de conservação evolutiva em mamíferos na isoforma canônica?

Após a realização das três metodologias desta análise, foram selecionadas 86 sequências de mamíferos para o alinhamento múltiplo na plataforma MEGAX. Nesse alinhamento, 132 posições mostraram-se totalmente conservadas entre as 86 espécies de mamíferos.

Entretanto, algumas posições apresentaram variação. Por exemplo, a Val66 da sequência canônica foi substituída, em diversas espécies, por leucina, fenilalanina, cisteína, serina ou isoleucina.

Observou-se também um padrão característico entre as espécies que possuem a isoforma 328. Um exemplo é o *Phascolarctos cinereus* (coala), que diverge em várias posições. Outras espécies com divergência acentuada incluem *Cricetulus griseus*, *Monodelphis domestica* e *Trichosurus vulpecula*, todos representantes da ordem dos roedores.

Além disso, verificou-se que, em alguns casos, famílias de espécies apresentam variações de forma correlacionada. Por exemplo, sempre que *Loxodonta africana* (elefante-africano) apresenta uma alteração, *Elephas maximus indicus* (elefante-indiano) também exibe a mesma modificação no mesmo aminoácido.

3.6 Qual o grau de profundidade da conservação evolutiva isoforma canônica de humanos em outras classes?

A isoforma canônica do gene *RAD51D* em humanos, com 328 aminoácidos, é predominantemente encontrada em mamíferos e apresenta alto grau de conservação evolutiva quando o alinhamento é realizado apenas dentro dessa classe. Entretanto, quando comparada a outras classes, o nível de conservação diminui. A classe Aves apresenta um padrão conservado, embora possua um gap significativo nas sequências de passeriformes, localizado entre as posições 161° e o final da proteína. Como mostrado na Figura 17, a conservação varia de acordo com a posição do resíduo. A isoforma canônica é pouco frequente em espécies de anfíbios e peixes. No alinhamento múltiplo que inclui aves, répteis, anfíbios e peixes, foram analisadas apenas 21 sequências, das quais 107 posições mostraram-se totalmente conservadas.

Parte da redução na conservação se deve à abertura de um gap entre as posições 161° e 177° nas sequências de aves. As regiões conservadas são semelhantes às observadas no alinhamento de mamíferos, especialmente nas posições iniciais da proteína. Seguindo o

mesmo critério aplicado aos mamíferos, verificou-se que existe um grau de conservação até peixes, porém ele não é integral: as poucas sequências de répteis (*Pogona vitticeps*), de anfíbios (*Rhinatrema bivittatum*) e de peixes (*Periophthalmus magnuspinnatus* e *Protopterus annectens*) apresentam múltiplas substituições de aminoácidos. Assim, conclui-se que, entre as demais classes, Aves é a que apresenta maior conservação em relação à isoforma canônica humana. A Figura 17 resume os resultados da análise da conservação evolutiva por resíduo.

FIGURA 17 Grau de profundidade da isoforma canônica

```
>sp|O75771|RA51D_HUMAN DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D PE=1 SV=1

MGVLRVGLCPGLTEMIQLLRSHRIKT---VVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ
FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLTGEVTEIVGGPGSGKTQVCLCM
AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAE-----ALRRIQVVHAFD
IFQMLDVLQELRGTVAQQVTGSSGTVKVVVDSVTAVPSLLGGQQREGLALMMQLA
RELKTLARDLGMAVVVTNHIRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAG-ASGGRR
MACLAKSSRQPTGFQEMVDIGTWGTSEQSATLOGDQT
```

Proteína *RAD51D* isoforma canônica (328), profundidade por posição de aminoácido.

FONTE: Feito pela autora com base de dados em alinhamentos MEGAX

LEGENDA

Letras em preto: Posições conservadas

Letras em vermelho: Posições não conservadas

Letras em rosa: Conservada até mamíferos

Letras em azul: Conservada até aves

Letras em roxo: Conservadas até réptil

Letras em amarelo: Conservadas até anfíbios

Letras em verde: Conservação em peixes

Letras sublinhadas: Tabela 3

Tabela 3

Posição	significado
16, 17 e 18	Conservado em aves e não em mamíferos
31	Não conservado em répteis e aves
36	Não conservado em répteis e peixes
38	Não conservado em aves e peixes
41	Não conservado em anfíbios
60	Não conservado em anfíbios
63	Não conservado em répteis
83	Não conservado em aves
123	Anfíbios conservado
129	Répteis conservado
133	Conservado em répteis e anfíbios
141	Não conservado em anfíbios
143	Não conservado em répteis
145	Não conservado em mamíferos
148	Conservado em répteis e anfíbios
162	Não conservado em anfíbios
173	Não conservado em aves
175	Peixes conservado
177 e 179	Não conservado em aves
180	Conservado em répteis e anfíbios
183	Não conservado em mamíferos
192	Não conservado em répteis
196	Conservado em anfíbios e répteis
197 e 198	Conservado em anfíbios e aves
199	Conservado em anfíbios
201	Não conservado em aves
217	Não conservado em répteis
221	Não conservado em anfíbios
228	não conservado em répteis
230	Não conservado em aves
231, 235, 238 e 239	Não conservado em répteis e peixes

236	Não conservado em anfíbios
240	Conservado em anfíbios
241	Conservado em répteis
243	Não conservado em anfíbios
246	Não conservado em répteis
249	Não conservado em répteis e aves
253 e 257	Não conservado em anfíbios
256	Conservado em anfíbios
275	Não conservado em peixes e aves
281	Conservado em anfíbios e aves
282	Conservado em anfíbios
285	Conservado em mamíferos
295	Conservado em aves
300	Não conservado em mamíferos
313	Conservado em anfíbios

3.7 Qual a classe tem a isoforma semelhante canônica de humanos mais bem conservada?

Realizamos também uma análise individual por classe, na qual verificamos que a classe mais bem preservada foi a das aves. Conforme ilustrado na Figura 18, dezesseis sequências apresentaram alto grau de conservação ao longo de todo o alinhamento, em contraste com as demais classes, que exibiram múltiplas substituições de aminoácidos e menor preservação fora das regiões correspondentes ao domínio funcional.

Apesar de ser a classe mais conservada, Aves é também uma das que apresentam maior diversidade de isoformas. Um exemplo é o galo-comum (*Gallus gallus*), que possui sete isoformas distintas, com diferentes comprimentos, a maioria não correspondendo às isoformas de 328 e 329 aminoácidos observadas em mamíferos — que são o foco deste estudo.

FIGURA 18 - Parte inicial do alinhamento do grupo de aves

```

CLUSTAL O(1.2.4) multiple sequence alignment

NWR48469.1      MVVLRAGLCPGLTEDMIQLLRDDGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXN03091.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NWR04180.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXR49687.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXO41541.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NWX65419.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NWX46519.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXM51848.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXX33842.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXW76011.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NWU23788.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXQ45909.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXI78041.1     MVVLRAGLCPGLTEDMIQLLRASGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXB20062.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXJ29801.1     MVVLRAGLCPGLTEDMIQLLRASGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
NXH32278.1     MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEHVAQKCSLSYKALVAVRRVLLAQ
*****

NWR48469.1      FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGELTELMGAPGSGKTQVCLGI
NXN03091.1     FSAFPANGADLYEELKGSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NWR04180.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXR49687.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXO41541.1     FSAFPTNGADLYEELKSSTAILPTGSPGLDKLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NWX65419.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NWX46519.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXM51848.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXX33842.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXW76011.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NWU23788.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXQ45909.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXI78041.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXB20062.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXJ29801.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
NXH32278.1     FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGCVTELMGAPGSGKTQVCLGI
*****

NWR48469.1      AASVSLGLKQHVLF.LDSTGGFTASRLYQMLRAQTEDEEEQARQVLGRL-RLCFVLQLEAL
NXN03091.1     AASVSLGLKQHVLF.LDSTGGFTASRLHQLRAQTEDEEEQARQVLGRL-RLCFVLQLEAL
NWR04180.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLRAQTEDEEEQARQVLGRL-RLCSVLQLEAL
NXR49687.1     AASVSLGLKQHVLF.LDSTGGFTASRLYEMLRQTEDEEEQARQVLGRL-RLCFVLQLEAL
NXO41541.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLRAQTEDEEEQARQVLGRL-RLCFVLQLEAL
NWX65419.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLRAQTEDEEEQARQVLGRL-GLCFVLQLEAL
NWX46519.1     AASVSLGLKQHVLF.LDSTAGFTASRLYQMLRAQTEDEEEQARQVLGRL-RLCFVPQLEAL
NXM51848.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLRAQTEDEEEQARQVLGRL-RLCFVLQLEAL
NXX33842.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQILRAQTEDEEEQARQVLGRL-RLCFVLQLEAL
NXW76011.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQILRAQTEDEEEQARQVLGRL-KLCFVLQLEAL
NWU23788.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLQAQTEDEEEQANCLILWTPPEPLCXPLQLEAL
NXQ45909.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLQAQTEDEEEQARQVLGRL-RLSFLVQLEAL
NXI78041.1     AASVSLGLKQHVLF.LDSTAGFTASRLYQMLQAQTEDEEEQARQVLGRL-RLCFVLQLEAL
NXB20062.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLQAQTEDEEEQARQVLGRL-RLCFVPQLEAL
NXJ29801.1     AASVSLGLKQHVLF.LDSTGGFTASRLYQMLQAQTEDEEEQARRVLGRL-RLCFVPQLEAL
NXH32278.1     AASVSLGLKQYVLF.LDSTGSFTASRLYQMLGAQTEDEEEQARQVLGRL-RLCFVPQLEAL
*****


```

FONTE: Captura de tela do alinhamento realizado pelo Clustal Ômega e adaptado pela autora

Sendo a classe mais bem conservada, também é umas das que mais possuem isoformas. Vejamos a espécie do galo comum *Gallus gallus*, possui 7 isoformas diferentes, com diferentes tamanhos e que não necessariamente possuem relação aos resíduos 328 e 329 dos mamíferos que estão presentes neste estudo.

- Alinhamento *Gallus gallus*

Neste alinhamento temos duas sequências de 402 aminoácidos, duas de 370 aminoácidos e duas de 327 aminoácidos. As isoformas maiores possuem um início distinto e precoce das de 327 aminoácidos. Em outras palavras, considera-se então, que as quatro maiores isoformas possuem um início alternativo, um éxon a mais em relação às sequências de 327 aminoácidos. (Figura 19)

FIGURA 19 Parte do alinhamento *Gallus gallus*

Alignments:

Query	1	MAAGVVWVALKGHPLGVTRHHCSVGGPMGMLRAVCPHGPPWYPRVRYGTPWVPIPSPHRS	60
Query_676558	1	MAAGVVWVALKGHPLGVTRHHCSVGGPMGMLRAVCPHGPPWYPRVRYGTPWVPIPSPHRS	60
Query_676559	1	MAAGVVWVALKGHPLGVTRHHCSVGGPMGMLRAVCPHGPPWYPRVRYGTPWVPIPSPHRS	60
Query_676560	1	MAAGVVWVALKGHPLGVTRHHCSVGGPMGMLRAVCPHGPPWYPRVRYGTPWVPIPSPHRS	60
Query_676561	1	MAAGVVWVALKGHPLGVTRHHCSVGGPMGMLRAVCPHGPPWYPRVRYGTPWVPIPSPHRS	60
Query	61	LRFHPRSA PSHSAGAMVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	120
Query_676558	61	LRFHPRSA PSHSAGAMVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	120
Query_676559	61	LRFHPRSA PSHSAGAMVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	120
Query_676560	61	LRFHPRSA PSHSAGAMVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	120
Query_676561	61	LRFHPRSA PSHSAGAMVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	120
Query_676562	1	MVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	45
Query_676563	1	MVVL RAGLCPGLTEEMIQLLRANNIRTVVDFVSSDLEDVAQSCSL	45

Alinhamento das isoformas *Gallus gallus*, onde observamos o início tardio em duas das isoformas presentes neste alinhamento que possuem 327 aminoácidos.

FONTE: Captura de tela do alinhamento realizado pelo Clustal Ômega

Também pode-se observar um pequeno *splicing* alternativo no meio do alinhamento nas isoformas de 370 aminoácidos (Figura 20)

FIGURA 20 Parte do alinhamento *Gallus gallus*.

Query	181	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEEQLEALQ	240
Query_676558	181	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEEQLEALQ	240
Query_676559	181	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEEQLEALQ	240
Query_676560	181	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEE-----	234
Query_676561	181	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEE-----	234
Query_676562	106	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEEQLEALQ	165
Query_676563	106	AGAPGSGKTQVCLSLIAASVSLGLRQHVFFLDSTGGFTASRLYQMLQARVEDKEEQLEALQ	165
Query	241	RVQVVRMFDIYEMLRALHEVRDCLSQQVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	300
Query_676558	241	RVQVVRMFDIYEMLRALHEVRDCLSQQVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	300
Query_676559	241	RVQVVRMFDIYEMLRALHEVRDCLSQQVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	300
Query_676560	235	-----QVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	268
Query_676561	235	-----QVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	268
Query_676562	166	RVQVVRMFDIYEMLRALHEVRDCLSQQVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	225
Query_676563	166	RVQVVRMFDIYEMLRALHEVRDCLSQQVESSAGPLKAVLIDSVSAVLSPLLGGGRQSEGLA	225

Alinhamento das isoformas *Gallus gallus*, onde observamos um GAP nas isoformas de 370 aminoácidos.

FONTE: Captura de tela do alinhamento realizado pelo Clustal Ômega

Mesmo que tenham sido encontradas estas divergências, o alinhamento é limpo e as isoformas são extremamente semelhantes entre si, mais um vez comprovando , que embora seja uma classe com variadas isoformas, poucas são as mudanças em relação a profundidade e conservação.

3.8 Como é o alinhamento da isoforma 329 de roedores e qual o grau de profundidade de conservação evolutiva?

O resultado das três metodologias, resultou em 134 sequências da isoforma de 329 aminoácidos. Realizando o alinhamento múltiplo no MEGAX, foram identificados gaps e apenas 34 posições que são totalmente conservadas até os peixes.

Foi visto a necessidade de realizar uma “limpeza” no alinhamento, logo algumas sequências foram retiradas para obtenção de um alinhamento mais coeso, em termos de profundidade evolutiva da isoforma 329.

Os critérios de exclusão utilizados foram os mesmos que utilizou-se para as amostras 328: retirada de gaps centrais, gaps únicos, início precoce e início tardio. Foram retiradas as seguintes sequências:

- #XP_028564325.1 Podarcis muralis (lagartixa-dos-muros)
- #XP_053223437.1 Podarcis raffonei (lagarto de parede elia)

- #XP_028838360.1 *Denticeps clupeioides* (peixe)
- #CAG07273.1 *Tetraodon nigroviridis* (baiacu verde pintado)
- #NXT65740.1 *Chaetops frenatus*- salta rochas do cabo (ave)
- #NWI42167.1 *Picathartes gymnocephalus* - picartatstes da guine (ave)

Após as amostras retiradas, foram selecionadas 128 sequências e um alinhamento mais qualificado para análise. Observou-se que 69 posições da isoforma 329 de *Rattus norvegicus* são conservadas até peixes.

3.9 Qual o grau de conservação entre os roedores?

Com base em alinhamentos feitos no Clustal ômega e no MEGAX obtivemos um alinhamento limpo e bem conservado, mais especificamente o grau de profundidade em roedores mostra que 189 posições há conservação total dos aminoácidos. Demonstrando que o grupo de roedores possui uma ótima conservação (Figura 21).

FIGURA 21 - Alinhamento múltiplo, usando-se a isoforma 329 do gene *RAD51D*

```

XP_055476374.1      MGKLRAGLCPGLSEEVRLRLDRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_051014088.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
CAH6785769.1        MGVLRVGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKGLVALRRVLLAQ      60
XP_051035720.1      MGVLRVGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKGLVALRRVLLAQ      60
EGV94391.1          MGVLRVGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_052615754.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_059125560.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_006990137.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_028722829.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAANLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_052051345.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_031209110.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_021068099.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
EDL15681.1          MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_021033215.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
NP_035365.1          MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
AAG42287.1          MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
NP_001100499.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_032769739.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_028615243.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCSLSYKALVALRRVLLAQ      60
XP_034363899.1      MGMLRAGLCPGLTEETVQLLRGRKIKTVADLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
XP_004594053.3      MGVLRVGLCPGLSQEMIELLRSCGIRTVVDLASADLEEVAQKSGLSYKALVALRRVLLAQ      60
XP_040857277.1      MGVLRVGLCPGLSQEMIELLRSCGIRTVVDLASADLEEVAQKSGLSYKALVALRRVLLAQ      60
XP_012868628.1      MGVLRAGLCPGLTEEMVQLLRSGVKTVMDLAAADLEEVAQKCSLSYKALVALRRVLLAQ      60
XP_042548222.1      MGVLRAGLCPGLTEEMVQLLRSGVKTVMDLAAADLEEVAQKCSLSYKALVALRRVLLAQ      60
XP_026267675.1      MGVLRAGLCPGLTQEMVQLLRSGIKTVVDLAAADLEEVAQKCSLSYKALVALRRVLLAQ      60
XP_047400674.1      MGVLRAGLCPGLTQEMVQLLRSGIKTVVDLAAADLEEVAQKCSLSYKALVALRRVLLAQ      60
KF025943.1          MGVLRAGLCPGLTEEMVQLLRSGAIRTVDLAAADLEEVAQRCSLSYKALVALRRVLLAQ      60
XP_005402921.1      MGVLRAGLCPGLTEEMVQLLRGGAIRTVVDLAAADLEEVAQKCGLSYKALVALRRVLLAQ      60
** ** *****:.* :.*:.. :.* ** **:*:*****:.. ***** *****

XP_055476374.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDNLDDAGLYTGEVTEFVGGP6SGKTQVCLCV      120
XP_051014088.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIAGCP6SGKTQVCLCV      120
CAH6785769.1        FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDPGLYTGVEITEVGGP6SGKTQVCLCV      120
XP_051035720.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDPGLYTGVEITEVGGP6SGKTQVCLCV      120
EGV94391.1          FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_052615754.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDRLDDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_059125560.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDRLDDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_006990137.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_028722829.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_052051345.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_031209110.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIAGGP6SGKTQVCLCV      120
XP_021068099.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
EDL15681.1          FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_021033215.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
NP_035365.1          FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
AAG42287.1          FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
NP_001100499.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_032769739.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_028615243.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_034363899.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_004594053.3      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGAP6SGKTQVCLCV      120
XP_040857277.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGAP6SGKTQVCLCV      120
XP_012868628.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_042548222.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_026267675.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_047400674.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
KF025943.1          FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
XP_005402921.1      FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGP6SGKTQVCLCV      120
**:* *****:***** ** ** *****:*** * *****:****

```

Parte inicial do alinhamento múltiplo em roedores. Foi usada na busca a sequência de 329 aminoácidos de *R. norvegicus* XP_055476374.1. Essa sequência não tem ortólogo em humanos, sendo mais parecida com a isoforma O75771-1 (canônica de humanos), que tem 328 aminoácidos.

FONTE: Captura de tela do alinhamento no Clustal ômega, adaptado pela autora.

3.10 Quais as principais posições suscetíveis a trocas de aminoácidos na RAD51D?

No alinhamento de mamíferos com a isoforma canônica humana, foram identificadas substituições em cerca de 30 posições, sendo que em cada uma delas ocorrem quatro ou mais aminoácidos distintos (posições destacadas em negrito na Tabela 4). Com base nos dados do AlphaMissense, elaborou-se uma tabela detalhando as principais substituições de aminoácidos por posição e as espécies em que essas alterações ocorrem. Além disso, utilizando o ConSurf, foi realizada a predição da porcentagem de substituições observadas para cada posição, incluindo a frequência de cada substituição específica. Ao analisar as cinco posições mais divergentes no alinhamento múltiplo da isoforma 328, observou-se que algumas substituições registradas não eram previstas pelo ConSurf ou apareciam com baixa frequência. Por exemplo, para a posição Lis42, verificou-se substituição por isoleucina (não prevista), glutamina (2% de frequência observada), arginina (12%) e treonina (2%) (Tabela 4).

Tabela 4 - Conservação e profundidade em mamíferos por aminoácido

Posição	Substituições observadas	Profundidade da conservação evolutiva
M1	-	
G2	-	
V3	L, M	não conservado em muitas espécies
L4	-	
R5	-	
V6	A, D, P, M	A maioria da espécies têm A
G7	A	não conservado em elefantes
L8	-	

C9	-	
P10	-	
G11	-	
L12	-	
T13	S	não conservado em muitas espécies
E14	Q,H,R,L,P	não conservado em muitas espécies
E15	D,N	não conservado em muitas espécies
M16	I,T,V	não conservado em muitas espécies
I17	V,A	não conservado em muitas espécies
Q18	H,L,G	não conservado em muitas espécies
L19	Q	não conservado em muitas espécies
L20	-	
R21	G,K	não conservado em roedor, primata e felino
S22	G,N	não conservado em muitas espécies
H23	R,A,C,M,I	não conservado em muitas espécies
R24	G	não conservado em muitas espécies
I25	V	<i>Canis lupus familiaris</i> (cachorro)
K26	R,S	não conservado em muitas espécies
T27	-	
V28	-	
V29	A	não conservado em hamster e morcego
D30	A	<i>Pipistrellus Kuhlii</i> (morcego-de-kuhl)

L31	-	
V32	A	não conservado em roedores e caninos
S33	A	não conservado em roedores
A34	S	<i>Phascolarctos cinereus</i> (coala)
D35	N	não conservado em roedores
L36	-	
E37	-	
E38	A,D	não conservado em muitas espécies
V39	A,I,L	não conservado em muitas espécies
A40	S	<i>Phascolarctos cinereus</i> (coala) <i>Monodelphis domestica</i> (catita-roedor) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes)
Q41	L,R	não conservado em muitas espécies
K42	I,Q,R,T,E,N	não conservado em muitas espécies
C43	G,R	<i>Carlito syrigha</i>
G44	A,C,S	<i>Phascolarctos cinereus</i> (coala)
L45	-	
S46	-	
Y47	-	
K48	-	
A49	V	<i>Desmodus rotundus</i> (morcego vampiro) <i>Globicephala melas</i> (baleia)
L50	-	

V51	-	
A52	S	<i>Phascolarctos cinereus</i> (coala)
L53	V,Q	não conservado em muitas espécies
R54	K	<i>Cavia porcellus</i> (porquinho da índia)
R55	-	
V56	-	
L57	-	
L58	-	
A59	V	<i>Phascolarctos cinereus</i> (coala) <i>Monodelphis domestica</i> (catita-roedor)
Q60	R	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)
F61	L	não conservado em muitas espécies
S62	A	<i>Cavia porcellus</i> (porquinho da índia) <i>Mesoplodon densirotrisaquatico</i> (baleia-bicuda-de-Blainville)
A63	S	<i>Phascolarctos cinereus</i> (coala) <i>Monodelphis domestica</i> (catita-roedor) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes)
F64	-	
P65	-	
V66	L,F,C,S,I	não conservado em muitas espécies
N67	S	<i>Mesoplodon densirotrisaquatico</i> (baleia-bicuda-de-Blainville)
G68	-	
A69	S	<i>Equus quagga</i> (zebra)

D70	-	
L71	V	<i>Loxodonta africana</i> (elefante-da-savana), <i>Elephas maximus indicus</i> (elefante-indiano) <i>Trichechus manatus latirostris</i> (peixe-boi-marinho)
Y72	-	
E73	-	
E74	D	<i>Galeopterus variegatus</i> (colugo)
L75	-	
K76	-	
T77	S	<i>Orycteropus afer afer</i> (porco formigueiro)
S78	T	<i>Elephantulus edwardii</i> (musaranho-elefante)
T79	-	
A80	-	
I81	-	
L82	-	
S83	A	<i>Tupaia chinensis</i> (musaranho chinês) <i>Phascolarctos cinereus</i> (coala) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes)
T84	-	
G85	-	
I86	L,M,V	não conservado em muitas espécies
G87	A,R,Q,E	não conservado em muitas espécies
S88	R	<i>Loxodonta africana</i> (elefante-da-savana), <i>Elephas maximus indicus</i> (<i>Elephas maximus indicus</i>)

		<i>Trichechus manatus latirostris</i> (peixe boi da Flórida) <i>Monodelphis domestica</i> (catita-roedor)
L89	-	
D90	-	
K91	R	<i>Elephantulus edwardii</i> (musaranho-elefante)
L92	-	
L93	-	
D94	-	
A95	S,T	não conservado em muitas espécies
G96	-	
L97	V	<i>Choloepus didactylus</i> (preguiça real)
Y98	-	
T99	-	
G100	-	
E101	-	
V102	M	<i>Jaculus jaculus</i> (jerboa-pequeno)
T103	-	
E104	-	
I105	F,L	não conservado em algumas espécies
V106	A,T,S,I	não conservado em algumas espécies
G107	-	
G108	A,S,V	não conservado em algumas espécies

P109	-	
G110	A,R	<i>Tupaia chinensis</i> <i>Phascolarctos cinereus</i> (coala) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes) <i>Monodelphis domestica</i> (catita-roedor)
S111	-	
G112	-	
K113	-	
T114	-	
Q115	-	
V116	I,L	não conservado em algumas espécies
C117	-	
L118	-	
C119	F,Y,R,H	não conservado em muitas espécies
M120	T,V	não conservado em muitas espécies
A121	S	Não conservado em camelos
A122	I	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)
N123	H	Não conservado em camelos e <i>Cavia porcellus</i> (porquinho da índia)
V124	M	<i>Carlito syichta</i>
A125	-	
H126	Y	não conservado em muitas espécies
G127	D,S	não conservado em muitas espécies

L128	-	
Q129	H,R,G	não conservado em algumas espécies
Q130	-	
N131	-	
V132	I	não conservado em algumas espécies
L133	V	não conservado em algumas espécies
Y 134	-	
V135	I	não conservado em muitas espécies
D136	-	
S137	-	
N138	S,T,K	não conservado em muitas espécies
G139	-	
G140	-	
L141	M,I,F	não conservado em elefantes, <i>Cricetulus griseus</i> (<i>Hamster chinês</i>) <i>Jaculus jaculus</i> (jerboa-pequeno)
T142	S,M	<i>Tupaia chinensis</i> (musaranho chinês) , <i>Phascolarctos cinereus</i> (coala) <i>Orycteropus afer afer</i> (porco formigueiro)
A143	G	<i>Ceratotherium simum simum</i> (rinoceronte branco do sul) <i>Equus caballus</i> (cavalo) <i>Equus guagga</i> (zebra)
S144	-	

R145	-	
L146	I,T	não conservado em muitas espécies
L147	-	
Q148	-	
L149	-	
L150	F	<i>Choloepus didactylus</i> (preguiça real)
A151	E	<i>Otelemur garnetti</i> (galago de orelhas pequenas) <i>Nycticebus coucang</i> (lóris lento)
K152	T,C	não conservado em algumas espécies
T153	R	não conservado em muitas espécies
T154		
Q155	E,P,R	não conservado em muitas espécies
D156	A,E	não conservado em muitas espécies
E157		
E158	G	<i>Desmodus rotundus</i> (Desmodus rotundus)
E159	K,V	não conservado em algumas espécies
Q160		
A161		
E162	G,R	não conservado em muitas espécies
A163	-	
L164	-	
R165	R	<i>Loxodonta africana</i> (elefante-da-savana),

		<i>Elephas maximus</i> (elefante-asiático) <i>Cavia porcellus</i> (porquinho da índia)
R166	K,S	<i>Gulo gulo luscus</i> (glutão) <i>Monodelphis domestica</i> (catita-roedor) <i>Meles meles</i> (texugo-europeu) <i>Mustela Lutreola</i> (doninha-europeia)
I167	-	
Q168	H	não conservado em algumas espécies
V169		
V170	A,T	não conservado em algumas espécies
H171	R,Y	não conservado em muitas espécies
A172	V	não conservado em algumas espécies
F173	Y	<i>Equus quagga</i> (zebra)
D174	-	
I175	V	não conservado em algumas espécies
F176	-	
Q177	E,H,R	não conservado em algumas espécies
M178	-	
L179	-	
D180	A,G,N,Q,S	não conservado em algumas espécies
V181	A,M,R,T	não conservado em algumas espécies
L182	I	não conservado em camelos
Q183	H	<i>Cavia porcellus</i> (porquinho da índia)

E184	D	não conservado em muitas espécies
L185	F,C	não conservado em algumas espécies
R 186	-	
G187	S	não conservado em algumas espécies
T188	S,A	não conservado em muitas espécies
V189	L,M,I	não conservado em muitas espécies
A190	S,T,V	não conservado em muitas espécies
Q191	-	
Q192	R	<i>Equus quagga</i> (zebra)
V193	S	<i>Crisetulus griseus</i> (Hamster chinês)
T194	I,N,S	não conservado em muitas espécies
G195	D,S,L,N	não conservado em muitas espécies
S196	A,G	não conservado em muitas espécies
S197	L	<i>Tupaia Chinesis</i> (musaranho-chinês)
G198	A,V,M	não conservado em muitas espécies
T199	A,I,P	não conservado em muitas espécie
V200	A,L,M,T	não conservado em muitas espécies
K201	R	Pongo abelii (Orangotango-de-sumatra) Carlito syrichta
V202	A,M,I	não conservado em algumas espécies
V203	L,I	<i>Meles Meles</i> (texugo-europeu), <i>Tupaia chinesis</i> (musaranho-chinês)

V204	I,M	não conservado em algumas espécies
V205	-	
D206	-	
S207	-	
V208	-	
T209	A,S	<i>Monodelphis domestica</i> (catita-roedor) <i>Phascolarctos cinereus</i> (coala) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes) <i>Cholopus didactylus</i> (preguiça real)
A210	-	
V211	-	
V212	I	<i>Cricetulus griseus</i> (hamster chinês) <i>Molossus molossus</i> (morcego de cauda grossa) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes)
S213	A	não conservado em algumas espécies
P214	-	
L215	I	<i>Tupaia chinensis</i> (musaranho chinês)
L216	-	
G217	-	
G218	-	
Q219	-	
Q220	Y, H	<i>Monodelphis domestica</i> (catita-roedor) <i>Phascolarctos cinereus</i> (coala) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes)

R221	S,T	não conservado em algumas espécies
E222	-	
G223	V	<i>Orycteropus afer afer</i> (porco formigueiro)
L224	F	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)
A225	-	
L226	-	
M227	-	
M228	I	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)
Q229	H,R	<i>Desmodus rotundus</i> (morcego vampiro) e <i>choloepus didactylus</i> (preguiça real)
L230	M	não conservado em felinos (gatos não domesticos)
A231	G,S	<i>Mesoplodon densirostris aquatico</i> (baleia) <i>Puma yagouaroundi</i> (jaguarundi) <i>Puma concolor</i> (onça-parda)
R232	L,Q	não conservado em algumas espécies
E233	-	
L234	-	
K235	-	
T236	I	<i>Cricetulus griseus</i> (hamster chinês)
L237	V	<i>Elephantulus edwardii</i> (musaranho-elefante)
A238	-	
R239	K, W, Q	não conservado em algumas espécies
D240	E	não conservado em algumas espécies

L241	I,V	não conservado em algumas espécies
G242	A,S	não conservado em algumas espécies
M243	I, L,T, V	não conservado em muitas espécies
A244	-	
V245	L	não conservado em algumas espécies
V246	L,M	não conservado em muitas espécies
V247	-	
T248	-	
N249	-	
H250	-	
I251	L,M,T,V	não conservado em muitas espécies
T252	I	não conservado em algumas espécies
R253	C	<i>Galeopterus variegatus</i> (colugo)
D254	V,E	não conservado em algumas espécies
R255	W	<i>Pipistrellus kuhlii</i> (morcego-de-kuhl)
D256	E	<i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes)
S257	G,N,T	não conservado em algumas espécies
G258	R	<i>Pongo abelii</i> (Orangotango-de-sumatra) <i>Cricetulus griseus</i> (hamster chinês)
R259	K,E,D,V,Q,N	não conservado em muitas espécies
L260	F	não conservado em algumas espécies
K261	R	não conservado em algumas espécies

P262	-	
A263	-	
L264	-	
G265	-	
R266	C,H,L	<i>Callithrix jacchus</i> (sagui-de-tufo-branco) , <i>Monodon monoceros</i> (narval) <i>Molossus molossus</i> (morcego de cauda grossa)
S267	-	
W268	-	
S269	-	
F270	-	
V271	-	
P272	-	
S273	N	<i>Elephantulus edwardii</i> (musaranho-elefante)
T274	-	
R275	-	
I276	L,V,F	não conservado em muitas espécies
L277	F,H,V	não conservado em algumas espécies
L278	-	
D279	A,E,G,N,S	não conservado em muitas espécies
T280	A,I,S,L,V	não conservado em muitas espécies
I281	S,T,V,,D,G,E,R,C	não conservado em muitas espécies

E282	D,G,K,Q,N	não conservado em muitas espécies
G283	E	<i>Carlito syrigha</i> <i>Pteropus alecto</i> (raposa voadora) <i>Pteropus giganteus</i> (morcego gigante)
A284	E,P,S,T,I	não conservado em algumas espécies
G285	E,R,Q	não conservado em algumas espécies
A286	T,V,M,L	não conservado em algumas espécies
S287	A,L,P	não conservado em algumas espécies
G288	A,D,S,H	não conservado em algumas espécies
G289	C,R,S	não conservado em muitas espécies
GAP	G,S	GAP
R290	-	
R291	C,H,Q,W,E	não conservado em muitas espécies
M292	-	
A293	T,V,R,I,S	não conservado em muitas espécies
C294	M,V	não conservado em muitas espécies
L295	Y,R,S	não conservado em algumas espécies
A296	-	
K297	T,M,I	não conservado em muitas espécies
S298	-	
S299	-	
R300	P	não conservado em muitas espécies

Q301	-	
P302	L	não conservado em muitas espécies
T303	A	<i>Orcinus orca</i> (orca) e <i>Elephantulus edwardii</i> (musaranho-elefante)
G304	-	
F305	-	
Q306	I,L,V,S	não conservado em muitas espécies
E307	R	<i>Carlito syrichta</i>
M308	K	<i>Callithix jacchus</i> (sagui-de-tufo-branco)
V309	T,V,P	não conservado em algumas espécies
D310	I	não conservado em algumas espécies
I311	E	<i>Monodelphis domestica</i> (catita-roedor) <i>Trichosurus vulpecula</i> (Cusu-de-orelhas-grandes) <i>Phascolarctos cinereus</i> (coala)
G312	T,L	não conservado em algumas espécies
T313	R,E	<i>Cavia porcellus</i> (porquinho da índia) <i>Molossus molossus</i> (morcego de cauda grossa)
W314	I,S	não conservado em algumas espécies
G315	L,C	não conservado em algumas espécies
T316	A	não conservado em algumas espécies
S317	P,I	não conservado em algumas espécies
E318	A,P,T,E	não conservado em algumas espécies
Q319	K,V,G,Q	não conservado em algumas espécies

S320	P,E,H,L,K	não conservado em algumas espécies
A321	G,R,N	não conservado em algumas espécies
T322	P,S,A,Q,M,E,R	não conservado em muitas espécies
L323	I,S,V	não conservado em muitas espécies
Q324	R,M,S	não conservado em muitas espécies
G325	D,E,P,K,H	não conservado em algumas espécies
D326	A,E,V,,R	não conservado em muitas espécies
Q327	E,H,K,N,Y,S,G	não conservado em muitas espécies
T328	K,P,H	não conservado em muitas espécies

NOTA: Finalizado dia 13/12/2024. Tabela referente a classe de mamíferos. Linhas em negrito representam as posições com possíveis trocas de aminoácidos com quatro opções ou mais)

Fonte: Feito pela autora

Tabela 5-Conservação e profundidade em aves, répteis, anfíbios e peixes por aminoácido

Posição	Substituições observadas	Profundidade da conservação evolutiva
M1	-	
G2	M, V	A maioria das espécies tem V
V3	F, I	<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
L4	-	
R5	-	
V6	E, A	A maioria das espécies tem A

G7	H	<i>Periophthalmus magnuspinnatus</i> (peixe)
L8	M	<i>Periophthalmus magnuspinnatus</i> (peixe)
C9	-	
P10	-	
G11	-	
L12	-	
T13	E	<i>Periophthalmus magnuspinnatus</i> (peixe)
E14	A, L, S	<i>Rhinatrema bivittatum</i> (cobra) <i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
E15	D, Q	A maioria das espécies tem D
M16	L, V	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
I17	V, M	não conservado em algumas espécies
Q18	C, R, A	não conservado em algumas espécies
L19	A, D	<i>Rhinatrema bivittatum</i> (cobra) <i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
L20	-	
R21	K	<i>Rhinatrema bivittatum</i> (cobra) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
S22	N, A, D, E, T	A maioria das espécies tem A

H23	N, S, D, A, Q	A maioria das espécies tem N
R24	S, G, H, D	A maioria das espécies tem G
I25	V	<i>Pogona vitticeps</i> (dragão barbudo)
K26	E, R, S	A maioria das espécies tem R
T27	-	
V28	-	
V29	E, L	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
D30	E	<i>Periophthalmus magnuspinnatus</i> (peixe)
L31	F	A maioria das espécies tem F
V32	-	
S33	A	<i>Protopterus annectens</i> (peixe pulmonado africano)
A34	S	A maioria das espécies tem S
D35	-	
L36	I, V	<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
E37	-	
E38	D, H, A	A maioria das espécies tem D
V39	I, L	A maioria das espécies tem L
A40	-	
Q41	R	<i>Rhinatrema bivittatum</i> (cobra)

K42	Q	<i>Protopterus annectens</i> (peixe pulmonado africano)
C43	S	<i>Rhinatrema bivittatum</i> (cobra) <i>Protopterus annectens</i> (peixe pulmonado africano)
G44	C, S	A maioria das espécies tem S
L45	V, I	não conservado em algumas espécies
S46	-	
Y47	-	
K48	-	
A49	M, T, G	<i>Rhinatrema bivittatum</i> (cobra) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
L50	-	
V51	L	<i>Periophthalmus magnuspinnatus</i> (peixe)
A52	-	
L53	V, I	A maioria das espécies tem V
R54	-	
R55	-	
V56	-	
L57	-	
L58	-	
A59	-	
Q60	E	<i>Rhinatrema bivittatum</i> (cobra)

F61	Y, H	<i>Rhinatrema bivittatum</i> (cobra)e <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Periophthalmus magnuspinnatus</i> (peixe)
S62	T	<i>Periophthalmus magnuspinnatus</i>
A63	S	<i>Pogona vitticeps</i>
F64	-	
P65	-	
V66	S, T, A	A maioria das espécies tem A
N67	S	<i>Periophthalmus magnuspinnatus</i> (peixe)
G68	-	
A69	-	
D70	E	<i>Rhinatrema bivittatum</i> (cobra)
L71	-	
Y72	-	
E73	-	
E74	D	<i>Periophthalmus magnuspinnatus</i> (peixe)
L75	-	
K76	L, M	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
T77	G, N, S	A maioria das espécies tem S
S78	-	
T79	V	<i>Pogona vitticeps</i> (dragão barbudo)

A80	-	
I81	-	
L82	-	
S83	P	A maioria das espécies tem P
T84	-	
G85	-	
I86	N, S, I, T, V	A maioria das espécies tem S
G87	R, P, K, S	A maioria das espécies tem P
S88	G, R, N	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Locustella ochotensis</i> (passarinho)
L89	-	
D90	-	
K91	T, Q, E	A maioria das espécies tem Q
L92	-	
L93	-	
D94	-	
A95	S	A maioria das espécies tem S
G96	-	
L97	-	
Y98	-	
T99	-	

G100	-	
E101	-	
V102	L	não conservado em algumas espécies
T103	-	
E104	-	
I105	F, L	A maioria das espécies tem L
V106	A, M, S, T	A maioria das espécies tem M
G107	-	
G108	A	A maioria das espécies tem A
P109	A	<i>Rhinatrema bivittatum</i> (cobra)
G110	-	
S111	T	<i>Periophthalmus magnuspinnatus</i> (peixe)
G112	-	
K113	-	
T114	-	
Q115	-	
V116	L	<i>Protopterus annectens</i> (peixe pulmonado africano)
C117	-	
L118	F	<i>Periophthalmus magnuspinnatus</i> (peixe)
C119	S, G, N	A maioria das espécies tem G
M120	I, V	A maioria das espécies tem I

A121	-	
A122	V	<i>Rhinatrema bivittatum</i> (cobra) <i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
N123	S, H	A maioria das espécies tem S
V124	I	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Pogona vitticeps</i> (dragão barbudo)
A125	S, T	A maioria das espécies tem S
H126	L, Y	A maioria das espécies tem L
G127	H, E, S	<i>Periophthalmus magnuspinnatus</i> <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
L128	-	
Q129	K, R	A maioria das espécies tem K
Q130	-	
N131	H, Y, G, D	A maioria das espécies tem H
V132	-	
L133	V, I	<i>Rhipidura dahli</i> (fantail de Bismarck) <i>Periophthalmus magnuspinnatus</i> (peixe)
Y 134	F	A maioria das espécies tem F
V135	I, L	A maioria das espécies tem L
D136	-	
S137	T	<i>Periophthalmus magnuspinnatus</i> (peixe)

N138	T, S	A maioria das espécies tem T
G139	A	<i>Rhipidura dahli</i> (fantail de Bismarck) <i>Daphoenositta chioptera</i> (trepadeirainha-australiana)
G140	S	<i>Platysteira castanea</i> (barbela-de-garganta-castanha)
L141	F, I	A maioria das espécies tem F
T142	M	<i>Rhinatrema bivittatum</i> (cobra)
A143	-	
S144	T	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Pogona vitticeps</i> (dragão barbudo)
R145	-	
L146	-	
L147	F, H, Y	A maioria das espécies tem Y
Q148	E	<i>Hippolais icterina</i> (felosa-icterina)
L149	M, I	A maioria das espécies tem M
L150	V	<i>Rhinatrema bivittatum</i> (cobra) <i>Pogona vitticeps</i> (dragão barbudo)
A151	K, R, G	A maioria tem R
K152	S, T, G	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) e <i>Rhinatrema bivittatum</i> (cobra)
T153	R, Q, E	A maioria tem Q
T154	S	<i>Periophthalmus magnuspinnatus</i> (peixe)
Q155	E, S, D	A maioria tem E
D156	E, S, C, N	<i>Periophthalmus magnuspinnatus</i> (peixe)

		<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) e <i>Rhinatrema bivittatum</i> (cobra)
E157	L, I	<i>Periophthalmus magnuspinnatus</i> (peixe) e <i>Protopterus annectens</i> (peixe pulmonado africano)
E158	-	
E159	D	<i>Protopterus annectens</i> (peixe pulmonado africano)
Q160	-	
A161	V, I, M	<i>Rhinatrema bivittatum</i> (cobra) <i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
E162	R, N, D	A maioria tem R
GAP		16 posições
A163	V	<i>Pogona vitticeps</i> (dragão barbudo)
L164	M	<i>Pogona vitticeps</i> (dragão barbudo)
R165	K, Q	A maioria tem Q
R166	-	
I167	V	<i>Rhagologus leucotigma</i> (passarinho) <i>Platysteira castanea</i> (barbela-de-garganta-castanha)
Q168	E, H	<i>Rhinatrema bivittatum</i> (cobra) <i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
V169	L, I	<i>Regulus satrapa</i> (estrelinha-de-coroa-dourada)

		<i>Hirundo rustica</i> (andorinha-das-chaminés) <i>Protopterus annectens</i> (peixe pulmonado africano)
V170	F	<i>Periophthalmus magnuspinnatus</i> (peixe)
H171	R, K, Q	A maioria tem R
A172	L, V	A maioria tem V
F173	-	
D174	N, S	Mudança apenas em aves
I175	V, A	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
F176	Y	A maioria tem Y
Q177	E, K, S	A maioria tem E
M178	L	A maioria tem L
L179	M	<i>Protopterus annectens</i> (peixe pulmonado africano)
D180	R, S, G	Mudança apenas em aves
V181	A, C	A maioria tem A
L182	F	<i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
Q183	C	<i>Periophthalmus magnuspinnatus</i> (peixe)
E184	F	<i>Pogona vitticeps</i> (dragão barbudo)
L185	W	<i>Hirundo rustica</i>
R 186	D, C, R, S	A maioria tem D

G187	R, S, N	A maioria tem R
T188	L, F, G	A maioria tem L
V189	L, S, P	A maioria tem S
A190	N	<i>Protopterus annectens</i>
Q191	P, E	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Rhagologus leucotigma</i> (passarinho)
Q192	I	<i>Pogona vitticeps</i> (dragão barbudo)
V193	L, V, A, M	A maioria é V
T194	S, V, N	A maioria é S
G195	P, A, G	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Regulus satrapa</i> (estrelinha-de-coroa-dourada)
S196	V, T, M, A, G, C	Variado
S197	A, T	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
G198	P, C	Maioria é P
T199	L, A	Maioria é L
V200	R	<i>Periophthalmus magnuspinnatus</i> (peixe)
K201	M, A	A maioria é M
V202	L	<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)

		<i>Rhinatrema bivittatum</i> (cobra)
V203	I, L	A maioria é L
V204	I, L	A maioria é L
V205	L	A maioria é L
D206	-	
S207	-	
V208	A, I	<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
T209	S, C, P	A maioria é S
A210	-	
V211	-	
V212	I	A maioria é I
S213	Y, A	A maioria é Y
P214	-	
L215	V	<i>Platysteira castanea</i> (barbela-de-garganta-castanha) <i>Protopterus annectens</i> (peixe pulmonado africano)
L216	-	
G217	-	
G218	T, S	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
Q219	R, K	A maioria tem R

Q220	H	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
R221	S, T, A, P	A maioria tem S
E222	D	<i>Rhinatrema bivittatum</i> (cobra)
G223	-	
L224	N	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
A225	S	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Rhinatrema bivittatum</i> (cobra)
L226	-	
M227	-	
M228	T	<i>Protopterus annectens</i> (peixe pulmonado africano)
Q229	H	<i>Pogona vitticeps</i> (dragão barbudo) <i>Sinosuthora webbiana</i> (passarinho)
L230	V, I	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
A231	S	Mudança apenas em aves
R232	A, L	<i>Periophthalmus magnuspinnatus</i> (peixe) e <i>Pogona vitticeps</i>
E233	L, D	<i>Periophthalmus magnuspinnatus</i> (peixe) e <i>Pogona vitticeps</i>
L234	M	<i>Hirundo rustica</i> (andorinha-das-chaminés)

K235	-	
T236	A, S, M	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
L237	M	<i>Rhinatrema bivittatum</i> (cobra)
A238	V	<i>Periophthalmus magnuspinnatus</i> (peixe)
R239	K	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
D240	E	A maioria E
L241	F	A maioria é F
G242	S, N	A maioria é S
M243	L, I, V	A maioria é L
A244	V	<i>Rhinatrema bivittatum</i> (cobra)
V245	A	<i>Periophthalmus magnuspinnatus</i> (peixe)
V246	L	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
V247	L	<i>Pogona vitticeps</i> (dragão barbudo)
T248	-	
N249	-	
H250	Q, Y	A maioria é Q
I251	V	A maioria é V
T252	-	

R253	H	<i>Protopterus annectens</i> (peixe pulmonado africano)
D254	N	<i>Rhinatrema bivittatum</i> (cobra)
R255	S, A, G	A maioria é S
D256	S, N, V	A maioria é S
S257	T, G, D	A maioria é T
G258	S	<i>Rhinatrema bivittatum</i> (cobra)
R 259	P, S, A, L	A maioria é P
L260	V, M	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
K261	R	<i>Rhinatrema bivittatum</i> (cobra) <i>Periophthalmus magnuspinnatus</i> (peixe)
P262	-	
A263	-	
L264	-	
G265	-	
R266	V	<i>Periophthalmus magnuspinnatus</i> (peixe)
S267	A	<i>Hippolais icterina</i> (felosa-icterina)
W268	-	
S269	-	
F270	H, Y	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)

V271	L	<i>Periophthalmus magnuspinnatus</i> (peixe)
P272	-	
S273	R, N	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
T274	V	<i>Protopterus annectens</i> (peixe pulmonado africano)
R275	-	
I276	L, V	A maioria é V
L277	-	
L278	-	
D279	E, Q	A maioria é Q
GAP NA CANONICA	S, G, H, Q	A maioria é S
T280	K, R, P, S	A maioria é K e R
I281	E, D, K	A maioria é E
E282	A, R, G, D	A maioria é A
G283	S, T, A, C	A maioria é T
A284	W, Q, T, R	A maioria é W
G285	E, D, H, S	A maioria é E
A286	K, N, S, V	A maioria é K
S287	A, G, F, R, T	A maioria é A e G
G288	T, S	A maioria é T

G289	T, S, N, A	A maioria é T
R290	Q, L, K, F	A maioria é Q
R291	-	
M292	A, V, L	A maioria é A e V
A293	-	
C294	S, T, Y	A maioria é S
L295	-	
A296	S, V, L, M	VARIADO
K297	-	
S298	-	
S299	P	A maioria é P
R300	-	
Q301	L	<i>Regulus satrapa</i> (estrelinha-de-coroa-dourada)
P302	-	
T303	A, C, I	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
G304	H, R	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)
F305	I, L, M	A maioria é I
Q306	C, R, S	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano)

		<i>Rhinatrema bivittatum</i> (cobra)
E307	V, K	A maioria é V
M308	E, A, D, V	A maioria E
V309	L, F	A maioria é L
D310	E, N	<i>Rhinatrema bivittatum</i> (cobra) <i>Pogona vitticeps</i> (dragão barbudo)
I311	L	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Hippolais icterina</i> (felosa-icterina)
G312	E, S	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
T313	S, W, D, M) INÍCIO GAP AVES)	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
W314	S, C	<i>Rhinatrema bivittatum</i> (cobra) <i>Protopterus annectens</i> (peixe pulmonado africano)
G315	S, D	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Pogona vitticeps</i> (dragão barbudo)
T316	V, R, F, Q	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
S317	L, M, T	<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)

		<i>Rhinatrema bivittatum</i> (cobra)
E318	K	<i>Periophthalmus magnuspinnatus</i> (peixe)
Q319	E, D	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
S320	N, V, H, T	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
A321	P, L	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
T322	V, S, F	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
L323	A, D, I	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo)
Q324	S, K, T	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
G325	K, R, D, C	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
D326	E, K, A	<i>Periophthalmus magnuspinnatus</i> (peixe)

		<i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
Q327	G, H	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annecten</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)
T328	L, R, S, P	<i>Periophthalmus magnuspinnatus</i> (peixe) <i>Protopterus annectens</i> (peixe pulmonado africano) <i>Pogona vitticeps</i> (dragão barbudo) <i>Rhinatrema bivittatum</i> (cobra)

Tabela finalizada dia 04/05/2025. Tabela referente a outras classes (aves, répteis, anfíbios e peixes) sem ser mamíferos

FONTE: criado pela autora

Tabela 6-Conservação e profundidade em primatas por aminoácido

Posição	Substituições	Substituições observadas em:	Profundidade evolutiva no alinhamento
M1	-		conservado em primatas
G2	-		conservado em primatas
V3	-		conservado em primatas
L4	-		conservado em primatas
R5	-		conservado em primatas
V6	-		
G7	-		conservado em primatas
L8	-		conservado em primatas
C9	-		conservado em primatas
P10	-		conservado em primatas
G11	-		conservado em primatas
L12	-		conservado em primatas
T13	-		conservado em primatas

E14	Q	A maioria é Q	
E15	-		
M16	V		
I17	V	<i>Nycticebus coucang</i> (lóris lento) <i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnetii</i> (galago-grande-do-norte)	
Q18	L, H	<i>Otelemur garnetii</i> (galago-grande-do-norte)	
L19	-		
L20	-		conservado em primatas
R21	G	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
S22	N	<i>Callithrix jacchus</i> (<i>sagui-de-tufo-branco</i>)	
H23	R	Na maioria é R	
R24	G	Na maioria é G	
I25	-		conservado em primatas
K26	-		
T27	-		conservado em primatas
V28	-		conservado em primatas
V29	-		conservado em primatas
D30	-		conservado em primatas
L31	-		conservado em primatas
V32	-		conservado em primatas
S33	-		conservado em primatas
A34	-		conservado em primatas

D35	N	<i>Otelemur garnetii</i> (galago-grande-do-norte)	
L36	-		conservado em primatas
E37	-		conservado em primatas
E38	-		
V39	L	<i>Pongo abelii</i> (Orangotango-de-sumatra)	
A40	-		conservado em primatas
Q41	-		
K42	-		conservado em primatas
C43	-		conservado em primatas
G44	-		conservado em primatas
L45	-		conservado em primatas
S46	-		conservado em primatas
Y47	-		conservado em primatas
K48	-		conservado em primatas
A49	-		conservado em primatas
L50	-		conservado em primatas
V51	-		conservado em primatas
A52	-		conservado em primatas
L53	-		conservado em primatas
R54	-		conservado em primatas
R55	-		conservado em primatas
V56	-		conservado em primatas
L57	-		conservado em primatas
L58	-		conservado em primatas
A59	-		conservado em primatas
Q60	R	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
F61	-		conservado em primatas

S62	-		conservado em primatas
A63	-		conservado em primatas
F64	-		conservado em primatas
P65	-		conservado em primatas
V66	I, L, F	Maioria F	
N67	-		<i>conservado em primatas</i>
G68	-		conservado em primatas
A69	-		<i>conservado em primatas</i>
D70	-		conservado em primatas
L71	-		conservado em primatas
Y72	-		conservado em primatas
E73	-		conservado em primatas
E74	-		conservado em primatas
L75	-		conservado em primatas
K76	-		conservado em primatas
T77	-		conservado em primatas
S78	-		conservado em primatas
T79	-		conservado em primatas
A80	-		conservado em primatas
I81	-		conservado em primatas
L82	-		conservado em primatas
S83	-		conservado em primatas
T84	-		conservado em primatas
G85	-		conservado em primatas
I86	-		conservado em primatas
G87	A	<i>Otelemur garnetii</i> (galago-grande-do-norte)	
S88	-		conservado em primatas
L89	-		conservado em primatas

D90	-		conservado em primatas
K91	-		conservado em primatas
L92	-		conservado em primatas
L93	-		conservado em primatas
D94	-		conservado em primatas
A95	T	<i>Otelemur garnettii</i> (galago-grande-do-norte), <i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Nycticebus coucang</i> (lóris lento)	
G96	-		conservado em primatas
L97	-		conservado em primatas
Y98	-		conservado em primatas
T99	-		conservado em primatas
G100	-		conservado em primatas
E101	-		conservado em primatas
V102	-		conservado em primatas
T103	-		conservado em primatas
E104	-		conservado em primatas
I105	-		conservado em primatas
V106	-		conservado em primatas
G107	-		conservado em primatas
G108	-		conservado em primatas
P109	-		conservado em primatas
G110	-		conservado em primatas
S111	-		conservado em primatas
G112	-		conservado em primatas
K113	-		conservado em primatas
T114	-		conservado em primatas
Q115	-		conservado em primatas

V116	I	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
C117	-		conservado em primatas
L118	-		conservado em primatas
C119	-		conservado em primatas
M120	V	A maioria é V	
A121	-		conservado em primatas
A122	I	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
N123	-		conservado em primatas
V124	-		conservado em primatas
A125	-		conservado em primatas
H126	-		
G127	S	<i>Saimiri boliviensis boliviensis</i> (Macaco-de-cheiro-boliviano) <i>Sapajus apella</i> (macaco prego das Guianas) <i>Cebus imitator</i> (macaco-prego-de-cara-branca do Panamá)	
L128	-		conservado em primatas
Q129	-		conservado em primatas
Q130	-		conservado em primatas
N131	-		conservado em primatas
V132	-		
L133	-		
Y 134	-		conservado em primatas
V135	I	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	

D136	-		conservado em primatas
S137	-		conservado em primatas
N138	S		
G139	-		conservado em primatas
G140	-		conservado em primatas
L141	-		conservado em primatas
T142	-		conservado em primatas
A143	-		conservado em primatas
S144	-		conservado em primatas
R145	-		conservado em primatas
L146	-		
L147	-		conservado em primatas
Q148	-		conservado em primatas
L149	-		conservado em primatas
L150	-		conservado em primatas
A151	E	<i>Otelemur garnettii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
K152	T	A maioria é T	
T153	R	<i>Otelemur garnettii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento) <i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
T154	-		conservado em primatas
Q155	-		conservado em primatas
D156	-		conservado em primatas
E157	-		conservado em primatas

E158	-		conservado em primatas
E159	-		conservado em primatas
Q160	-		conservado em primatas
A161	-		conservado em primatas
E162	R, G	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento), <i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
A163	-		conservado em primatas
L164	-		conservado em primatas
R165	Q	A maioria é Q	
R166	-		conservado em primatas
I167	-		conservado em primatas
Q168	-		conservado em primatas
V169	-		conservado em primatas
V170	A, T	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
H171	Y, R	A maioria é Y	
A172	V	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
F173	-		conservado em primatas
D174	-		conservado em primatas
I175	-		conservado em primatas
F176	-		conservado em primatas
Q177	-		conservado em primatas
M178	-		conservado em primatas
L179	-		conservado em primatas

D180		<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
V181	-		conservado em primatas
L182	-		conservado em primatas
Q183	-		conservado em primatas
E184	D	<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
L185	-		conservado em primatas
R 186	-		conservado em primatas
G187	S	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
T188	-		conservado em primatas
V189	-		conservado em primatas
A190	S, V	A maioria é V	
Q191	-		conservado em primatas
Q192	-		conservado em primatas
V193	-		conservado em primatas
T194	-		conservado em primatas
G195	S	<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
S196	-		conservado em primatas

S197	-		conservado em primatas
G198	-		conservado em primatas
T199	S	<i>Nycticebus coucang</i> (lóris lento)	
V200	M		
K201	R	<i>Pongo abelli</i> (Orangotango-de-sumatra)	
V202	-		conservado em primatas
V203	-		conservado em primatas
V204	-		conservado em primatas
V205	-		conservado em primatas
D206	-		conservado em primatas
S207	-		conservado em primatas
V208	-		conservado em primatas
T209	-		conservado em primatas
A210	-		conservado em primatas
V211	-		conservado em primatas
V212	-		conservado em primatas
S213	-		conservado em primatas
P214	-		conservado em primatas
L215	-		conservado em primatas
L216	-		conservado em primatas
G217	-		conservado em primatas
G218	-		conservado em primatas
Q219	-		conservado em primatas
Q220	-		conservado em primatas
R221	-		conservado em primatas
E222	-		conservado em primatas

G223	-		conservado em primatas
L224	F	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
A225	-		conservado em primatas
L226	-		conservado em primatas
M227	-		conservado em primatas
M228	I	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
Q229	-		conservado em primatas
L230	-		conservado em primatas
A231	-		conservado em primatas
R232	-		conservado em primatas
E233	-		conservado em primatas
L234	-		conservado em primatas
K235	-		conservado em primatas
T236	-		conservado em primatas
L237	-		conservado em primatas
A238	-		conservado em primatas
R239	K	<i>Nycticebus coucang</i> (lóris lento)	
D240	-		conservado em primatas
L241	-		conservado em primatas
G242	-		conservado em primatas
M243	T	<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnettii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
A244	-		conservado em primatas
V245	-		conservado em primatas
V246	-		conservado em primatas

V247	-		conservado em primatas
T248	-		conservado em primatas
N249	-		conservado em primatas
H250	-		conservado em primatas
I251	T, M	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento) <i>Pongo abellii</i> (Orangotango-de-sumatra)	
T252	-		conservado em primatas
R253	-		conservado em primatas
D254	-		conservado em primatas
R255	-		conservado em primatas
D256	-		conservado em primatas
S257	-		conservado em primatas
G258	R	<i>Pongo abelii</i> (Orangotango-de-sumatra)	
R 259	-		conservado em primatas
L260	-		conservado em primatas
K261	-		conservado em primatas
P262	-		conservado em primatas
A263	-		conservado em primatas
L264	-		conservado em primatas
G265	-		conservado em primatas
R266	H	<i>Callithrix jacchus</i> (<i>sagui-de-tufo-branco</i>)	
S267	-		conservado em primatas
W268	-		conservado em primatas
S269	-		conservado em primatas
F270	-		conservado em primatas

V271	-		conservado em primatas
P272	-		conservado em primatas
S273	-		conservado em primatas
T274	-		conservado em primatas
R275	-		conservado em primatas
I276	L	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
L277	F	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
L278	-		conservado em primatas
D279	-		conservado em primatas
T280	I	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento) <i>Cholorocebus sabaesus</i> (macaco do velho mundo)	
I281	-		conservado em primatas
E282	K, D, N	A maioria é D	
G283	-		conservado em primatas
A284	-		conservado em primatas
G285	E, R	<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
A286	-		conservado em primatas
S287	-		conservado em primatas
G288	A	A maioria é A	
G289	S, R	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	

		<i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
R290	Q, L, W	A maioria é Q	
R291	-		conservado em primatas
M292	T, R	<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>coucang</i> (lóris lento)	
A293	V	<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnetii</i> (galago-grande-do-norte) <i>Nycticebus coucang</i> (lóris lento)	
C294	-		conservado em primatas
L295	-		conservado em primatas
A296	T, I	A maioria é T	
K297	-		conservado em primatas
S298	-		conservado em primatas
S299	P	A maioria é P	
R300	-		conservado em primatas
Q301	-		conservado em primatas
P302	-		conservado em primatas
T303	-		conservado em primatas
G304	-		conservado em primatas
F305	L	<i>Propithecus coquereli</i> (Sifaka-de-coquerel)	
Q306	-		conservado em primatas

E307	K	<i>Callithrix jacchus</i> (sagui-de-tufo-branco)	
M308	-		conservado em primatas
V309	-		conservado em primatas
D310	-		conservado em primatas
I311	T	<i>Pongo abelii</i> (Orangotango-de-sumatra)	
G312	-		conservado em primatas
T313	S	<i>Saimiri boliviensis boliviensis</i> (Macaco-de-cheiro-boliviano)	
W314	-		conservado em primatas
G315	-		conservado em primatas
T316	-		conservado em primatas
S317	P, A	<i>Propithecus coquereli</i> (Sifaka-de-coquerel) <i>Otelemur garnettii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>cougang</i> (lóris lento)	
E318	-		conservado em primatas
Q319	P, H	<i>Otelemur garnettii</i> (galago-grande-do-norte) <i>Nycticebus</i> <i>cougang</i> (lóris lento) <i>Pan troglodytes</i> (chimpanzés)	
S320	-		conservado em primatas
A321	P, T		
T322	A, M, E		
L323	M	<i>Nycticebus cougang</i> (lóris lento)	
Q324	-		conservado em primatas

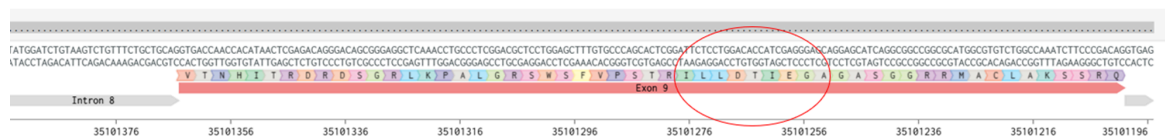
G325	V		
D326	-		conservado em primatas
Q327	-		conservado em primatas
T328	M	<i>Callithrix jacchus</i> (<i>sagui-de-tufo-branco</i>)	

Nota : Tabela finalizada dia 04/05/2025

Fonte: criado pela autora

Com base nas tabelas 4,5 e 6 e nos alinhamentos realizados, observou-se que a partir de posições específicas da proteína, a conservação evolutiva diminui. Este ponto para primatas é na posição 276, para roedores, 279 (Figuras 22 e 22A e 23). Já na tabela 5 e no alinhamento com outras classes, devido a presença dos exemplares de répteis, peixes e anfíbios, a conservação vai até aproximadamente a posição 118 (FIG 24 E 24A).

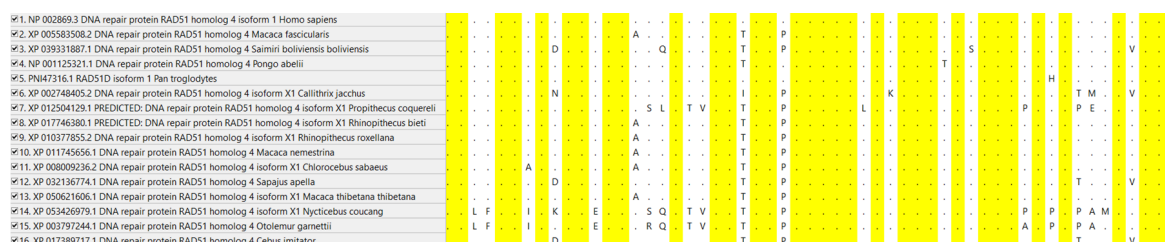
FIGURA 22 Localização do éxon número 9



Região no éxons onde a ocorre a perda da conservação evolutiva em primatas e roedores

Fonte: captura de tela do (<https://www.thermofisher.com/br/en/home/life-science/genome-editing/invitrogen-truedesign-genome-editor.html>), adaptado pela autora. Acesso em 22/05/2025

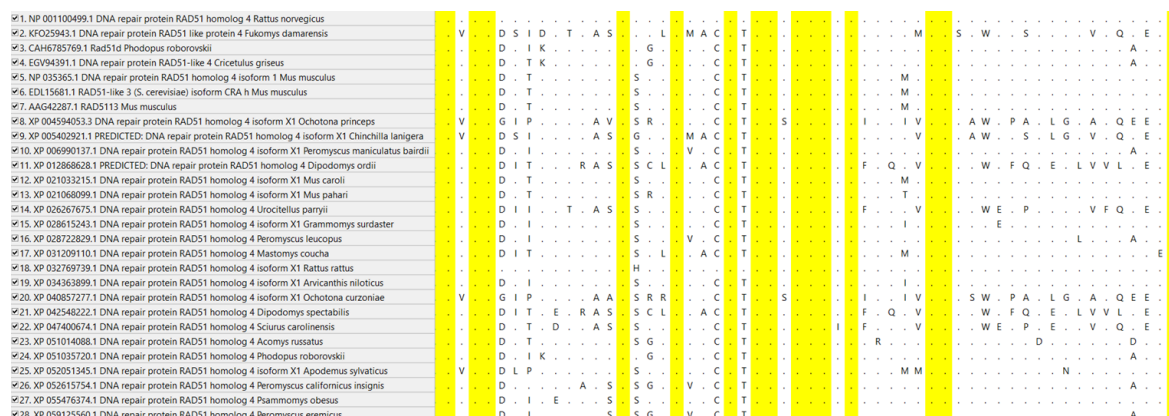
FIGURA 22A) Ponto de diminuição de conservação evolutiva em primatas



Captura de tela do programa MEGA X, que mostra no alinhamento múltiplo da proteína RAD51D em primatas, na qual diminui a conservação evolutiva a partir da posição 276 da proteína.

FONTE: Captura de tela do programa MEGA X

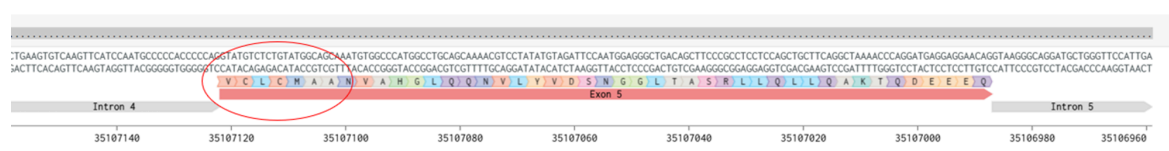
FIGURA 23 - Ponto de diminuição de conservação evolutiva em roedores



Captura de tela do programa MEGA X que mostra no alinhamento múltiplo da proteína RAD51D em roedores, porção final da proteína na qual se perde parte da conservação evolutiva a partir da posição 279 da proteína.

FONTE: Captura de tela do programa MEGA X

FIGURA 24 Localização no éxon número 5



Região no exons onde a perda da conservação evolutiva de aves, répteis, anfíbios e peixes ocorre

Fonte: captura de tela do <https://www.thermofisher.com/br/en/home/life-science/genome-editing/invitrogen-truedesign-genome-editor.html>, adaptado pela autora. Acesso em 22/05/2025

FIGURA 24A Ponto de início da diminuição da conservação evolutiva em aves, répteis, anfíbios e peixes

Para a análise filogenética, na tentativa de utilizar um software que pudesse contornar o problema das diferentes isoformas, empregamos o TaxOnTree, uma ferramenta de construção de árvores taxonômicas. O objetivo foi testar se a sequência canônica obtida no UniProt poderia gerar uma árvore filogenética confiável, excluindo demais isoformas. No entanto, concluímos que o programa não é adequado para nossos objetivos, pois apresentou duas isoformas distintas — uma com 297 aminoácidos e outra com 296 aminoácidos (Figura 26). Apesar da alta similaridade entre essas duas isoformas e outras sequências obtidas pela plataforma, observamos que o alinhamento gerado cobre apenas até o aminoácido 312 da sequência canônica. Dessa forma, o TaxOnTree não será utilizado para avaliação filogenética nesta pesquisa. Ainda assim, vale destacar que, na região alinhada, foi observada uma homologia de cerca de 95% em relação à isoforma canônica.

FIGURA 26- Isoforma 297

Sequence ID: **Query_7674805** Length: **297** Number of Matches: **1**

Range 1: 1 to 297 [Graphics](#)

[▼ Next Match](#) [▲ Previous Matr](#)

Score	Expect	Method	Identities	Positives	Gaps
555 bits(1430)	0.0	Compositional matrix adjust.	297/312(95%)	297/312(95%)	15/312(4%)
Query 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Sbjct 1		MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ			60
Query 61		FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM			120
Sbjct 61		FS--PVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKT-VCLCM			117
Query 121		AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD			180
Sbjct 118		AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQ-EALRRIQVVHAFDIFQMLD			176
Query 181		VLQELRGTVAQQVTGSSGTVKVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLD			240
Sbjct 177		VLQELRGTVAQQVTG-GTVKVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLD			234
Query 241		LGMVVVVTNHITRDRDSGRKLPALGRSWSFVPSTRILLDTIEGAGASGRRMACLAKSSR			300
Sbjct 235		LGMVVVVTNHITRD-DS-RLKLPALGRSWSFVPSTRILLDTIEGAGA--GRRMACLAKSSR			290
Query 301		QPTGFQEMVDIG 312			
Sbjct 291		QP-----MVDIG 297			

Alinhamento pairwise da isoforma canônica com um exemplar das amostras dadas pelo *tax on tree*.

FONTE: Captura de tela do alinhamento realizado pelo BLASTP

Até o momento parece não haver uma plataforma com a finalidade de fazer filogênia excluindo por isoforma. Logo a necessidade da criação de uma plataforma que considere isoforma e que consiga realizar a exclusão delas, para análise filogenética apenas das isoformas canônicas das espécies. Isto facilitaria a análise de conservação evolutiva para um resultado preciso da classificação de variantes genéticas.

3.12 Conservação E Classificação

Mutação de ponto e polimorfismo são amplamente usados para definir qualquer alteração que ocorra no DNA, uma vez que mutação é uma mudança definitiva no DNA e o polimorfismo trata-se de uma variante com frequência populacional acima de 1%. Por esta razão, recomenda-se o uso de variante para definir tais trocas de aminoácidos nas proteínas. (RICHARDS *et al.*, 2015)

A classificação destas variantes referentes a herança mendeliana podem ser patogênicas, provavelmente patogênicas, significado incerto ou VUS, provavelmente benigna ou benigna. O que irá definir isto são critérios rigorosos. Os relatórios laboratoriais que abordam essa classificação devem conter a sequência da variante no nível de DNA, também a codificação e nomenclatura de proteínas (RICHARDS *et al.*, 2015).

Todas essas isoformas complicam a análise de conservação evolutiva, uma vez que não sabemos se posições não conservadas são derivadas de outra isoforma ou apenas são região realmente não conservadas, por esta razão focamos em apenas duas, a 328 e 329.

A conservação evolutiva é um método de análise para classificação de variantes, quando uma região é bem conservada e uma variação ocorre, esta variante pode ser considerada mais patogênica.

Pelo *AphaMissense* é possível observar por posição, as possíveis trocas de aminoácido que favorecem o surgimento de uma variante patogênica. Mas de acordo com as tabelas 4,5 e 6, foi observado que estas posições cuja a troca podem originar uma variante patogênica, são na maioria conservados na isoforma canônica em seres humanos.

Na figura 27 podemos ver que as isoformas presentes em seres humanos em um alinhamento múltiplo para observar as regiões não conservadas são derivados de *splicing* alternativo.

Figura 27 - Alinhamento múltiplo de todas as isoformas presentes em seres humanos

Isoforma 1	1	MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRV-LLA	59
Isoforma 5	1-.....	59
Isoforma 4	1-.....	59
Isoforma 8	1TWR.HSSG-N.G	59
Isoforma 7	1-.....	59
Isoforma 3	1-E....IQVVH	59
Isoforma 6	1-.....	59
Isoforma 2	1S	49
Isoforma 1	60	QFSAFPVNGADLYEELKTSTAILSTGIGS-----LDKLLDAGL	97
Isoforma 5	60R-----Q..SGGSR	96
Isoforma 4	60	97
Isoforma 8	60	GLQLPQ.PAGRSWSGVR--N.LKKA.L.HGGTDGLSLNAFDERGTAVSTSR.....	117
Isoforma 7	60	97
Isoforma 3	60	A.DI.QM--L.VLQ..RGTV.QQV..	83
Isoforma 6	60	87
Isoforma 1	98	YTGE-VTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQD	156
Isoforma 5	97	WCMHL.....	156
Isoforma 4	98	115
Isoforma 8	118	176
Isoforma 7	98	156
Isoforma 1	157	EEEQAEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLL	216
Isoforma 5	157	216
Isoforma 4	116	171
Isoforma 8	177	236
Isoforma 7	157	192
Isoforma 3	48	K.....	104
Isoforma 1	217	GGQQREGLALMMQLARELKT LARDLGMVVVVTNHITRDRDSGRLKPALGRSWSFVPSTRI	276
Isoforma 5	217	276
Isoforma 4	172	231
Isoforma 8	237	296
Isoforma 3	105	164
Isoforma 1	277	LLDTIEGAGASGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT	328
Isoforma 5	277	328
Isoforma 4	232	283
Isoforma 8	297	348
Isoforma 3	165	216

4. Conclusão

Com base em todos os alinhamentos e tabelas analisados neste trabalho, concluí-se que a existência de isoformas da proteína RAD51D existentes afeta a avaliação da conservação evolutiva da proteína e dos seus resíduos individualmente. Observamos que a isoforma canônica varia conforme a classe e dentro das classes, conforme a espécie. Por exemplo, a isoforma canônica em humanos e outros primatas é a 328. Já, em roedores, a isoforma canônica é a 329, mas há espécies que apresentam a 328 e a 329. A isoforma 329 não é vista em primatas. Em aves, as isoformas encontradas com maior frequência têm 313, 314 e 315. A partir de anfíbios, répteis e peixes a conservação evolutiva é menor e mais restrita ao início da proteína.

5. Referências eletrônicas e bibliográficas

ADOLPH, Madison B.; CORTEZ, David. Mechanisms and regulation of replication fork reversal. **Dna Repair**, [S.L.], v. 141, p. 103731, set. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.dnarep.2024.103731>.

BALDOCK, Robert A.; PRESSIMONE, Catherine A.; BAIRD, Jared M.; KHODAKOV, Anton; LUONG, Thong T.; GRUNDY, McKenzie K.; SMITH, Chelsea M.; KARPENSHIF, Yoav; BRATTON-PALMER, Dominique S.; PRAKASH, Rohit. RAD51D splice variants and cancer-associated mutations reveal XRCC2 interaction to be critical for homologous recombination. **Dna Repair**, [S.L.], v. 76, p. 99-107, abr. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.dnarep.2019.02.008>.

BONILLA, Braulio *et al.* **RAD51 Gene Family Structure and Function**. 2020. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7703940/>. Acesso em: 26 nov. 2023.

CASAUBON, Jesse T *et al.* **Mutações BRCA1 e BRCA2**. 2023. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK470239/>. Acesso em: 26 fev. 2025.

CHO, Kathleen R. *et al.* OVARIAN CANCER. Annual Reviews, [s. l], v. 4, n. 1, p. 287-313, 2009. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC2679364/>. Acesso em: 26 fev. 2025.

GRUNDY, McKenzie K; BUCKANOVICH, Ronald J; A BERNSTEIN, Kara. Regulation and pharmacological targeting of RAD51 in cancer. **Nar Cancer**, [S.L.], v. 2, n. 3, p. 1-14, 1 set. 2020. Oxford University Press (OUP). <http://dx.doi.org/10.1093/narcan/zcaa024>. Disponível em: <https://academic.oup.com/narcancer/article/2/3/zcaa024/5911784?login=false>. Acesso em: 11 fev. 2025.

<https://www.gov.br/inca/pt-br> acesso em 05 mar de 2025.

HANAHAN, Douglas. Hallmarks of cancer: new dimensions. *Cancer discovery*, v. 12, n. 1, p. 31-46, 2022. Doi: <https://doi.org/10.1158/2159-8290.CD-21-1059>

HANAHAN, Douglas; WEINBERG, Robert A. Hallmarks of cancer: the next generation. *cell*, v. 144, n. 5, p. 646-674, 2011. Doi: <https://doi.org/10.1016/j.cell.2011.02.013>

Invitrogen true design genome editor Disponível em :
<https://www.thermofisher.com/br/en/home/life-science/genome-editing/invitrogen-true-design-genome-editor.html> acesso 22/05/2025

JIANG, Shan *et al.* Network Analysis of RAD51 Proteins in Metazoa and the Evolutionary Relationships With Their Archaeal Homologs. **Frontiers In Genetics**, China, v. 9, n. 383, p. 1-9, 26 set. 2018. Frontiers Media SA. <http://dx.doi.org/10.3389/fgene.2018.00383>. Disponível em: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00383/full>. Acesso em: 26 nov. 2023.

KATSURA, Chie; OGUNMWONYI, Innocent; KANKAM, Hadyn Kn; SAHA, Sunita. Breast cancer: presentation, investigation and management. **British Journal Of Hospital Medicine**, [S.L.], v. 83, n. 2, p. 1-7, 2 fev. 2022. Mark Allen Group. <http://dx.doi.org/10.12968/hmed.2021.0459>.

KONTOMANOLIS, Emmanuel N. *et al.* Role of Oncogenes and Tumor-suppressor Genes in Carcinogenesis: a review. **Anticancer Research**, [S.L.], v. 40, n. 11, p. 6009-6015, 27 out. 2020. Anticancer Research USA Inc.. <http://dx.doi.org/10.21873/anticancer.14622>.

LIN, Zhenguo *et al.* Origins and evolution of the recA / RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. **Proceedings Of The National Academy of Sciences**, [S.L.], v. 103, n. 27, p. 10328-10333, 5 jul. 2006. Proceedings of the National Academy of Sciences. <http://dx.doi.org/10.1073/pnas.0604232103>. Disponível em: <https://www.pnas.org/doi/full/10.1073/pnas.0604232103>.

LI, Jinbao *et al.* **Pathways and assays for DNA double-strand break repair by homologous recombination.** 2019. Disponível em:

<https://www.sciengine.com/ABBS/doi/10.1093/abbs/gmz076;JSESSIONID=bc01d4e9-2f65-4ed8-86f1-958825f29ec8>.

MENON, Gopal *et al.* **Breast Cancer**. 2024. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK482286/>. Acesso em: 05 mar. 2025.

MOYERS, Julie S.; BILAN, Philip J.; ZHU, Jianhua; KAHN, C. Ronald. Rad and Rad-related GTPases Interact with Calmodulin and Calmodulin-dependent Protein Kinase II. **Journal Of Biological Chemistry**, [S.L.], v. 272, n. 18, p. 11832-11839, maio 1997. Elsevier BV. <http://dx.doi.org/10.1074/jbc.272.18.11832>.

OBEAGU, Emmanuel; Ifeanyi; OBEAGU, Getrude Uzoma. 2024. Breast cancer: a review of risk factors and diagnosis. **Medicine**, [S.L.], 103(3):1-6. <http://dx.doi.org/10.1097/md.00000000000036905>.

RICHARDS, Sue *et al.* Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. **Genet Med.**, [s. l], p. 405-424, maio 2015.

SAKAMOTO, Tetsu *et al.* **TaxOnTree: a tool that generates trees annotated with taxonomic information**. 2020. Disponível em: <https://www.biorxiv.org/content/10.1101/2020.12.24.424364v1>. Acesso em: 05 ago. 2024.

The American Cancer Society. **Genes and Cancer**. 2023. Disponível em: <https://www.cancer.org/cancer/understanding-cancer/genes-and-cancer/oncogenes-tumor-suppressor-genes.html>. Acesso em: 26 fev. 2025.

TORRE, Lindsey A. et al. Ovarian cancer statistics, 2018. **CA: a cancer journal for clinicians**, v. 68, n. 4, p. 284-296, 2018. Doi: <https://doi.org/10.3322/caac.21456>

TOH, Mingren; NGEOW, Joanne. Homologous Recombination Deficiency: cancer predispositions and treatment implications. **The Oncologist**, [S.L.], v. 26, n. 9, p. 1526-1537, 2 jun. 2021. Oxford University Press (OUP). <http://dx.doi.org/10.1002/onco.13829>. Disponível em: <https://academic.oup.com/oncolo/article/26/9/e1526/6512109?login=false>. Acesso em: 10 dez. 2023.

Varsome. Disponível em: <https://varsome.com/gene/hg38/Rad51d> Acesso em 02/04/2025

WASSING, Isabel E.; ESASHI, Fumiko. RAD51: beyond the break. **Seminars In Cell & Developmental Biology**, [S.L.], v. 113, p. 38-46, maio 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.semcd.2020.08.010>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1084952120301014?via%3Dihub>. Acesso em: 27 nov. 2023.

XIONG, Xin; ZHENG, Le-Wei; DING, Yu; CHEN, Yu-Fei; CAI, Yu-Wen; WANG, Lei-Ping; HUANG, Liang; LIU, Cui-Cui; SHAO, Zhi-Ming; YU, Ke-Da. Breast cancer: pathogenesis and treatments. **Signal Transduction And Targeted Therapy**, [S.L.], v. 10, n. 1, p. 1-33, 19 fev. 2025. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41392-024-02108-4>.

YOSHIDA, Kiyotsugu; MIKI, Yoshio. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. **Cancer Science**, [S.L.], v. 95, n. 11, p. 866-871, nov. 2004. Wiley. <http://dx.doi.org/10.1111/j.1349-7006.2004.tb02195>.

6.0 Sequências

>sp|O75771|RA51D_HUMAN DNA repair protein RAD51 homolog 4 OS=Homo sapiens OX=9606 GN=RAD51D PE=1 SV=1

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ
 FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM
 AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD
 VLQELRGTVAAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLD
 LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSRILLDTIEGAGASGGRRMACLAKSSR
 QPTGFQEMVDIGTWGTSEQSATLQGDQT

>sp|O75771-2|RA51D_HUMAN Isoform 2 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKS

>sp|O75771-3|RA51D_HUMAN Isoform 3 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKAEALRRIQVVHA
FDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAR
ELKTLARDLGMVVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRM
ACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT

>sp|O75771-4|RA51D_HUMAN Isoform 4 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ
FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQAEALR
RIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLA
LMMQLARELKTALARDLGMVVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAG
ASGGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT

>sp|O75771-5|RA51D_HUMAN Isoform 5 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ
FSAFPVNGADLYEELKTSTAILSTGIGRQKLSGGSRWCMHLVTEIVGGPGSGKTQVCLCM
AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD
VLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTALARD
LGMVVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSSR
QPTGFQEMVDIGTWGTSEQSATLQGDQT

>sp|O75771-6|RA51D_HUMAN Isoform 6 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ
FSAFPVNGADLYEELKTSTAILSTGIGRHGGRTQVGTWEDCSCLRSPQDGRGVGSGML

>sp|O75771-7|RA51D_HUMAN Isoform 7 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ
FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM
AANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD
VLQELRGTVAQQDGIPEHLNHIPHCLHVHLPC

>sp|O75771-8|RA51D_HUMAN Isoform 8 of DNA repair protein RAD51 homolog 4 OS=Homo sapiens
OX=9606 GN=RAD51D

MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKTWRAHSSGNLGG
LQLPQVPAGRSWSGVRNALKKAGLGHGGTDGLSLNAFDERGTAVSTSRLDKLLDAGLYTG
EVTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNGGLTASRLQLLQAKTQDEEEQ
AEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDVAVVSPLLGGQQ
REGLALMMQLARELKTLDLGMVAVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDT
IEGAGASGGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT