UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

Thalía Rodrigues de Souza Zózimo

O IMPACTO DAS FERRAMENTAS *IN SILICO* E RECURSOS DE BIOINFORMÁTICA ESTRUTURAL NA DETERMINAÇÃO DA PATOGENICIDADE DE VARIANTES GENÉTICAS ENCONTRADAS EM *RAD51D*

Belo Horizonte
2023

Thalía Rodrigues de Souza Zózimo

O IMPACTO DAS FERRAMENTAS *IN SILICO* E RECURSOS DE BIOINFORMÁTICA ESTRUTURAL NA DETERMINAÇÃO DA PATOGENICIDADE DE VARIANTES GENÉTICAS ENCONTRADAS EM *RAD51D*

Dissertação apresentada ao Programa de Pós-Graduação em Genética do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Mestre em Genética.

Orientadora: Profa. Dra. Maria Raquel Santos Carvalho

Co-orientadora: Profa. Dra. Letícia da Conceição Braga

Belo Horizonte
2023

043 Zózimo, Thalía Rodrigues de Souza.
O impacto das ferramentas in silico e recursos de bioinformática estrutural na determinação da patogenicidade de variantes genéticas encontradas em RAD51D [manuscrito] / Thalía Rodrigues de Souza Zózimo. – 2023.
243 f. : il. ; 29,5 cm.

Orientadora: Profa. Dra. Maria Raquel Santos Carvalho. Co-orientadora: Profa. Dra. Letícia da Conceição Braga.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética. 2. Bioinformática. 3. Síndrome Hereditária de Câncer de Mama e Ovário. I. Carvalho, Maria Raquel Santos. II. Braga, Letícia da Conceição. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 575

Ficha catalográfica elaborada pela bibliotecária Fabiane C M Reis – CRB 6 – 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética

ATA DE DEFESA DE DISSERTAÇÃO

ATA DA DEFESA DE DISSERTAÇÃO	**339/2023** **entrada**
Thalía Rodrigues de Souza	**1º/2021** **CPF: 124.716.936-79**

Às quatorze horas do dia **28 de abril de 2023**, reuniu-se, no Instituto de Ciências Biológicas da UFMG, a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **"O impacto das ferramentas in silico e recursos de bioinformática estrutural na determinação da patogenicidade de variantes genéticas encontradas em RAD51D"**, requisito para obtenção do grau de Mestre em **Genética.** Abrindo a sessão, a Presidente da Comissão, **Maria Raquel 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	CPF	Indicação
Maria Raquel Carvalho	UFMG	349.651.730-15	APROVADA
Letícia da Conceição Braga	FUNDEP	864.990.846-20	APROVADA
Vasco Ariston Carvalho de Azevedo	UFMG	283.171.225-49	APROVADA
Michele Araújo Pereira	Hospital Israelita Albert Einstein	066.507.726-23	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, 28 de abril de 2023.

Maria Raquel Carvalho

Letícia da Conceição Braga

Vasco Ariston Carvalho de Azevedo

Michele Araújo Pereira

AGRADECIMENTOS

Primeiramente, dedico minha vida a Deus e agradeço pelas oportunidades a mim oferecidas, por essa conquista e pelo apoio que me deu nos momentos de desânimo, desespero e insegurança, que foram muitos. Agradeço, não só por ser meu apoio em momentos difíceis, mas acima de tudo, por me motivar, proporcionar momentos tão felizes como este e colocar pessoas tão incríveis ao meu lado.

À minha família dedico meu trabalho. Queridos pais, Leila e Valdeci, obrigada por serem meu exemplo, ensinando-me a ser responsável, honesta, determinada, sem jamais esquecer de ser amiga e companheira. Aos meus padrinhos Simone e Carlos, agradeço imensamente pelo suporte, energias positivas e por me lembrarem que sou mais forte do que penso. Meu marido, Richard, meus irmãos e irmãs, Haleff, Ketlen, Amanda e Thiago, agradeço o apoio e compreensão. Essa conquista não seria possível sem vocês ao meu lado, confiando no meu potencial e me incentivando a ser alguém melhor.

À professora Maria Raquel, cujo olhar aguçado e conhecimento admirável encontrou muitas incongruências que eu poderia ter deixado passar, agradeço por ter confiado em mim em seu grupo de pesquisa e incentivo para o desenvolvimento deste projeto. Agradeço também ao Laboratório de Genética Humana e Médica e a todos os colegas pelas oportunidades, troca de conhecimento e parceria.

À minha coorientadora, Letícia Braga, obrigada pelo apoio e auxílio em todas as etapas do projeto. Aproveito para agradecer a todos os colaboradores do laboratório do Instituto Mário Pena, pois durante minhas visitas ao Instituto todos foram receptivos e muito atenciosos.

À bioinformata Carolina Guimarães, quero agradecer por me ajudar na execução dos programas, por se colocar à disposição para sanar minhas dúvidas e por ser uma grande amiga que com certeza quero levar comigo para o resto da vida.

À Doutora Anisse Ferraz, parceira de trabalho e grande amiga, gostaria de agradecer todo suporte, incentivo e troca de conhecimento ao longo do projeto.

Agradeço a todos os meus amigos que, sempre acreditaram em mim, me incentivaram a seguir esse caminho e estiveram ao meu lado sempre que precisei. Em especial, quero agradecer a Lorna Samantha, Eduardo Utsch, Thamiris Matias e Flávia Santiago, por me apoiarem e me ajudarem em diversos momentos. Agradeço também a todos os colegas que fiz durante as disciplinas, obrigada por todo aprendizado.

Agradeço à Universidade Federal de Minas Gerais, ao Programa de Pós-graduação em Genética, aos professores com quem fiz as disciplinas ao longo do mestrado, que me ensinaram muito e à equipe administrativa, todos muito prestativos.

Por fim, gostaria de agradecer às agências de fomento à pesquisa, CNPq e CAPES, que financiaram minha bolsa, permitindo-me continuar no projeto com total dedicação.

Recebam o meu “muito obrigada”, repleto de amor e carinho!

RESUMO

O câncer é uma das principais causas de morbimortalidade. No Brasil, os tipos de câncer mais comuns são: mama, câncer colorretal, próstata e estômago. A principal causa do câncer são as alterações genéticas e epigenéticas em oncogenes, genes supressores de tumor, genes miRNAS, entre outros. Pelo menos 291 genes humanos estão implicados na oncogênese, sendo todos eles codificadores de proteínas, que atuam em regulação da proliferação celular, reparo do DNA, diferenciação e morte celular. Consequentemente, a inclusão testes genéticos se tornou fundamental no manejo clínico e no aconselhamento genético de indivíduos afetados e seus familiares. O número de genes incluídos nos painéis de triagem de variantes genéticas em pacientes com cânceres aumentou e existe a tendência para substituição destes painéis pelo exoma. Entretanto, variantes genéticas são muito comuns e é necessário avaliar o seu potencial de patogenicidade considerando-se o gene, no qual a variante foi detectada, e o fenótipo. Neste contexto, a Bioinformática fornece muitos recursos para a avaliação de variantes, incluindo conservação evolutiva do resíduo, impacto na estrutura e função da proteína ou RNA, impacto sobre modificações pós-síntese etc. Entretanto, a quantidade de recursos usados varia conforme o laboratório. Além disto, evidências vindas da Bioinformática frequentemente não são suficientes para permitir a classificação de uma variante, sendo necessários dados clínicos e familiares. Consequentemente, existe muitas variantes genéticas classificadas como de significado incerto (VUS). O gene *RAD51D* foi recentemente implicado na suscetibilidade aos cânceres de mama e ovário. O objetivo deste trabalho foi avaliar a eficiência de uma análise bioinformática pormenorizada na reclassificação de 20 variantes germinativas identificadas no gene *RAD51D* em indivíduos brasileiros com câncer de mama, ovário ou outras neoplasias. Estas variantes já foram depositadas em bancos de dados, estando classificadas como VUS, provavelmente benignas ou patogênicas. No presente estudo, elas foram avaliadas quanto à frequência alélica, impacto na proteína (conservação evolutiva do resíduo, localização em domínios funcionais da proteína) ou no RNA (*splicing* e *nonsense mediated decay*). Algoritmos também foram executados para prever o efeito de substituições de sentido trocado sobre estrutura terciária e função da proteína. Ao final das análises, foi executada a reclassificação das variantes segundo os critérios do ACMG/AMP. Foi possível reclassificar para patogênica/provavelmente patogênica as variantes *RAD51D* c.26G>C; (p.Cys9Ser), *RAD51D* c.137C>G; (p.Ser46Cys), *RAD51D* c.140A>G; (p.Tyr47Cys), *RAD51D* c.493C>T; (p.Arg165Trp). RAD51D c.355T>C; (p.Cys119Arg) permanece classificada como VUS, embora a conservação evolutiva em mamíferos e provável formação de pontes de dissulfeto com a Cys117, sugiram patogenicidade. As demais variantes analisadas mantém a classificação atual como VUS ou benignas/provavelmente benignas. A principal dificuldade encontrada, especificamente em relação à proteína *RAD51D* foi a existência de grande número de isoformas, que dificultam a análise da conservação evolutiva. Estes resultados mostram, que a avaliação *in silico* detalhada pode ajudar a resolver a classificação de determinadas variantes ou a selecionar quais seriam alvos prioritários para estudos funcionais ou estudos de associação nas famílias.

Palavras-chaves: Síndromes de Câncer Hereditárias. Câncer de mama. Câncer de ovário. *RAD51D*. Bioinformática.

ABSTRACT

Cancer is one of the main causes of morbidity and mortality. In Brazil, the most common types of cancer are breast, colorectal, prostate, and stomach cancer. The main causes of cancer are genetic and epigenetic alterations in oncogenes, tumor suppressor genes, miRNA genes, among others. At least 291 human genes are implicated in oncogenesis, all of which encode proteins that regulate cell proliferation, DNA repair, differentiation, and cell death. Consequently, the inclusion of genetic testing has become fundamental in the clinical management and genetic counseling of affected individuals and their families. The number of genes included in screening panels for genetic variants in cancer patients has increased and there is a tendency to replace these panels with the exome. However, genetic variants are very common, and it is necessary to evaluate their pathogenicity potential considering the gene, in which the variant was detected, and the phenotype. In this context, Bioinformatics provides many resources for the evaluation of variants, including evolutionary conservation of the residue, impact on protein or RNA structure and function, impact on post-synthesis modifications, etc. However, the number of resources used varies by laboratory. In addition, evidence from Bioinformatics is often not sufficient to allow the classification of a variant, requiring clinical and family data. Consequently, there are many genetic variants classified as of uncertain significance (VUS). The *RAD51D* gene has recently been implicated in susceptibility to breast and ovarian cancers. The objective of this work was to evaluate the efficiency of a detailed bioinformatics analysis in the reclassification of 20 germline variants identified in the *RAD51D* gene in Brazilian individuals with breast, ovarian or other cancers. These variants have already been deposited in databases, being classified as VUS, probably benign or pathogenic. In the present study, they were evaluated for allele frequency, impact on the protein (evolutionary conservation of the residue, location in protein functional domains) or on the RNA (splicing and nonsense mediated decay). Algorithms were also run to predict the effect of missense substitutions on tertiary protein structure and function. At the end of the analyses, the variants were reclassified according to the ACMG/AMP criteria. It was possible to reclassify *RAD51D* c.26G>C; (p.Cys9Ser), *RAD51D* c.137C>G; (p.Ser46Cys), and *RAD51D* c.140A>G; (p.Tyr47Cys) as pathogenic/probably pathogenic. *RAD51D* c.355T>C; (p.Cys119Arg) remains classified as a VUS, although evolutionary conservation in mammals and probable formation of disulfide bonds with Cys117 suggest pathogenicity. The other variants analyzed maintain the current classification as VUS or benign/probably benign. The main difficulty encountered, specifically in relation to the RAD51D protein, was the existence of many isoforms, which make the analysis of evolutionary conservation difficult. These results show that detailed *in silico* evaluation can help to solve the classification of certain variants or to select which would be priority targets for functional studies or association studies in families.

Keywords: Cancer Syndromes Hereditary. Breast Cancer. Ovary Cancer. RAD51D protein, human. Bioinformatics.

LISTA DE ILUSTRAÇÕES

Figura 1 - Marcadores de câncer 16
Figura 2- Principais causas de morte no mundo em 2020 17
Figura 3 - Distribuição proporcional dos dez tipos de câncer mais incidentes no triênio 2020-2022 19
Figura 4 - Origem de diferentes tipos de câncer de ovário 24
Figura 5 - Mecanismos de reparo de danos no DNA 28
Figura 6 - Etapas e recursos usados na avaliação do impacto funcional das variantes neste estudo 43
Figura 7 - Alinhamento das isoformas funcionais de RAD51D, produzidas por splicing alternativo 48
Figura 8 - Nível de expressão de cada transcrição de RAD51D em diferentes tecidos 49
Figura 9 - Estrutura tridimensional e predição de desordem intrínseca para a proteína RAD51D 50
Figura 10 - Representação esquemática dos sítios de modificações pós-síntese da RAD51D 51
Figura 11 - Representação esquemática da organização dos domínios e motifs na estrutura da proteína RAD51D 52
Figura 12 - Localização das variantes em relação ao gene e aos transcritos, ENST00000345365.11 (O75771-1/canônico) e ENST00000590016.5 (O75771-8) 66
Figura 13 - Representação esquemática dos éxons que compõem as isoformas O75771-1/canônica e O75771-8 67
Figura 14 - Impacto da RAD51D c.26G>C;(p.Cys9Ser) nas sequências auxiliares regulatórias de sítios de splicing 77
Figura 15 - Impacto da variante RAD51D c.26G>C;(p.Cys9Ser) sobre pontes de dissulfeto da RAD51D usando Dianna 78
Figura 16- Dinâmica molecular da variante RAD51D c.26G>C; (p.Cys9Ser) 79
Figura 17- Impacto da variante RAD51D c.26G>C; (p.Cys9Ser) na desordem proteica 80
Figura 18 - Predição de impacto de RAD51D c.26G>C; (p.Cys9Ser) nos motivos proteicos 81
Figura 19- Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.26G>C; (p.Cys9Ser)) 82
Figura 20- Impacto da variante RAD51D c.137C>G; (p.Ser46Cys) sobre pontes de dissulfeto da RAD51D usando Dianna 86
Figura 21- Dinâmica molecular da variante RAD51D c.137C>G; (p.Ser46Cys) 87
Figura 22 - Figura esquemática do efeito da RAD51D c.137C>G; (p.Ser46Cys) sobre o sítio de modificações pós-síntese 88
Figura 23 - Predição de impacto de RAD51D c.137C>G; (p.Ser46Cys) nos motivos proteicos 89
Figura 24- Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.137C>G; (p.Ser46Cys)) 90

Figura 25- Impacto da RAD51D c.140A>G;(p.Tyr47Cys) nas sequências auxiliares regulatórias de sítios de splicing 93
Figura 26 - Dinâmica molecular da variante RAD51D c.140A>G; (p.Tyr47Cys) 94
Figura 27 - Impacto da variante RAD51D c.140A>G;(p.Tyr47Cys) sobre pontes de dissulfeto da RAD51D usando Dianna 95
Figura 28 - Figura esquemática do efeito da variante sobre o sítio de modificações pós-síntese .. 96
Figura 29 - Predição de impacto de RAD51D c.140A>G; (p.Tyr47Cys) sobre os motivos proteicos 97
Figura 30 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.140A>G; (p.Tyr47Cys) . 98
Figura 31 - Impacto da RAD51D c.190C>T; (p.Pro64Ser) no sítio de splicing 101
Figura 32 - Dinâmica molecular da variante RAD51D c.190C>T; (p.Pro64Ser) 102
Figura 33 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.190C>T; (p.Pro64Ser)) usando a isoforma 8 como modelo 103
Figura 34 - Dinâmica molecular da variante RAD51D c.196G>A; (p.Val66Met) 106
Figura 35 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.196G>A; (p.Val66Met)) 107
Figura 36- Dinâmica molecular da variante RAD51D c.286G>A; (p.Ala96Thr)..... 110
Figura 37- Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.286G>A; (p.Ala96Thr)) 111
Figura 38 - Dinâmica molecular da variante RAD51D c.355T>C; (p.Cys119Arg)..... 114
Figura 39 - Impacto da variante RAD51D c.355T>C; (p.Cys119Arg) sobre pontes de dissulfeto da RAD51D usando Dianna 115
Figura 40 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.355T>C; (p.Cys119Arg)) 116
Figura 41 - Dinâmica molecular da variante RAD51D c.413A>G; (p.Asn138Ser)..... 119
Figura 42 - Predição de impacto de RAD51D c.413A>G; (p.Asn138Ser) nos motivos proteicos120
Figura 43- Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.413A>G; (p.Asn138Ser)) 121
Figura 44 - Dinâmica molecular da variante RAD51D c.493C>T; (p.Arg165Trp)..... 124
Figura 45 - Impacto da variante RAD51D c.493C>T; (p.Arg165Trp) na desordem proteica 125
Figura 46 - Predição de impacto de RAD51D c.493C>T; (p.Arg165Trp) nos motivos proteicos126
Figura 47 - Estrutura 3D da RAD51D selvagem e da RAD51D c.493C>T; (p.Arg165Trp)..... 127
Figura 48 - Dinâmica molecular da variante RAD51D c.628G>A; (p.Ala210Thr)..... 131
Figura 49 - Predição de impacto de RAD51D c.628G>A; (p.Ala210Thr) nos motivos proteicos133
Figura 50 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.628G>A; (p.Ala210Thr)) 134

Figura 51 - Dinâmica molecular da variante RAD51D c.695G>A; (p.Arg232Gln) 137
Figura 52 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.695G>A; (p.Arg232Gln)) 138
Figura 53 - Impacto da RAD51D c.698A>G; (p.Glu233Gly) nas sequências auxiliares regulatórias de sítios de splicing 140
Figura 54 - Dinâmica molecular da variante RAD51D c.698A>G; (p.Glu233Gly) 142
Figura 55 - Predição de impacto de RAD51D c.698A>G; (p.Glu233Gly) nos motivos proteicos143
Figura 56 - Estrutura 3D da RAD51D selvagem e mutada RAD51D c.698A>G; (p.Glu233Gly)144
Figura 57 - Impacto da RAD51D c.715C>T (p.Arg239Trp) nas sequências auxiliares regulatórias de sítios de splicing 148
Figura 58 - Dinâmica molecular da variante RAD51D c.715C>T (p.Arg239Trp) 150
Figura 59 - Predição de impacto de RAD51D c.715C>T (p.Arg239Trp) nos motivos proteicos 151
Figura 60 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.715C>T (p.Arg239Trp)) 152
Figura 61 - Dinâmica molecular da variante RAD51D c.862G>C (p.Gly288Arg) 155
Figura 62 - Predição de impacto de RAD51D c.862G>C (p.Gly288Arg) nos motivos proteicos 156
Figura 63 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.862G>C (p.Gly288Arg)) 157
Figura 64 - Dinâmica molecular da variante RAD51D c.919G>A; (p.Glu307Lys) 160
Figura 65 - Predição de impacto de RAD51D c.919G>A; (p.Glu307Lys) nos motivos proteicos161
Figura 66 - Estrutura 3D da RAD51D selvagem e mutada (RAD51D c.919G>A; (p.Glu307Lys)) 162
Figura 67 - Impacto da RAD51D c.973G>A (p.Gly325Ser) nas sequências auxiliares regulatórias de sítios de splicing 164
Figura 68 - Dinâmica molecular da variante RAD51D c.973G>A (p.Gly325Ser) 165
Figura 69 - Impacto da variante RAD51D c.973G>A (p.Gly325Ser) na desordem proteica 166
Figura 70 - Estrutura 3D da RAD51D selvagem e mutada RAD51D c.973G>A (p.Gly325Ser) 167
Figura 71- Impacto da RAD51D c.666A>G; (p.Glu222=) nas sequências auxiliares regulatórias de sítios de splicing 169
Figura 72 - Impacto da RAD51D c.694C>T; (p.Arg232Ter) nas sequências auxiliares regulatórias de sítios de splicing 172
Figura 73 - Estrutura 3D da RAD51D selvagem e mutada RAD51D c.694C>T; (p.Arg232Ter) 173

LISTA DE TABELAS E QUADROS

Tabela 1 - Variantes germinativas em RAD51D avaliadas no presente estudo 40
Tabela 2 - Descrição da consequência molecular, frequência populacional, localização molecular e classificação de cada variante RAD51D analisada neste estudo 67
Tabela 3 - Conservação evolutiva dos resíduos afetados pelas variantes de sentido trocado em relação aos mamíferos, aves, anfíbios e répteis e peixes 72
Tabela 4 - Conservação evolutiva de resíduos afetados por variantes de sentido trocado, que ocorrem na isoforma 8, considerando as espécies de mamíferos, aves, anfíbios e répteis e peixes 75
Tabela 5 - Comparação da frequência de (p.Cys9Ser) entre casos e controles 83
Quadro 1 - Descrição da classificação das variantes e nível de patogenicidade (P) e benignidade (B) de acordo com ACMG/AMP.....31
Quadro 2 - Regras do ACMG/AMP para definição da interpretação das variantes como patogênica (P) e provavelmente patogênica (PP).....35
Quadro 3 - Isoformas funcionais codificadas por *RAD51D* e potenciais de isoformas mapeadas computacionalmente 46
Quadro 4 - Descrição funcional e esquematização da localização dos motivos funcionais preditos para RAD51D com o Eukaryotic Linear Motif.....52
Quadro 5 - Predição de impacto molecular da *RAD51D* c.666A>G; (p.Glu222=) sobre a estrutura de RNA, usando ViennaRNA 170
Quadro 6 - Descrição dos resultados obtidos na análise in silico, evidenciando os efeitos no DNA/RNA e proteína RAD51D.....178

LISTA DE SIGLAS E ABREVIATURAS

ACMG - *American College of Medical Genetics and Genomics*

AMP - *Association for Molecular Pathology dos EUA*

BCL - *Basal cell-like*

BER – *Base excision repair/*reparo por excisão de base

CA-125 - Antígeno de câncer 125

DDR - *DNA damage response*/Resposta ao dano do DNA

DSBs - *Double-strand breaks/*quebras de fita dupla de DNA

ER - Receptores de estrógenos

HBOC - Síndrome hereditária do câncer de mama e ovário

HER2 - Receptor 2 do fator de crescimento epidérmico humano

INCA - Instituto Nacional do Câncer

MMR - *Mismatch repair/*reparo de despareamento

NGS - *Next Generation Sequencing*/Sequenciamento de Próxima Geração

NHEJ - *Non-homologous ending joining/*junção de extremidades não homólogas

PR - Receptores de progesterona

RR - Risco Relativo

SSBs - *Single-strand DNA breaks/*quebras de fita simples de DNA

TNBC - *Triple-negative breast cancer/*Câncer de mama triplo-negativo

VUS - *Variant uncertain significance*/Variante de significado incerto

WES - *Whole exome sequencing*/sequenciamento completo de exoma

WGS - *Whole genome sequencing*/sequenciamento completo do genoma

SUMÁRIO

1 INTRODUÇÃO 16
1.1 Câncer de mama 20
1.2 Câncer de ovário 23
1.3 Câncer e instabilidade genômica 24
1.4 Testes genéticos e bioinformática na medicina genômica 30
1.5 Justificativa 36
1.6 Objetivo Geral 38
1.7 Objetivos Específicos 39
2. MATERIAIS E MÉTODOS 39
2.1 Aprovação do COEP 39
2.2 Casuística 39
2.3 Metodologia de análise *in silico* das variantes 42
2.3.1 Anotação funcional 43
2.3.2 Variantes de *RAD51D* 44
2.3.2.1 Conservação evolutiva 44
2.3.2.2 Análise das variantes germinativas (sentido trocado, sinônimas e intrônicas) em nível do DNA, RNA e proteína (primária, secundária e terciária) 44
3. RESULTADOS E DISCUSSÃO 45
3.1 - Anotação funcional do gene, transcritos e isoformas de *RAD51D* 45
3.2 - Variantes no gene *RAD51D* 66
3.2.1 Conservação evolutiva 72
3.2.2 Estudo das variantes germinativas de sentido trocado 75
3.2.2.1 *RAD51D* c.26G>C;(p.Cys9Ser) 76
3.2.2.1.1 Reclassificação da variante *RAD51D* c.26G>C; (p.Cys9Ser) 82
3.2.2.1.2 Conclusão sobre a reclassificação da variante *RAD51D* c.26G>C; (p.Cys9Ser) 85
3.2.2.2 *RAD51D* c.137C>G;(p.Ser46Cys) 86
3.2.2.2.1. Reclassificação da variante *RAD51D* c.137C>G; (p.Ser46Cys) 90
3.2.2.2.2 Conclusão sobre a reclassificação da variante *RAD51D* c.137C>G; (p.Ser46Cys) 92
3.2.2.3 *RAD51D* c.140A>G;(p.Tyr47Cys) 92
3.2.2.3.1 Reclassificação da variante *RAD51D* c.140A>G; (p.Tyr47Cys) 98
3.2.2.3.2 Conclusão sobre a reclassificação da variante *RAD51D* c.140A>G; (p.Tyr47Cys) 99
3.2.2.4 *RAD51D* c.190C>T; (p.Pro64Ser) 100
3.2.2.4.1 Reclassificação da variante *RAD51D* c.190C>T; (p.Pro64Ser) 103

3.2.2.4.2 *Conclusão sobre a reclassificação da variante RAD51D c.190C>T; (p.Pro64Ser)* 104
3.2.2.5 *RAD51D* c.196G>A; (p.Val66Met) 104
3.2.2.5.1. Reclassificação da variante *RAD51D* c.196G>A; (p.Val66Met).....107
3.2.2.5.2 Conclusão sobre a reclassificação da variante *RAD51D* c.196G>A; (p.Val66Met).....108
3.2.2.6 *RAD51D* c.286G>A; (p.Ala96Thr).....109
3.2.2.6.1 Reclassificação da variante *RAD51D* c.286G>A; (p.Ala96Thr) 111
3.2.2.6.2 Conclusão sobre a reclassificação da variante *RAD51D* c.286G>A; (p.Ala96Thr) 112
3.2.2.7 *RAD51D* c.355T>C; (p.Cys119Arg).....112
3.2.2.7.1 Reclassificação da variante *RAD51D* c.355T>C; (p.Cys119Arg) 116
3.2.2.7.2 Conclusão sobre a reclassificação da *RAD51D* c.355T>C; (p.Cys119Arg)117
3.2.2.8 *RAD51D* c.413A>G; (p.Asn138Ser).....118
3.2.2.8.1 Reclassificação da variante *RAD51D* c.413A>G; (p.Asn138Ser) 121
3.2.2.8.2 Conclusão sobre a reclassificação da variante *RAD51D* c.413A>G; (p.Asn138Ser).....122
3.2.2.9 *RAD51D* c.493C>T; (p.Arg165Trp).....123
3.2.2.9.1 Reclassificação da variante *RAD51D* c.493C>T; (p.Arg165Trp).....127
3.2.2.9.2 Conclusão sobre a reclassificação da variante *RAD51D* c.493C>T; (p.Arg165Trp).....128
3.2.2.10 *RAD51D* c.628G>A;(p.Ala210Thr).....130
3.2.2.10.1 Reclassificação da variante *RAD51D* c.628G>A; (p.Ala210Thr) 134
3.2.2.10.2 Conclusão sobre a reclassificação da variante *RAD51D* c.628G>A; (p.Ala210Thr) 135
3.2.2.11 *RAD51D* c.695G>A;(p.Arg232Gln) 135
3.2.2.11.1 Reclassificação da variante *RAD51D* c.695G>A; (p.Arg232Gln).....138
3.2.2.11.2 Conclusão sobre a reclassificação da variante *RAD51D* c.695G>A; (p.Arg232Gln).....139
3.2.2.12 *RAD51D* c.698A>G; (p.Glu233Gly) 140
3.2.2.12.1 Reclassificação da variante *RAD51D* c.698A>G; (p.Glu233Gly).....144
3.2.2.12.2 Conclusão sobre a reclassificação da variante *RAD51D* c.698A>G; (p.Glu233Gly).....146
3.2.2.13 *RAD51D* c.715C>T (p.Arg239Trp) 148
3.2.2.13.1 Reclassificação da variante *RAD51D* c.715C>T (p.Arg239Trp).....152
3.2.2.13.2 Conclusão sobre a reclassificação da variante *RAD51D* c.715C>T (p.Arg239Trp).....153
3.2.2.14 RAD51D c.862G>C (p.Gly288Arg).....154
3.2.2.14.1 Reclassificação da variante *RAD51D* c.862G>C (p.Gly288Arg) 157
3.2.2.14.2 Conclusão sobre a reclassificação da variante RAD51D c.862G>C (p.Gly288Arg).....158
3.2.2.15 *RAD51D* c.919G>A; (p.Glu307Lys) 158

3.2.2.15.1 Reclassificação da variante RAD51D c.919G>A; (p.Glu307Lys) 162

3.2.2.15.2 Conclusão sobre a reclassificação da variante *RAD51D* c.919G>A; (p.Glu307Lys).....163

3.2.2.16 *RAD51D* c.973G>A (p.Gly325Ser) 164

3.2.2.16.1 Reclassificação da variante *RAD51D* c.973G>A (p.Gly325Ser).....167

3.2.2.16.2 Conclusão sobre a reclassificação da variante *RAD51D* c.973G>A (p.Gly325Ser) 168

3.2.3 Estudo da variante germinativa sinônima *RAD51D* c.666A>G; (p.Glu222=).....169

3.2.3.1 Reclassificação da variante *RAD51D* c.666A>G; (p.Glu222=).....170

3.2.3.2 Conclusão sobrea a reclassificação da variante *RAD51D* c.666A>G; (p.Glu222=) 171

3.2.4 Estudo da variante germinativa perda de sentido *RAD51D* c.694C>T; (p.Arg232Ter) 172

3.2.4.1 Reclassificação da variante RAD51D c.694C>T; (p.Arg232Ter) 174

3.2.4.2 Conclusão sobre a reclassificação da variante *RAD51D* c.694C>T; (p.Arg232Ter).....174

3.2.5 Estudo das variantes germinativas intrônicas *RAD51D* c.345+5A>G e a *RAD51D* c.481-7G>A 175

3.2.5.1 Reclassificação da variante *RAD51D* c.345+5A>G 175

3.2.5.1.2 Conclusão sobre a reclassificação da variante *RAD51D* c.345+5A>G.....176

3.2.5.2 Reclassificação da variante *RAD51D* c.481-7G>A 176

3.2.5.2.1 Conclusão sobre a reclassificação da variante *RAD51D* c.481-7G>A 176

3.2.6 Síntese dos resultados 177

4. CONSIDERAÇÕES FINAIS..... 186

6. REFERÊNCIAS 193

ANEXO 1 - COEP UFMG.....211

ANEXO 2 - COEP FACULDADE DE MEDICINA DA UFMG 215

ANEXO 3 - Resultado da *RAD51D* c.666A>G; (p.Glu222=).....219

ANEXO 4 - Alinhamento múltiplo da isoforma canônica (O75771-1).....221

ANEXO 5 - Alinhamento múltiplo da isoforma 8 (O75771-8).....240

1 INTRODUÇÃO

O termo câncer é utilizado para um conjunto de mais de 100 tipos de neoplasias, que podem acometer diversos órgãos ou tecidos (National Cancer Institute, 2022). Sua natureza está intimamente relacionada ao acúmulo de mutações gênicas, cromossômicas ou epimutações, que ocasionam fenótipos incompatíveis com as funções celulares normais e a organização estrutural e fisiológica dos tecidos (Weinberg, 2008). Em conjunto, alterações genéticas e epigenéticas colaboram para o fenótipo patogênico das células cancerígenas da formação à manutenção dos tumores. Um modelo conceitual foi proposto por Hanahan e Weinberg (2000) acerca das capacidades distintas e complementares, que possibilitam a sobrevivência, proliferação e disseminação de células cancerígenas. Denominado como habilidades características do câncer, esse modelo inclui: sinalização mitogênica sustentada, evasão dos supressores de proliferação, resistência à morte celular, imortalidade replicativa celular, indução da angiogênese, capacidade ativa de invasão e metástase. Em 2011, esses mesmos autores adicionaram outras duas habilidades emergentes: a reprogramação do metabolismo energético celular e a evasão do sistema imune (Figura 1) (Hanahan & Weinberg, 2000; Hanahan & Weinberg, 2011).

Figura 1 - Marcadores de câncer



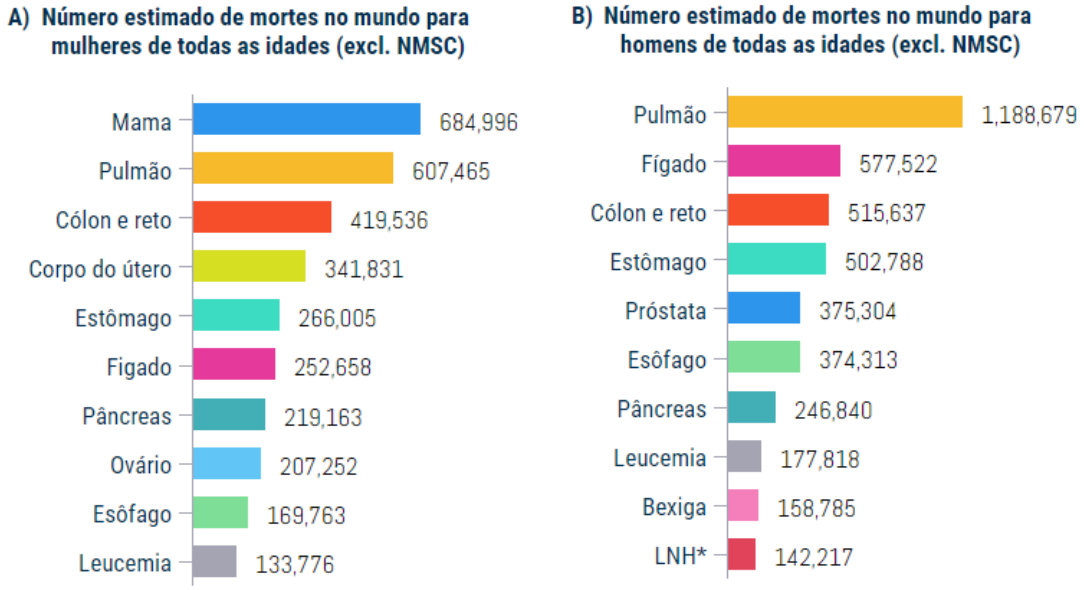
Nota: A maioria das células tumorais adquire um conjunto de habilidades funcionais, que favorecem seu desenvolvimento, escapando dos mecanismos compensatórios, que regulam diversas vias metabólicas que mantêm a homeostase celular. Essas novas habilidades celulares podem ser consideradas marcadores, pois são específicas das células tumorais.

Fonte: Elaborada pela autora, adaptada de Hanahan & Weinberg, 2000 e Hanahan & Weinberg, 2011.

Em termos epidemiológicos, nas últimas décadas, o número de casos de câncer aumentou consideravelmente, tornando-se o maior problema de saúde pública mundial e, atualmente, é uma das principais causas de morte nas Américas (Organização Mundial da Saúde, 2022).

Em2020, quase 20 milhões de novos casos de câncer e 10 milhões de mortes por esta doença foram relatados no mundo. Desses novos casos, cerca de 2,26 milhões (11,7%) referem-se a novos diagnósticos de câncer de mama, seguidos do câncer de pulmão (11,4%), que é representado por 2,21 milhões de novos casos (Ferlay *et al.*, 2021). Neste período, as principais causas de morte no mundo foram por câncer de mama, câncer de pulmão (traqueia e brônquios), câncer colorretal, colo de útero e estômago para a população feminina (Figura 2A).

Figura 2 - Principais causas de morte no mundo em 2020



Nota: As taxas de mortalidade foram estimadas para mulheres (A) e homens (B) de todas as idades, excluindo o câncer de pele não melanoma. (*) Linfoma Não-Hodgkin. *NMSC: Non melanoma skin cancer* (câncer de pele não melanoma).

Fonte: Elaborada pela autora a partir de dados do GLOBOCAN 2020, *Graph Production: IARC* (http://gco.iarc.fr/today) e *World Health Organization*. Acesso em 20/10/2022.

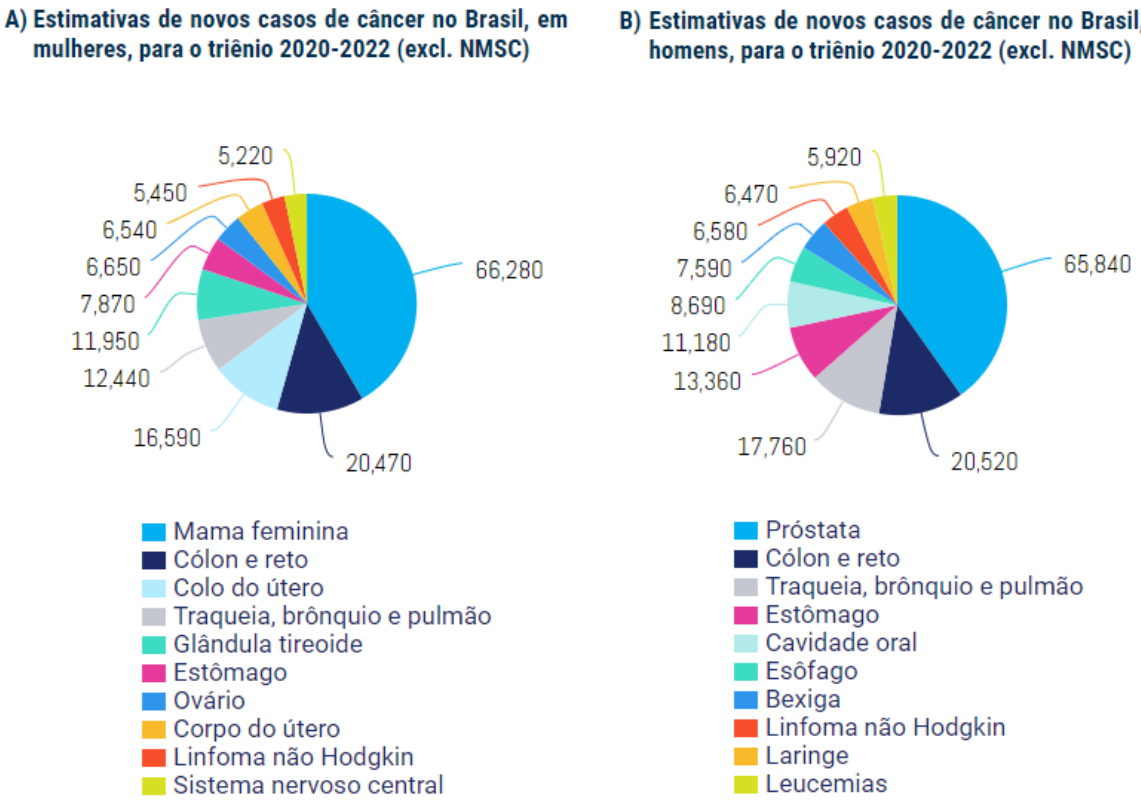
Para o ano de 2040, espera-se que esses números cresçam, atingindo uma projeção de 30,2 milhões de novos casos e 16,3 milhões de óbitos (GLOBOCAN, 2022). Apesar da incidência desses cânceres ser semelhante entre países desenvolvidos e menos desenvolvidos, a taxa de mortalidade é maior nos países mais pobres (INCA, 2022; Bray *et al.*, 2018). Por exemplo, países de baixa e média renda respondem por cerca de 70% das mortes por câncer. Um dos fatores que contribui para a mortalidade mais alta é que apenas 26% desses países de baixa renda usufruem de serviços de patologia no setor público, em contraste com países de alta renda, que mais de 90% fornecem esses serviços (Organização Mundial da Saúde, 2022).

Deste modo, a transição epidemiológica torna o combate ao câncer mais desafiador entre os países de economia emergente (Lima *et al.*, 2020).

No Brasil, o aumento do número de casos tem gerado grande preocupação, pois o crescimento exponencial de idosos possibilita estimar que o aumento da prevalência de câncer vai ampliar a demanda no Sistema Único de Saúde (SUS) para a prestação de cuidados adequados aos doentes (Ministério da Saúde, 2007). Portanto, compreender a situação da doença permite estabelecer prioridades e alocar recursos, que mudem positivamente o cenário da população brasileira com diagnóstico e tratamento mais direcionados (INCA, 2014).

O Instituto Nacional do Câncer (INCA) registrou mais de 224 mil mortes por câncer em 2018 e a expectativa era de que 625 mil novos casos de câncer fossem diagnosticados a cada ano no período 2020-2022 no País (INCA, 2020; Romagna *et al.*, 2022). Para esse período, também foram estimados os principais tipos de câncer com maior incidência na população brasileira, sendo os cânceres de mama, colorretal, colo de útero e pulmão os mais prevalentes na população feminina (Figura 3A). Dos cânceres ginecológicos, o câncer de colo do útero e o câncer de ovário ocupam a quarta e a sétima posição de incidência, respectivamente (INCA, 2022). Apesar dos avanços no manejo clínico de pacientes com câncer de mama, esta doença ainda é a principal causa de morte entre as mulheres (Ministério da Saúde, 2022). Enquanto, entre os homens, predominam os cânceres de próstata, colorretal, pulmão, estômago e cavidade oral (Figura 3B), sendo o câncer de pulmão com maior taxa de mortalidade, seguido pelo câncer de fígado e colorretal (GLOBOCAN, 2020). O câncer de mama também é observado entre os homens, contudo representa apenas 1% do total de casos da doença (Ministério da Saúde, 2022).

Figura 3 - Distribuição proporcional dos dez tipos de câncer mais incidentes no triênio 2020-2022



Nota: As taxas de incidência foram estimadas por sexo, mulheres (A) e homens (B), exceto pele não melanoma. Uma cor foi atribuída a cada tipo de câncer de localização primária, conforme mostrado na legenda ao lado do gráfico de pizza. *NMSC: Non melanoma skin cancer* (câncer de pele não melanoma).

Fonte: Elaborada pela autora a partir de dados do Instituto Nacional de Câncer (INCA, 2020). Data de acesso: 17/10/2022.

Segundo o INCA, a taxa de mortalidade por câncer de mama em relação à população mundial foi superior a 11 mil óbitos/100 mil mulheres, em 2020, com as maiores taxas nas regiões Sudeste e Sul, com taxa acima de 12 mil óbitos/100 mil mulheres, nas duas regiões. Apesar do grande número de óbitos por câncer de mama, o prognóstico é bom, quando tratado adequadamente e em tempo hábil (INCA, 2021, Guerrero *et al.*, 2017). A estimativa de novos casos de câncer de mama diagnosticados no Brasil a cada ano no período 2020-2022 era mais de 65 mil casos (Ministério da Saúde, 2022).

O câncer de ovário possui baixa taxa de incidência e ocupa a oitava posição como o câncer mais comum no mundo (GLOBOCAN, 2020). A maior incidência desta neoplasia é nos países europeus e na América do Norte (Bray *et al.*, 2018; Ferlay *et al.*, 2018). Nos EUA, por

exemplo, o câncer de ovário ocupa o sexto lugar nas taxas de mortalidade e é o segundo câncer ginecológico mais comum, atrás apenas do câncer do colo do útero (GLOBOCAN, 2020). Inclusive, a estimativa de novos casos de câncer de ovário nos EUA para o ano de 2022 é superior a 21 mil e quase 14 mil mortes. No Brasil, eram esperados mais de 6 mil novos casos de câncer de ovário no triênio 2020-2022 (INCA, 2022).

1.1 Câncer de mama

O câncer de mama pode acometer diferentes regiões mamárias, uma vez que a mama feminina é composta por diferentes tipos celulares e tecidos, incluindo lóbulos (glândulas produtoras de leite); ductos lactíferos (pequenos tubos que transportam o leite dos lóbulos até o mamilo); e estroma (tecido adiposo e tecido conjuntivo que circunda os ductos e lóbulos - além de vasos sanguíneos e linfáticos) (Luz *et al.,* 2021). Fenotipicamente, o câncer de mama não é uma doença de etiologia única, apresentando diferentes manifestações clínicas, progressão e resposta terapêutica. Geralmente, o câncer de mama acomete tecidos de origem epitelial e tecidos de origem não epitelial (sarcomas), sendo que o carcinoma de origem epitelial, carcinoma de células epiteliais ductais infiltrantes, carcinoma lobular invasivo e o ductal lobular infiltrante são os tumores malignos mais comuns (Luz *et al.,* 2021).

Considerando a heterogeneidade histológica desta doença, o câncer de mama pode ser classificado de acordo com as características histopatológicas e morfológicas observadas. Características histopatológicas são definidas pelo grau de diferenciação do tecido tumoral, no qual os tumores grau I são mais diferenciados com bom prognóstico; tumores grau II são moderadamente diferenciados, apresenta características intermediárias ao grau I e III; e os tumores grau III são pouco diferenciados e com prognóstico ruim. Em relação às características morfológicas, os cânceres de mama se subdividem em três tipos: formação de túbulos, pleomorfismo nuclear e índice mitótico (Boasquevisque, 2020). Imuno-histoquimicamente, o câncer de mama pode ser categorizado em quatro subtipos moleculares primários: câncer luminal A e B, subtipo basal-like (BCL)/triplo-negativo e receptor 2 do fator de crescimento epidérmico humano (HER2+) (Park *et al.,* 2016). Neste contexto, a detecção desses diferentes tipos de câncer de mama é mediada através da detecção de marcadores imuno-histoquímicos clássicos, tais como receptores de estrógenos (ER), receptores de progesterona (PR) e HER2, além da avaliação das características clínico-patológicas dos tumores, como o tamanho, o grau de diferenciação e o envolvimento linfonodal (Spitale *et al.,* 2009; Dai *et al.,* 2015).

Os tumores do tipo luminal são subdivididos em luminal A e B, apresentando perfis semelhantes (Makki *et al.,* 2015). O luminal A, que é o subtipo mais comum, é caracterizado pela alta expressão de genes relacionados ao ER e PR e baixa expressão de HER2 e marcadores relacionados à proliferação (Ki67 <20%) (Inic *et al.,* 2014). Em contraste, o luminal B compreende baixa expressão de ER e PR, expressão variável de genes relacionados a HER2- e alta expressão de genes relacionados à proliferação (ER/PR + < 20%, HER2−, Ki67 + ≥ 20%). O tumor luminal B, embora tenha um bom prognóstico, normalmente é um tumor mais volumoso e pouco diferenciado quando comparado ao luminal A (Hashmi *et al.,* 2018; Yin *et al.,* 2020). Tanto os tumores classificados como luminal A e luminal B, por expressarem o ER, podem ser tratados com terapia endócrina (Inic *et al.,* 2014). O subtipo basal-like (BCL) está presente em cerca de 10-20% dos casos de câncer de mama e é classificado como triplo-negativo (TNBC) porque não expressa ER, PR e HER2. O tratamento dos TNBCs limitam-se à quimioterapia sistêmica à base de platina (Huo *et al.,* 2018). O subtipo ER/PR-; HER2+ representa cerca de 15-20% dos cânceres de mama. Por ser caracterizado por altos níveis de expressão de genes HER2, podem ser tratados com agentes anti-HER2 (Zielińska *et al.,* 2020).

De fato, os tumores luminais A e B são menos agressivos e, portanto, apresentam melhor prognóstico quando comparados aos tumores basais ou Her2+. Inclusive, o subtipo BCL tem sido associado à presença de variantes em *BRCA1* e prognóstico ruim, devido ao comportamento agressivo do tumor, falta de resposta às terapias endócrinas usuais, e consequentemente, menor sobrevida das pacientes (Garrido-Castro *et al.,* 2019). Vale ressaltar ainda que, os carcinomas ductal e lobular se enquadram em qualquer um desses subtipos moleculares, enquanto os tipos histológicos raros (como o carcinoma tubular e metaplásico) são tipicamente classificados em apenas um dos subtipos moleculares (luminal A ou B) (Dieci *et al.,* 2014; Dai *et al.,* 2015). A melhor compreensão desses subtipos garante um direcionamento do tratamento mais adequado à paciente (Luz *et al.,* 2021).

Os fatores de risco são os maiores contribuintes para o desenvolvimento e progressão do câncer de mama. Dentre eles, a idade acima de 50 anos é o fator mais importante. Contudo, fatores como menopausa tardia (fatores de história reprodutiva e hormonal), obesidade, sedentarismo e exposição frequente à radiação ionizante (fatores ambientais e comportamentais) têm se mostrado fortes contribuintes para a tumorigênese (Guerrero *et al.,* 2017). Além disso, a hereditariedade é outro fator cada vez mais importante para a incidência do câncer de mama. Cerca de 15% dos casos de câncer de mama hereditário são decorrentes

da segregação de variantes germinativas raras em genes de suscetibilidade de alta penetrância como o *BRCA1, BRCA2, TP53, PTEN, STK11* e *CDH1*. O Risco Relativo (RR) para estes genes é > 10. Por outro lado, os genes *ATM, CHEK2, BARD1, NBN, NF1, RAD51C, RAD51D* e *PALB2* (sugeridos por alguns como sendo de alta penetrância), *BRIP1* são genes de penetrância moderada (ou baixa), responsáveis por cerca de 8% do câncer de mama hereditário, podendo apresentar um RR de ≤ 10. Já os genes *FGFR2, CASP8, MAP3K1, LSP1* e outros são considerados como genes de baixa penetrância, representando cerca de 15% dos casos de câncer de mama hereditário e um RR < 3 (Moslehi *et al.,* 2020).

Alterações genéticas nesses genes fomentam o desenvolvimento da síndrome hereditária do câncer de mama e ovário (HBOC) (Lucci-Cordisco *et al.,* 2022, Tsaousis *et al.,* 2019, Zheng *et al.,* 2020). A HBOC é uma síndrome autossômica dominante, cujos indivíduos acometidos tendem a desenvolver câncer em idade precoce e, na maioria dos casos, de alto grau, sendo submetidos a medidas diagnósticas e terapêuticas que reduzem o risco, como ressonância magnética (geralmente acima de 30 a 40 anos), quimioprevenção e/ou cirurgia profilática. Essas medidas reduziram consideravelmente a morbidade e mortalidade associadas ao HBOC (Kraemer *et al.,* 2019).

Na maioria dos casos, as variantes patogênicas em ambos os genes (*BRCA1* e *BRCA2*) estão principalmente associadas ao câncer de adenocarcinoma ductal invasivo. Os cânceres de mama com apenas *BRCA1* mutado são classificados como basais, enquanto que os câncer de mama com apenas *BRCA2* mutado são classificados principalmente como subtipo luminal B, devido à superexpressão do ER, PR e citoqueratinas CK8 e CK18 (proteínas de filamento intermediário que são expressas em muitos órgão epiteliais, incluindo a mama) (Aiad *et al.,* 2014). Variantes patogênicas em *RAD51C* estão geralmente associadas ao carcinoma ductal invasivo de grau intermediário, que pertence ao subtipo luminal A, enquanto que os câncer de mama associados ao gene *CHECK2* pertencem ao subtipo luminal B (Aiad *et al.,* 2014). Normalmente, genes de suscetibilidade ao câncer de mama atuam principalmente em vias de recombinação homóloga, reparo de despareamento e genes causadores da Anemia de Fanconi.

Os cânceres de mamas hereditários e esporádicos, embora semelhantes nos fenótipos, na maioria das vezes, apresentam características marcantes no acometimento, que são consideradas para a classificação clínica da origem do câncer de mama. Nos casos hereditários, os indivíduos afetados costumam manifestar um fenótipo precoce (<35 anos) (Figueiredo, 2014). Outra característica específica do câncer de mama hereditário é o aumento

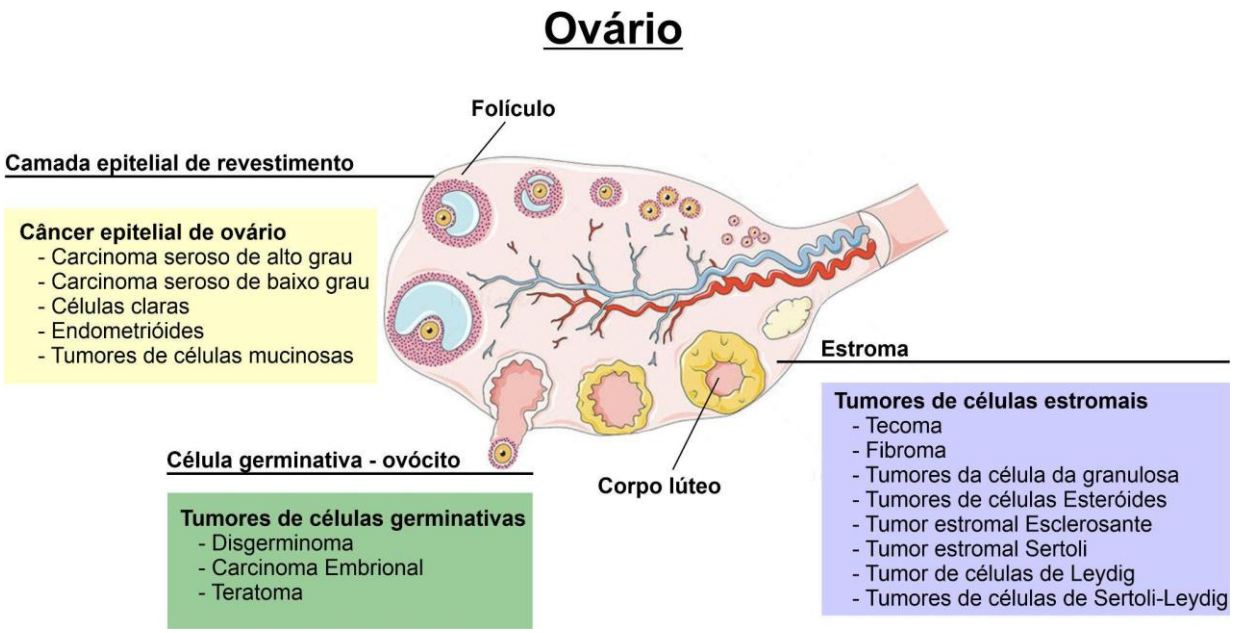
da prevalência do carcinoma de mama bilateral sincrônico (CMBS) (presença simultânea de dois tumores primários), e a associação do câncer de mama com outros tipos de tumor, como câncer de ovário e o colorretal (Figueiredo, 2014).

1.2 Câncer de ovário

Dentre os cânceres ginecológicos, o câncer de ovário é o mais letal, devido principalmente à ausência de sintomas nos estágios iniciais e, consequentemente, diagnóstico em estágio avançado da doença (III-IV). Neste estádio, medidas terapêuticas, como citorredução e quimioterapia sistêmica, são parcialmente eficientes, e cerca de 75% das pacientes acabam apresentando recidiva tumoral nos linfonodos pélvicos/retroperitoneais e peritônio. No entanto, quando diagnosticado precocemente, a paciente tem uma sobrevida superior a 90% e a cirurgia costuma ser o único tratamento necessário (Esteves *et al.*, 2018).

Histologicamente, o câncer de ovário é uma neoplasia com grande heterogeneidade. Cerca de 95% dos cânceres de ovário são carcinomas formados a partir de células epiteliais, que revestem o ovário, e o restante dos cânceres é formado por tumores de células germinativas e células estromais, que produzem maioria dos hormônios femininos (Bodelon *et al.*, 2019) (Figura 4). Os carcinossarcomas e os carcinomas de células indiferenciadas são ainda mais agressivos do que os carcinomas epiteliais, embora menos frequentes. Os tumores de células germinativas costumam ser benignos, apresentando bom prognóstico, com 90% de sobrevida dos pacientes em cinco anos. No entanto, os tumores de células germinativas são fatais, quando se tornam malignos. São vários tipos de tumores de células germinativas, incluindo os teratomas, disgerminomas, tumores do seio endodérmico e coriocarcinomas. Contrário aos tumores de células germinativas, os tumores estromais cancerosos apresentam um bom prognóstico por serem diagnosticados ainda no estágio inicial. Os tumores de células da granulosa (o tipo mais comum), tumores da granulosa-teca e tumores de células de Sertoli-Leydig (raramente encontradas em cânceres de ovário, Castro *et al.*, 2019) são exemplos de tumores estromais malignos. Já, os tumores estromais benignos, incluem os tecomas e fibromas. Além disso, os tumores de células estromais são geralmente encontrados em meninas (cerca de 5%) ou em mulheres com mais de 50 anos (Yoshida *et al.*, 2021) (Figura 4).

Figura 4 - Origem de diferentes tipos de câncer de ovário



Nota: Células de Leydig e Sertoli são encontradas raramente em tumores ovarianos, onde representam cerca de 0,5% dos tumores.

Fonte: Elaborada pela autora, adaptada de Chen *et al.*, 2003 e Algeciras-Schimnich, 2013, usando llllline, 2022.

Molecularmente, os carcinomas epiteliais de ovário são divididos em cinco diferentes tipos: tumores serosos de alto e baixo grau (responsáveis por 60%-70% dos cânceres de ovário epitelial maligno), tumores de células mucinosas (10%), endometrióides (15%) e células claras (5%) (Kaku *et al.*, 2003; Bodelon *et al.*, 2019). Os carcinomas de alto grau têm desfechos clínicos diferentes devido à existência de diferentes tipos moleculares como o imunorreativo, o diferenciado, o proliferativo e o mesenquimal.

Para simplificar, foi proposta ainda uma classificação molecular para carcinomas de ovário como Tipo I e II (Kurman e Shih, 2010). Os tumores do tipo I incluem carcinomas serosos de baixo grau (LGSCs), carcinomas endometrióides, carcinomas de células claras, carcinomas seromucinosos e mucinosos e tumores de Brenner, nos quais são menos frequentes que os tumores do tipo II e geralmente são diagnosticados em estágio inicial. Enquanto os tumores do tipo II incluem carcinomas serosos de alto grau (HGSCs), carcinossarcoma (carcinoma + sarcoma) e carcinoma de células indiferenciadas, nos quais são mais frequentes e diagnosticados tardiamente, quando são mais agressivos e evoluem rapidamente. Consequentemente, os tumores tipo I têm melhor prognóstico quando comparados aos tumores tipo II, que estão associados a 90% dos óbitos (Kaku *et al.*, 2003; JosahKian *et al.*,

2016; Bodelon *et al.,* 2019). Geralmente, os tumores do tipo I apresentam mutações específicas nos genes *KRAS, BRAF* e *ERBB2*, sendo tais mutações presentes em aproximadamente dois terços dos carcinomas serosos de baixo grau. Já mutações em *TP53* são raras em tumores do tipo I. Enquanto que os carcinomas Tipo II resultam de variantes patogênicas presentes no gene *TP53*, estando associados a alterações na via de reparo do DNA por recombinação homóloga ou à amplificação do *CCNE1*. Como esperado, tumores do tipo II geralmente, são diagnosticados quando a doença está em estágio avançado (Shih *et al.*, 2021).

Exames físicos ou moleculares viabilizam o rastreio precoce do câncer de ovário, entretanto, somente três exames são eficientes no diagnóstico e, portanto, frequentemente usados na prática clínica: exame pélvico completo, ultrassonografia transvaginal e dosagem sérica de CA-125 (antígeno de câncer 125). Normalmente, as células ovarianas afetadas expressam altos níveis de CA-125, podendo atuar como um marcador tumoral antigênico no diagnóstico de câncer de ovário (Gandhi & Bhatt, 2022). A sensibilidade e especificidade desse marcador na detecção do câncer de ovário pode variar conforme o estágio da doença e o tipo histológico do tumor. Para o estágio I, a sensibilidade está entre 27 e 66%, para o estágio II está entre 65 e 100%, para o estágio III entre 87 e 90% e para o estágio IV em torno de 94%. A especificidade, por outro lado, é menor em mulheres na pré-menopausa, tornando-se mais específica em mulheres na pós-menopausa (Reis, 2005; Gandhi & Bhatt, 2022).

Existem alguns fatores de risco, que podem estar associados à incidência do câncer de ovário, como por exemplo, a idade acima de 60 anos, a nuliparidade, o atraso na idade fértil ou a menarca precoce, o atraso na menopausa e/ou no histórico familiar. Os principais fatores de risco para o desenvolvimento do câncer de ovário são hormonais e reprodutivos, porque são responsáveis pela manutenção e função ovariana, respectivamente. Inclusive, estudos apontam que a síndrome do ovário policístico pode contribuir para a tumorigênese, porque estimula a desregulação hormonal (Zore *et al.,* 2017). Outros fatores de risco têm efeito moderado, como ambiente e estilo de vida (obesidade e tabagismo) (Friedenreich *et al.,* 2021). A história familiar também é um fator de risco significativo para câncer de ovário, pois ter um parente de 1º grau com câncer endometrial, câncer de ovário isolado ou associado ao de mama, ou ainda de cólon, aumenta em cerca de 85-95% o risco de câncer de ovário (American Cancer Society, 2022). Estima-se que os fatores genéticos e hereditários contribuem cerca de 15-23% do risco de desenvolvimento de tumor e vários genes já foram implicados na predisposição ao câncer de ovário. O *BRCA1* e *BRCA2* foram pioneiros na

hereditariedade ao câncer de ovário, responsáveis por cerca de 25% do risco familiar. Variantes patogênicas em *BRCA1* levam a um risco estimado de câncer de ovário ao longo da vida entre 35% e 70% e, em *BRCA2*, entre 10% e 30% (American Cancer Society, 2022). Outros genes e síndromes, embora de penetrância baixa ou moderada, como *RAD51C, RAD51D, PALB2, BRIP1, BARD1, TP53* e Síndrome de Lynch, também se associam a um risco significativamente aumentado de câncer de ovário. Por exemplo, a Síndrome de Lynch é responsável por 10 a 15% de todos os cânceres de ovário hereditários, com risco cumulativo de câncer de ovário estimado ao longo da vida de 6-10% em portadores comprovados ou prováveis das mutações *MSH2* e *MLH1*, com idade média de início de 51 anos em famílias associadas a mutações *MLH1* e 45 anos em famílias associadas a mutações *MSH2* (American Cancer Society, 2022; Toss *et al.*, 2015).

Em contrapartida, a paridade, a amamentação, uso de anticoncepcionais orais, a ooforectomia profilática (remoção de ovários) e a laqueadura tubária são considerados como fatores de proteção, pois levam a inibição da função ovariana (Elshami *et al.*, 2022). Inclusive, os contraceptivos orais representam uma redução de risco em 30%, sendo o único método de quimioprevenção. Em relação às portadoras das variantes genéticas patogênicas *BRCA1* e *BRCA2*, o uso de anticoncepcionais orais foi associado a uma redução de risco de 20%, por até 3 anos, ou de até 60%, por 6 anos ou mais (Kathawala *et al.*, 2018). A salpingorrafia (popularmente chamada de ligadura de trompas) reduz em cerca de 30% o risco para câncer de ovário epitelial, teoricamente por interferir em fluxo sanguíneo e acesso de fatores carcinogênicos ao ovário pela via ascendente (Wang *et al.*, 2016).

1.3 Câncer e instabilidade genômica

A marca registrada do desenvolvimento do câncer é a instabilidade genômica, que ocorre por meio de fatores intrínsecos ou extrínsecos, que interferem na conformação do genoma (Sullivan & Bernstein, 2018). Os fatores intrínsecos podem ser observados durante a manutenção genômica, quando, por exemplo, ocorrem erros na incorporação de nucleotídeos, estresse replicativo e até mesmo ataque de substâncias químicas endógenas. Por outro lado, os fatores extrínsecos incluem exposição à luz solar, radiação ionizante e muitos produtos químicos encontrados no ambiente, alimentos e água (Sullivan & Bernstein, 2018; Wilson III *et al.*, 2021). Esses fatores desencadeiam uma variedade de danos genéticos, podendo promover quebras cromossômicas, translocações, rearranjos, amplificação gênica e deleções. Algumas dessas lesões são mais complexas e prejudiciais ao bom funcionamento das células (Senter *et al.*, 2022).

Existem muitos mecanismos de preservação do genoma que atuam coletivamente na manutenção da integridade genômica, criando um sistema global chamado resposta ao dano do DNA (do Inglês, *DNA damage response - DDR*) (Catalano *et al.*, 2022, Nathans *et al.*, 2021; Wilson III *et al.*, 2021). O DDR é composto por várias etapas, que acontecem durante o ciclo celular ou fora deste. Além disso, os diferentes mecanismos vão ter diferentes componentes moleculares e eficiência. Durante o ciclo celular, os checkpoints permitem a avaliação da integridade do DNA (de outras moléculas e organelas), seguido de tomada de decisão sobre o destino da célula, ou seja, se o ciclo prossegue ou se a célula é encaminhada para apoptose ou senescência celular.

Os *checkpoints* levam à interrupção do ciclo celular para corrigir danos por meio de mecanismos de sinalização como o complexo enzimático MRN (Mre11/Rad50/Nbs1), que recruta proteínas específicas, como serina/treonina proteína quinases (ATM e ATR). Nessa etapa, pequenas inserções e deleções são reparadas principalmente pela via do reparo de despareamento (do Inglês, *mismatch repair - MMR*), que ocorre imediatamente após a replicação. Já em casos de alterações como oxidação, desaminação e alquilação, a principal via ativada é o reparo por excisão de base (do Inglês, *base excision repair - BER*) (Vogelstein & Kinzler, 2004). Porém, se o dano for irreparável, a célula estimula o processo de apoptose (morte celular programada) ou senescência celular (crescimento interrompido), evitando a instabilidade genética (Wilson III *et al.*, 2021; Nathans *et al.*, 2021).

Eventos mutagênicos podem resultar em quebras de fita simples de DNA (do Inglês *single-strand DNA breaks - SSBs*) e/ou quebras de fita dupla de DNA (do Inglês *double-strand breaks - DBS*) (citotóxicas), que comumente são reparadas por duas vias principais: recombinação homóloga (do Inglês *homologous recombination - HR*) e junção de extremidades não homólogas (do Inglês, *non-homologous end joining, NHEJ*) (Caracciolo *et al.*, 2021). Resumidamente, a via HR inicia-se com proteínas sensoras, ATM e complexo MRN, que reconhecem danos no DNA, permitindo a fosforilação da histona H2AX. Isso permite o recrutamento das proteínas NBS1, TP53, CHEK2 e BRCA1 em locais específicos para mediar a ativação adequada do *checkpoint* e facilitar a parada do ciclo celular (Moshous *et al.*, 2001; Andres *et al.*, 2012; Rosa, 2016). Por outro lado, o complexo MRN inicia o processamento final através da ressecção de nucleotídeos da fita na direção 5'-3', gerando a 3'ssDNA (Somyajit *et al.*, 2010). Com a extremidade 3' livre, as proteínas RAD54 e RPA são recrutadas e se ligam a fita simples, impedindo a formação de estruturas secundárias na região 3' terminal, bem como a degradação enzimática (Figura 5A).

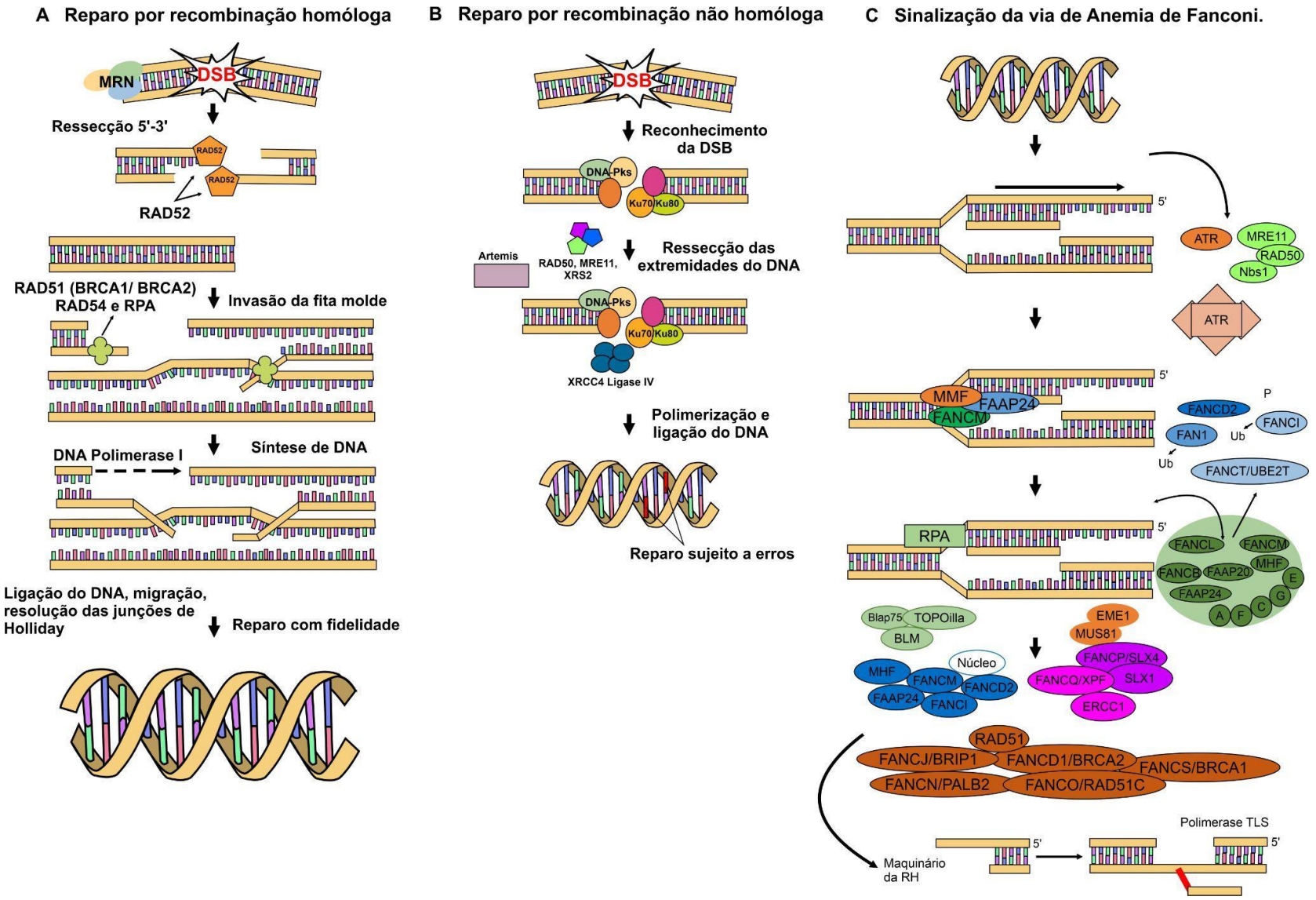
Simultaneamente, a BRCA2 recruta a recombinase RAD51, que se liga às proteínas mediadoras RAD51B, RAD51C, RAD51D, XRCC2 e XRCC3, ligando-as a quebra de fita dupla para formar uma fita de nucleoproteína. Essas proteínas catalisam a invasão da fita e o emparelhamento com a molécula de DNA homóloga intacta para formar DNA heteroduplex (D-loop) (Oliveira, 2016).

Em seguida, a DNA polimerase insere nucleotídeos na região 3'terminal da fita danificada (Rosa, 2016). Uma vez que a etapa de extensão da fita está completa, a DNA ligase funciona unindo a fita recém-sintetizada, permitindo a formação de junções chamadas "junções de Holliday", que precisam ser clivadas para reparo completo (Rosa, 2016). A clivagem dessas junções ou resolução das junções de Holliday é mediada pela atividade das endonucleases Mus81 e Mus4 e RAD51C, que clivam essas regiões e ligam as fitas para a restauração completa da molécula de DNA. O resultado final da resolução dessas junções pode ou não ser a troca de fitas (*crossing-over*), principalmente durante o processo de recombinação meiótica (Somyajit *et al.,* 2010; Rosa, 2016). Vale ressaltar, que na via de HR a cromátide irmã homóloga é usada como molde para alinhar as pontas quebradas e inserir as informações que faltam. Um detalhe importante da HR é que ela também é uma via essencial na recuperação de bifurcações de replicação paralisadas (Moshous *et al.,* 2001; Andres *et al.,* 2012; Rosa, 2016).

Já o reparo pela via NHEJ tem início com a atividade do complexo proteico dimérico KU (Ku70-Ku80), que se liga às extremidades de quebra de fita dupla do DNA e recruta a proteína quinase dependente de DNA (DNA-PKcs) (Fell & Schild-Poulter, 2012). Em seguida, a proteína de reparo Artemis (proteína nuclear) interage com a subunidade catalítica da DNA-PKcs, formando um complexo Artemis-DNA-PKcs, que cliva as extremidades do DNA danificado, preparando-o para a ligação das extremidades. Simultaneamente, a XRCC4 (proteína de reparo de DNA) interage com o fator XRCC4-like (XLF), formando complexo XRCC4-XLF, que permite que a Ligase IV (LigIV) interaja com o DNA, gerando uma ponte de DNA (Figura 5B) (Moshous *et al.,* 2001; Andres *et al.,* 2012). Nesta via, as bases nitrogenadas podem ser adicionadas ou perdidas, o que geralmente se manifesta como variantes de sentido trocado, de mudança de fase ou de inserções (Andres *et al.,* 2012).

Por fim, lesões de ligações cruzadas de DNA (do Inglês *Interstrand DNA crosslinks - ICLs*) que, semelhantes aos *DSBs*, requerem a atividade simultânea de duas ou mais vias reparadoras, como ativação das proteínas da via de Anemia de Fanconi (Figura 5C), reparo por excisão de nucleotídeos, HR e síntese de translesão (TLS) (Rosa, 2016).

Figura 5 - Mecanismos de reparo de danos no DNA



Fonte: desenvolvida pela autora adaptado de Rosa, 2016

1.4 Testes genéticos e bioinformática na medicina genômica

Os testes genéticos utilizados na prática clínica desempenham um papel crítico na medicina personalizada, garantindo tratamento de precisão e prevenção de doenças (Mersch *et al.,* 2018). Embora o método tradicional de Sequenciamento de Sanger seja frequentemente utilizado na investigação de 'mutações pontuais' e de pequena escala, o Sequenciamento de Próxima Geração (do Inglês *Next Generation Sequencing - NGS*) estende essa abordagem, proporcionando maior capacidade de geração de informações (milhões de leituras curtas em paralelo) e com uma saída de sequenciamento detectada automaticamente sem a necessidade de eletroforese, o que o torna essencial na área clínica e de pesquisa (Slatko *et al.,* 2018; Kchouk *et al.,* 2017). Hoje, as metodologias de NGS permitem a avaliação de sequenciamento completo (do Inglês, *whole genome sequencing - WGS*), sequenciamento completo de exoma (do Inglês, *whole exome sequencing - WES*) (restrito à parte codificadora de proteínas e parte dos RNAs do genoma) e sequenciamento focado (target sequencing) de regiões selecionadas do genoma (painéis de genes específicos) (Kumar *et al.,* 2019). A principal vantagem da utilização do WGS é a grande quantidade de regiões que o teste abrange, incluindo regiões não codificantes, o que vem se tornando cada vez mais importante na área clínica devido à grande quantidade de variação de pequena e grande escala, que podem ser a resposta para doenças desconhecidas ou não resolvidas (Kumar *et al.,* 2019). Entretanto, a cobertura em profundidade e em extensão requerida numa triagem de variantes genéticas ainda é muito cara, de forma que o WGS ainda tem seu uso limitado à situações específicas de pesquisa. Painéis multigênicos e WES, embora não tão completos quanto o WGS, têm sido amplamente utilizados na prática clínica, por solidificar o diagnóstico de condições genéticas de predisposição à doenças ou mesmo o diagnóstico de doenças hereditárias (Tsaousis *et al.,* 2019). Entretanto, estes testes, tipicamente, não cobrem os introns inteiros, nem regiões regulatórias anteriores ao promotor, o que pode ocasionar resultados falso-negativos.

Esses testes genéticos por NGS permitem a triagem de variantes classificadas como: *missense* (sentido trocado), *nonsense* (códon de parada prematuro), sinônimas (ou silenciosas, que não alteram o aminoácido nem a sequência), *frameshift* (mudança de fase de leitura), e variantes em regiões promotoras, regulatórias ou intrônicas (Scheid, 2018). A interpretação da patogenicidade dessas variantes genéticas é feita utilizando-se os critérios estabelecidos conjuntamente pela *American College of Medical Genetics and Genomics* (ACMG) e pela *Association for Molecular Pathology* dos EUA (AMP), doravante referidos como critérios

ACMG/AMP. Eles foram publicados pela primeira vez em 2013, contendo diretrizes sobre a classificação de variantes para laboratórios que realizam testes genéticos, sendo sujeitos a revisões constantes (Rehder *et al.*, 2021). Essas diretrizes incluem critérios de benignidade e patogenicidade baseados em: frequência populacional, estudos experimentais, tipo de variantes, análise computacional (predição de impacto funcional e conservação evolutiva) e evidências clínicas (Quadro 1) (Rehder *et al.*, 2021). Variantes de sentido trocado são muito frequentes e, na maioria das vezes, apresentam conflito na interpretação da patogenicidade, resultando na classificação de variante de significado incerto (VUS), quando não é possível obter informações suficientes para classificá-la como benigna (B) ou patogênica (P) (Quadro 2).

Quadro 1 - Descrição da classificação das variantes e nível de patogenicidade (P) e benignidade (B) de acordo com ACMG/AMP

Critério	Descrição	Interpretação
PVS1	Variante nula (*nonsense*, *frameshift*, canônico ± 1 ou 2 locais de *splice*, códon de iniciação, deleção única ou multiéxon) em um gene, no qual a *LOF* é um mecanismo conhecido de doença	Evidência muito forte ou forte de patogenicidade
PS1	Mesma mudança de aminoácido, que uma variante patogênica previamente estabelecida, independentemente da mudança de nucleotídeo	
PS2	*De Novo* confirmado	
PS3	Estudos funcionais *in vitro* ou *in vivo* bem estabelecidos, que apoiam um efeito prejudicial no gene ou produto gênico	
PS4	Prevalência em pessoas afetadas maior do que nos controles	
PM1	Localizado em uma região *hotspot* mutacional ou domínio funcional crítico e bem estabelecido (por exemplo, sítio ativo de uma enzima) sem variantes benignas frequentes	Evidência moderada de patogenicidade
PM2	Ausente do populacional, ou com baixa frequência em bancos de dados populacionais	
PM3	Para recessivos - detectável com uma mutação patogênica em trans	
PM4	O comprimento da proteína muda como resultado de deleções/inserções *in-frame* em uma região não repetida ou variantes de *stop-loss*	
PM5	*Missense* que altera um resíduo de aminoácido, no qual uma *missense* diferente é previamente estabelecida como patogênica	
PM6	*De Novo* não confirmado	
PP1	Variante segrega com doença em múltiplos membros da família afetados	Evidência de suporte de

PP2	Variante *missense* em um gene que tem uma baixa taxa de variantes *missense* benigna e em que as variantes *missense* são um mecanismo comum de doença	patogenicidade
PP3	Múltiplas linhas de evidência computacional com um efeito deletério sobre o gene ou produto do gene (conservação, evolução, impacto no *splicing*, etc.)	
PP4	O fenótipo do paciente ou história familiar é altamente específico para uma doença com um único gene etiologia	
PP5	Fonte confiável relatou recentemente a variante como patogênica, mas a evidência não está disponível para o laboratório realizar uma avaliação independente.	
BP1	Variante *missense* em um gene, sendo que, principalmente, as variantes truncadas são conhecidas por causar doenças	Evidência forte de benignidade
BP2	Observado em trans com uma variante patogênica para um gene/distúrbio dominante totalmente penetrante ou observado em cis com uma variante patogênica em qualquer padrão de herança	
BP3	Deleções/inserções *in-frame* em uma região repetitiva sem uma função conhecida	
BP4	Múltiplas linhas de evidência computacional não sugerem qualquer impacto no gene ou produto do gene (conservação, evolução, impacto de *splicing*, dentre outras)	
BP5	Causa alternativa (variante encontrada em um caso com uma base molecular alternativa para doença)	
BP6	Fonte confiável relatou recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente	

BP7	Uma variante sinônima (silenciosa) para a qual algoritmos de previsão de *splicing* não preveem nenhum impacto na sequência de consenso de *splice* nem a criação de um novo sítio de *splice*. E o nucleotídeo não é altamente conservado	
BS1	Alta frequência (comum)	Evidência de suporte de benignidade
BS2	Presente no controle saudável	
BS3	Sem consequência funcional	
BS4	Falta de segregação nos membros afetados de uma família	
BA1	Muito comum (frequência do alelo é> 5%)	

Nota: LOF - Loss of function. Do Inglês, perda de função

Quadro 2 - Regras do ACMG/AMP para definição da interpretação das variantes como patogênica (P) e provavelmente patogênica (PP)

Variante patogênica	**Variante provavelmente patogênica**
1- 1 Muito fortemente (PVS1) e *a) ≥ 1 fortemente (PS1-PS4) ou* *b) ≥2 moderada (PM1-PM6) ou* *c) ≥1 moderada (PM1-PM6) e 1 a favor de patogenicidade (PP1-PP5) ou* *d) 2 a favor de patogenicidade (PP1-PP5)* *2- ≥2 fortemente (PS1-PS4) ou* *3- 1 fortemente (PS1-PS4) e* *a) ≥3 moderada (PM1-PM6)* *b) 2 moderada (PM1-PM6) e ≥2 a favor de patogenicidade (PP1-PP5) ou* *c) 1 moderada (PM1-PM6) e ≥ 4 a favor de patogenicidade (PP1-PP5)*	*1- 1 muito fortemente (PVS1) e 1 moderada (PM1-PM6)* *2- 1 forte (PS1-PS4) e ≥1 moderadas (PM1-PM6)* *3- 1 forte (PS1-PS4) e 2 a favor de patogenicidade (PP1-PP5) ou* *4- ≥3 moderada (PM1-PM6) ou* *5- 2 moderadas (PM1-PM6) e ≥2 a favor de patogenicidade (PP1-PP5) ou* *6- 1 moderada (PM1-PM6) e ≥4 a favor de (PP1-PP5)*
Variante benigna	**Variante provavelmente benigna**
1- 1 critério independente (BA1) *ou* 2- *≥2 fortemente (BS1-BS5)*	1- 1 fortemente (BS1-BS4) e 1 a favor de benignidade (BP1-BP7) *ou* 2- *≥ 2 a favor de (BP1-BP7)*

Fonte: adaptado de (Henson & Resta, 2021)

Durante muito tempo, os testes genéticos priorizaram a busca de variantes genéticas nos genes de alta penetrância, *BRCA1* e *BRCA2*, pioneiros na herdabilidade do câncer de mama e ovário. No entanto, como consequência dos avanços no sequenciamento, outros genes envolvidos na via de reparo por recombinação homóloga têm sido associados à suscetibilidade aos cânceres de mama e ovário, como *ATM, CHEK2*, *RAD51C* e *RAD51D* (Alenezi *et al.*, 2022; Lucci-Cordisco et al., 2022).

Um desses genes, o *RAD51D* foi recentemente incluído em painéis multigênicos devido a suas vartes genéticas estarem associadas a um risco aumentado de câncer de mama e ovário. O risco estimado de portadores de variantes patogênicas em *RAD51D* desenvolverem câncer de mama é de 3,1–8,3, enquanto para câncer de ovário é de 6,3 a 12,0, com idade ao diagnóstico superior a 50 anos. Os riscos cumulativos estimados de desenvolvimento de câncer de mama

aos 80 anos para portadores da variante patogênica em *RAD51D* foram de 20% e 13% para câncer de ovário. Além disso, a frequência populacional estimada de uma variante patogênica no gene *RAD51D* é equivalente a 0,00026 (26:100.000) (Yang et al., 2020).

O *RAD51D* é considerado de fato um gene de penetrância moderada, o que dificulta o seu uso no manejo clínico devido a dados limitados, como terapias direcionadas e eficácia da quimioterapia e status mutacional de genes de risco não *BRCA1* e *BRCA2*. Assim, estudos sobre *RAD51D* são ainda necessários para entender os mecanismos de reparo envolvidos, além de contribuir para o entendimento de como defeitos nesses mecanismos de reparo podem afetar a estabilidade genômica e desenvolver doenças como o câncer (Garcin *et al.,* 2019; Berti *et al.,* 2020). Esses estudos auxiliariam no aconselhamento genético de indivíduos afetados e não afetados e orientariam as opções de gerenciamento de risco para esses indivíduos (Graffeo *et al.,* 2022).

1.5 Justificativa

A popularização dos testes genéticos por sequenciamento na triagem de variantes genéticas germinativas em pacientes e famílias com câncer de mama e/ou câncer de ovário, bem como para outros tipos de tumores, tem gerado um aumento no número de variantes, incluindo uma grande quantidade de VUS, e na incidência de achados secundários ou incidentais. Isso vem tornando o manejo clínico de pessoas e famílias com riscos de ocorrência ou recorrência familial de anomalias genéticas mais desafiador para os profissionais (Berliner *et al.,* 2021; Tsaousis *et al.,* 2019; Carmo, 2018).

Devido a insuficiência de dados para classificar assertivamente como patogênicas ou benignas, o número de VUS identificadas é crescente, até mesmo em genes estabelecidos como de alto risco. Entretanto, a classificação de uma variante como VUS não é permanente, podendo ser reclassificada para patogênica ou benigna à medida que novas informações forem encontradas (Tsaousis *et al.,* 2019; Carmo, 2018). Além disso, deve-se atentar aos achados incidentais (ou secundários), que são representados por genótipos e/ou fenótipos que não estejam relacionados ao propósito original do exame, e podem ou não estar associados à significância clínica (Berliner *et al.,* 2021).

Com isso, quanto mais testes genéticos, maiores são as dificuldades enfrentadas pelos médicos para gerenciarem o cuidado do paciente. Embora existam abordagens de classificação de variantes com base nas diretrizes do ACMG/AMP, a interpretação é

complicada devido à ausência de dados clínicos e à falta de padronização na interpretação e divulgação. Consequentemente, existe um esforço periódico de reclassificação de variantes, particularmente daquelas classificadas como VUS. Em um estudo com mais de 1,5 milhão de resultados, avaliando a reclassificação de variantes, foram emitidos cerca de 60 mil laudos de reclassificação. Considerando somente as SNVs, foi encontrado que apenas cerca de 8% (2048 ou 26.670) das VUS puderam ser reclassificadas. Quando a reclassificação foi obtida, cerca de 90% das VUS foram atualizadas para benignas e cerca de 9% para patogênicas, corroborando que a reinterpretação dessas variantes embora essencial, é difícil (Mersch *et al.*, 2018). A reclassificação de VUS para benignas frequentemente reflete o aumento de informação sobre sua frequência em bases de dados populacionais. Com menor frequência, é baseada em estudos funcionais ou em análise de segregação.

Uma questão importante, que tem sido considerada para reduzir as incertezas da classificação de variantes, é a obtenção de mais dados clínicos (incluindo história familiar e relatos de caso), estudos funcionais e populacionais (frequências alélicas em casos e controles), alternativa recomendada pelo ACMG/AMP, mas que é extremamente laboriosa e onerosa. Outra alternativa usualmente adotada, de maneira mais empírica e menos dispendiosa, tem sido a colaboração e comunicação entre grupos médicos, laboratórios e pacientes (Hoffman-Andrews *et al.,* 2018). Neste contexto, as análises *in silico* empregadas na avaliação aprofundada do impacto das VUS em nível de DNA, mRNA e proteína, e na caracterização da patogenicidade dessas variantes, representam uma alternativa relevante para ajudar a resolver esse problema, pode fornecer evidências adicionais para a reclassificação de VUS (US Preventive Services Task Force, 2019; Berliner *et al.,* 2021). Um problema aqui é a falta de padronização das análises desenvolvidas pelos diferentes laboratórios e a não descrição dos resultados, que dão suporte à classificação da variante em algumas das bases de dados. Outro problema são os achados incidentais, que estão presentes em cerca de 1-2% dos indivíduos submetidos a WES (Kraemer *et al.,* 2019).

Considerando relevância da hereditariedade como fator de risco para a incidência do câncer de mama e câncer de ovário e o impacto dessas neoplasias na saúde da mulher, os profissionais que lidam com aconselhamento genético precisam estar cada vez mais capacitados para lidar com as demandas impostas durante a triagem de seus pacientes, ou ter acesso às ferramentas tecnológicas capazes de apoiar essa demanda na prática clínica. A interpretação adequada dos testes genéticos é fundamental para ampliar as possibilidades de o paciente obter um diagnóstico mais assertivo em relação à estratificação de risco, manejo

clínico e identificação de familiares potencialmente em risco por meio de triagem em cascata (Bains *et al.*, 2021). Entretanto, o achado de VUS não autoriza a triagem de outros familiares, afetados ou não, de forma que a principal informação, que poderia ajudar na classificação da variante, vai frequentemente faltar.

Este estudo buscou estudar variantes genéticas de *RAD51D*, que foram identificadas em indivíduos brasileiros portadoras de câncer de mama, câncer de ovário e outras neoplasias. As variantes *RAD51D* passíveis de reclassificação foram repassadas aos médicos assistentes e laboratórios parceiros que re-contatam as pacientes em um serviço abrangente de aconselhamento genético de câncer. O gene *RAD51D* foi incluído em painéis multigênicos mais recentemente, e pouco se sabe sobre o impacto de suas variantes genéticas, bem como quais são mais frequentes no Brasil. Por isso, vamos nos concentrar em algumas questões relevantes para a interpretação adequada de variantes genéticas encontradas em *RAD51D*:

1) Quais características específicas do gene *RAD51D* precisam ser levadas em consideração ao se avaliar a patogenicidade potencial de variantes genéticas?

2) Quais variantes genéticas em *RAD51D* já foram descritas/inseridas em bancos de dados em pacientes com câncer da população brasileira?

3) É possível reclassificar variantes genéticas usando uma estratégia de análise de bioinformação mais abrangente? Avaliar corretamente o impacto funcional das variantes genéticas é a principal estratégia atualmente disponível para definir o cuidado de familiares de pessoas com câncer;

Este projeto de pesquisa está vinculado a um projeto assistencial, que prevê o retorno dos resultados obtidos em termos de reclassificação de variantes genéticas aos médicos assistentes desses pacientes.

1.6 Objetivo Geral

Empregar um conjunto de análises *in silico* e bioinformática estrutural na caracterização de variantes germinativas identificadas no gene *RAD51D* em indivíduos brasileiros portadoras de câncer de mama, câncer de ovário e outras neoplasias, especificando seu impacto a nível funcional e estrutural do gene, transcrito e proteína.

1.7 Objetivos Específicos

1) Avaliar a classificação atual das variantes germinativas encaminhadas para avaliação em bases de dados de câncer públicas;

2) Identificar as características moleculares específicas do gene e da proteína *RAD51D*, que possam afetar a predição do impacto de variantes genéticas (por exemplo, éxons, íntrons, *splicing* e *splicing* alternativo, estrutura secundária e terciária da proteína, domínios funcionais etc.);

3) Determinar o impacto das variantes em nível de DNA, mRNA e proteína;

4) Avaliar a reclassificação das variantes segundo os critérios ACMG/AMP.

2. MATERIAIS E MÉTODOS

2.1 Aprovação do COEP

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais sob os protocolos CAAE.01758418.0.0000.5149 e CAAE.48770621.6.0000.5149.

2.2 Casuística

O primeiro estudo do grupo "*Germinative mutations landscape in Minas Gerais, Brazil patients undergone to genetic counseling for gynecological and breast cancer*" foi desenvolvido por Camila Martins de Carvalho e colaboradores, que realizaram o rastreamento de variantes germinativas em 382 pacientes em aconselhamento genético. Este trabalho foi apresentado em outubro de 2020, em formato de pôster, no Congresso Internacional da Sociedade Brasileira de Oncologia em parceria com a *American Association of Cancer Research* e submetido na revista Brasileira de Ginecologia e Obstetrícia. A partir desse trabalho, foi possível identificar novas variantes genéticas (raras), que despertaram o interesse de colaboradores: Anisse Marques Chami, Leticia da Conceição Braga, Agnaldo Lopes, Henrique Galvão, Instituto Hermes Pardini e Clínica Personal, que propuseram uma parceria com outros profissionais com experiência na análise e interpretação de variantes usando ferramentas de Bioinformática, incluindo a Dra. Prof. Maria Raquel Santos Carvalho (Médica), Dra. Carolina Guimarães Matozinho (Bióloga), a mestranda em genética Thalía

Rodrigues de Souza Zózimo (Biomédica) e a bolsista de iniciação científica Thamiris Matias Alves (graduanda em Ciências Biológicas da UFMG).

Como parte do trabalho de doutorado da Dra. Anisse Marques Chami Ferraz e colaboradores desenvolveram dois estudos de caso e uma proposta de reclassificação de variantes genéticas encontradas em mulheres brasileiras no gene *RAD51C*. Dois estudos já foram submetidos. O primeiro estudo de caso foi intitulado *"Evidence for pathogenicity of MSH2 c.1894_1898del; (p.Ile633Lysfs*9) mutation in Lynch Syndrome: genotype-phenotype correlation and bioinformatic approaches"* que foi submetido em *Hereditary Cancer in Clinical Practice*; o segundo estudo de caso intitulado *"In-deep Bioinformatic characterization of a novel fumarate hydratase mutation (FH c.199T>G; (p. Tyr67Asp) in hereditary leiomyomatosis and renal cell carcinoma"* que foi submetido na revista *Familial Cancer.* Um terceiro estudo vai relatar uma reanálise e reclassificação de variantes nos genes *RAD51C* e *RAD51D* (desenvolvida no presente estudo), está em fase de redação.

Com a evolução desses trabalhos, foi possível a obtenção e interpretação de variantes genéticas em *RAD51D* identificadas em indivíduos brasileiros portadoras de neoplasias de mama e/ou ovário e outras, compartilhadas por médicos/instituições parceiras voltadas para o aconselhamento genético. De modo que, nesta reanálise das variantes genéticas em *RAD51D*, uma série de análises *in silico* foram realizadas nos níveis de DNA, mRNA e proteína na tentativa de elucidar os principais mecanismos de patogenicidade de cada variante em análise. Foram avaliadas as 20 variantes apresentadas na Tabela 1.

Tabela 1 - Variantes germinativas em RAD51D avaliadas no presente estudo

Variantes *RAD51D* rs/ID	**Consequência molecular**	**HGVS***	**Dados clínicos observados nos pacientes**	**Classificação reportada pelo laboratório de referência****
c.26G>C; (p.Cys9Ser) rs140825795	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	Câncer de mama	VUS
c.137C>G; (p.Ser46Cys) rs587780102	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	Câncer de mama, cólon e polipose	VUS
c.140A>G; (p.Tyr47Cys) rs776528342	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	VUS

c.190C>T; (p.Pro64Ser) rs201506572	Variante de sentido trocado	NM_001142571/ ENST00000590016.5	-	Provavelmente benigna
c.196G>A; (p.Val66Met) rs56026142	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna
c.286G>A; (p.Ala96Thr) rs147933658	Variante de sentido trocado	NM_001142571/ ENST00000590016.5	-	Provavelmente benigna
c.345+5A>G rs878854562	Variante intrônica	NM_002878.4/ ENST00000345365.11	-	VUS
c.355T>C; (p.Cys119Arg) rs201313861	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	Linfoma de Hodgkin	VUS
c.413A>G; (p.Asn138Ser) rs201676898	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	VUS
c.481-7G>A rs145832514	Variante intrônica	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna
c.493C>T; (p.Arg165Trp) rs544654228	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	Câncer de pâncreas	VUS
c.628G>A; (p.Ala210Thr) rs80116829	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna
c.666A>G; (p.Glu222=) rs114012742	Variante sinônima	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna
c.694C>T; (p.Arg232Ter) rs587780104	Variante *Frameshift*	NM_002878.4/ ENST00000345365.11	-	Patogênica
c.695G>A; (p.Arg232Gln) rs28363283	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna
c.698A>G; (p.Glu233Gly) rs28363284	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna
c.715C>T; (p.Arg239Trp) rs770250516	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	VUS
c.862G>C; (p.Gly288Arg) rs1391505912	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	VUS
c.919G>A; (p.Glu307Lys) rs115031549	Variante de sentido trocado	NM_002878.4/ ENST00000345365.11	-	Provavelmente benigna

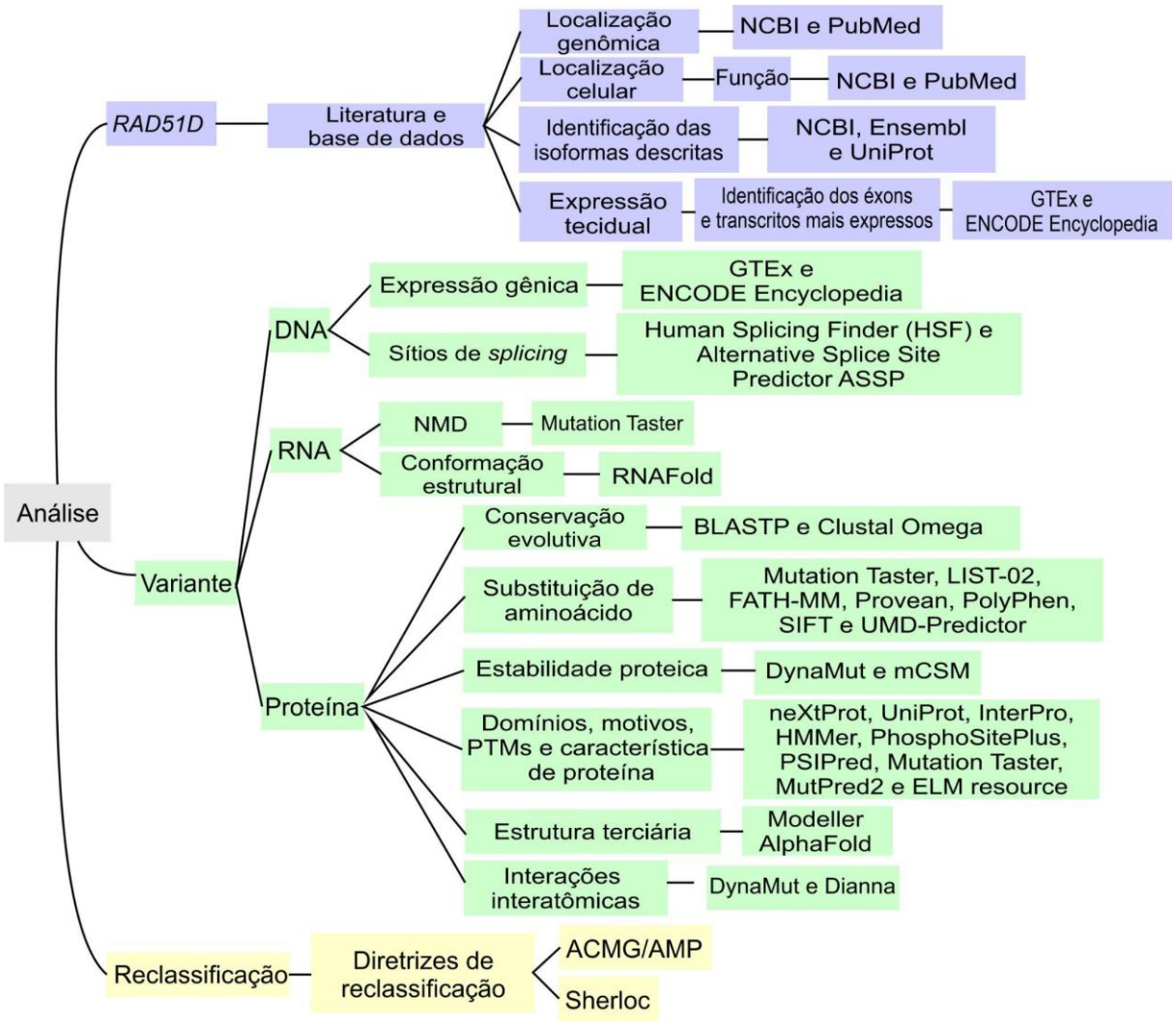
c.973G>A; (p.Gly325Ser) rs587780106	Variante sentido trocado	de NM_002878.4/ ENST00000345365.11	-	VUS

Nota: (*) A nomenclatura (HGVS) das variantes é apresentada conforme o transcrito envolvido. (**) a classificação atual das variantes conforme os laboratórios parceiros. NM_002878.4/ENST00000345365.11 codifica a isoforma canônica da proteína. NM_001142571/ENST00000590016.5 codifica a isoforma 6 da proteína, de acordo com o NCBI, e com a isoforma 8 da proteína, de acordo com outras bases de dados (UniProt, SwissPro). rs/ID: identificação da variante de acordo com dbSNP (https://www.ncbi.nlm.nih.gov/snp/). Neste estudo, optamos por reavaliar todas as variantes encaminhadas pelos laboratórios parceiros, ainda que a classificação não fosse VUS.

2.3 Metodologia de análise *in silico* das variantes

A estratégia adotada neste estudo para a interpretação da patogenicidade dessas variantes foi realizada de acordo com o fluxograma apresentado na Figura 6.

Figura 6 - Etapas e recursos usados na avaliação do impacto funcional das variantes neste estudo



Nota: PTM - *Post-translational modifications*.

Fonte: Elaborada pela autora

2.3.1 Anotação funcional

Buscou-se obter informação sobre as características do gene *RAD51D*, relevantes para a avaliação das variantes genéticas, tais como: a localização genômica, localização celular e identificação dos transcritos e proteínas codificadas pelo *RAD51D* em bases de dados e na literatura. A avaliação da expressão tecidual dos éxons e transcritos foi realizada no GTEx (v8) (GTEx Consortium, 2015) e no ENCODE Encyclopedia (v1) (ENCODE Project Consortium, 2012); a identificação das características das proteínas e informações específicas de domínios foram recuperadas do banco de dados UniProt (Consortium U, 2019), InterPro

(Apweiler *et al.,* 2001), Mutation Taster (Steinhaus *et al.,* 2021), HMMER (v3.3.2) (Potter *et al.,* 2018) e PSI-Pred (v4.0) (Buchan *et al.,* 2019).

2.3.2 Variantes de *RAD51D*

2.3.2.1 Conservação evolutiva

A avaliação da conservação evolutiva de resíduos afetados por variantes de sentido trocado foi desenvolvida usando-se BLASTP (Altschul *et al.,* 1990) e Clustal Omega (Madeira *et al.,* 2022). Todas as sequências foram extraídas dos bancos de dados não redundantes de proteínas disponíveis na ferramenta BLASTP e no NCBI (NCBI Resource Coordinators, 2018). Para o alinhamento múltiplo, sequências preditas, hipotéticas ou com similaridade abaixo de 70% foram excluídas.

2.3.2.2 Análise das variantes germinativas (sentido trocado, sinônimas e intrônicas) em nível do DNA, RNA e proteína (primária, secundária e terciária)

O significado clínico das variantes benignas, provavelmente benignas, VUS e patogênica e as informações específicas da doença associada as variantes foram examinadas no ClinVar (Landrum *et al.,* 2020), VarSome (Kopanos *et al.,* 2019), dbSNP (Sherry *et al.,* 2001), Ensembl (Howe *et al.,* 2018l) e ABraOM (Naslavsky *et al.,* 2017). As frequências alélicas foram obtidas em bancos de dados populacionais como o Genome Aggregation Database (gnomAD) (Karczewski *et al.,* 2020), o 1000 Genomes Project (Fairley *et al.,* 2020) e Exome Variant Server (Exome Variant Server consortium, 2021).

Foi avaliado *in silico* o efeito das variantes localizadas em sítios de *splicing* através do Alternative Splice Site Predictor (ASSP - 2011) (Wang *et al.,* 2006) e, para comparar as possibilidades de alteração de *splicing* para sequências selvagens e mutadas, utilizou-se Human Splicing Finder (HSF - v3.0) (Desmet *et al.,* 2009). Para a predição do potencial efeito das variantes germinativas em relação à função proteica, aplicou-se os algoritmos UMD-Predictor (v1) (Salgado *et al.,* 2016), SIFT (Vaser *et al.,* 2016), Polyphen-2 (v2.0) (Adzhubei *et al.,* 2010), Mutation Taster (v2) (Steinhaus *et al.,* 2021), LIST-S2 (Malhis *et al.,* 2020), FATH-MM (v2.3) (Rogers *et al.,* 2018) e Provean (v1.1) (Choi *et al.,* 2012). Para a predição dos efeitos das variantes de sentido trocado sobre a estabilidade da estrutura proteica foram usados DynaMut (v2.0) (Rodrigues *et al.,* 2018) e mCSM (Pires *et al.,* 2013). Para a predição dos efeitos das variantes de sentido trocado sobre domínios, motivos, modificações pós-síntese e características moleculares da RAD51D (por exemplo, sítios de interação), foram

usados UniProt, InterPro, MutPred2 (Pejaver *et al.*, 2020), Eukaryotic Linear Motif (Kumar *et al.*, 2022), ProSite e PhosphoSitePlus (Hornbeck *et al.*, 2012).

A modelagem da estrutura 3D de RAD51D foi desenvolvida com programas Modeller (Webb e Sali, 2016) e ColabFold v1.5.2: AlphaFold2 (Mirdita *et al.*, 2022). As estruturas previstas foram validadas com programas PROCHECK (Laskowski *et al.*, 1993) e PDBSum (Laskowski *et al.*, 2021). O PyMol (Zhu *et al*., 2014) foi utilizado para analisar as estruturas e gerar as figuras. Ao final das análises, foi aplicado o critério ACMG/AMP para a classificação de cada variante quanto à potencial patogenicidade (Richards *et al*., 2015). Estes critérios consideram principalmente os efeitos de variantes em nível populacional, análises *in silico*, estudos funcionais, dados de segregação das variantes nas famílias, a existência de relatos de caso descrevendo os dados clínicos em pacientes com determinada variante, entre outros.

3. RESULTADOS E DISCUSSÃO

3.1 - Anotação funcional do gene, transcritos e isoformas de *RAD51D*

O *RAD51D* se localiza no braço longo do cromossomo 17 (17q12; GRCh 38 p13 35092221..35119860), tem 27.640 bp e 15 éxons (Pittman *et al.*, 1998), sendo os éxons 3, 4, 6, 7, 8, 10, 11, 12 e 13 os mais expressos nos tecidos de acordo com o GTEx . Os éxons 4, 5, 6, 7, 10, 11 e 12 aparecem com a mesma sequência em todos os transcritos, que os apresentam. Por outro lado, os éxons 1, 2 e 3 são considerados éxons alternativos, pois apresentam diferentes inícios e tamanhos nos transcritos. Inclusive, todas as isoformas funcionais se iniciam no éxon 3 do gene, saltando os éxons 1 e 2.

O *RAD51D* tem 23 transcritos anotados, sendo que apenas nove deles deles são codantes; os demais são transcritos processados, transcritos com introns retidos e transcritos sujeitos a NMD (*nonsense mediated decay*, decaimento do transcrito devido à presença de códons de parada) (Ensembl, 2022). Além disso, há oito isoformas produzidas por *splicing* alternativo descritas em bases de dados (UniProt, NCBI e Ensembl) e que, de acordo com MaxQB (v. 6.2) (Schaab *et al.*, 2012) (base de dados de proteômica), são verdadeiras e funcionais (Quadro 3, Fig. 7). Outras oito potenciais isoformas, preditas computacionalmente (Uniprot, 2022 e Ensembl, 2022), não são identificadas como variantes proteicas pelo MaxQB, que prevê o NMD dos transcritos processados (Quadro 3). Entretanto, no NCBI, somente três

transcritos e proteínas para o gene *RAD51D* são descritos: NM_002878.4/NP_002869.3, (ENST00000345365.11/O75771-1);NM_133629.3/NP_598332.1,(ENST00000335858.11/O75771-3); NM_001142571.2/NP_001136043.1, (ENST00000590016.5/O75771-8).

Quadro 3 - Isoformas funcionais codificadas por *RAD51D* e potenciais de isoformas mapeadas computacionalmente

Isoformas produzidas por *splicing* alternativo preditas como funcionais	
Isoformas	**Descrição**
O75771-1	A isoforma O75771-1 é a canônica, utilizada como sequência de referência nas análises *in silico*. Possui 328 aminoácidos e é codificada pelos transcritos ENST00000345365.11 (éxon 3a, 3b, 4, 6, 7, 8, 10, 11, 12 e 13) e ENST00000394589.8 (éxon 2, 3a, 3b, 4, 6, 7, 8, 10, 11, 12, 13)
O75771-2	A isoforma O75771-2 é codificada pelos transcritos NMDs ENST00000586044.5 (éxon 3a, 3b, 6, 7, 8, 10, 11, 12, 13) e ENST00000588594.5 (éxon 3a, 3b, 6, 8, 10, 11, 12 e 13), portanto, apresenta apenas 49 aminoácidos. Nesta isoforma, perderam-se 278 aminoácidos (50-328) e, em comparação com a isoforma canônica, a posição 49 tem uma alanina em vez de uma serina
O75771-3	A isoforma O75771-3 é codificada pelo transcrito ENST00000335858.11 (éxon 3a, 3b, 8, 10, 11, 12 e 13). Possui 216 aminoácidos e sua diferença está na ausência dos aminoácidos 50-161 da isoforma canônica
O75771-4	O transcrito canônico ENST00000345365.11 sofre alterações nos sítios de *splicing* levando à perda do éxon 7. A isoforma possui 283 aminoácidos e, em comparação com a isoforma canônica O75771-1, a isoforma O75771-4 perde resíduos 116-160
O75771-5	O transcrito ENST00000345365.11 sofre alterações nos sítios de *splicing*, codificando a isoforma O75771-5, que embora apresente 328 aminoácidos, se difere da isoforma canônica O75771-1 nos resíduos 88-101 (SLDKLLDAGLYTGE → RQKLSGGSRWCMHL)
O75771-6	A isoforma O75771-6, com 118 aminoácidos, é codificada pelo transcrito ENST00000587977.5 (éxons 3a, 3b, 4, 5, 7, 8, 10, 11, 12, 13). A diferença em relação à isoforma canônica está nos aminoácidos 88-118 (SLDKLLDAGLYTGEVTEIVGGPGSGKTQVCL → RHGGRTQVGTWEDCSCLRSPQGD RGVGSGML) e na ausência dos aminoácidos 119-328

O75771-7	O transcrito ENST00000345365.11 sofre alterações nos sítios de *splicing*, codificando a isoforma O75771-7, que apresenta apenas 212 aminoácidos, perdendo os aminoácidos 213-328. Ela difere da sequência da isoforma canônica nos aminoácidos 193-212 (VTGSSGTVKVVVVDSVTAVV → DGIPEHLNHIPHCLHVHLPC)
O75771-8	A isoforma O75771-8 é codificada pelo transcrito ENST00000590016.5 (éxon 3a, 3b, 5, 6, 7, 8, 10, 11, 12 e 13). Ela apresenta 348 aminoácidos e difere da sequência da isoforma canônica pelos aminoácidos 49-88 (ALVALRRVLLAQFSAFPVNGADLYEELKTSTAILSTGIGS → TWRAHSSGNLGGLQLPQVPAGRSWSGVRNALKKAGLGHGGTDGLSLNAFDERGTAVSTSR)
Isoformas potenciais computacionalmente mapeadas e preditas como não funcionais	
K7EN92	Têm 249 resíduos e é codificada pelo transcrito ENST00000592577.5 (PT)
K7EMF1	Têm 141 resíduos e é codificada pelo transcrito ENST00000587405.5 (PT)
K7ESL4	Têm 93 resíduos e é codificada pelo transcrito ENST00000588372 (NMD)
H0UID0	Têm 151 resíduos e é codificada pelo transcrito ENST00000460118.6 (PT)
K7ERM7	Têm 96 resíduos é codificada pelo transcrito ENST00000586186.2 (PT)
K7EJ58	Têm 24 resíduos e é codificada pelo transcrito ENST00000590631.1 (PT)
K7EKG7	Têm 107 resíduos e é codificada pelo transcrito ENST00000586210.5 (NMD)
K7ESI1	Têm 86 resíduos e é codificada pelo transcrito ENST00000585343.5 (NMD)

Nota: todas as informações de proteínas foram extraídas dos bancos de dados UniProt, Ensembl, GTEx e NCBI. O software MaxQB armazena e exibe coleções de grandes projetos de proteômica, permitindo análises e comparações conjuntas, e de acordo com seus resultados, as isoformas potenciais (mapeadas computacionalmente) são preditas como não funcionais. OBS: A isoforma 3 apresenta duas sequências diferentes, em nível de proteína, geradas por *splicing* alternativo, dando origem aos éxons 3a e 3b. *PT: processed transcript* (transcrito processado). *NMD: nonsense mediated decay*.

Figura 7 - Alinhamento das isoformas funcionais de *RAD51D*, produzidas por *splicing* alternativo

```
O75771-1  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQFSAFPVNGADLYEELKTSTA  80
O75771-5  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQFSAFPVNGADLYEELKTSTA  80
O75771-4  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQFSAFPVNGADLYEELKTSTA  80
O75771-8  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKTWRAHSSGNLGGLQLPQVPAGRSWSGVRNALK  80
O75771-7  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQFSAFPVNGADLYEELKTSTA  80
O75771-3  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYK--------------------------------  48
O75771-6  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQFSAFPVNGADLYEELKTSTA  80
O75771-2  MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKS--------------------------------  49
          **********************************************:

O75771-1  ILSTGIG--------------------SLDKLLDAGLYTGE-VTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNG  139
O75771-5  ILSTGIG--------------------R-QKLSGGSRWCMHLVTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNG  139
O75771-4  ILSTGIG--------------------SLDKLLDAGLYTGE-VTEIVGGPGSGKT-------------------------  114
O75771-8  KAGLGHGGTDGLSLNAFDERGTAVSTSRLDKLLDAGLYTGE-VTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNG  159
O75771-7  ILSTGIG--------------------SLDKLLDAGLYTGE-VTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNG  139
O75771-3  ------------------------------------------------------------------------------  
O75771-6  ILSTGIGRHGGRTQVGTWEDCSCLRSPQGDRGVGSGML---------------------------------------------  118
O75771-2  ------------------------------------------------------------------------------  

O75771-1  GLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQ  219
O75771-5  GLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQ  219
O75771-4  --------------------QAEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQ  174
O75771-8  GLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQ  239
O75771-7  GLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLDVLQELRGTVAQQ---DGIPEHLNHIPHCLHVHLPC----  212
O75771-3  ----------------------AEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQ  107
O75771-6  ------------------------------------------------------------------------------  
O75771-2  ------------------------------------------------------------------------------  

O75771-1  QREGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSS  299
O75771-5  QREGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSS  299
O75771-4  QREGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSS  254
O75771-8  QREGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSS  319
O75771-7  ------------------------------------------------------------------------------  
O75771-3  QREGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSS  187
O75771-6  ------------------------------------------------------------------------------  
O75771-2  ------------------------------------------------------------------------------  

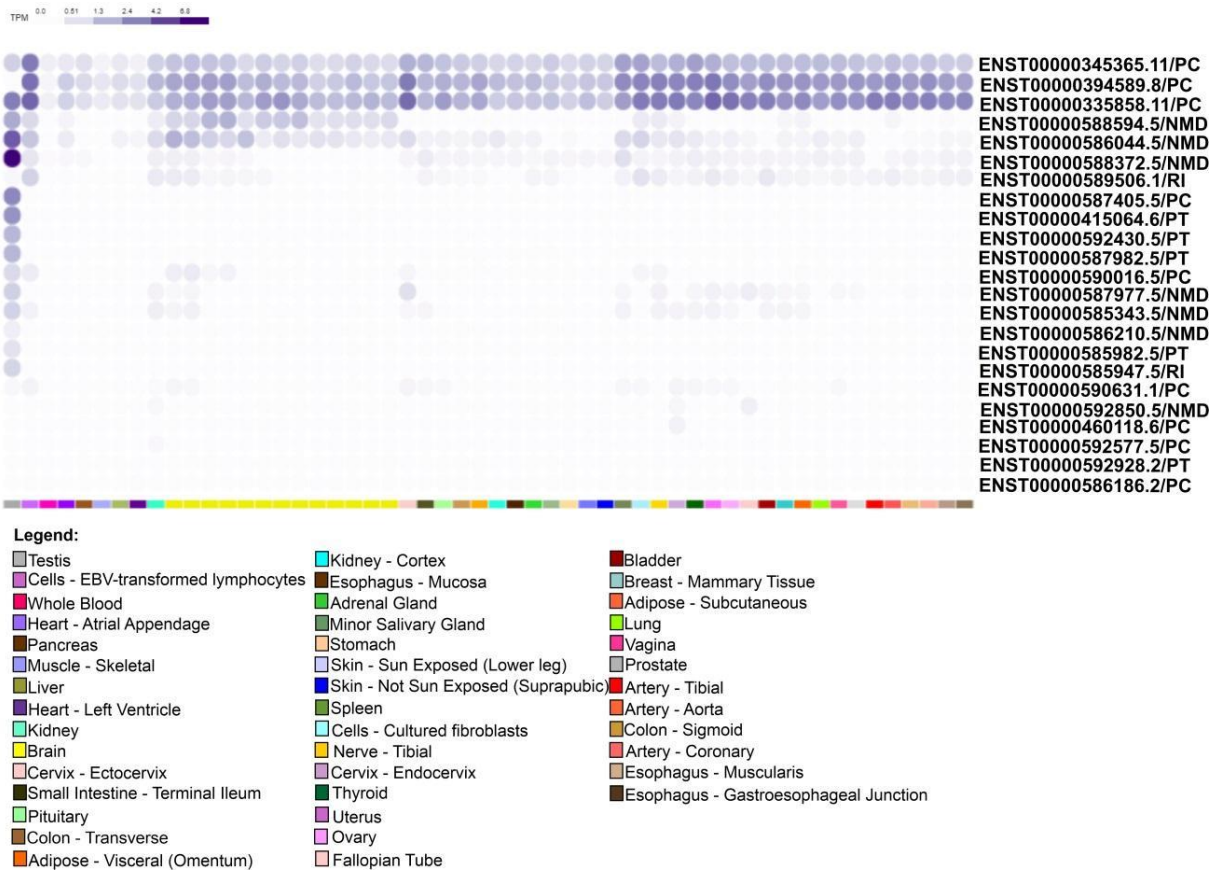
O75771-1  RQPTGFQEMVDIGTWGTSEQSATLQGDQT  328
```

Nota: As cores das letras representam as características físico-químicas de cada aminoácido: vermelho - aminoácidos apolares; azul - aminoácidos ácidos; verde - aminoácidos polares; rosa - aminoácidos básicos. Os *gaps* (--) representam a ausência de aminoácidos nesta região. Os asteriscos (*) referem-se aos aminoácidos homólogos e os dois pontos (:) simbolizam a existência de similaridade entre as características físico-química dos aminoácidos.

Fonte: Elaborada pela autora, usando o alinhamento múltiplo realizado no software Clustal Omega.

Para a análise do padrão de expressão dos transcritos, foram utilizadas a ENCODE Encyclopedia e o GTEx. Apenas as isoformas ENST00000345365.11 (canônica), ENST00000394589.8 e ENST00000335858.11 apresentam um padrão de expressão elevado; os outros 20 transcritos têm baixo padrão de expressão tecidual. Notou-se que o transcrito canônico é comumente expresso em graus variados no cólon, próstata, mama, baço, testículos, ovários, útero, colo do útero, bexiga, timo, tireóide, intestino delgado e leucócitos (Fig. 8) (GTEx Consortium, 2022; Kawabata & Saeki, 1999).

Figura 8 - Nível de expressão de cada transcrição de RAD51D em diferentes tecidos



Nota: Até agora, foram descritos para *RAD51D* os transcritos que codificam proteínas (*protein coding - PC*), transcritos processados (*processed transcript - PT*), transcritos com introns retidos (*retained introns - RI*) e transcritos submetidos a decaimento mediado sem sentido (*nonsense mediated decay - NMD*). A expressão dos transcritos nos tecidos é representada em tons graduais de roxo, variando do escuro (maior expressão) ao claro.

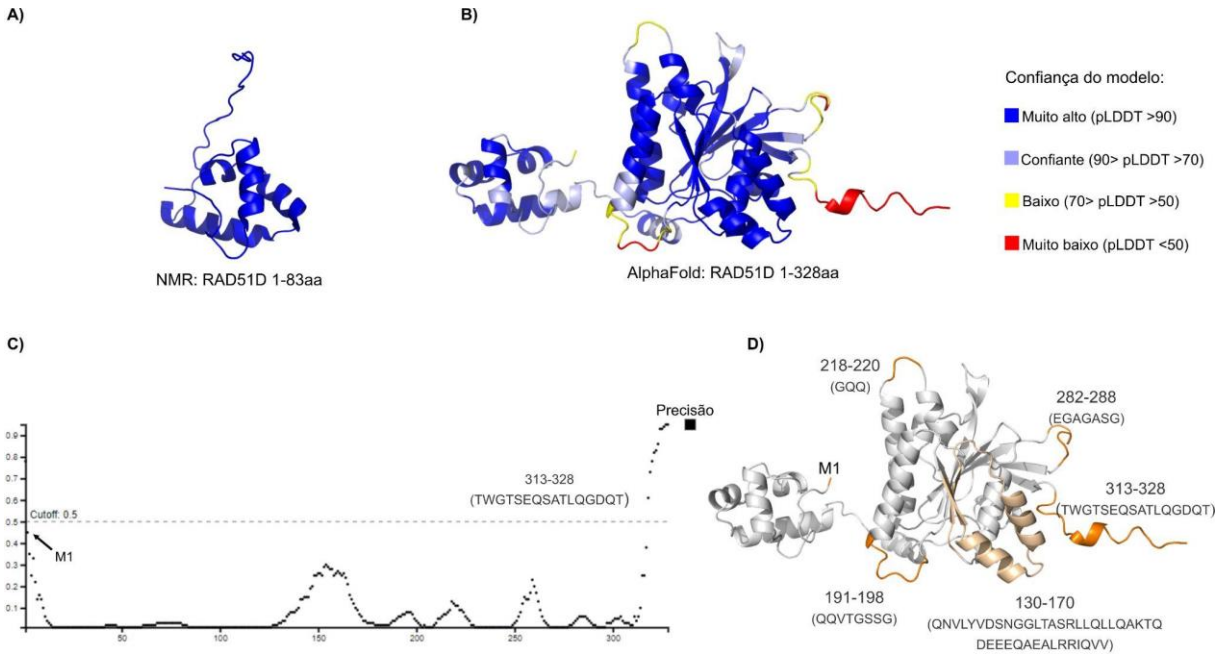
Fonte: Elaborada pela autora a partir de dados do GTEx (acesso em Outubro, 2022).

Nos bancos de dados de proteínas, apenas os primeiros 83 aminoácidos do N-terminal da proteína RAD51D foram definidos tridimensionalmente por ressonância nuclear magnética (Fig. 9A); os resíduos restantes (84-328) até agora não foram definidos. Mas, recentemente a estrutura 3D prevista da RAD51D tornou-se disponível em todos os bancos de dados de proteínas predita com o AlphaFold. No entanto, a estrutura predita apresentou algumas regiões/resíduos de confiança baixa (1, 191-193, 197-198, 218-220, 282, 283, 286-288, 313, 315, 316) ou muito baixa (194-196, 284, 285, 317-328) (Fig. 9B).

Geralmente, regiões de baixa confiança correspondem a regiões intrinsecamente desordenadas (IDRs) devido à heterogeneidade conformacional (David *et al.*, 2021; Ruff & Pappu, 2021).

No presente estudo, foram usados DISOPRED e MobiDB para prever as regiões intrinsecamente desordenadas da RAD51D. De acordo com DISOPRED, são intrinsecamente desordenados o primeiro aminoácido da cadeia N-terminal (Met1) e os resíduos 313 a 328 (TWGTSEQSATLQGDQT). Também previmos regiões em toda a proteína que, embora não previstas como desordenadas, têm um alto potencial para desordem intrínseca (130 a 170 aa - regiões não totalmente estáveis) que entretanto não atingem significância (Fig. 9C). Além dessas regiões previstas pelo DISOPRED, outras regiões com desordem são previstas no MobiDB, compreendem os resíduos 191 a 198 (QQVTGSSG), 218 a 220 (GQQ) e 282 a 288 (EGAGASG) (Fig. 9D).

Figura 9 - Estrutura tridimensional e predição de desordem intrínseca para a proteína RAD51D



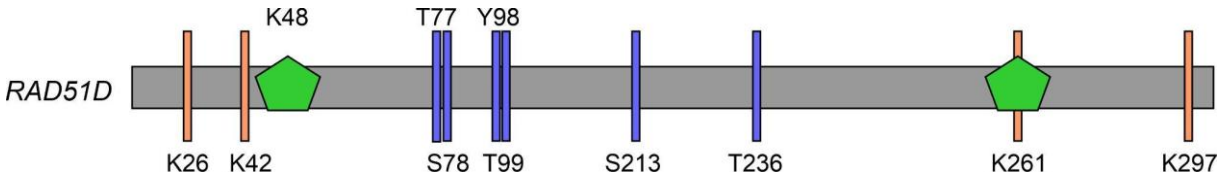
Nota: Em A, está o modelo 3D curto da RAD51D, correspondendo apenas os resíduos do 1 ao 83, que foram definidos por ressonância nuclear magnética (RMN). Em B, está o modelo 3D da proteína RAD51D matura previsto usando-se AlphaFold. AlphaFold produz uma pontuação de confiança por resíduo (pLDDT) entre 0 e 100, que varia, para cada resíduo, de alta confiança, confiança, baixa confiança a confiança muito baixa. Em C, é mostrado o gráfico da desordem por resíduo, conforme predita com DISOPRED. A primeira região desordenada da proteína é a metionina inicial (posição 1 - indicada pela primeira seta). Embora ela não esteja em uma região desordenada (acima do *cutoff* 0.5), ela apresenta alto potencial para a desordem; há outras regiões ao longo da proteína que também apresentam esse potencial para desordem por aproximarem-se do limítrofe, como os resíduos de 130 a 170. No final da região C-terminal, os aminoácidos 318 a 328 (a sequência de aminoácidos estão entre parênteses) ultrapassam o ponto de corte para desordem, estando caracterizados como desordenados. Em D, é mostrada a predição da desordem da RAD51D, tridimensionalmente, considerando-se

simultaneamente os resultados do DISOPRED e MobiDB. Em laranja escuro, estão os resíduos preditos como desordenados e, em laranja claro, estão os resíduos com alto potencial para desordem.

Fonte: Elaborada pela autora.

Foram previstos três tipos de modificações pós-síntese para a RAD51D usando-se o PhosPhoSitePlus (v6.6.0.4): ubiquitinação, fosforilação e acetilação, que são mostradas na figura 10.

Figura 10 - Representação esquemática dos sítios de modificações pós-síntese da RAD51D

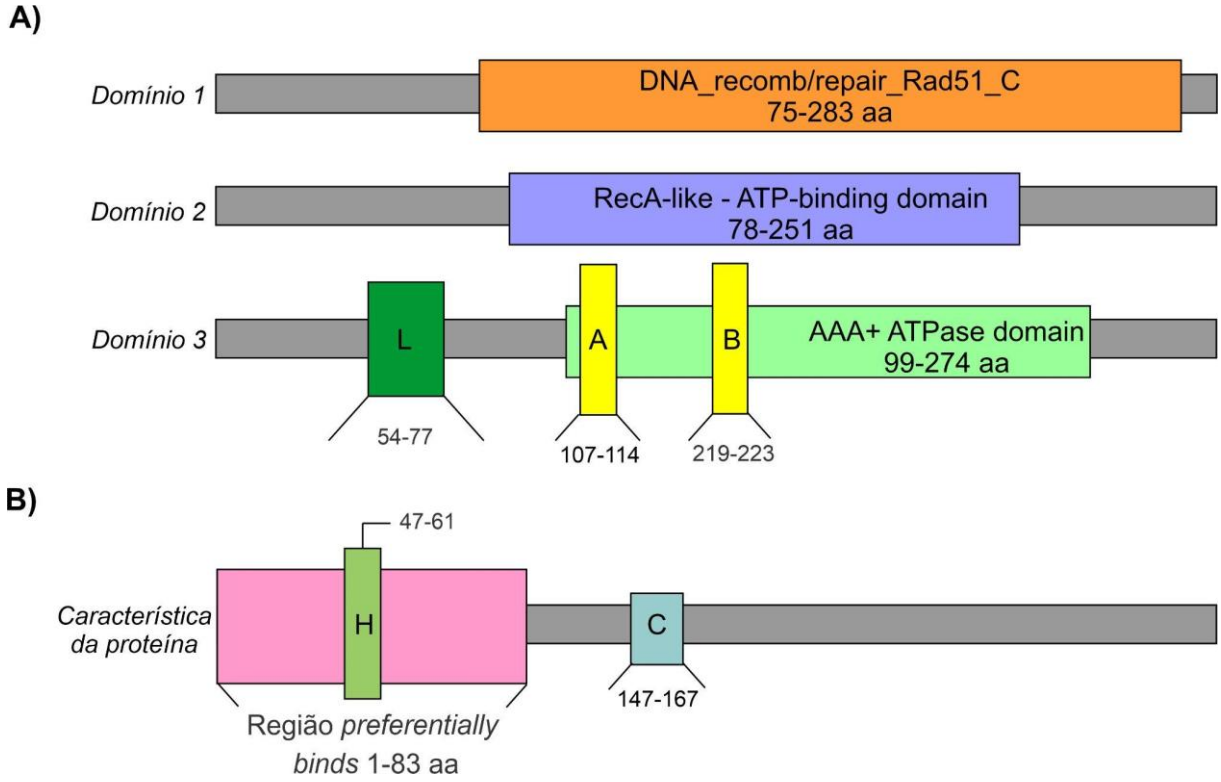


Nota: As barras laranja representam os sítios de ubiquitinação, que correspondem aos resíduos lisina nas posições 26, 42, 261 e 297; as barras azuis são os sítios de fosforilação, representando os resíduos treonina das posições 77, 99 e 236, tirosina na posição 98 e serina nas posições 78 e 213; e, os pentágonos verdes representam os sítios de acetilação, nos resíduos 48 e 261.

Fonte: Elaborada pela autora, usando PhosphoSitePlus.

Identificou-se no UniProt, Pfam e InterPro, os três domínios funcionais que se sobrepõem ao longo da proteína, que incluem *DNA recombination and repair protein* (75-283), *RecA_ATP-biding* (78-251) e *AAA+ATPase* (99-274) (Fig. 11). Dentro do domínio ATPase existem dois motivos (*do Inglês, motifs*), *Walker A* (composto por oito resíduos - 107 a 114) e *Walker B* (composto por cinco resíduos - 219 a 223), altamente conservados, que potencializam a ligação de ATP (Pittman *et al.*, 1998; Cartwright *et al*., 1998). Além disso, há na extremidade aminoterminal um sítio de *preferentially binds ssDNA*, composto pelos resíduos 1 a 83; um sítio *Helix* composto por 15 aminoácidos (47-61); uma "região de ligação" localizada dos resíduos 54-77, que auxilia nas interações adequadas com XRCC2; na extremidade carboxiterminal, há uma região abrangendo os aminoácidos 77-329, que auxilia na interação da proteína RAD51D com a RAD51C, e por fim, um *coiled coil* composto por 21 aminoácidos (147-167) (Fig. 11).

Figura 11 - Representação esquemática da organização dos domínios e motifs na estrutura da proteína RAD51D



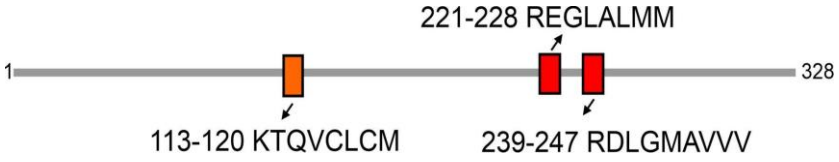
Nota: O primeiro domínio (laranja) é maior, composto por 209 resíduos (75 a 283); o segundo domínio é composto por 174 resíduos (78 a 251; azul) e aminoácidos; o terceiro domínio (verde claro) é composto por 185 resíduos (99-274). Dentro deste domínio, existem dois motivos altamente conservados, *Walker* A e *Walker* B, que potencializam a ligação de ATP. No mesmo esquema, mostramos uma região de ligação (L; verde escuro), que promove interações proteína-proteína (54 a 77). Características proteicas observadas em RAD51D: aa região de ligação de fita dupla (rosa), composta por 83 aminoácidos; a região *Helix* (H; verde), resíduos 47 a 61; e, por fim, a região helicoidal (*coiled coil*) (C; azul envelhecido), composta por 21 resíduos (147 a 167).

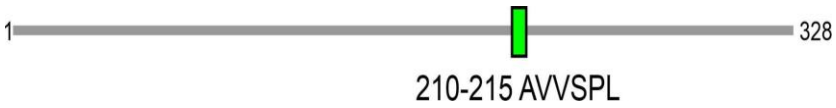
Fonte: Elaborada pela autora, usando NextProt.

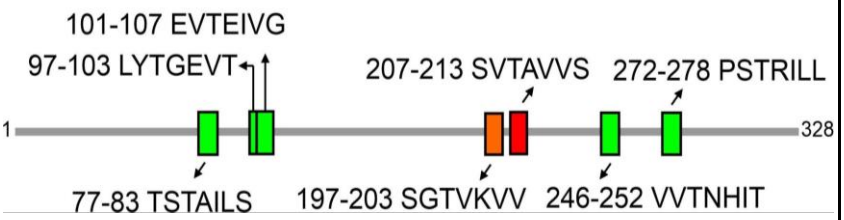
Outros motivos estruturais presentes ao longo da conformação da RAD51D foram identificados no Eukaryotic Linear Motif Resource (ELM) e no Prosite. A função e a localização desses motivos são mostradas no Quadro 4. Embora esses motivos tenham sido previstos para a RAD51D, pouco se sabe sobre como eles afetam as funções da proteína.

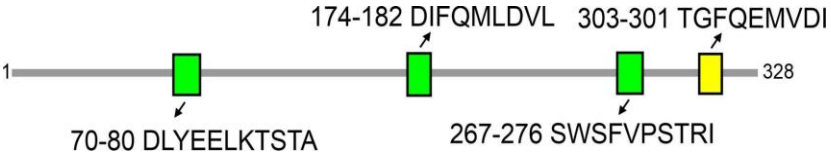
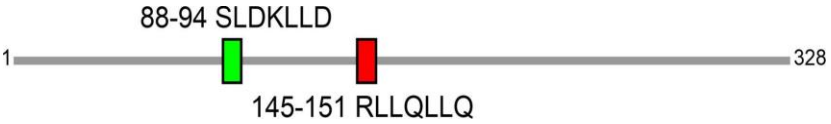
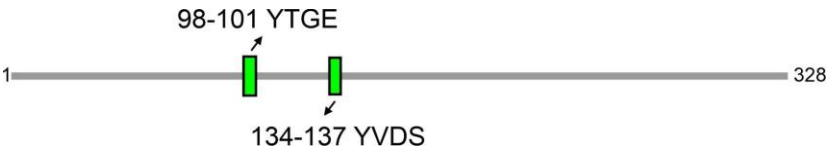
Quadro 4 - Descrição funcional e esquematização da localização dos motivos funcionais preditos para RAD51D com o Eukaryotic Linear Motif

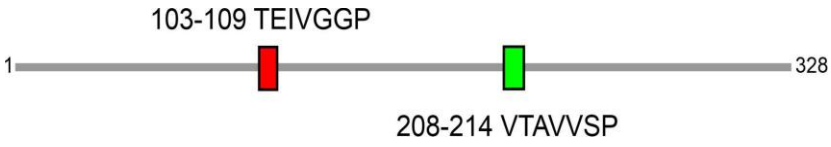
Motivos preditos para RAD51D usando Eukaryotic Linear Motif	
DEG_APCC_DBOX_1 Motivos de destruição de ligação APC/C Compartimento celular: Núcleo e citoplasma	Os motivos de destruição de ligação APC/C são sítios curtos e muito específicos, que permitem o recrutamento de alvos para o complexo APC/C. O complexo promotor de anáfase/ciclossoma (APC/C) é uma ubiquitina ligase, que regula o ciclo celular através da degradação mediada por ubiquitina de proteínas alvo. CDH1 e CDC20 são exemplos de proteínas que reconhecem alvos através desses "motivos de destruição", que são em seguida poliubiquitinados (Peters, 2006; Castro *et al.*, 2005; Petersen *et al.*, 2000) 
DOC_CYCLIN_RxL_1 Motivos de ancoragem (*docking)* de domínio ciclina N-terminal Compartimento celular: Núcleo e citoplasma	Os motivos de ancoragem são sítios, que permitem a interação de alvos com o complexo ciclina/CDKs (quinases dependentes de ciclina). As ciclinas se ligam a CDKs e fornecem aos substratos um sítio de ancoragem, modulando a especificidade do sítio ativo de CDK. As CDKs, assim como os motivos de ancoragem, coordenam centenas de eventos moleculares durante o ciclo celular, incluindo a via de fosforilação (Örd *et al.*, 2019) 
DOC_MAPK_gen_1 Motivo *MAPK-docking* Compartimento celular: Núcleo e citoplasma	O motivo *MAPK-docking* consiste em um ou mais resíduos básicos e dois a quatro resíduos hidrofóbicos em grupos adjacentes. Esse motivo permite a interação de moléculas-alvo com a MAPK (*mitogen-activated protein kinase*). *MAPK* é uma subfamília de proteínas-quinases específicas de serina/treonina, que respondem a estímulos extracelulares e regulam várias atividades celulares, como expressão gênica, mitose, diferenciação, sobrevivência celular e apoptose (Biondi & Nebreda, 2003; Bardwell, 2006) 
DOC_MAPK_MEF2A_6 Motivo *MAPK-docking* Compartimento celular:	Esses sítios *docking* também são mediadores da interação de alvos com as subfamílias ERK1/2 e p38 de MAP quinases (Bardwell, 2006)

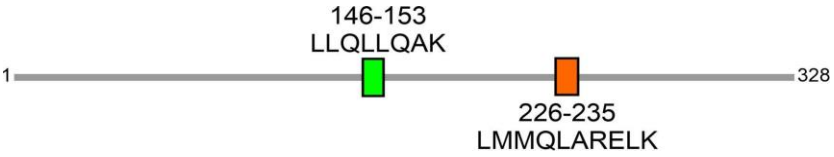
Núcleo e citoplasma	
DOC_MAPK_NFAT4_5 Motivo *MAPK-docking* Compartimento celular: Núcleo e citoplasma	Sítio *docking* reconhecido pelas quinases JNK (c-Jun N-terminal cinase), uma proteína que também pertence a família MAPK e atua na regulação do processo de autodestruição celular (Bardwell, 2006) 
DOC_PP2A_B56_1 Motivo B56-*docking* da Holoenzima PP2A Compartimento celular: Núcleo e citoplasma	A proteína fosfatase 2A (PP2A) pertence à superfamília das fosfoproteínas fosfatases (PPPs) e catalisa a desfosforilação de proteínas na progressão do ciclo celular, na dinâmica do citoesqueleto e na sinalização do fator de crescimento. Três subunidades são vistas nas proteínas PP2: A, B e C. A subunidade B tem atividade reguladora, direcionando substratos para B56 (a maior família da subunidade B). O motivo B56-*docking* da Holoenzima PP2A é um sítio de ancoragem necessário para a subunidade reguladora B56 para a desfosforilação de proteínas. Este motivo modula a afinidade para B56, que por sua vez pode determinar o estado de fosforilação de substratos associados. Em resumo, as isoformas do tipo B56 se ligam aos seus alvos por meio desse motivo (McCright *et al.*, 1996; Hertz *et al.*, 2018) 
DOC_USP7_MATH_1 Motivo de ligação USP7 Compartimento celular: Núcleo	O motivo de ligação USP7 é um domínio encontrado em moléculas-alvo, que interagem com a USP7. USP7 é uma deubiquitinase, que cliva ubiquitinas de seus substratos. Em sua estrutura possuem um domínio MATH N-terminal, que é responsável pelo reconhecimento do substrato e localização nuclear; também possuem um domínio catalítico central, que é necessário para a atividade de ubiquitinação. Por fim, o domínio C-terminal Ubl, responsável por várias interações de substrato USP7, incluindo ICP0, GMPS, DNMT1 e UHRF1, que medeiam à estabilização do substrato, translocação ou ativação de USP7 (Holowaty *et al.*, 2003; Amerik & Hochstrasser, 2004; Cheon & Baek, 2006) 

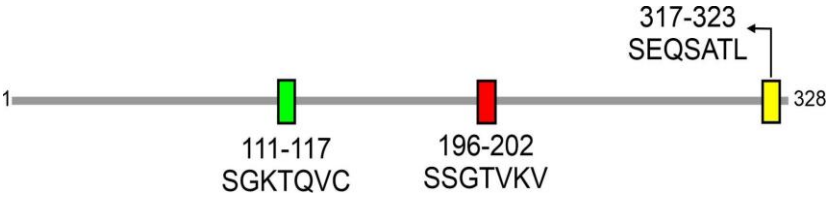
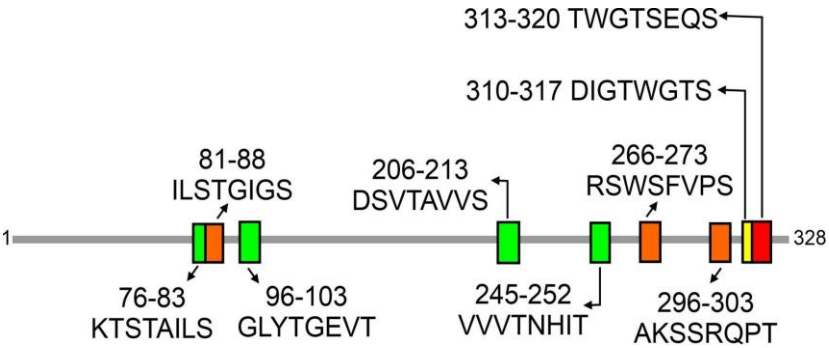
DOC_WW_Pin1_4 Motivos de ligação ao domínio WW Compartimento celular: Núcleo e citoplasma	Os motivos de ligação ao domínio WW medeiam a interação de moléculas peptídicas com o domínio WW de proteínas sinalizadoras e regulatórias. Esses motivos são curtos e apresentam pelo menos um resíduo de prolina na sequência. A fosforilação dirigida por prolina é um importante mecanismo de sinalização celular. As quinases dirigidas por prolina incluem Map quinases e quinases dependentes de ciclina (Sudol *et al.*,, 2001; Lu *et al.*, 1999) 
LIG_14-3-3_CanoR_1 Motivo de fosfopeptídeo de ligação 14-3-3 Compartimento celular: Núcleo e citoplasma	Os motivos de fosfopeptídeo de ligação 14-3-3 são motivos que contêm fosfoserina ou treonina. Esses fosfo-motivos permitem que as proteínas 14-3-3 formem homo- ou heterodímeros com seus alvos. As proteínas 14-3-3 são uma família de moléculas reguladoras que estão envolvidas em diversos processos celulares, incluindo transdução de sinal e regulação do ciclo celular. Existem fosfo-motivos canônicos, que contém arginina, e um grupo de motivos não canônicos, que são difíceis de classificar, mas utilizam interações hidrofóbicas adicionais (Muslin *et al.*, 1996; Obsil & Obsilova, 2011) 
LIG_Actin_WH2_2 Motivos de ligação à actina Compartimento celular: Citoplasma	O domínio WH2 (domínio de homologia WASP-2) é um dos motivos de ligação do monômero de actina. Este motivo interage com a G-actina ligando-se à fenda hidrofóbica formada pelos subdomínios 1 e 3 e promove ou inibe a montagem do filamento. WH2 atua regulando a atividade da actina, que têm um papel central em muitos processos biológicos, incluindo o processo de nucleação de filamentos (Paunola *et al.*, 2002) 
LIG_BRCT_BRCA1_1 Motivo fosfopeptídeo de ligação ao BRCT	O motivo de fosfopeptídeo de ligação de BRCT é um domínio de ligação ao domínio BRCT da BRCA1. Os domínios BRCT estão presentes em proteínas, que estão associadas à resposta ao dano do DNA. Este domínio pode reconhecer duas sequências de consenso deste motivo: motivo de baixa afinidade e motivo de alta afinidade. O motivo de baixa afinidade contém pelo menos uma serina e uma fenilalanina em sua sequência e o motivo de alta afinidade contém uma uma serina, seguida por quaisquer um ou mais resíduos, e uma fenilalanina seguida por uma lisina (S..FK) (Glover *et al.*, 2004;

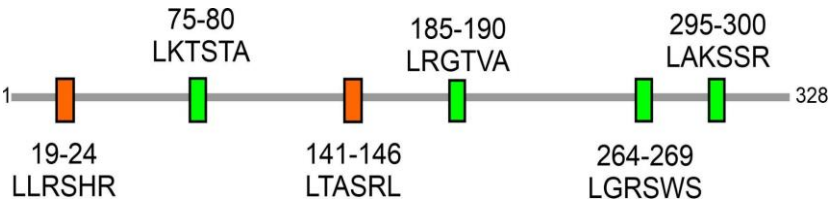
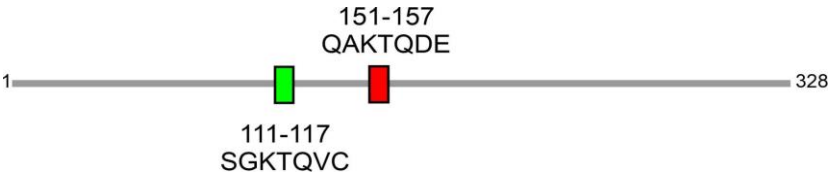
Compartimento celular: Núcleo	Clapperton *et al.*, 2004) 
LIG_FHA_1 Motivo fosfopeptídeo de ligação ao *forkhead-associated* (FHA) Compartimento celular: Núcleo	Motivos do tipo *LIG_FHA_1* são fosfomotivos, que se ligam a domínio *forkhead-associated* (FHA). O domínio FHA reconhece epitopos fosfotreonina de proteínas, com especificidade por aminoácidos alifáticos (possuem hidrocarbonetos saturados como cadeia lateral) na posição pT+3. O motivo tem a sequência de consenso de TXX[IVL]. Domínios FHA estão associados a proteínas nucleares, que participam de muitos processos, incluindo sinalização, *checkpoints* do ciclo celular, reparo do DNA e regulação da transcrição. Está presente em muitas proteínas (e.g., quinases, fosfatases, cinesinas, fatores de transcrição, proteínas ligadoras de RNA e enzimas (Durocher & Jackson, 2002; Li *et al.*, 2000; Yongkiettrakul *et al.*, 2004) 
LIG_FHA_2 Motivo fosfopeptídeo de ligação ao *forkhead-associated* (FHA) Compartimento celular: Núcleo	Semelhante a *LIG_FHA_1*, motivos do tipo *LIG_FHA*_2 são fosfomotivos, que são reconhecidos por domínios FHA (ver acima). Entretanto, nesta situação, o domínio FHA reconhece o peptídeo contendo fosfotreonina com especificidade para aminoácidos ácidos na posição pT+3. O motivo tem a sequência de consenso com uma sequência de consenso de TXX[ED] 
LIG_LIR_Gen_1 Motivos de ligação à família de proteínas Atg8 Compartimento celular: Citoplasma	O motivo LIR canônico é encontrado em proteínas adaptadoras específicas ligadas a proteínas ubiquitinadas, organelas ou patógenos marcados para degradação. Esse motivo se liga a membros da família de proteínas Atg8/LC3 para mediar processos envolvidos na autofagia e uma série de processos catabólicos, que mantêm a homeostase celular removendo componentes intracelulares danificados e disfuncionais (Wesch *et al.*, 2020)

	
LIG_NRBOX Motivo receptor nuclear Compartimento celular: Núcleo	O motivo receptor nuclear (NR) permite a ligação de proteínas coativadoras (e.g., fatores de transcrição) a receptores nucleares. Este motivo é hidrofóbico e apresenta uma sequência de consenso LXXLL (Heery *et al.*, 1997; Nolte *et al.*, 1998; Leo & Chen, 2000) 
LIG_SH2_CRK Motivos de fosfotirosina de ligação aos domínios SH2 Compartimento celular: Citoplasma	*LIG_SH2_CRK* são motivos que contêm pelo menos um resíduo de tirosina fosforilado (motivos de fosfotirosina), conectando Receptor Tirosina Quinases (RTKs) ativados, bem como TKs solúveis como SRC e ABL, com proteínas citoplasmáticas. Esses motivos são reconhecidos pelo domínio SH2 *(Src Homologia 2),* que são geralmente encontrados em famílias CRK (proteínas sinalizadoras). SH2 é o principal domínio de ligação para eventos de sinalização de fosforilação de TK. Este motivo atua no citoesqueleto de actina e no citosol (Kaneko *et al.*, 2012; Pawson *et al.*, 2001) 
LIG_SH2_SRC Motivos de fosfotirosina de ligação aos domínios SH2 Compartimento celular: Citoplasma	O motivo de fosfotirosina *LIG_SH2_SRC* é semelhante ao *LIG_SH2_CRK*, cuja atividade é permitir a ligação dos alvos ao domínio SH2 da família SRC. Este motivo atua no citosol (Kaneko *et al.*, 2012; Pawson *et al.*, 2001) 
LIG_SH2_STAT5 Motivos de fosfotirosina de ligação aos domínios SH2 Compartimento celular:	*LIG_SH2_STAT5* é também um motivo de fosfotirosina, que se liga ao domínio STAT5 SRC Homology 2 (SH2). Atua também no citosol (Kaneko *et al.*, 2012; Pawson *et al.*, 2001) 

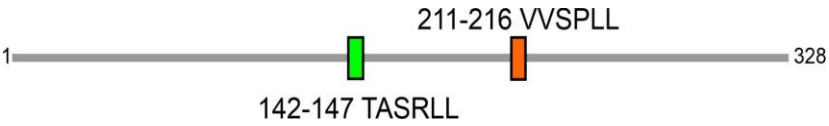
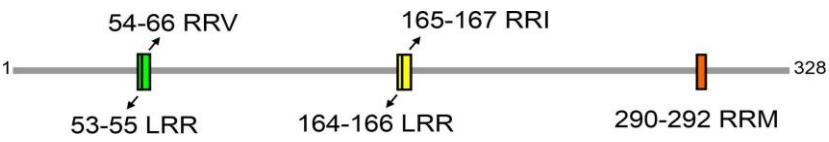
Citoplasma	
LIG_SH3_3 Motivo de ligação ao SH3 Compartimento celular: Citoplasma	Motivo *LIG_SH3_3* permite a interação de moléculas alvos com outras proteínas através da interação com o domínio SH3. Domínio SH3 (*SRC Homolgy 3*) são encontrados em proteínas envolvidas em diversos processos biológicos como vias de transdução de sinal, organização do citoesqueleto, tráfego de membrana ou montagem de organelas. Domínio SH3 têm afinidade por prolina e o motivo *LIG_SH3_3* contém pelo menos um resíduo de prolina (Rámon, 2008) 
LIG_SUMO_SIM_anti_2 Motivo de interação SUMO Compartimento celular: Núcleo	Motivo de interação SUMO (*small ubiquitin-like modifiers*) ou SIM (*SUMO-interacting motif*) permite a ligação não covalente de moléculas nucleares alvos à proteína *SUMO*. O motivo de interação SUMO se dobra como uma folha-beta e pode se ligar à proteína SUMO em uma orientação paralela ou antiparalela. O motivo de ligação antiparalela tem menor afinidade, podendo tolerar sequências mais diversas. Por outro lado, o motivo de ligação paralela tem maior afinidade, tendo preferência por sequências de consenso como (I/V)DLT. O motivo *SUMO_anti_2* favorece a interação da folha-beta na orientação antiparalela). Proteínas contendo o domínio SUMO estão envolvidas em uma variedade de processos celulares, incluindo regulação transcricional, localização subnuclear, montagem dos *nuclear bodies* e resposta antiviral. Este motivo têm regiões hidrofóbicas consistindo de 3 a 4 resíduos hidrofóbicos (I, V ou L) (Namanja *et al.*, 2012; Hecker *et al.*, 2006; Song *et al.*, 2004) 
LIG_SUMO_SIM_par_1 Motivo de interação SUMO	*LIG_SUMO_SIM_par_1* é também um motivo de interação *SUMO* que, contrário ao *SUMO_anti_2*, favorece a ligação da folha-beta na orientação paralela à proteína SUMO (Namanja *et al.*, 2012) 
LIG_TYR_ITIM Motivo inibidor baseado em tirosina imunorreceptor	ITIM são motivos essenciais para a ativação e terminação das vias de transdução de sinal. Este motivo é encontrado na cauda citoplasmática de alguns receptores inibitórios (KIRs) presentes nas células natural *killer*. O motivo ITIM é fosforilado, ligando-se ao MHC Classe I, permitindo o recrutamento e ativação de proteínas

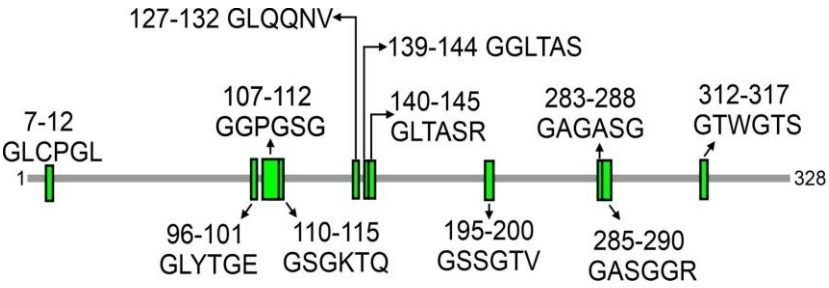
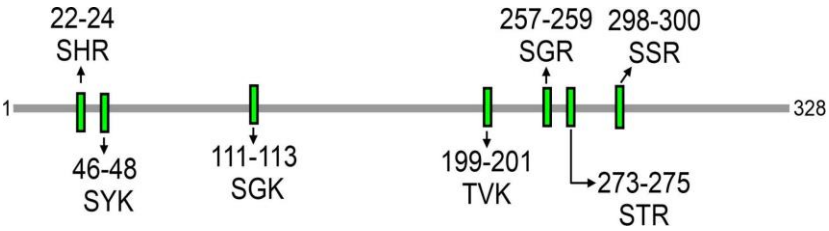
(ITIM) Compartimento celular: Citoplasma	tirosinafosfatase (Long, 1999; Unkeless & Jin, 1997) 
LIG_UBA3_1 Motivo de ligação para o domínio de adenilação UBA3 Compartimento celular: Núcleo	*LIG_UBA3_1* é uma variante do motivo de ligação do domínio de adenilação UBA3 com base nas interações UBE2M e UBE2F. Este motivo é encontrado nas regiões N-terminais de NEDD8 E2s (semelhante à ubiquitina), UBE2M e UBE2F, que medeiam a ligação a UBA3-NAE1 em seu domínio de adenilação UBA3. A proteína NEDD8 (*neural-precursor-cell-expressed developmentally down-regulated 8*) está envolvida na regulação do crescimento, viabilidade e desenvolvimento celular, e é transferida para seus alvos por cascatas multienzimáticas distintas e paralelas, que envolvem a ação sequencial das enzimas NEDD8 E1, E2 e E3 e são conjugadas aos seus alvos através da nedilação (Huang *et al.*, 2004; Brown & Jackson, 2015) 
LIG_WD40_WDR5_VDV_1 Motivo de ligação de repetição WDR5 WD40 (lâmina 5,6) Compartimento celular: Núcleo	Este motivo de ligação a WDR5 liga-se a uma fenda entre as lâminas 5 e 6 do domínio de repetição WD40 de WDR5. O domínio de repetição WD40 de WDR5 é encontrado em proteínas, que desempenham um papel fundamental na metilação de H3K4 e acetilação de H4K16, atuando como uma proteína de suporte para a montagem do respectivo complexo de metilação e acetilação de histonas, que são conservados através da evolução (Odho *et al.*, 2010) 
MOD_CK1_1 Sítio de fosforilação de *Casein kinase 1* (CK1) Compartimento celular: Núcleo e citoplasma	Este motivo é um sítio de fosforilação de CK1 (caseína quinase 1). CK1 é uma família de serina/treonina quinases que regula diversos processos celulares, tais como vias de sinalização, ciclo celular, tráfico de membranas, manutenção do citoesqueleto, resposta ao dano do DNA entre outras atividades (Rachidi *et al.*, 2014; Schittek & Sinnberg, 2014)

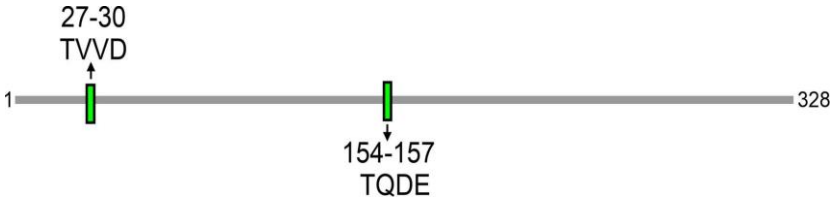
	
MOD_CK1_2 Sítio de fosforilação de *Casein kinase* 2 (CK2) Compartimento celular: Núcleo e citoplasma	Este motivo é um sítio de fosforilação de caseína quinase 2 (CK2), uma serina e treonina quinase expressa de forma ubíqua, altamente conservada presente em todos os eucariotos. A CK2 promove a fosforilação de substratos envolvidos em muitos processos celulares, incluindo progressão do ciclo celular, apoptose, regulação transcricional, transporte de proteínas, metabolismo, infecção por vírus, morfologia celular e transdução de sinal. A desregulação da função CK2 está associada a patologias humanas, como vários tipos de câncer, infecção viral e doenças neurodegenerativas. Geralmente, os sítios de fosforilação reconhecidos por CK2 tem aminoácidos ácidos próximos ao resíduo fosfoaceptor (posição +3) (Meggio & Pinna, 2003; Ahmad *et al.*, 2008; Niefind et al., 2001) 
MOD_GSK3_1 Sítio de fosforilação GSK3 Compartimento celular: Núcleo e citoplasma	Este motivo é um sítio reconhecido por GSK3 para fosforilação de Ser/Thr. A família GSK3 é composta por duas proteínas, GSK3-alfa e GSK3-beta, que reprimem a função da proteína-alvo através da fosforilação. Glicogênio sintase, ATP citrato liase, c-Jun, c-Myc, L-Myc e proteína tau são exemplos de proteínas-alvo de GSK3, que atuam em diferentes processos bioquímicos 
MOD_NEK2_1 Sítio de fosforilação NEK2 Compartimento celular:	Este motivo é um sítio de fosforilação da proteína NEK2. NEK2 é uma quinase mitótica identificada em muitos eucariotos cuja atividade é controlar o ciclo celular, especificamente na fase S e G2, através da fosforilação de proteínas centrossomais. Assim, NEK2 medeia a manutenção e modulação da arquitetura do centrossomo, atuando na progressão do ciclo celular, formação do polo do fuso, ancoragem de microtúbulos, coesão centriolar, formação de cílios e

Núcleo e citoplasma	condensação da cromatina (Fry, 2002; Jeong *et al.*, 2007) 
MOD_PIKK_1 Sítio de fosforilação de PIKK Compartimento celular: Núcleo e citoplasma	Este motivo é um sítio de fosforilação de membros da família PIKK (*phosphoinositide-3-kinase-related protein kinase*), que possui a sequência consenso [S/T]Q. Esta glutamina adjacente à serina ou treonina é importante para o reconhecimento do substrato. PIKK são proteínas eucariotas únicas com atividade de Ser/Thr quinase, que desempenham papéis importantes no reparo do DNA e nos pontos de verificação de danos ao DNA. São encontradas três proteínas PIKK em células de mamíferos, com funções de reparo e *checkpoint*: DNA-PK (proteína quinase dependente de DNA), ATM (mutada em ataxia telangiectasia) e ATR (relacionada à ATM e Rad3) (Falck *et al.*, 2005) 
MOD_PKA_2 Sítio de fosforilação PKA Compartimento celular: Núcleo e citoplasma	Este motivo é um sítio de fosforilação de um subconjunto de quinases basofílicas PKA, PKG e PKC, ou seja, quinases A, G e C combinadas (AGC), todas com especificidade de sequência semelhante. Especificamente, a quinase PKA tem preferência por resíduos de arginina na posição -2 (por exemplo, XRX[T/S]XX) (Zhu et al., 2005) 
MOD_Plk_4 Sítio de fosforilação Plk4 Compartimento celular:	Este motivo é um sítio de fosforilação de membros da família Plks4. A família Plks (quinases tipo polo) é composta por quatro membros Plk1-4, que regulam a progressão do ciclo celular através de papéis funcionais chave na mitose. Elas são consideradas quinases acidófilas, mas Plk4 têm preferência ácida mais fraca e preferência hidrofóbica aumentada. O sítio de fosforilação, ao qual ela se liga é composto por resíduos hidrofóbicos ao redor do resíduo S/T (Leung *et al.*, 2007; Kettenbach *et al.*, 2012)

Núcleo e citoplasma	
MOD_ProDKin_1 Sítio de fosforilação MAPK Compartimento celular: Núcleo e citoplasma	Este motivo é onde a MAPK fosforila suas moléculas-alvo. As proteínas *Mitogen-activated protein kinases (MAPK)* são quinases de sinalização, ancoradas por resíduos de prolina e que fosforilam um resíduo Ser/Thr em várias vias de transdução de sinal. Este motivo normalmente tem uma sequência de consenso (S/T)P, presente em muitas proteínas. Curiosamente, proteínas quinases MAPK, ERKK e ELK-1 apresentam maior afinidade pelo motivo quando a prolina está presente, favorecendo a fosforilação da serina ou treonina (Fantz *et al.*, 2001) 
MOD_SUMO_rev_2 Sítio de sumoilação Compartimento celular: Núcleo	Este motivo é alvo de sumoilação pela proteína SUMO (*small ubiquitin-like modifier*). As proteínas SUMO são pequenos modificadores do tipo Ubiquitina. Elas se ligam covalentemente em resíduos de lisina dentro de proteínas alvo promovendo a sumoilação, uma modificação pós-síntese comum entre as proteínas nucleares (fatores de transcrição, proteínas da cromatina, poros nucleares, entre outras). A sumoilação é conhecida por causar rearranjos dramáticos da localização subnuclear de proteínas modificadas (Tang *et al.*, 2008) 
TRG_DiLeu_BaEn_1 Motivos de Dileucina Compartimento celular: Citoplasma	Motivos de dileucina são comuns em eucariotos e ligam-se preferencialmente aos complexos de adaptina ativados AP-1 ou AP-2. Adaptina é uma proteína que medeia a formação de vesículas revestidas de clatrina, através da interação com receptores ligados à membrana. Os motivos de dileucina são variáveis (especialmente em suas posições carregadas negativamente e nos resíduos hidrofóbicos) e os vários subtipos de motivos tendem a ter funções ligeiramente diferentes (Meyer *et al.*, 2018; Mattera, 2011) 

TRG_DiLeu_BaLyEn_6 Motivos de Dileucina Compartimento celular: Citoplasma	Existem vários subtipos de motivos de dileucina com atividades específicas. A variedade desses motivos se deve à composição da sequência, por exemplo, diferentes posições de aminoácidos hidrofóbicos ou presença ou não de resíduos ácidos. O motivo *TRG_DiLeu_BaLyEn_6* não inicia com um ácido glutâmico na primeira posição tendo preferência por prolina ou arginina na posição +4 da sequência 
TRG_ENDOCYTIC_2 Motivo de sinal de classificação endocítica baseado em Y Compartimento celular: Citoplasma	Motivo de sinal de classificação baseado em tirosina determinam qual via de tráfego vesicular será usada para transportar uma determinada molécula. A sequência do motivo baseada em tirosina é então reconhecida pela subunidade *mu* do complexo de proteínas adaptadoras (PA) (Boll *et al.*, 2002). 
TRG_ER_diArg_1 Motivo di-Arg Compartimento celular: Núcleo e citoplasma	O motivo di-Arg funciona como um domínio de recuperação e retenção no RE (retículo endoplasmático), cuja função é controlar o dobramento correto e a montagem do complexo proteico, que governa a saída do RE. Normalmente, este motivo é composto por dois resíduos de arginina situados lado a lado ou pode conter entre duas argininas um resíduo de aminoácido não conservado (RR ou RXR), seguido por um resíduo hidrofóbico/arginina adjacente, que pode estar em ambos os lados do par RR. Este motivo é comumente encontrado em proteínas de membrana (algumas multiméricas) e, portanto, deveria-se desconsiderar a correspondência de motivo di-Arg em proteínas não membranares (Michelsen *et al.*, 2005; Boulaflous *et al.*, 2009; Uemura *et al.*, 2009) 
Motivos preditos para RAD51D usando Prosite	
MYRISTYL Sítio de N-myristoylation	A miristoilação é uma modificação de acilação (adição covalente de miristato (um ácido graxo saturado em C14) que acontece em proteínas eucarióticas. A sequência consenso é iniciada por uma glicina na N-terminal, seguida de um resíduo não carregado. Nas

	posições 3 e 4, todos os resíduos são permitidos. Na posição 5, são permitidos pequenos resíduos não carregados como Ala, Ser (resíduo favorecido), Thr, Cys, Asn e Gly). E na posição 6, a prolina não é permitida 
PKC_PHOSPHO_SITE Sítio de fosforilação da proteína quinase C	A proteína quinase C exibe uma preferência pela fosforilação de resíduos de serina ou treonina encontrados próximos a um resíduo básico C-terminal. A sequência de consenso [ST]-x-[RK] 
CK2_PHOSPHO_SITE Sítio de fosforilação da caseína quinase II	A caseína quinase II (CK-2) é uma proteína serina/treonina quinase cuja atividade é independente de nucleotídeos cíclicos e cálcio. A CK-2 fosforila muitas proteínas diferentes. A sequência de consenso é composta pelos resíduos [ST]-x(2)-[DE], onde S ou T é o sítio de fosforilação. A especificidade do substrato desta enzima pode ser resumida da seguinte forma: (1) Ser é favorecido em relação a Thr; (2) resíduos ácidos como, ácido aspártico ou ácido glutâmico, devem estar próximos e a C-terminal do resíduo receptor de fosfato; (3) resíduos ácidos adicionais nas posições +1, +2, +4 e +5 aumentam a taxa de fosforilação. A maioria dos substratos fisiológicos tem pelo menos um resíduo ácido nestas posições; (4) Asp é preferível a Glu como fornecedor de determinantes ácidos; (5) Um resíduo básico na N-terminal do resíduo receptor de fosfato diminui a taxa de fosforilação, enquanto um ácido irá aumentá-la

	
TYR_PHOSPHO_SITE_1 Sítio 1 de fosforilação da tirosina quinase	Este motivo geralmente têm um aminoácido básico sete resíduos antes do receptor de fosfato tirosina. Duas sequências de consenso são descritas para este motivo. A primeira sequência de consenso é composta pelos resíduos [RK]-x(2)-[DE]-x(3)-Y, onde Y é o sítio de fosforilação. Nesta sequência, os resíduos ácidos (Asp ou Glu) são encontrados três resíduos antes da tirosina a ser fosforilada. A segunda sequência de consenso é composta pelos resíduos [RK]-x(3)-[DE]-x(2)-Y, e os resíduos ácidos estão dois resíduos antes da tirosina a ser fosforilada 
ATP_GTP_A Motivo A do sítio de ligação ATP/GTP (p-loop)	Proteínas, que se ligam a ATP ou GTP, compartilham vários motivos de sequência mais ou menos conservados. O mais bem conservado desses motivos é uma região rica em glicina, que normalmente forma uma alça flexível entre uma folha β e uma α-hélice. Essa alça interage com um dos grupos fosfato do nucleotídeo. Este motivo de sequência é geralmente referido como sequência de consenso 'A'. Sequência consenso é [AG]-x(4)-GK-[ST] 
AMIDATION Sítio de amidação	O precursor de hormônios e outros peptídeos ativos, que são amidados no terminal C, é sempre seguido diretamente por um resíduo de glicina, que fornece o grupo amida e, mais frequentemente, por pelo menos dois resíduos básicos consecutivos (Arg ou Lys), que geralmente funcionam como um sítio de clivagem para ativação do peptídeo precursor: G-[RK]-[RK] 

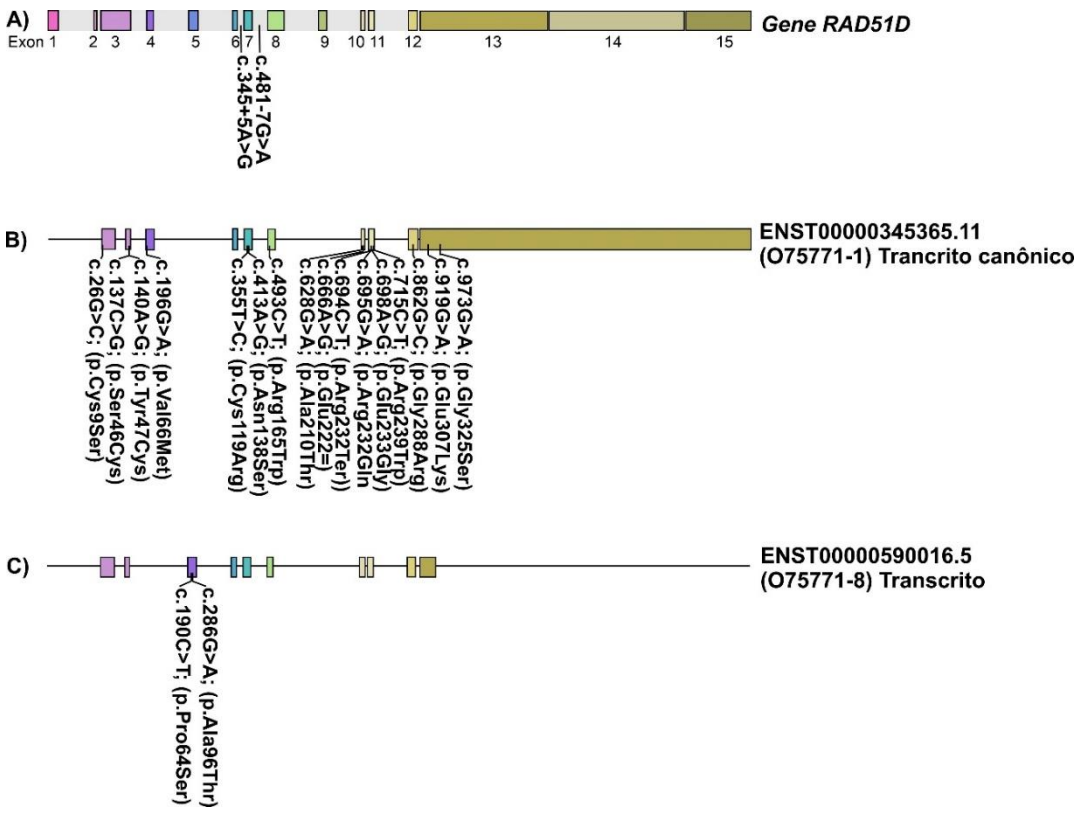
Nota: Consulta realizada em 18 de Setembro de 2022. Predição do score de conservação dos motivos: verde são os motivos de alta conservação, amarelo são os motivos de média conservação e laranja e vermelho são os de baixa conservação. Os motivos estruturais estão espalhados ao longo da proteína e desempenham diferentes funções. Na RAD51D, a maioria dos motivos se sobrepõem. As informações de cada motivo foram extraídas do Eukaryotic Linear Motif resource (ELM) e do Prosite. Fonte: Elaborado pela autora.

3.2 - Variantes no gene *RAD51D*

Foram avaliadas 20 variantes germinativas, apresentadas na Tabela 1. Destas, 16 são variantes de sentido trocado, uma variante sinônima, uma variante *frameshift* e duas variantes intrônicas, estando classificadas nas bases de dados como benigna/provavelmente benigna, VUS e patogênica.

A maioria das variantes ocorrem ao longo do transcrito ENST00000345365.11 (canônica/O75771-1) e duas acontecem no transcrito ENST00000590016.5 (O75771-8) (Fig. 12, Tab. 2). Devido à alta expressão tecidual do transcrito canônico ENST00000345365.11, as variantes que impactam a função desta isoforma têm potencial para afetar o padrão de expressão. As frequências populacionais das variantes foram avaliada no gnomAD e são apresentadas na Tabela 2, juntamente, com a localização molecular (éxons e íntrons) e a atual classificação, de acordo com as bases de dados ClinVar, VarSome, ABraOM e HGVD. Na figura 12, são mostradas a composição dos éxons em relação às isoformas (O75771-1/canônica) e (O75771-8), em nível de proteína.

Figura 12 - Localização das variantes em relação ao gene e aos transcritos, ENST00000345365.11 (O75771-1/canônico) e ENST00000590016.5 (O75771-8)



Nota: Em A, estrutura do gene RAD51D, composto por 15 éxons, cada um com suas respectivas cores. Duas variantes, que afetam introns, também estão sendo representadas. Em B, representação esquemática dos éxons, que compõem o transcrito canônico (ENST00000345365.11). As variantes que

afetam este transcrito também são representadas, respeitando sua localização exônica. Em C, representação esquemática do transcrito ENST00000590016.5. As variantes, que afetam este transcrito, também são representadas de acordo com suas posições exônicas.

Fonte: Elaborada pela autora.

Figura 13 - Representação esquemática dos éxons que compõem as isoformas O75771-1/canônica e O75771-8



Nota: Em A, a isoforma O75771-1/canônica é composta pelos éxons 3, 4, 6, 7, 8, 10, 11, 12 e 13, cada um com suas respectivas cores. Em B, a isoforma O75771-8 é composta pelos éxons 3, 5, 6, 7, 8, 9, 10, 11, 12 e 13. Esta isoforma tem a ausência do éxon 4 quando comparada com a isoforma canônica.

Fonte: Elaborada pela autora.

Na tabela 2, são mostradas as variantes analisadas neste estudo, incluindo tipo de variante, localização, frequência e classificação conforme diferentes ferramentas ou bases de dados. Observa-se que 80% (16) das variantes são de sentido trocado a, seguidas de 5% (1) de variantes sinônimas, 5% (1) de variantes com ganho de códon parada e 10% (2) de variantes intrônicas. A frequência populacional destas variantes variou de 0.01 a 0.000008. Cerca de 35% (7) das variantes já são classificadas como benignas pelo ClinVar (Tab. 2).

Tabela 2 - Descrição da consequência molecular, frequência populacional, localização molecular e classificação de cada variante *RAD51D* analisada neste estudo

Variantes RAD51D	Consequência Molecular	HGVS	Localização	Frequência	ClinVar	VarSome	ABraOM	Citações
c.26G>C; (p.Cys9Ser) rs140825795	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E3 mRNA:E1	0.0004	VUS	B	VUS	Wickramanayake *et al.*, 2012 Gutierrez-Enrıquez *et al.*, 2014; Song *et al.*, 2015; Yurgelun *et al.*, 2017; Hauke *et al.*, 2018 Madeddu *et al.*, 2019; Taghizadeh *et al.*, 2020
c.137C>G; (p.Ser46Cys) rs587780102	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E3 mRNA:E1	0.00008	VUS	PB	VUS	Wickramanayake *et al.*, 2012 Alenezi *et al.*, 2022
c.140A>G; (p.Tyr47Cys) rs776528342	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E3 mRNA:E1	0.000004	VUS	VUS	-	Não encontrada na literatura
c.190C>T; (p.Pro64Ser) rs201506572	Sentido trocado	NM_001142571/ ENST00000590016.5	Gene:E5 mRNA:E2	0.0004	PB	B	PB	Não encontrada na literatura
c.196G>A; (p.Val66Met) rs56026142	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E4 mRNA:E2	0.0003	PB	PB	VUS	Yurgelun *et al.*, 2017; Hauke *et al.*, 2018, Wei *et al.*, 2019 Tomioka *et al.*, 2021
c.286G>A; (p.Ala96Thr) rs147933658	Sentido trocado	NM_001142571/ ENST00000590016.5	Gene: E5 mRNA: E2	0.0003	B	PB	B	Não encontrada na literatura

c.345+5A>G rs878854562	Intrônica	NM_002878.4/ ENST00000345365.11	Gene:I6	0.000004	VUS	VUS	-	Não encontrada na literatura
c.355T>C; (p.Cys119Arg) rs201313861	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E7 mRNA:E4	0.000053	VUS	B	VUS	Osher *et al.*, 2012; Gutierrez-Enrıquez *et al.*, 2014; Song *et al.*, 2015; Penkert *et al.*, 2018
c.413A>G; (p.Asn138Ser) rs201676898	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E7 mRNA:E4	0.0001	VUS	B	VUS	Johnson *et al.*, 2014; Tung *et al.*, 2014; Ring *et al.*, 2016; Shindo *et al.*, 2017; Sánchez-Bermúdez *et al.*, 2018; Tsaousis *et al.*, 2019
c.481-7G>A rs145832514	Intrônica	NM_002878.4/ ENST00000345365.11	Gene:I7	0.0005	B	B	B	Tung *et al.*, 2014 Yurgelun et al., 2015
c.493C>T; (p.Arg165Trp) rs544654228	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E8 mRNA:E5	0.00002	VUS	PB	VUS	Loveday *et al.*, 2011; Gutierrez-Enrıquez *et al.*, 2014; Song *et al.*, 2015; Dominguez-Valentin *et al.*, 2018
c.628G>A; (p.Ala210Thr) rs80116829	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E10 mRNA:E6	0.0005	VUS	B	VUS	Song *et al.*, 2015; Tung *et al.*, 2015; Yadav *et al.*, 2016; Ricker *et al.*, 2017; Yurgelun *et al.*, 2017

c.666A>G; (p.Glu222=) rs114012742	Sinônima	NM_002878.4/ ENST00000345365.11	Gene:E10 Trans:E6	0.0004	B	B	B	Não encontrada na literatura
c.694C>T; (p.Arg232Ter) rs587780104	Sem sentido	NM_002878.4/ ENST00000345365.11	Gene:E11 mRNA:E7	0.00003	P	P	P	Wickramanayake *et al.*, 2012; Pennington *et al.*, 2013; Gutierrez-Enrıquez *et al.*, 2014; Baker *et al.*, 2015; Susswein *et al.*, 2016; Tedaldi *et al.*, 2017; Sánchez-Bermúdez *et al.*, 2018; Suszynska *et al.*, 2020 Suszynska *et al.*, 2020; Tedaldi *et al.*, 2020; Alenezi *et al.*, 2022
c.695G>A; (p.Arg232Gln) rs28363283	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E11 mRNA:E7	0.002	B	B	B	Loveday *et al.*, 2011; Wickramanayake *et al.*, 2012; Gutierrez-Enrıquez *et al.*, 2014; Pratz *et al.*, 2016; Yurgelun *et al.*, 2017
c.698A>G; (p.Glu233Gly) rs28363284	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E11 mRNA:E7	0.01	B	B	B	Rodríguez-López *et al.*, 2004 Dowty *et al.*, 2008; Nadkarni *et al.*, 2009; Jara *et al.*, 2010; Loveday *et al.*, 2011; Wickramanayake *et al.*, 2012

								Osher *et al.*, 2012; Gutierrez-Enrıquez *et al.*, 2014; Song *et al.*, 2015; Frimer *et al.*, 2016 Maxwell *et al.*, 2016; Dominguez-Valentin *et al.*, 2018
c.715C>T; (p.Arg239Trp) rs770250516	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E11 mRNA:E7	0.00006	VUS	B	VUS	Wickramanayake *et al.*, 2012 Gutierrez-Enrıquez *et al.*, 2014; Song *et al.*, 2015; Sánchez-Bermúdez *et al.*, 2018 Hauke *et al.*, 2018
c.862G>C; (p.Gly288Arg) rs1391505912	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E12 mRNA:E8	0.000008	VUS	PB	-	Song *et al.*, 2015
c.919G>A; (p.Glu307Lys) rs115031549	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E13 mRNA:E9	0.0003	VUS	B	-	Tung *et al.*, 2015
c.973G>A; (p.Gly325Ser) rs587780106	Sentido trocado	NM_002878.4/ ENST00000345365.11	Gene:E13 mRNA:E9	0.000008	VUS	PB	VUS	Tung *et al.*, 2015 Yurgelun *et al.*, 2015

Nota: NM= Nomenclatura; Trans=Transcrito; E=éxon; I=Intron; B=Benigna; PB=Provavelmente benigna; VUS= Variante de significado incerto; P=Patogênica. Para a análise da frequência populacional foram consideradas apenas dados de população latino-americano e semelhantes.

3.2.1 Conservação evolutiva

Para a análise da conservação evolutiva dos resíduos afetados por variantes de sentido trocado, alinhamentos múltiplos foram realizados entre a isoforma canônica (O75771-1) de humanos e isoformas de especies relacionadas, como mamíferos, aves, anfíbios e répteis e peixes. Das variantes identificadas neste estudo, 14 afetam a isoforma canônica e seus resultados em relação à conservação evolutiva são apresentados na tabela 3 e nos alinhamentos 1, 2, 3 e 4.

Aqui, é importante ressaltar que a presença de múltiplas isoformas da RAD51D, assim como a existência de ortólogos e parálogos dificulta o alinhamento, distorcendo a avaliação da conservação evolutiva de cada resíduo. Para evitar estas distorções, cada sequência usada nos alinhamentos foi conferida por inspeção visual e alinhamento par a par com a sequência canônica, só sendo incluída nos alinhamentos se efetivamente correspondesse à isoforma canônica. Em espécies mais distantes do que vertebrados, a RAD51D apresenta, geralmente, homologia menor do que 30%. Em vertebrados, abaixo de 70 % de conservação começaram a surgir sequências de outros genes da família RAD, pelo que, sequências com menos de 70% de homologia não foram incluídas nos alinhamentos múltiplos.

Tabela 3 - Conservação evolutiva dos resíduos afetados pelas variantes de sentido trocado em relação aos mamíferos, aves, anfíbios e répteis e peixes

Variantes *RAD51D*	**Espécies**			
	Mamíferos	**Aves**	**Anfíbios e répteis**	**Peixes**
Isoforma canônica				
c.26G>C; (p.Cys9Ser)	Altamente conservada	Altamente conservada	Altamente conservada	Altamente conservada
c.137C>G; (p.Ser46Cys)	Altamente conservada	Altamente conservada	Altamente conservada	Altamente conservada
c.140A>G; (p.Tyr47Cys)	Altamente conservada	Altamente conservada	Altamente conservada	Altamente conservada (uma única vez substituída por uma fenilalanina)

c.196G>A; (p.Val66Met)	Não conservada, substituída por Phe, Ile, Ser, Leu e Cys	Não conservada, substituída por Ala, Ile e Thr	Altamente conservada	Altamente conservada
c.355T>C; (p.Cys119Arg)	Conservada, substituída por Phe, Tyr, Arg e His	Não conservada, substituída por Gly e Ser	Não conservada, substituída por Ser	Não conservada Substituída por Gly e Arg
c.413A>G; (p.Asn138Ser)	Pouco conservada Substituída por Thr e Ser	Não conservada Substituída por Thr	Não conservada Substituída por Thr	Não conservada Substituída por Thr
c.493C>T; (p.Arg165Trp)	A Gln é o resíduo típico nesta posição. Quando substituído, entra a Arg	A Gln é o resíduo típico nesta posição	A Gln é o resíduo típico nesta posição	A Gln é o resíduo típico nesta posição
c.628G>A; (p.Ala210Thr)	Altamente conservada	Altamente conservada (raramente substituída por um serina)	Altamente conservada	Altamente conservada
c.695G>A; (p.Arg232Gln)	Conservada. Substituída por Leu e Gln	Altamente conservada (raramente substituída por um lisina)	Altamente conservada	Não conservada, substituída por Gly
c.698A>G; (p.Glu233Gly)	Altamente conservada (uma única vez substituída por um ácido aspártico)	Altamente conservada	Altamente conservada	Não conservada, substituída por Val, Ala e Ile
c.715C>T; (p.Arg239Trp)	Conservada, substituída por Leu, Trp e Gln	Conservada, substituída por Lys	Altamente conservada	Não conservada, substituída por Lys
c.862G>C; p.Gly288Arg)	Conservada, substituída por Ala, Ser, Hys e Asp	Não conservada, substituída por Thr, Ser, Cys, Asn, Pro e Met	Não conservada, substituída por Ser	Não conservada, substituída por Ser e Pro
c.919G>A; (p.Glu307Lys)	Altamente conservada	Não conservada, substituída por Val	Não conservada, substituída por Val	Altamente conservada (raramente substituída por uma alanina)

c.973G>A; (p.Gly325Ser)	Conservada, substituída por Ala,Gln, Arg, Val, Glu e Asp	Não conservada, substituída por Thr e Ala	Altamente conservada (uma única vez substituída por uma alanina)	Não conservada, substituída por Arg

Nota: Altamente conservados: resíduos que são raramente substituídos por outros aminoácidos (raramente = até três espécies apresentaram aminoácidos divergentes); conservados: resíduos que, mesmo substituídos por outros aminoácidos, estavam presentes na maioria das sequências alinhadas; pouco conservados: resíduos frequentemente substituídos por outros e presentes na minoria das sequências alinhadas; e, não conservados: resíduos que não foram encontrados em nenhuma das sequências alinhadas.

Como está sendo mostrado nos alinhamentos 1, 2, 3 e 4 (Anexo 4), os resíduos das variantes c.26G>C; (p.Cys9Ser), c.137C>G; (p.Ser46Cys) e c.140A>G; (p.Tyr47Cys) foram classificados como altamente conservados em todas espécies avaliadas. Os resíduos das variantes c.628G>A; (p.Ala210Thr), c.695G>A; (p.Arg232Gln), 698A>G; (p.Glu233Gly) e c.715C>T; (p.Arg239Trp) também foram conservados em mamíferos, aves e anfíbios e répteis. Entretanto, os resíduos das variantes c.196G>A; (p.Val66Met), c.413A>G; (p.Asn138Ser), c.493C>T; (p.Arg165Trp) e c.862G>C; p.Gly288Arg) foram pouco ou não conservados em mamíferos. Entretanto, características específicas da conservação evolutiva vão ser abordadas em maior detalhe na descrição de cada variante, abaixo.

Das 16 variantes de sentido trocado, duas variantes *RAD51D* c.190C>T; (p.Pro64Ser) e *RAD51D* c.286G>A; (p.Ala96Thr) afetam a isoforma 8 (O75771-8) em vez da isoforma canônica (O75771-1). A isoforma 8 (O75771-8) é produzida por *splicing* alternativo e, nos bancos de dados de proteínas, só foi encontrada em humanos e algumas espécies de primatas, como *Pan troglodytes* (chimpanzés), *Nomascus leucogenys* (gibão de bochecha branca), *Hylobates moloch* (gibão prateado). Contudo, a limitação de dados sobre a isoforma 8 em bancos de dados, resulta em uma análise de conservação evolutiva pouco aprofundada, não se podendo afirmar de forma assertiva se esses resíduos são ou não profundamente conservados (Tabela 4; alinhamento 5 no anexo 5).

Tabela 4 - Conservação evolutiva de resíduos afetados por variantes de sentido trocado, que ocorrem na isoforma 8, considerando as espécies de mamíferos, aves, anfíbios e répteis e peixes

Variantes *RAD51D*	**Espécies**			
	Mamíferos	**Aves**	**Anfíbios e répteis**	**Peixes**
Isoforma 8 (não canônica)				
c.190C>T; (p.Pro64Ser)*	Conservada*	Isoforma não encontrada nessas espécies	Isoforma não encontrada nessas espécies	Isoforma não encontrada nessas espécies
c.286G>A; (p.Ala96Thr)*	Conservada*	Isoforma não encontrada nessas espécies	Isoforma não encontrada nessas espécies	Isoforma não encontrada nessas espécies

Nota: O (*) indica as variantes que afetam a isoforma 8; Conservados: os resíduos que, mesmo substituídos por outros aminoácidos, estavam presentes na maioria das sequências alinhadas.

3.2.2 Estudo das variantes germinativas de sentido trocado

O objetivo principal deste estudo é a reclassificação das variantes de RAD51D identificadas nesta amostra. No processo de reclassificação são avaliadas características de cada variante assim como características gerais do gene em estudo, que afetam a classificação de todas as variantes desse gene. Os critérios influenciados por características gerais do gene são o *PP2* e o BP1. A seguir, vamos avaliar a aplicação destes critérios no contexto do *RAD51D*.

PP2 - Variante missense em um GENE, que tem uma baixa taxa de variação missense benigna e na qual variantes missense são um mecanismo comum de doença: O gene *RAD51D* possui 1307 variantes descritas no ClinVar, incluindo sinônimas, não sinônimas, intrônicas e truncantes. Das 1307, 749 são VUS, das quais 553 são variantes não sinônimas. São descritas 388 variantes benignas ou provavelmente benignas para *RAD51D*, mas a maioria delas são intrônicas ou sinônimas, e apenas 12 são não-sinônimas (acesso em novembro 2022). Geralmente, no ClinVar e em outras bases de dados de classificação de variantes, apenas as variantes truncantes (*fsm* e sem sentido) são classificadas como patogênicas. Este viés faz com que a taxa de variantes de sentido trocado classificadas como benignas seja baixa. Isso vale para as variantes patogênicas e isto impediria que se use o critério alternativo ao PP2, que

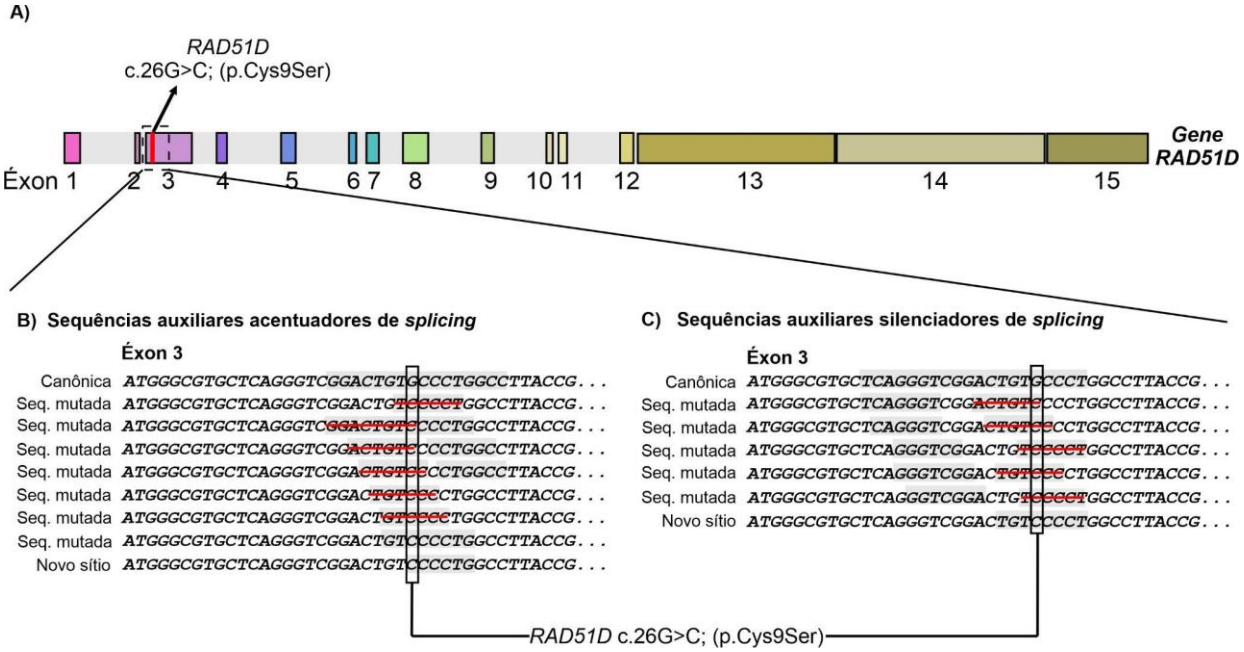
seria o **BP1** (Variantes de sentido trocado em um gene caracterizado, primariamente, por variantes truncantes serem causadores de doenças).

Além disso, a maioria as variantes não sinônimas classificadas como VUS de acordo com ClinVar, para este gene, foram relatadas em associação com câncer de mama e/ou ovário, familiar ou isolado, outros tipos de câncer (e.g., câncer colorretal) ou com síndromes hereditárias de predisposição ao câncer. Outro fator importante é que as 12 de 14 variantes de sentido trocado avaliadas neste estudo foram descritas na literatura também associadas ao câncer de mama e/ou ovário ou a outros tipos de cânceres (Wickramanayake *et al.,* 2012, Gutierrez-Enrıquez *et al.,* 2014; Song *et al.,* 2015; Yurgelun *et al.,* 2017; Hauke *et al.,* 2018 Madeddu *et al.,* 2019; Wei *et al.,* 2019; Taghizadeh *et al.,* 2020; Alenezi *et al.,* 2022).
Com base nos argumentos mencionados acima, consideramos indicado usar o critério PP2 na avaliação das variantes do *RAD51D*.

3.2.2.1 *RAD51D* c.26G>C;(p.Cys9Ser)

Localizada no éxon 3 do gene e no éxon 1 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2) esta variante resulta de uma substituição de G para C na posição nucleotídica 26. Para essa substituição, foram previstos efeitos significativos nas sequências auxiliares reguladoras de *splicing*, sugerindo a perda e criação de alguns sítios acentuadores e silenciadores (Fig. 14).

Figura 14 - Impacto da *RAD51D* c.26G>C;(p.Cys9Ser) nas sequências auxiliares regulatórias de sítios de *splicing*

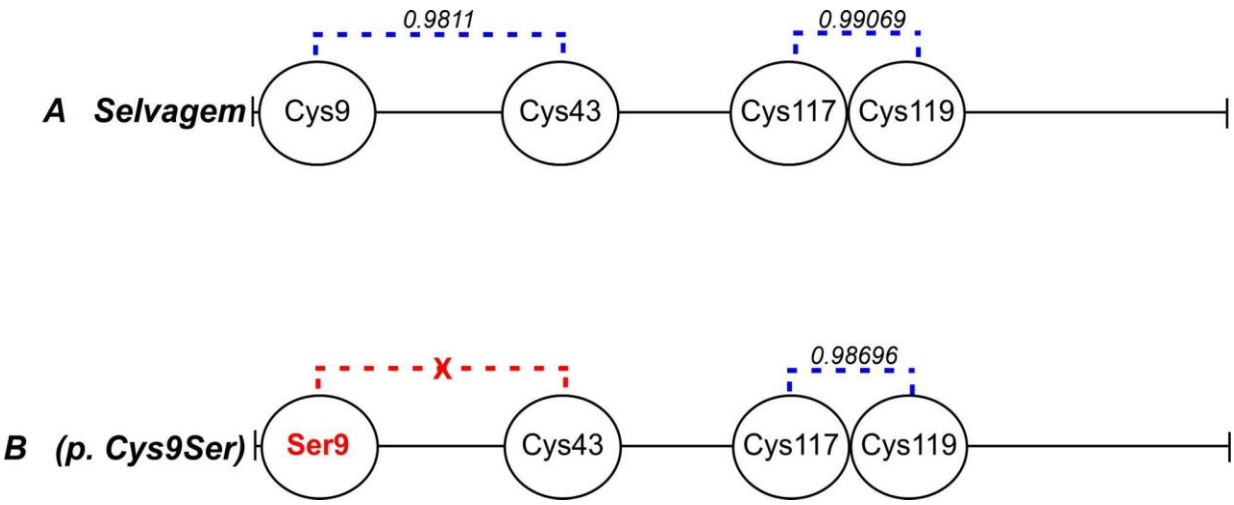


Nota: Em A, está sendo mostrado o arranjo dos éxons ao longo do gene RAD51D. *RAD51D* tem 15 éxons (cada um com uma cor atribuída). A variante *RAD51D* c.26G>C;(p.Cys9Ser) ocorre no éxon 3 do gene e sua localização pode ser vista por meio da seta indicando a variante e por uma linha vertical vermelha. Nas figuras B e C são mostradas as sequências auxiliares, acentuadores ou silenciadores, que foram afetados após a troca de nucleotídeos. Em B, está sendo mostrado o impacto da variante em relação aos acentuadores de *splicing*. A sequência canônica contém o sítio acentuador de *splicing* de tipo selvagem sublinhado em cinza e abaixo dele estão as sequências mutadas. Alguns sítios não foram afetados mesmo após a troca de nucleotídeos (sequência intacta), outros foram abolidos (representados por uma linha vermelha sobre a sequência) ou criados (novo sítio). Por outro lado, em C, está sendo mostrado o impacto da variante em relação aos silenciadores de *splicing*. A sequência canônica contém o sítio silenciador de *splicing* de tipo selvagem e abaixo estão as sequências mutadas. Na imagem é mostrado que sítios próximos à variante não foram afetados após a troca de nucleotídeos, mas alguns sítios envolvendo a variante foram abolidos e um novo sítio foi gerado no local.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

A cisteína no resíduo 9 é substituída por serina, um aminoácido com as mesmas propriedades físico-químicas, sendo ambos polares não carregados. Os preditores *in silico* LIST-02, Provean, Polyphen, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos deletérios** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Previu-se que Cys9 seria um potencial ponte dissulfeto e poderia interagir com Cys43. Contudo, a perda de Cys9 leva a perda dessa interação (Fig. 15).

Figura 15 - Impacto da variante RAD51D c.26G>C;(p.Cys9Ser) sobre pontes de dissulfeto da RAD51D usando Dianna

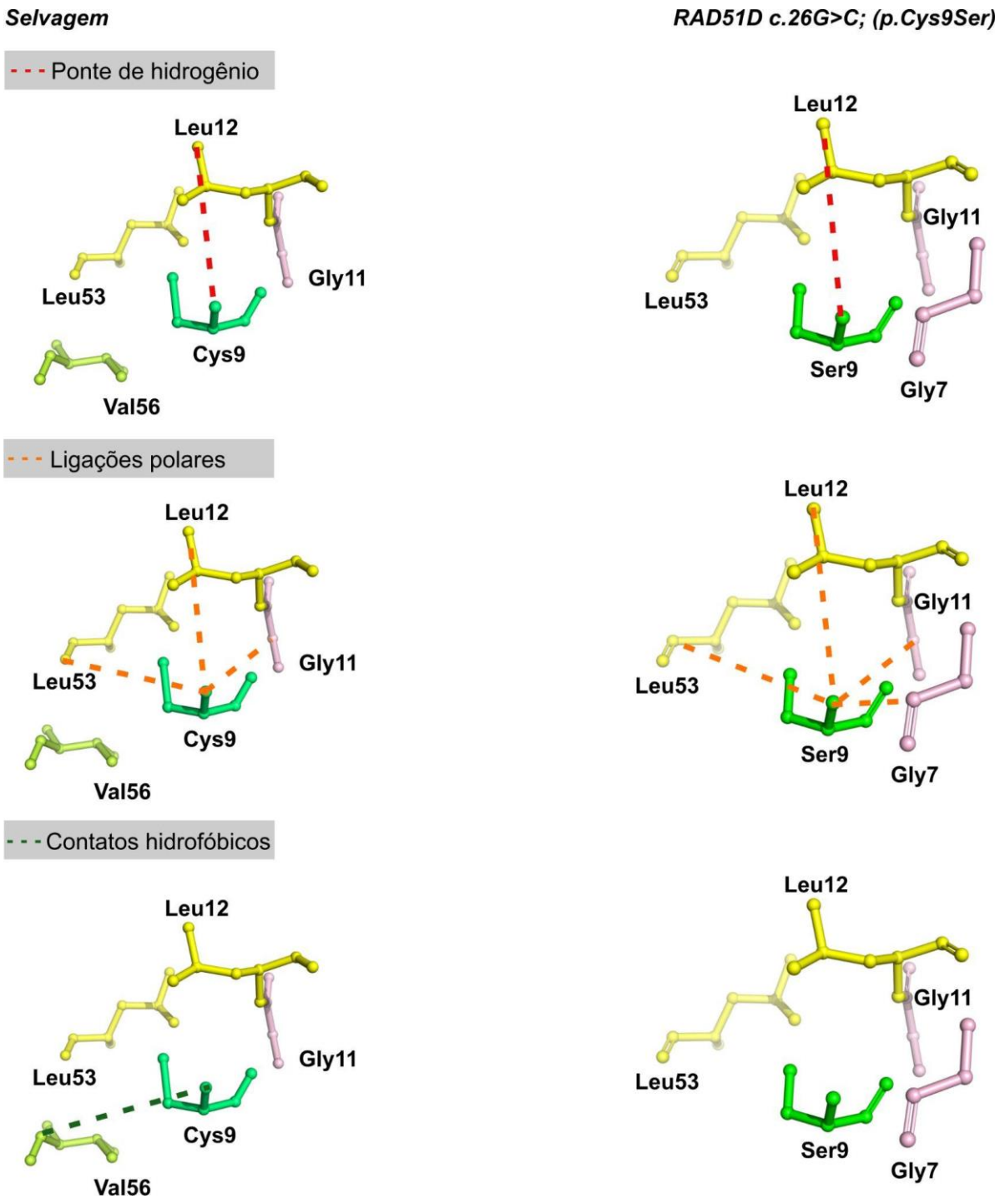


Nota: Em A, estão sendo mostrados os aminoácidos Cys9, Cys43, Cys117, Cys119 e Cys294, resíduos potenciais nas interações de pontes de dissulfeto. As interações são representadas pelas linhas pontilhadas azuis, por exemplo, a Cys9 interage com a Cys43 e assim por diante. Em B, é mostrada a consequência da variante *RAD51D* c.26G>C; (p.Cys9Ser) sobre as pontes dissulfeto, na qual a linha pontilhada vermelha e o **X** indica a perda da ligação covalente.

Fonte: Elaborada pela autora.

Foi previsto que a Cys9 faça ponte de hidrogênio (Cys9 → Leu12), contato hidrofóbico (Cys9 → Val56), ligações polares (Cys9 → Gly11, Cys9 → Leu12, Cys9 → Leu53) (Fig. 16). As interações, tais como ponte de hidrogênio e interações polares, se mantiveram mesmo após a troca de aminoácidos (Cys9Ser) (Fig. 16). Porém, a entrada da Ser9 conduz a uma nova interação polar (Ser9 → Gly7). Também levou à perda do contato hidrofóbico (Cys9 → Val56) (Fig. 16). Essas alterações são previstas para desestabilizar a estrutura da proteína e desestabilizar a habilidade de interação com o DNA e proteínas (Quadro 6, na seção 3.2.6).

Figura 16- Dinâmica molecular da variante *RAD51D* c.26G>C; (p.Cys9Ser)

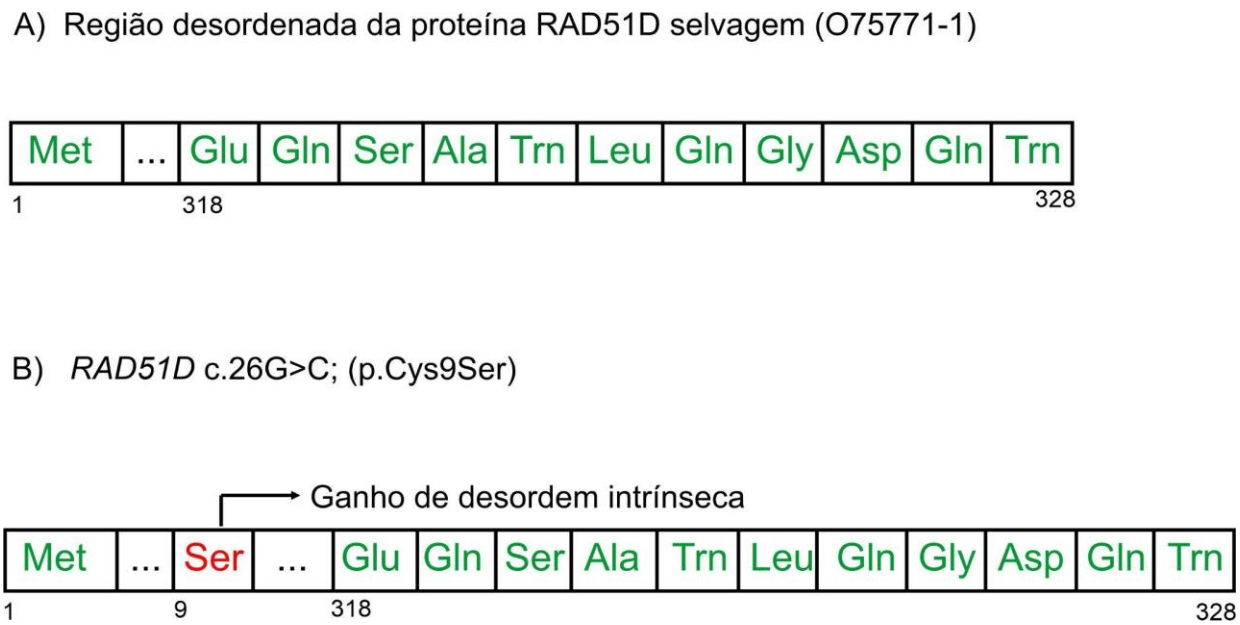


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Cys9, Leu12 e Leu53, Gly11 e Val56. Na estrutura mutada, além destes resíduos citados anteriormente, também são vistos novos aminoácidos: Gly7 e Ser9.

Fonte: Elaborada pela autora, usando DynaMut.

A perda de Cys9 também promove o ganho de desordem intrínseca da proteína (Fig. 17). Além disso, ela desestabiliza a estrutura da proteína e desestabiliza a habilidade de interação com o DNA e proteínas (Quadro 6, na seção 3.2.6).

Figura 17- Impacto da variante RAD51D c.26G>C; (p.Cys9Ser) na desordem proteica



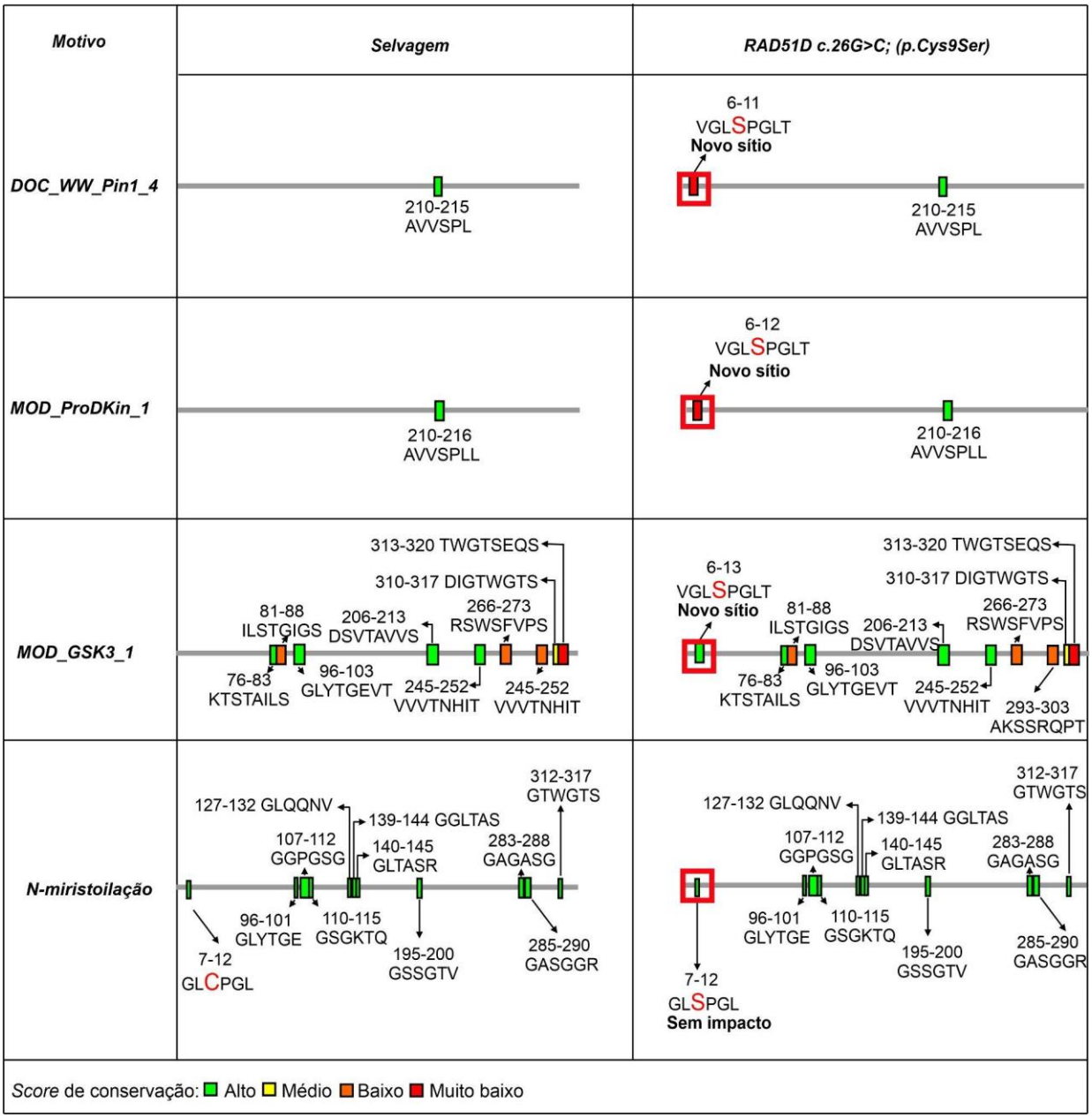
Nota: As informações sobre desordem intrínseca são mostradas com detalhes na parte de anotação funcional do gene e proteína. Aqui, está sendo mostrado o impacto da variante sobre a desordem, onde em A estão alguns dos aminoácidos preditos como intrinsecamente desordenados na sequência selvagem. Em B, está sendo mostrado que a entrada da Ser9 promove um ganho de desordem intrínseca (aminoácido vermelho na sequência) para essa região.

Fonte: Elaborada pela autora, usando PSIPred.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Cys9 está localizada em um sítio de ligação preferencial ao ssDNA (resíduos 1-83) (Quadro 6, na seção 3.2.6), bem como no motivo *MYRISTYL (sítio de N-miristoilação - mecanismo de acilação)* na posição 7-12 [GLCPGL] (Fig. 18). Para esse motivo, não foi previsto impacto, pois a Cys9 está situada na terceira posição do motivo, uma posição que permite a entrada de qualquer aminoácido, especialmente a serina (Prosite, 2022). Por outro lado, a entrada da Ser9 cria um novo sítio de baixa conservação neste motivo *DOC_WW_Pin1_4 (Motivos de ligação ao domínio WW - medeiam a interação de proteínas sinalizadoras e regulatórias)* na posição 6-11 [VGLSPG]; um novo sítio de baixa

conservação do motivo *MOD_ProDKin_1* (*sítio de fosforilação MAPK*) na posição 6-12 [VGLSPGL] (Fig. 18); e, um novo sítio de alta conservação do motivo *MOD_GSK3_1 (sítio de fosforilação GSK3) na posição* 6-13 [VGLSPGLT] (Fig. 18). Nenhuma evidência foi encontrada na literatura de uma possível repercussão funcional desses motivos selvagens e/ou mutados em RAD51D.

Figura 18 - Predição de impacto de *RAD51D* c.26G>C; (p.Cys9Ser) nos motivos proteicos

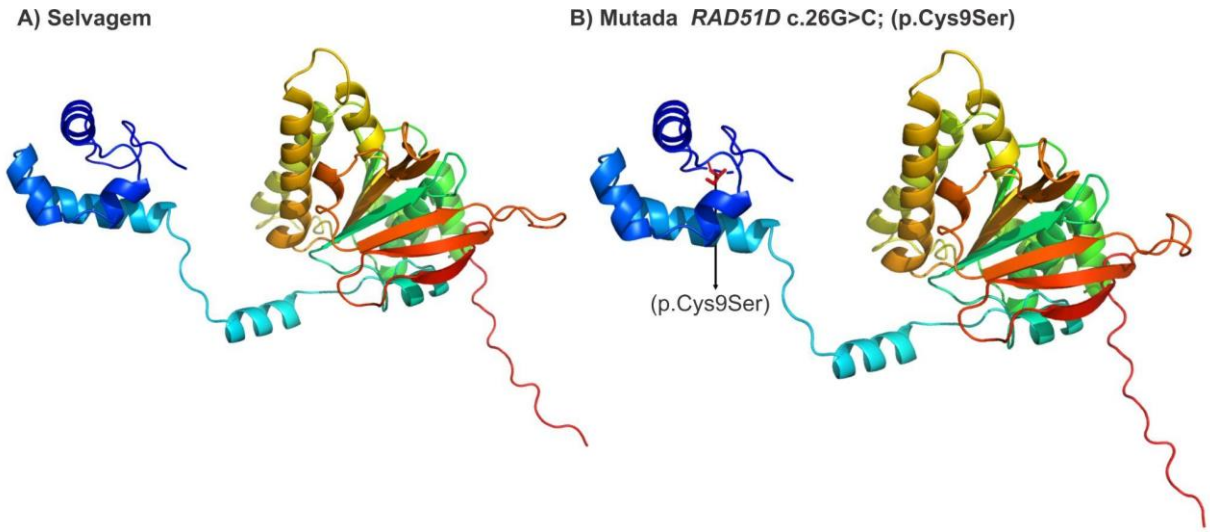


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde

a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo. Usando ELM motifs.
Fonte: Elaborada pela autora, usando ELM motifs.

Considerando-se a estrutura 3D da proteína, previu-se que Cys9 está localizada em uma alça e, após a troca de aminoácidos, não foram encontrados impactos significativos da variante (Fig. 19).

Figura 19- Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.26G>C; (p.Cys9Ser))



Nota: Em A, mostra a estrutura da proteína 3D selvagem em relação a estrutura mutada que é indicada na figura B. A *RAD51D* c.26G>C; (p.Cys9Ser) está localizada em uma alça, indicada por uma seta. A estrutura 3D não mostrou efeitos significativos.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.1.1 Reclassificação da variante *RAD51D* c.26G>C; (p.Cys9Ser)

Considerando as informações acima, propomos a reclassificação da variante *RAD51D* c.26G>C; (p.Cys9Ser) de VUS para provavelmente patogênica, segundo os critérios ACMG/AMP: PS4, PM1, PP2 (ver na Seção 3.2.2), PP3.

PS4 - Prevalência da variante em indivíduos afetados é significativamente aumentada em comparação com a prevalência em controles

Como esperado para variantes em genes de penetrância moderada, (p.Cys9Ser) tem sido descrita tanto em afetados quanto em controles. Além disto, tem sido descrita em pacientes com diferentes tipos de câncer. Para avaliar a diferença de frequência entre casos e controles, buscamos nos artigos científicos que descrevem esta variante (Tab. 5).

Tabela 5 - Comparação da frequência de (p.Cys9Ser) entre casos e controles

Variável / **Artigo/Tipo de câncer**	**Frequências alélicas**	
	Casos	**Controles**
Wickramanayake *et al.*, 2012*/ *ovarian carcinoma*	2alelos/1822 sujeitos ou seja 2 alelos/3644 alelos 0,00055 (frequência alélica)	1/2120 ou seja 1 alelo/4240 alelos 0,000236
Gutierrez-Enríquez *et al.*, 2013 *breast and ovarian Spanish families testing*	1 alelo/ 842 indivíduos 1 alelo /1684 alelos 0,0006 (frequência alélica)	-
Song *et al.*, 2015	5 alelos/3429 pacientes 5 alelos/6858 alelos 0,00073 (frequência alélica)	2 alelos/2772 indivíduos 2/5544 alelos 0,00036
Yurgelun *et al.*, 2017	1 alelo/ 1058 pacientes com Câncer colorretal 1 alelo /2116 alelos 0,00047 (frequência alélica)	-
Hauke *et al.*, 2018	1 alelo/5589 casos 1 alelo/11178 alelos 0,000089 (frequência alélica)	?/2089 controles**

Total	15/25480	3/9784

Nota: *Dados de Loveday e colaboradores não foram incluídos, porque os controles aparentemente se sobrepõem aos de Wickramanayake *et al.*, 2012. **Os dados de Hauke e colaboradores sobre os controles não foram incluídos, porque não informam especificamente que nenhum alelo (p.Cys9Ser) foi identificado no grupo controle.

Para avaliarmos se a diferença das frequências alélicas entre casos e controles é estatisticamente significativa, foi feita a estimativa do tamanho necessário da amostra usando-se o samplesize.aspx (www.clincalc.com/stats/samplesize.aspx), com os seguintes parâmetros: frequência nos casos; 0,0006; frequência nos controles: 0,0003; alfa: 0,05, beta: 0,2; poder: 0,8. O tamanho necessário estimado foi de 4.352.266 indivíduos. Este tamanho é inatingível, considerando-se que os maiores estudos envolvem cerca de 1 milhão de pacientes (Mersch et al., 2018), todo são caso-controle, a seleção dos controles busca representar uma amostra semelhante à dos casos, o que frequentemente não é atingível em populações híbridas.

Dentro dos dados que conseguimos averiguar, com descrições específicas nos artigos científicos, a frequência alélica nos casos (15/25062=0,0006) é o dobro da descrita nos controles (3/9784=0,0003).

Além dos estudos caso-controle, investigamos a frequência desta variante em bases de dados populacionais. Encontramos as seguintes frequências: 0.0004 (gnomAD); 0.0007 (Exome Project Genome); variante não encontrada (1000 Genomes Project). Entretanto, é difícil usar esta informação, uma vez que estas amostras são misturas de casos e controles.

Esta variante também foi descrita em dois relatos de caso. Madeddu e colaboradores (2019) descreveram esta variante em uma paciente de 48 anos com câncer cervical. Taghizadeh e colaboradores (2020) identificaram esta variante em 2/50 tumores de glândula salivar e adenocarcinoma.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em *hotspot* de mutações, mas afeta sítio de interação transitória um "sítio de ligação preferencialmente ao ssDNA (resíduos 1-83)". Além disso, afeta domínios estruturais curtos, criando novos sítios dos motivos "GSK3 phosphorylation site", WW domain ligands e MAPK Phosphorylation Site. Conforme avaliado na base de dados do

ClinVar, essa região da variante não contém variantes de sentido trocado benignas. Para essa região (resíduos 1-83), as variantes benignas ou provavelmente benignas são sinônimas.

PP3 - Múltiplas linhas de evidência computacional suportam um efeito deletério no gene ou produto gênico

- A Cys9 é altamente conservada. Nenhum outro aminoácido é visto nesta posição em mamíferos, aves, anfíbios e répteis e peixes;
- A maior parte dos preditores *in silico* utilizados (LIST-02, Provean, Polyphen, SIFT e UMD-Predictor) avaliam que esta substituição tenha numerosos efeitos deletérios na proteína, podendo afetar não só a função e a estabilidade da estrutura da proteína, mas também as diferentes características moleculares da RAD51D que se prevê serem funcionais (sítios de interação e pequenos domínios de proteínas);
- Foram previstas alterações nas interações moleculares, como perda de uma ponte dissulfeto e de contatos hidrofóbicos;
- Foi previsto que a troca de nucleotídeos ativa novas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores). Entretanto, ainda não existem estudos na literatura descrevendo o impacto destas modificações.

3.2.2.1.2 Conclusão sobre a reclassificação da variante *RAD51D* c.26G>C; (p.Cys9Ser)

Uma questão relevante aqui é o que causa associação desta variante com os diferentes tipos de câncer, conforme relatado na literatura. A questão que se coloca é o que de fato estaria causando o efeito funcional, se é a entrada de um novo aminoácido ou a perda da cisteína nesta posição. Foi previsto que a substituição da Cys9 conduz a perda de um potencial de pontes de dissulfeto, além de criar novos sítios fosforilação (*MOD_ProDKin_1* e *MOD_GSK3_1*) e um motivo de ligação ao domínio WW *(DOC_WW_Pin1_4)*, que medeiam a interação transitórias entre proteínas sinalizadoras e regulatórias. Essas mudanças podem afetar a conformação 3D da RAD51D, levando à desestabilização da proteína ou mesmo a alterações na conformação, o que pode prejudicar a interação transitória da proteína com outras proteínas, bem como com o DNA. A presença de outras variantes não sinônimas, previamente classificadas como patogênica ou provavelmente patogênica no mesmo resíduo da variante, auxilia na interpretação da patogenicidade. No entanto, a variante patogênica encontrada nesta posição é truncante *RAD51D* c.27C>A;(p.Cys9Ter) (patogênica) e as outras três variantes encontradas neste resíduo foram classificadas como benigna (*RAD51D* c.27C>T;(p.Cys9=) ou VUS (*RAD51D* c.26G > A;(p.Cys9Tyr) e *RAD51D*

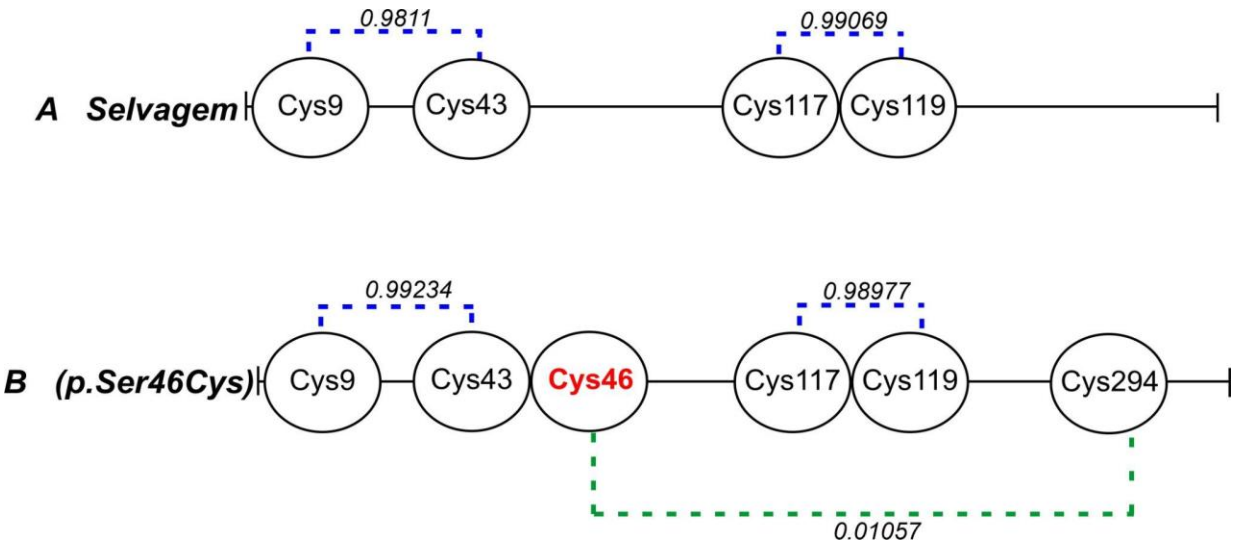
c.26G>T;(p.Cys9Phe). Contudo, embora se espere que a variante tenha efeitos deletérios significativos sobre a proteína, ela só passou a ser reclassificada quando estudos de caso-controle foram encontrados na literatura, contribuindo para a compreensão do risco, já que se trata de uma variante de um gene de penetrância moderada.

3.2.2.2 *RAD51D* c.137C>G;(p.Ser46Cys)

Localizada no éxon 3 do gene e no éxon 1 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), esta variante resulta de uma substituição de C para G na posição nucleotídica 137. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores).

A serina no códon 46 é substituída por uma cisteína, um aminoácido com as mesmas propriedades físico-químicas, sendo ambos polares não carregados. Os preditores *in silico* LIST-02, Provean, Polyphen e UMD-Predictor preveem **efeitos deletérios** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). É importante salientar que a Cys46 (aminoácido substituto) é predita, usando-se o Dianna, como tendo potencial para formar ponte de dissulfeto com a Cys294 (Fig. 20).

Figura 20- Impacto da variante *RAD51D* c.137C>G; (p.Ser46Cys) sobre pontes de dissulfeto da RAD51D usando Dianna



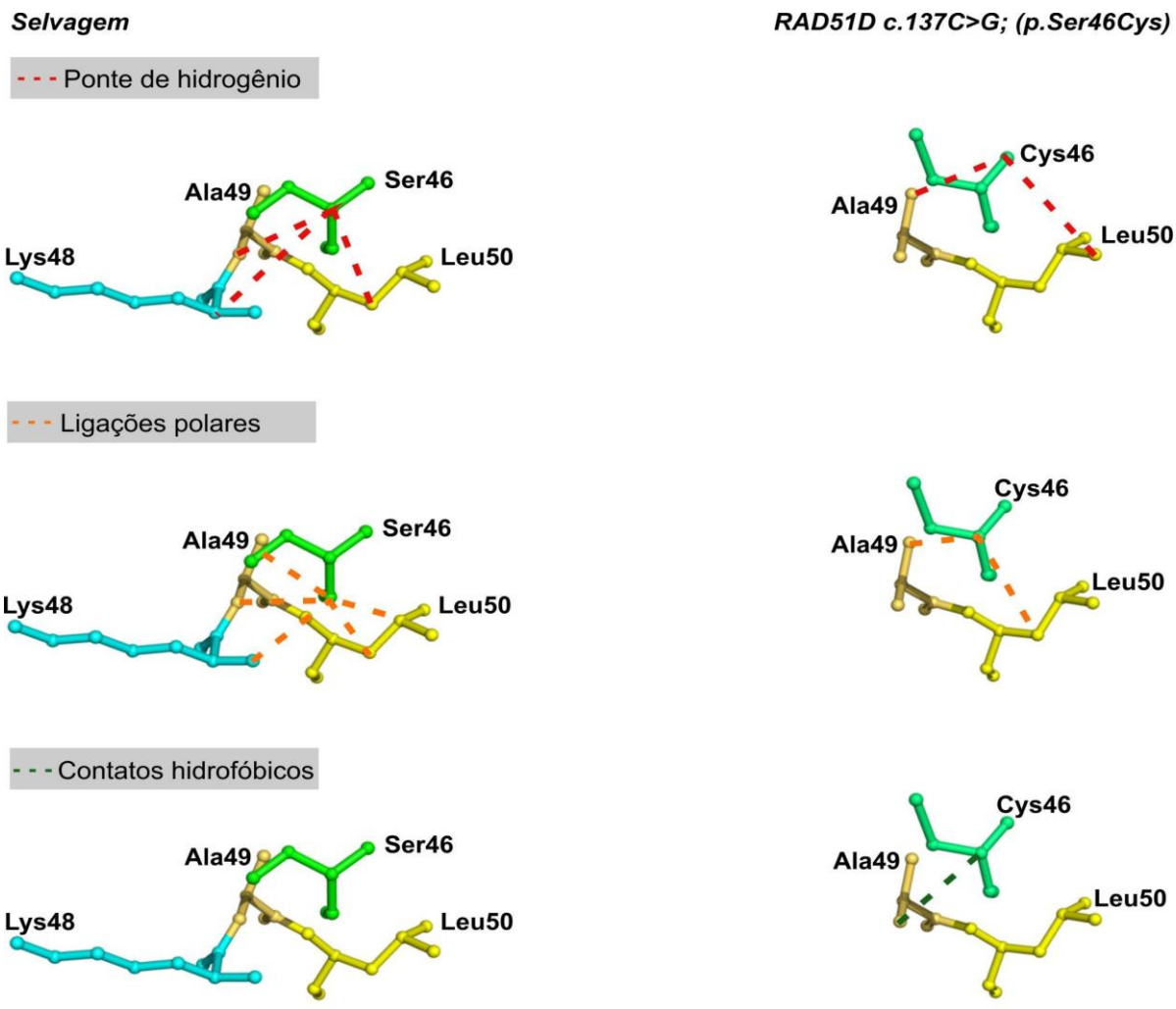
Nota: Em A, estão sendo mostrados os resíduos Cys9, Cys43, Cys117 e Cys119, assim como as potenciais pontes de dissulfeto formado na proteína selvagem. As pontes de dissulfeto são representadas pelas linhas pontilhadas azuis, por exemplo, a Cys9 interage com a Cys43 e assim por diante. Em B, é mostrada a consequência da variante *RAD51D c.137C>G;(p.Ser46Cys)* sobre as

pontes de dissulfeto, na qual a linha pontilhada verde sugere uma nova ligação covalente entre a Cys46 e Cys294. Os números junto às linhas pontilhadas indicam os escores estimados para cada ponte de dissulfeto predita usando-se o Dianna.

Fonte: Elaborada pela autora.

Foi previsto que a Ser46 faça pontes de hidrogênio (Ser46 → Lys48, Ser46 → Ala49 e Ser46 → Leu50) e interações polares (Ser46 → Lys48, Ser46 → Ala49 e Ser46 → Leu50) (Fig. 21). Após a troca de uma serina para uma cisteína, são perdidas uma ponte de hidrogênio (Ser46 → Lys48) e uma interação polar (Ser46 → Lys48); além disso, é gerado um novo contato hidrofóbico entre os resíduos Cys46 → Ala49 (Fig. 21). Estes achados sugerem, que a presença da Cys46 desestabilize a estrutura da proteína, afetando a habilidade de interação com o DNA e proteínas (Quadro 6, na Seção 3.2.6)

Figura 21- Dinâmica molecular da variante *RAD51D* c.137C>G; (p.Ser46Cys)

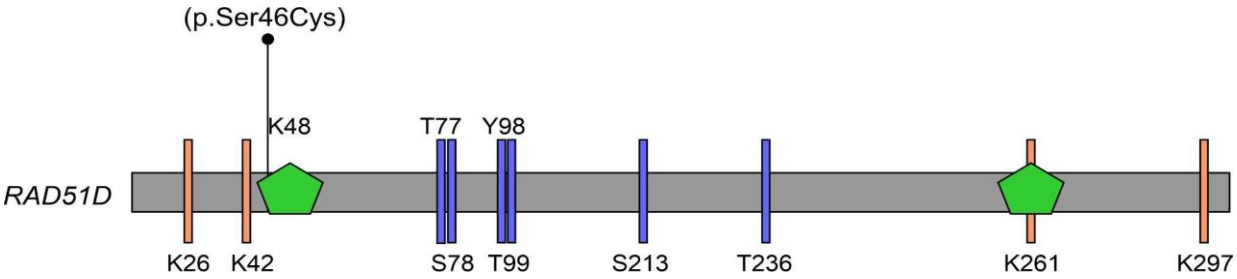


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Ser46, Lys48, Ala49 e Leu50. Na estrutura mutada, perde-se a interação com a Lys48, com consequente afrouxamento da alça.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Ser46 está localizada em um sítio de ligação preferencial ao ssDNA (resíduos 1-83) (Quadro 6, na Seção 3.2.6). Foi predito que a perda da Ser46 conduziria a perda de um sítio de acetilação na Lys48 (Fig. 22). Não encontramos na literatura a descrição dos efeitos da perda da acetilação na Lys48.

Figura 22 - Figura esquemática do efeito da *RAD51D* c.137C>G; (p.Ser46Cys) sobre o sítio de modificações pós-síntese

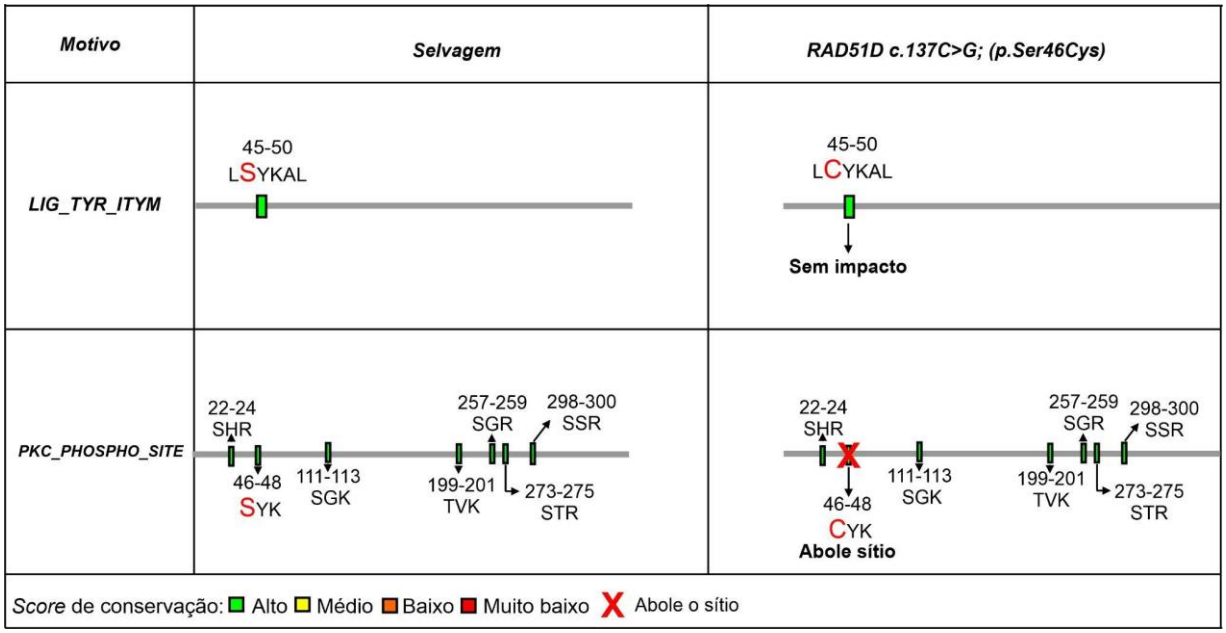


Nota: Foram previstos três tipos de modificações pós-síntese: ubiquitinação (laranja), fosforilação (azul) e acetilação (verde). Os sítios de ubiquitinação acontecem nas lisinas (K); os sítios de fosforilação acontecem em treonina (T), serina (S) e tirosina (Y); e os sítios de acetilação acontecem em duas lisinas. Previu-se que após a troca de aminoácidos, ocorra uma perda do sítio de acetilação da lisina na posição 48 (K48). Usando PhosphoSitePlus.

Fonte: Elaborada pela autora, usando PhosphoSitePlus.

A Ser46 está localizada no motivo *LIG_TYR_ITYM (Motivo inibidor baseado em tirosina imunorreceptor),* na posição 45-50 [LSYKAL], mas não abole este sítio do motivo. A entrada da Cys46 leva a perda do motivo *PKC_PHOSPHO_SITE (sítio de fosforilação da proteína quinase C)* situado na posição 46-48 [SYK] (Fig. 23). Nenhuma evidência foi encontrada na literatura de uma possível repercussão funcional da perda deste motivo desses motivos na RAD51D.

Figura 23 - Predição de impacto de *RAD51D* c.137C>G; (p.Ser46Cys) nos motivos proteicos

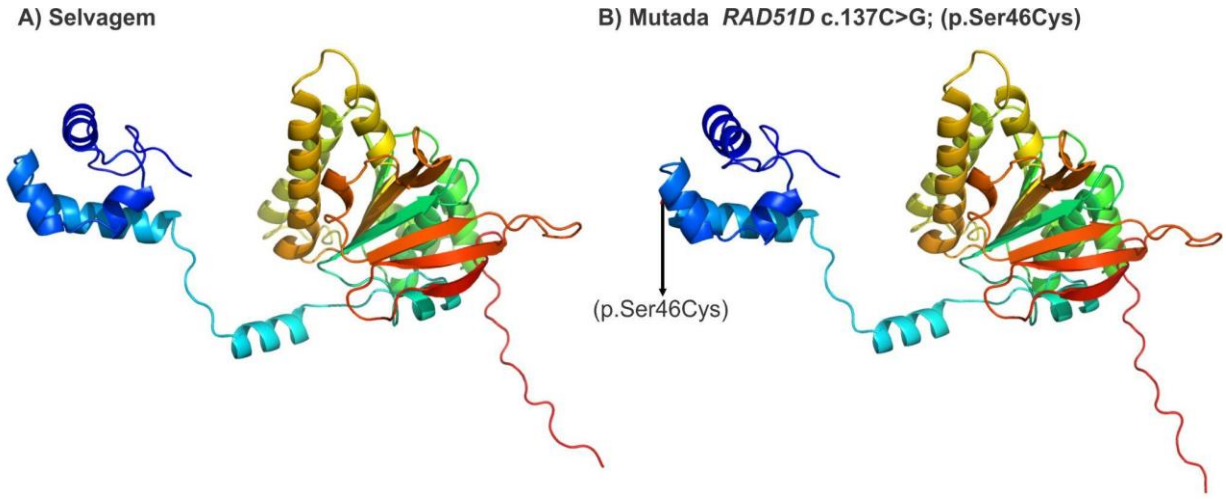


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM Motifs.

A Ser46 está localizada em uma alça e, embora seja substituída por uma cisteína, não foi previsto impacto na estrutura 3D (Fig. 24). Aqui, é importante salientar que os dois *softwares* usados (AlphaFold e Modeller) não permitem informar onde ocorrem as pontes dissulfeto, que foram preditas pelo Dianna e, consequentemente, não conseguem usar esta informação na avaliação do impacto da variante na estrutura 3D da proteína.

Figura 24- Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.137C>G; (p.Ser46Cys))



Nota: Em A, estrutura 3D da proteína selvagem; em B, estrutura mutada, conforme predito usando-se AlphaFold. A *RAD51D* c.137C>G; (p.Ser46Cys) está localizada em uma alça, indicada por uma seta. A estrutura 3D não mostra impactos significativos; entretanto, o programa usado (AlphaFold) não avalia o impacto em pontes de dissulfeto.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.2.1. Reclassificação da variante *RAD51D* c.137C>G; (p.Ser46Cys)

Com base nas informações acima, propomos a reclassificação da variante *RAD51D* c.137C>G; (p.Ser46Cys) de VUS para patogênica, segundo os critérios ACMG/AMP: PS3, PM1, PM2, PP2 (ver na Seção 3.2.2) e PP3.

PS3 - Estudos funcionais bem estabelecidos in vitro ou in vivo que suportam um efeito prejudicial no gene ou produto gênico

Alenezi et al., 2022 investigaram o impacto biológico da variante *RAD51D* (p.Ser46Cys). Por meio da cultura celular, eles descobriram que as células que expressam *RAD51D* (p.Ser46Cys) eram mais sensíveis ao olaparibe (fármaco que impede reparos do DNA em célula neoplásica, levando à sua morte), sendo incapazes de resgatar a viabilidade celular. Além disso, eles viram que as células transfectadas com a variante apresentavam uma expressão menor da RAD51D quando comparadas ao tipo selvagem, isso se deve à instabilidade da proteína. Também avaliaram os efeitos biológicos quanto à formação de complexos RAD51 e, como esperado, observaram uma diminuição no número de focos RAD51 por célula expressando a variante. Em resumo, a variante *RAD51D* (p.Ser46Cys) é

instável, prejudica a recombinação homóloga em resposta ao aumento das quebras de fita dupla do DNA e afeta a sensibilidade celular ao olaparibe e ao talazoparibe.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em hotspot de mutações, mas afeta sítio de interação transitória um "sítio de ligação preferencialmente ao ssDNA (resíduos 1-83)". Além disso, afeta domínios estruturais curtos, abolindo o sítio do motivo de fosforilação da *Protein kinase C phosphorylation*. Ser46 também está localizada perto do sítio de acetilação que ocorre na Lys48. Após a substituição Ser46Cys, foi prevista a perda deste sítio de acetilação. No ClinVar, não há variantes de sentido trocado benignas próximas ao resíduo 46 da RAD51D. Para toda a região em torno deste resíduo (resíduos 1-83), as variantes descritas como benignas ou provavelmente benignas são sinônimas.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica no gnomAD (0.00008) e não foi encontrada no 1000 Genomes Project e no Exome Sequencing Project (acesso em Novembro, 2022).

PP3 – Múltiplas linhas de evidência computacional suportam um efeito deletério no gene ou produto gênico:

- A Ser46 é altamente conservada entre espécies de mamíferos, aves, anfíbios e répteis e peixes, sem outros resíduos vistos nesta posição;
- Os preditores *in silico* utilizados sugerem que esta substituição tenha numerosos efeitos deletérios na proteína, que podem afetar não só a função e a estabilidade da estrutura da proteína, mas também as diferentes características moleculares da RAD51D, que se prevê serem funcionais (sítios de interação e pequenos domínios de proteínas);
- No presente estudo, foram previstas alterações nas interações moleculares, como perda de pontes de hidrogênio e interações polares, além de criar um novo sítio potencial para pontes dissulfeto. Essas mudanças podem contribuir para diferentes conformações da estrutura 3D da RAD51D.

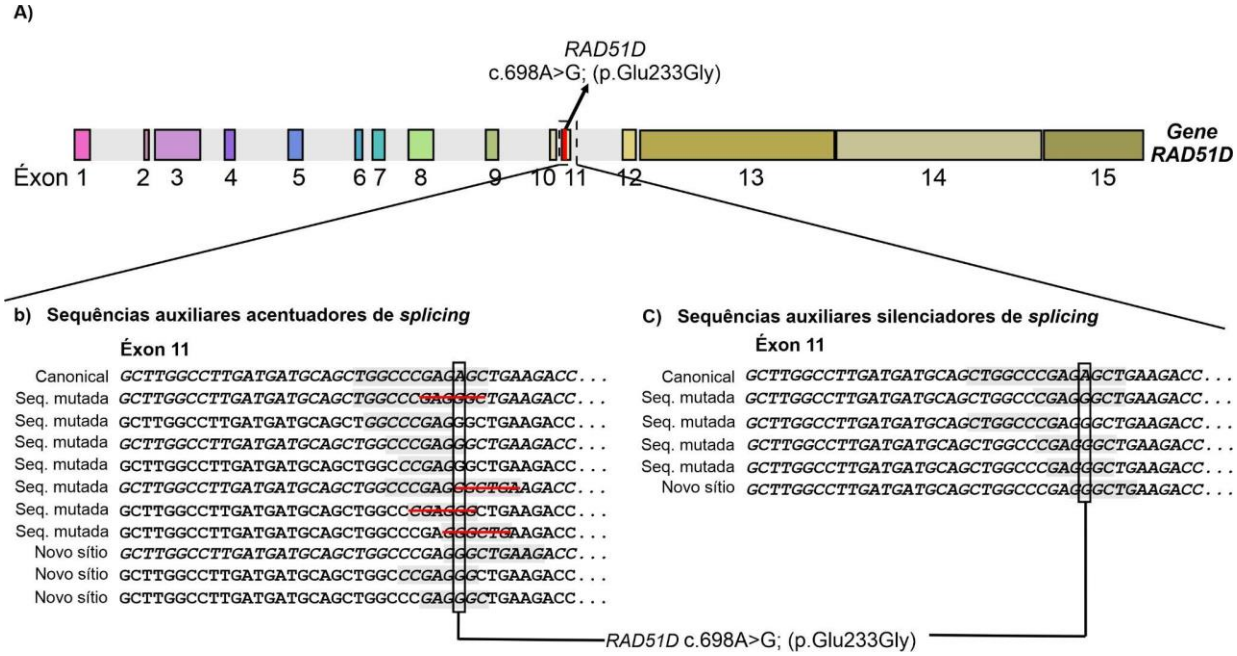
3.2.2.2.2 Conclusão sobre a reclassificação da variante *RAD51D* c.137C>G; (p.Ser46Cys)

Para esta variante foram previstos impactos em sítios de interação, motivos e interações intermoleculares, que podem prejudicar a estabilidade estrutural, a capacidade de interação transitória ou mesmo alterações na dinâmica conformacional. O ganho de uma cisteína no códon 46 pode levar ao surgimento de uma nova ponte de dissulfeto, resultando em mudanças conformacionais que podem ter diferentes desfechos moleculares, incluindo a perda da capacidade de interação. Contudo, a capacidade da proteína de interagir com outras moléculas é crítica para RAD51D, pois ela interage com suas parálogas e outras proteínas para promover o reparo de danos ao DNA. Corroborando este achado, o estudo de Alenezi *et al.,* 2022, demonstra que células, que expressaram *RAD51D* c.137C>G; (p.Ser46Cys), apresentaram diminuição da atividade de interação com outras proteínas, resultando na instabilidade da via de reparo por recombinação homóloga e, consequentemente, no aumento das quebras de fita dupla de DNA (Alenezi *et al.,* 2022). Não encontramos na literatura ou em bases de dados informação sobre dados clínicos ou análise de segregação para esta variante. Ainda assim, há evidências de patogenicidade, incluindo: a variante é muito rara ou ausente em bancos de dados populacionais; existem estudos in vitro descrevendo impacto funcional e há evidência in silico de efeitos na estrutura da molécula. Com base nestes achados e nos critérios elencados acima, propomos a reclassificação desta variante de VUS para patogênica.

3.2.2.3 *RAD51D* c.140A>G;(p.Tyr47Cys)

Localizada no éxon 3 do gene e no éxon 1 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), resulta de uma substituição de A para G na posição nucleotídica 140. Para essa substituição, foram previstos efeitos significativos nas sequências auxiliares reguladoras de *splicing*, sugerindo a perda de algumas sequências acentuadoras e criação de novas sequências silenciadoras (Fig. 25).

Figura 25- Impacto da *RAD51D* c.140A>G;(p.Tyr47Cys) nas sequências auxiliares regulatórias de sítios de *splicing*



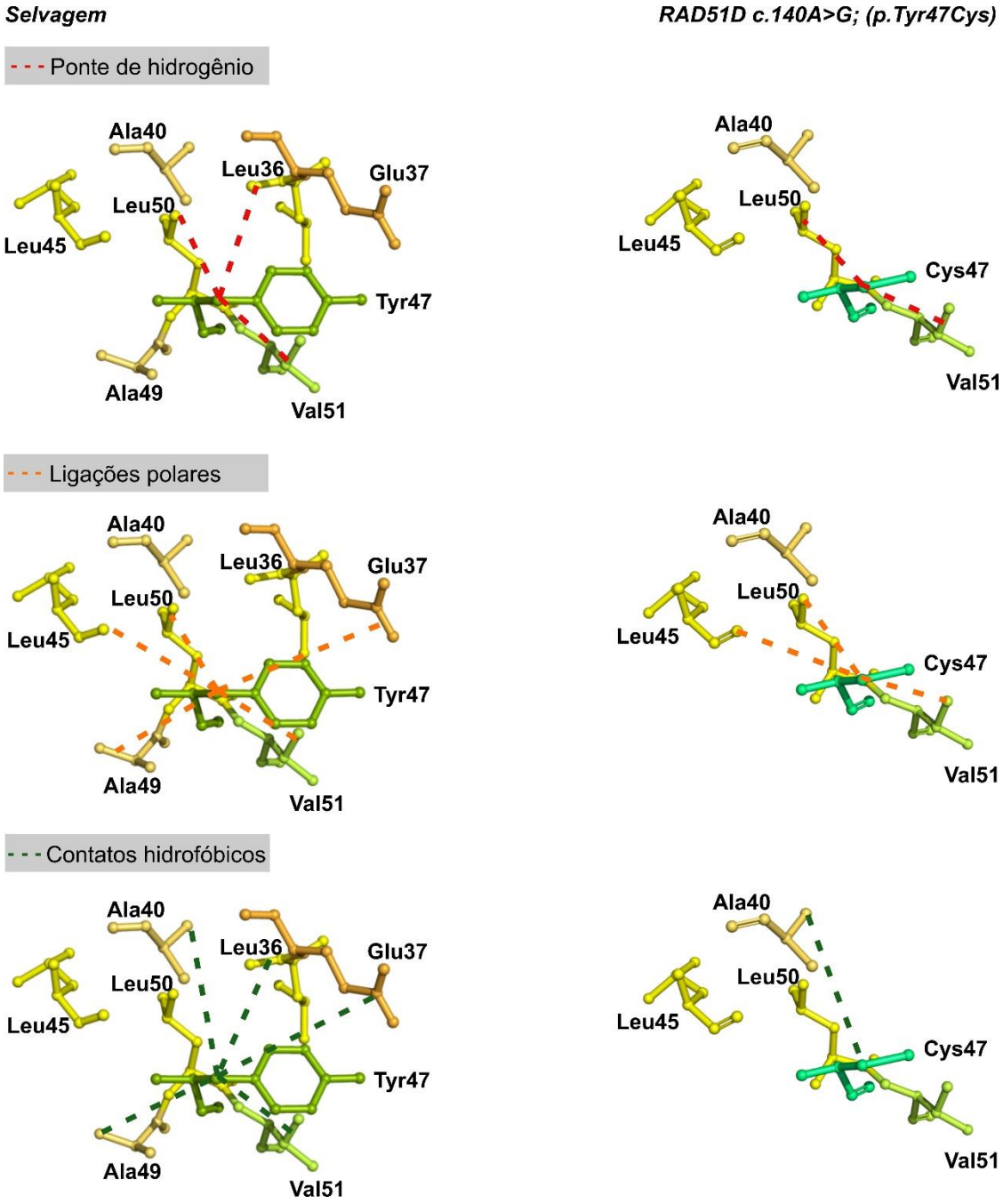
Nota: Em A, o arranjo dos éxons ao longo do gene RAD51D. O *RAD51D* tem 15 éxons (barras verticais). A variante *RAD51D* c.140A>G;(p.Tyr47Cys) ocorre no éxon 3 do gene (linha vertical vermelha). Nas figuras B e C, são mostradas as sequências auxiliares, acentuadores ou silenciadores, que foram afetados pela substituição de nucleotídeos. Em B, está sendo mostrado o impacto da variante em relação aos acentuadores de *splicing*. A sequência canônica contém o sítio acentuador de *splicing* de tipo selvagem, sublinhado em cinza, e abaixo dele estão as sequências mutadas. Alguns sítios não foram afetados mesmo após a substituição de nucleotídeos (sequência intacta), outros foram abolidos (linha vermelha cruzando a sequência). Em C, é mostrado o impacto da variante em relação aos silenciadores de *splicing*. A sequência canônica contém o sítio silenciador de *splicing* de tipo selvagem e abaixo estão as sequências mutadas. Os sítios próximos à variante não foram afetados após a troca de nucleotídeos, nem mesmo o sítio que envolve a variante. Contudo, um novo sítio silenciador foi gerado no local.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

A tirosina no códon 47 é substituída por uma cisteína, um aminoácido com as mesmas propriedades físico-químicas, sendo ambos polares não carregados. Os preditores *in silico* Mutation Taster, LIST-02, Provean, Polyphen e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica predisseram **efeitos deletérios** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Foi previsto que a Tyr47 faça ponte de hidrogênio (Tyr47 → Leu36, Tyr47 → Leu50 e Tyr47 → Val51), contatos hidrofóbicos (Tyr47 → Leu36, Tyr47 → Glu37, Tyr47 → Ala40, Tyr47 → Val51), polares (Tyr47 → Glu37, Tyr47 → Ala45, Tyr47 → Ala49, Tyr47 → Val51) (Fig. 26). Após troca de

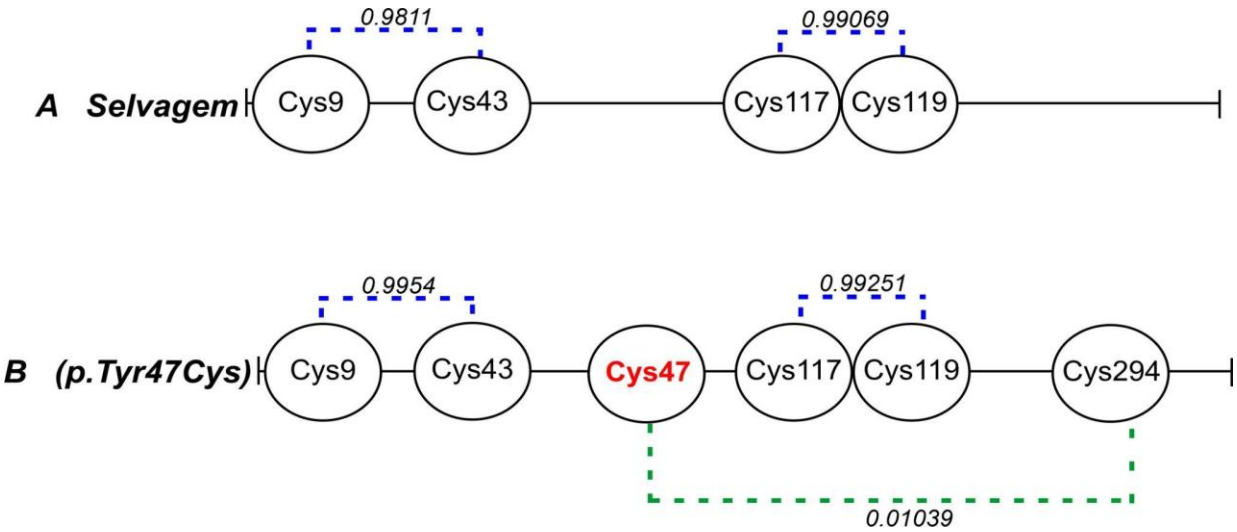
aminoácidos, ligações como ponte de hidrogênio (Tyr47 → Leu36), contatos hidrofóbicos (Tyr47 → Leu36, Tyr47 → Glu37, Tyr47 → Val51), polares (Tyr47 → Ala49 e Tyr47 → Glu37) foram perdidas (Fig. 26). A Cys47 (aminoácido substituto) é predita como um potencial para pontes de dissulfeto, potencialmente formando ponte dissulfeto com a Cys294 (Fig. 27). Além disso, a Cys47 também foi predita como desestabilizante da estrutura proteica, podendo afetar a habilidade de interação com o DNA e proteínas (Quadro 6, na seção 3.2.6)

Figura 26 - Dinâmica molecular da variante *RAD51D* c.140A>G; (p.Tyr47Cys)



Nota: Interações interatômicas observadas nas estruturas selvagens e mutadas. O traço vermelho, representa a ponte de hidrogênio; em laranja, interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Leu36, Glu37, Ala40, Leu45, Tyr47, Ala49, Leu50 e Ala51. Na estrutura mutada, perde-se a interação com a Leu36, Glu37 e Ala49.
Fonte: Elaborada pela autora, usando DynaMut.

Figura 27 - Impacto da variante *RAD51D* c.140A>G;(p.Tyr47Cys) sobre pontes de dissulfeto da RAD51D usando Dianna

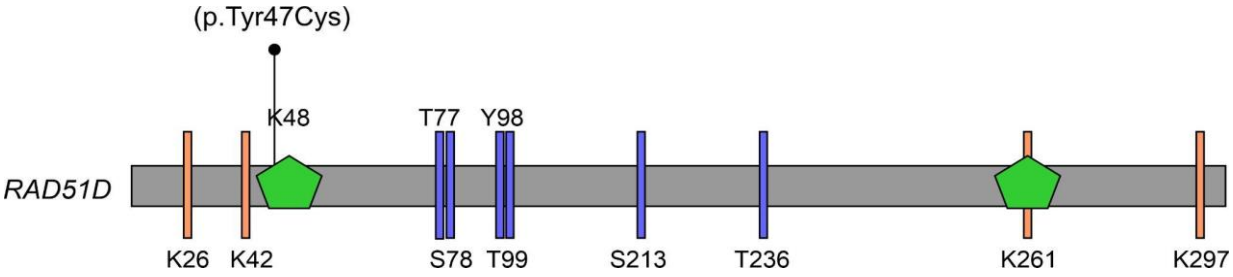


Nota: Em A, são mostrados os resíduos Cys9, Cys43, Cys117, Cys119 e Cys294, e as pontes de dissulfeto preditas usando-se Dianna para a proteína selvagem. As pontes de dissulfeto preditas são mostradas em azul, estão presente na proteína selvagem. O número próximo a cada linha tracejada indica a probabilidade de ocorrência de cada ponte de dissulfeto. Em B, é mostrada a consequência da variante *RAD51D* c.140A>G; (p.Tyr47Cys) sobre as pontes de dissulfeto preditas. Para essa substituição, foi previsto que a entrada da Cys 47 forma uma ligação covalente com a Cys 294, que é representada pela linha pontilhada verde.

Fonte: Elaborada pela autora.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Tyr 47 está localizada em um sítio de ligação preferencial ao ssDNA (resíduos 1-83) e ao sítio *HELIX 47-61* (Quadro 6, na seção 3.2.6). No presente estudo, foi predito que, a mesma forma que a variante *RAD51D* c.137C>G; (p.Ser46Cys), a *RAD51D* c.140A>G; (p.Tyr47Cys) também levaria a perda do sítio de acetilação na Lys 48 (Fig. 28).

Figura 28 - Figura esquemática do efeito da variante sobre o sítio de modificações pós-síntese

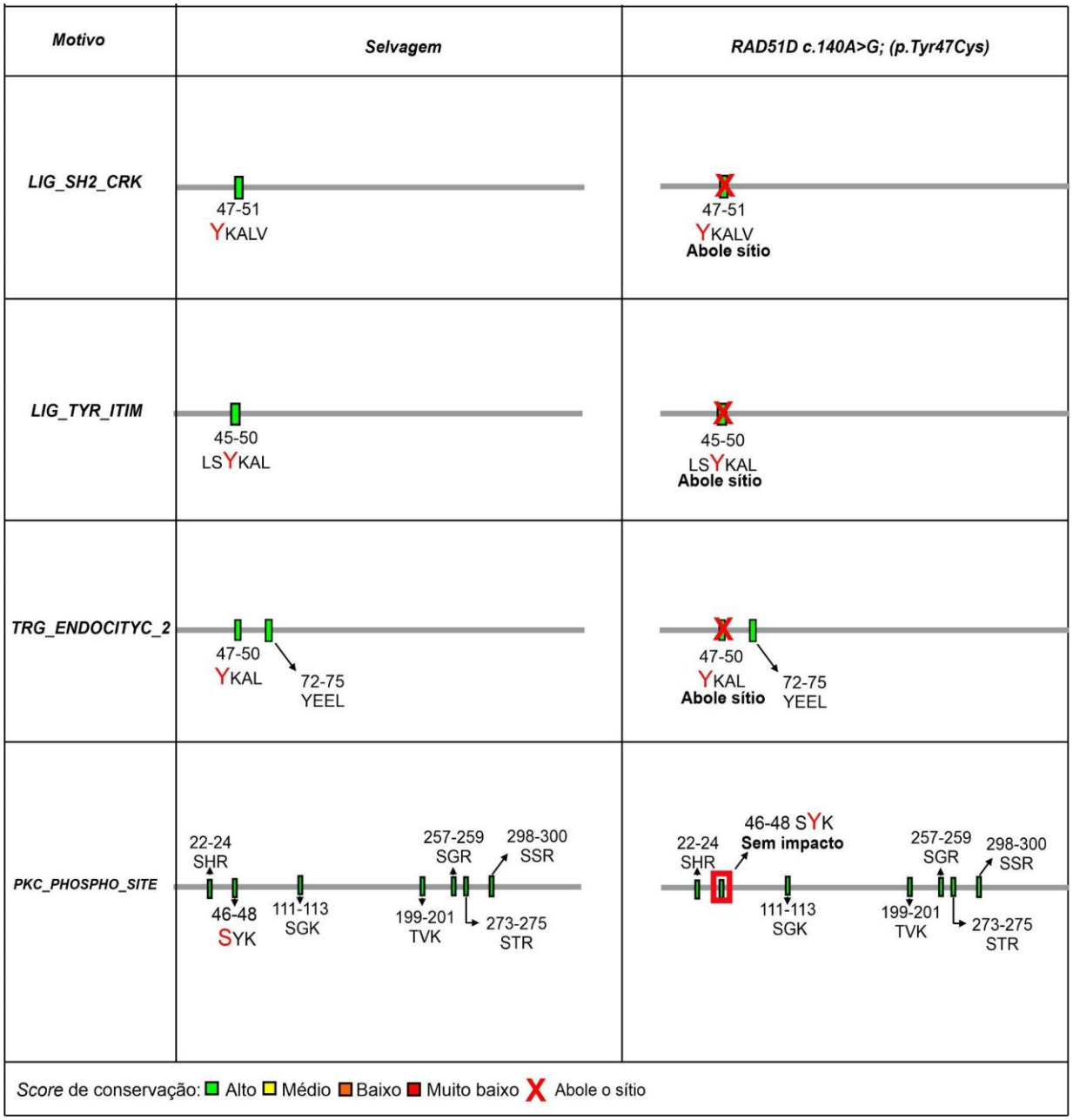


Nota: Foram preditos três tipos de modificações pós-síntese: ubiquitinação (laranja), fosforilação (azul) e acetilação (verde). Os sítios de ubiquitinação acontecem nas lisinas (26 e K42); os sítios de fosforilação acontecem nos aminoácidos treonina (T), serina (S) e tirosina (Y); e os sítio de acetilação acontecem em duas lisinas (K48 e K261). Previu-se que, após substituição Tyr47Cys, ocorre perda do sítio de acetilação da K48.

Fonte: Elaborada pela autora, usando PhosphoSitePlus

A Tyr 47 está localizada no motivo *PKC_PHOSPHO_SITE (sítio de fosforilação da proteína quinase C)* situado na posição 46-48 [SYK-SCK], mas a substituição não a leva à perda do motivo. Contudo, a perda da Tyr47 abole o motivo *LIG_SH2_CRK (Motivos de fosfotirosina de ligação aos domínios SH2) da posição* 47-51 [YKALV]; abole o motivo *LIG_TYR_ITYM (Motivo inibidor baseado em tirosina imunorreceptor)* da posição 45-50 [LSYKAL] (Fig. 29); e, abole o motivo *TRG_ENDOCITYC_2 (Motivo de sinal de classificação endocítica baseado em Y*) 47-50 [YKAL], todos motivos de alta conservação. A entrada da Cys47 cria um novo sítio de baixa conservação do motivo *MOD_NEK_2_2 (sítio de fosforilação de NEK2)* na posição 43-48 [CGLSCK] (Fig. 29). Estudos funcionais seriam importantes para avaliar o impacto destas modificações *in vivo*.

Figura 29 - Predição de impacto de *RAD51D* c.140A>G; (p.Tyr47Cys) sobre os motivos proteicos



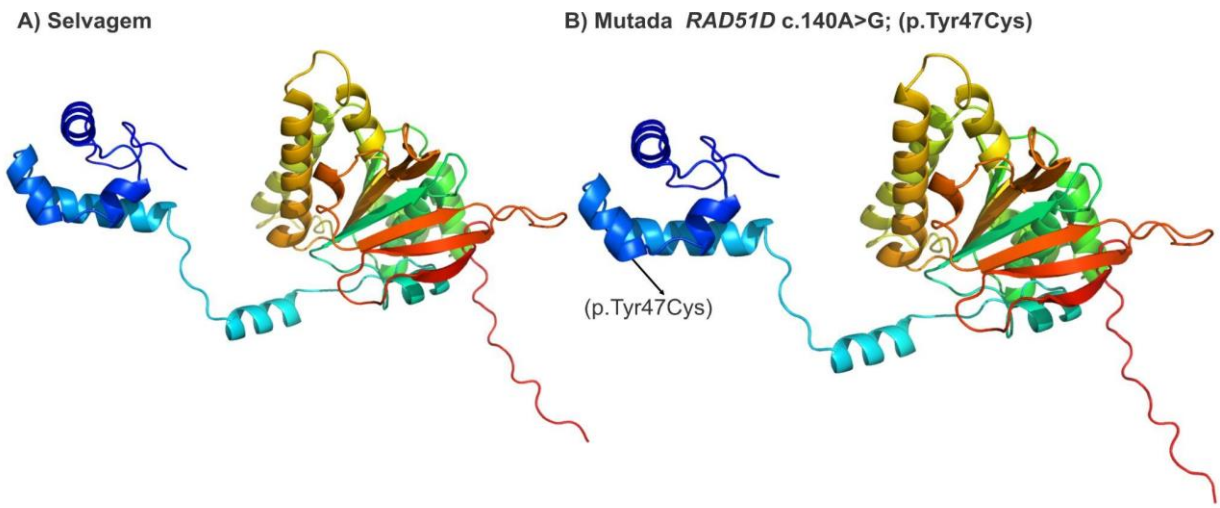
Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM motifs.

Considerando a estrutura 3D da proteína, previu-se que a Tyr47 está localizada em uma alfa-hélice e, embora seja substituída por uma cisteína, não foi previsto impacto na estrutura 3D

(Fig. 30). Entretanto, é relevante destacar que os programas usados nesta predição não usam a informação sobre pontes de dissulfeto.

Figura 30 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.140A>G; (p.Tyr47Cys)



Nota: Em A, está a estrutura da proteína 3D; em B, a estrutura mutada. A *RAD51D* c.140A>G; (p.Tyr47Cys) está localizada em uma alça, indicada pela seta. A estrutura 3D não mostra possíveis efeitos da formação de novas pontes de dissulfeto.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.3.1 Reclassificação da variante *RAD51D* c.140A>G; (p.Tyr47Cys)

Com base nas informações acima, a *RAD51D* c.140A>G; (p.Tyr47Cys) é classificada como provavelmente patogênica, segundo os critérios ACMG/AMP: PM1, PM2, PP2 (ver na seção 3.2.2) e PP3.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em *hotspot* de mutações, mas afeta um sítio de interação transitória, o "sítio de ligação preferencial ao ssDNA (resíduos 1-83)". Além disso, afeta domínios estruturais curtos, abolindo os sítios dos motivos *LIG_SH2_CRK* (Motivos de fosfotirosina de ligação aos domínios SH2), *LIG_TYR_ITYM* (Motivo inibidor baseado em tirosina imunorreceptor) e o *TRG_ENDOCITYC_2* (Motivo de sinal de classificação endocítica baseado em Y). Após a troca de aminoácidos, foi prevista a perda do sítio de acetilação (Lys48), um aminoácido próximo a Tyr47. Conforme avaliado no ClinVar, esta região da variante não contém variantes de sentido trocado benignas. Para essa região (resíduos 1-83), as variantes benignas ou provavelmente benignas são sinônimas.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante não foi encontrada em bancos de dados populacionais (gnomAD, 1000 Genomes Project ou Exome Sequencing Project). Além disso, não há nenhuma publicação associada a esta variante.

PP3 – Múltiplas linhas de evidência computacional suportam um efeito deletério no gene ou produto gênico

- A Tyr47 é altamente conservada entre espécies de mamíferos, aves, anfíbios e répteis e peixes, sem outros aminoácidos vistos nesta posição;
- Esta variante não está descrita em bancos de dados populacionais;
- Os preditores *in silico* utilizados sugerem que esta substituição tem numerosos efeitos deletérios na proteína, que podem afetar não só a função e a estabilidade da estrutura da proteína, mas também as diferentes características moleculares da RAD51D, que se prevê serem funcionais (sítios de interação e pequenos domínios de proteínas).
- Foram previstas alterações nas interações moleculares, como perda de ponte de hidrogênio e contatos hidrofóbicos, além de criar um novo sítio potencial para ponte de dissulfeto;
- Foi predito impacto em modificações pós-síntese da proteína RAD51D, devido a potenciais efeitos na fosforilação e na acetilação da molécula;
- Previu-se que esta variante alterasse elementos regulatórios de *splicing*.

3.2.2.3.2 Conclusão sobre a reclassificação da variante *RAD51D* c.140A>G; (p.Tyr47Cys)

Como previsto, a substituição de aminoácidos ocorre em um resíduo altamente conservado da proteína (PP3), afeta pequenos domínios funcionais (PM1) e gera, potencialmente, uma nova ponte de dissulfeto (PP3). A substituição de aminoácidos abole os sítios dos motivos *LIG_SH2_CRK*, *LIG_TYR_ITYM* e o *TRG_ENDOCITYC_2* . O motivo *LIG_SH2_CRK* são motivos que contêm pelo menos um resíduo de tirosina fosforilado (motivos de fosfotirosina), possuem Receptor Tirosina Quinases (RTKs). Esses motivos são reconhecidos pelo domínio SH2 *(Src Homologia 2)* que são geralmente encontrados em famílias CRK (proteínas sinalizadoras). SH2 é o principal domínio de ligação para eventos de sinalização de fosforilação de TK (Kaneko *et al.*, 2012; Pawson *et al.*, 2001); O motivo *LIG_TYR_ITYM*

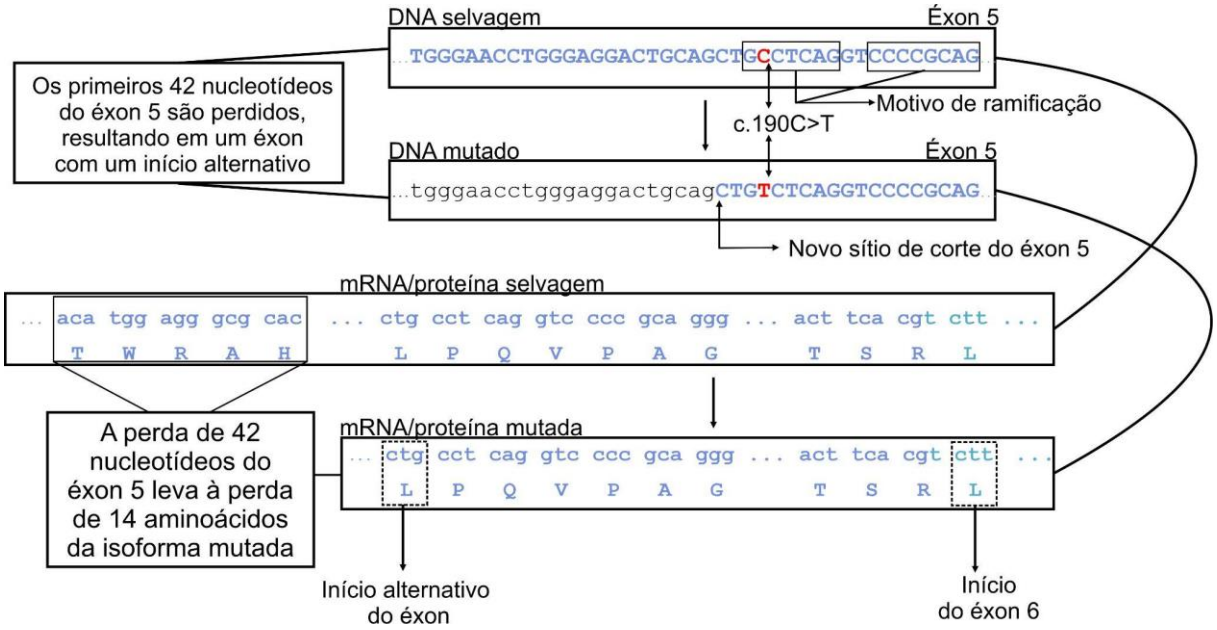
ITIM são motivos essenciais para a ativação e terminação das vias de transdução de sinal. Este motivo é encontrado na cauda citoplasmática de alguns receptores inibitórios (KIRs) como células natural *killer*. O motivo ITIM é fosforilado, ligando-se ao MHC Classe I, permitindo o recrutamento e ativação de uma proteína tirosina fosfatase (Long, 1999; Unkeless & Jin, 1997)*;* e, por fim, o motivo *TRG_ENDOCITYC_2* é um motivo de sinal de classificação baseado em tirosina, e determinam qual via de tráfego vesicular é usada para transportar uma determinada molécula. A sequência do motivo baseada em tirosina é então reconhecida pela subunidade *mu* do complexo de proteínas adaptadoras (PA) (Boll *et al.*, 2002). Esses resultados fornecem evidências de um papel patogênico para essa variante, uma vez que a inatividade de domínios curtos e novas ligações covalentes entre os resíduos podem alterar a dinâmica conformacional da proteína e interferir nas interações proteína-proteína e proteína-ssDNA. A frequência alélica da *RAD51D* c.140A>G; (p.Tyr47Cys) é outra evidência que apoia a patogenicidade (PM2), visto que a variante encontra-se ausente em estudos caso-controles, bem como na PubMed. No ClinVar, não foram encontradas variantes de sentido trocado previamente classificadas como patogênicas neste resíduo. As quatro variantes encontradas envolvendo o resíduo Tyr47 são uma benigna (*RAD51D* c.141C>T (p.Tyr47=)) e três patogênicas (*RAD51D* c.141C>G (p.Tyr47Ter), *RAD51D* c.141C>A (p.Tyr47Ter) e *RAD51D* c.140_141insAA (p.Tyr47Ter)). Conforme discutido acima, este gene preencheria o critério PP2. Considerando essas evidências de patogenicidade, PM1 propõe-se a classificação da *RAD51D* c.140A>G; (p.Tyr47Cys) de VUS para provavelmente patogênica.

3.2.2.4 *RAD51D* c.190C>T; (p.Pro64Ser)

Em função do *splicin*g alternativo, esta variante afeta apenas uma isoforma não canônica (isoforma 8). Ela está localizada no éxon 5 do gene e no éxon 2 do transcrito ENST00000590016.5 (Fig. 12, na Seção 3.2) e resulta de uma substituição de C para T na posição nucleotídica 190. Foi previsto que esta substituição leve à ativação de um sítio aceptor críptico no éxon 5, causando a perda dos primeiros 42 nucleotídeos do éxon, o que resultaria em um éxon com um início alternativo. O éxon 5 selvagem tem 179 nucleotídeos de comprimento, após a iniciação alternativa, o éxon passa a ter 137 nucleotídeos de comprimento. O transcrito predito possui os éxons 3, 5 (mais curto), 6, 7, 8, 10, 11, 12 e 13 do gene, gerando uma proteína de 334 aminoácidos, com códon de parada na posição 335. Isso resulta em 14 aminoácidos a menos que a isoforma 8 do tipo selvagem (Fig. 31). Vale ressaltar que a perda de 42 nucleotídeos não afeta a sequência de nucleotídeos dos outros éxons (apenas do éxon 5), mantendo-os com o mesmo comprimento e sequência dos éxons do

transcrito selvagem. Deve-se notar que não foram encontrados sítios de ramificação à 5' deste novo aceptor, na distância típica, cerca de 40 nucleosídeos. Desta forma, é possível que este novo sítio aceptor não seja usado no *splicing* (figura 31).

Figura 31 - Impacto da *RAD51D* c.190C>T; (p.Pro64Ser) no sítio de *splicing*

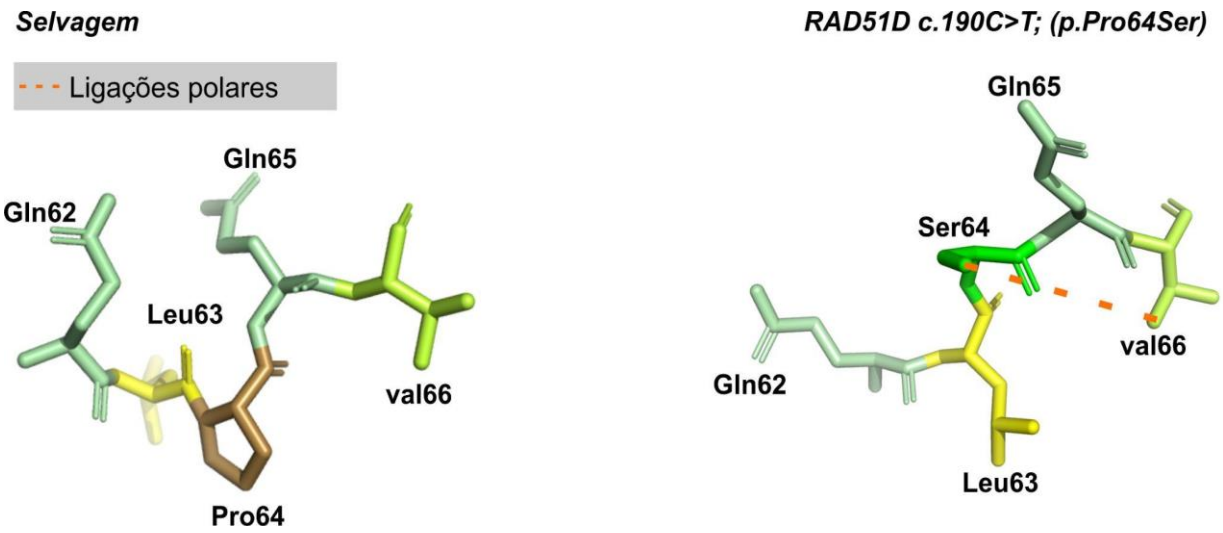


Nota: Na imagem, a troca de citosina por timina (nucleotídeos em vermelho) levaria à ativação de um sítio aceptor críptico próximo ao local onde ocorre a variante. Dois sítios de ramificação foram preditos à 3' deste novo aceptor, próximo ao sítio de corte. Se este novo aceptor fosse ativado, 42 nucleotídeos exônicos seriam perdidos e passariam a fazer parte do íntron 4. Essa alteração não mudaria o quadro de leitura da proteína, pois os demais éxons não seriam alterados com a perda dos 42 nucleotídeos.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

A prolina no códon 64 é substituída por uma serina, um aminoácido com propriedades físico-químicas diferentes, visto que a prolina é apolar e a serina é polar não carregada. Os preditores *in silico* Mutation Taster, LIST-02, FATH-MM, Provean, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Também foi previsto que essa substituição desestabilize a estrutura da proteína, podendo afetar a habilidade de interação com o DNA e outras proteínas (Quadro 6, na seção 3.2.6). Não foram previstas interações intermoleculares da prolina com outros resíduos, porém, após a troca de aminoácidos, uma nova ligação polar surgiu entre a serina e a valina na posição 66 (Fig. 32).

Figura 32 - Dinâmica molecular da variante *RAD51D* c.190C>T; (p.Pro64Ser)

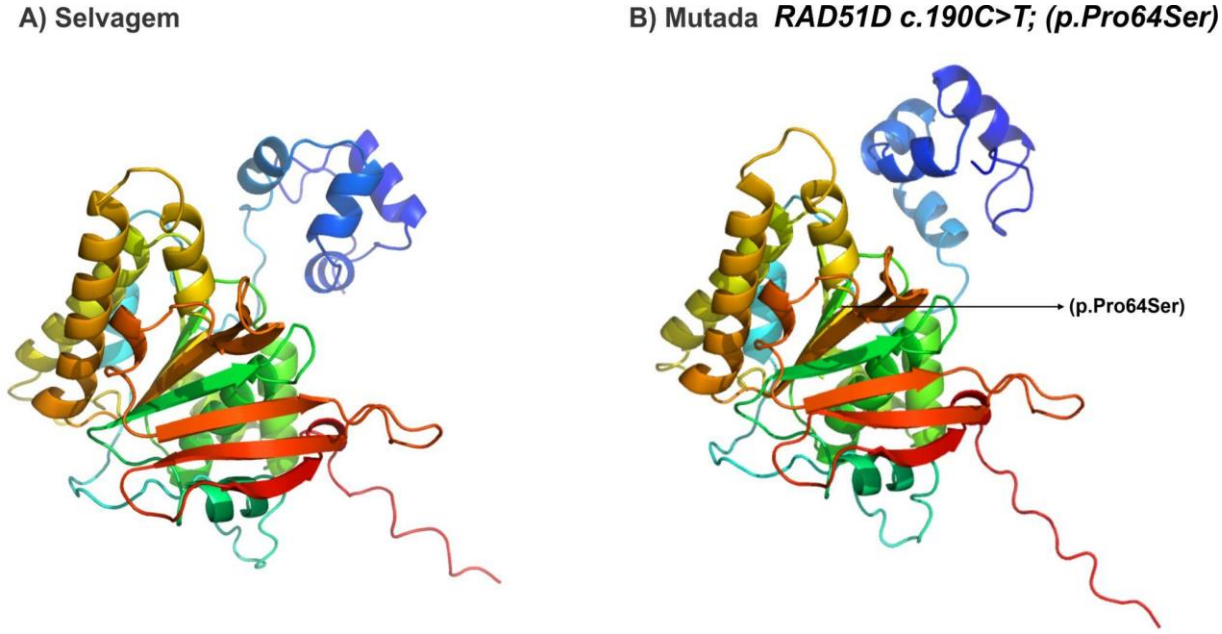


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço laranja representa a interação polar. Na estrutura selvagem, não foram previstas interações interatômicas envolvendo a prolina na posição 64 (Pro64). Na estrutura mutada, foi prevista uma nova interação polar com a Ser64 e Val66. A predição foi desenvolvida utilizando a sequência da isoforma 8.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios funcionais e características moleculares, a isoforma 8, não canônica, possui dois sítios que medeiam a ligação do ATP conhecidos como *Walker* A (127-134) e *Walker* B (239-243), mas que não são afetados pela variante. Pro64 não está localizado nos domínios funcionais da proteína *DNA recombination/repair_Rad51_C-terminal (107-303)*, *AAA+ATPase (119-294)* e *ReacA_ATP-biding (98-271)*. Considerando a estrutura 3D da proteína, previu-se que a Pro64 está localizada em uma alça e, embora seja substituída por uma serina, não foi previsto impacto na estrutura 3D (Fig. 33).

Figura 33 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.190C>T; (p.Pro64Ser)) usando a isoforma 8 como modelo



Nota: Em A está sendo mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada que é indicada na figura B. A *RAD51D* c.190C>T; (p.Pro64Ser) está localizada em uma alça, indicada por uma seta vermelha. A estrutura 3D não mostra impactos significativos. O modelo foi construído utilizando a sequência da isoforma 8.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.4.1 Reclassificação da variante *RAD51D* c.190C>T; (p.Pro64Ser)

Com base nas informações acima, a *RAD51D* c.190C>T; (p.Pro64Ser*)* é classificada como provavelmente benigna, segundo os critérios ACMG/AMP: PP2 (ver na seção 3.2.2), BP4 e BP6.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A *RAD51D* c.190C>T; (p.Pro64Ser*)* afeta a isoforma 8 não canônica, que é conservada apenas em humanos e primatas, impedindo uma análise aprofundada da conservação evolutiva desse resíduo;
- Os preditores *in silico* utilizados sugerem efeitos benignos para esta variante;
- Não foram previstas alterações maiores nas interações intermoleculares;

- Prevê-se que a variante desestabiliza a estrutura da proteína;
- Foi previsto que a substituição de nucleotídeos ativa um sítio aceptor críptico, mas a apreciação do impacto dessa modificação depende de estudos funcionais.

BP6 - Fonte respeitável relata recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente:

A variante é reportada como provavelmente benigna em bancos de dados públicos como ClinVar, VarSome e ABraOm.

3.2.2.4.2 *Conclusão sobre a reclassificação da variante RAD51D c.190C>T; (p.Pro64Ser)*

Em resumo, no presente estudo, foi predito que a *RAD51D* c.190C>T; (p.Pro64Ser) ative um sítio de *splicing* aceptor críptico. É classificada como benigna pelos algoritmos utilizados, pois há uma grande possibilidade de que o aceptor críptico não seja ativado, considerando a falta de sítio de ramificação à 5' do novo aceptor. A avaliação da profundidade da conservação evolutiva deste resíduo é limitada, pois a variante afeta a isoforma não canônica O75771-8, que em bancos de dados de proteínas está presente apenas em humanos e alguns primatas. Outra limitação iminente é a falta de dados clínicos ou estudos funcionais bem estabelecidos que investiguem o real impacto dessa variante na estrutura do DNA e da proteína. A variante está reportada como provavelmente benigna nos bancos de dados ClinVar e VarSome, sem estudos que associam essa variante a câncer de mama e/ou ovário no PubMed. Portanto, devido à falta de dados mais conclusivos na investigação da patogenicidade dessa variante, bem como a frequência alélica dessa variante, por exemplo no gnomAD (0.0004), 1000 Genomes Project (0.00100) e Exome Project Genome (0.00019), este estudo mantém a variante como provavelmente benigna.

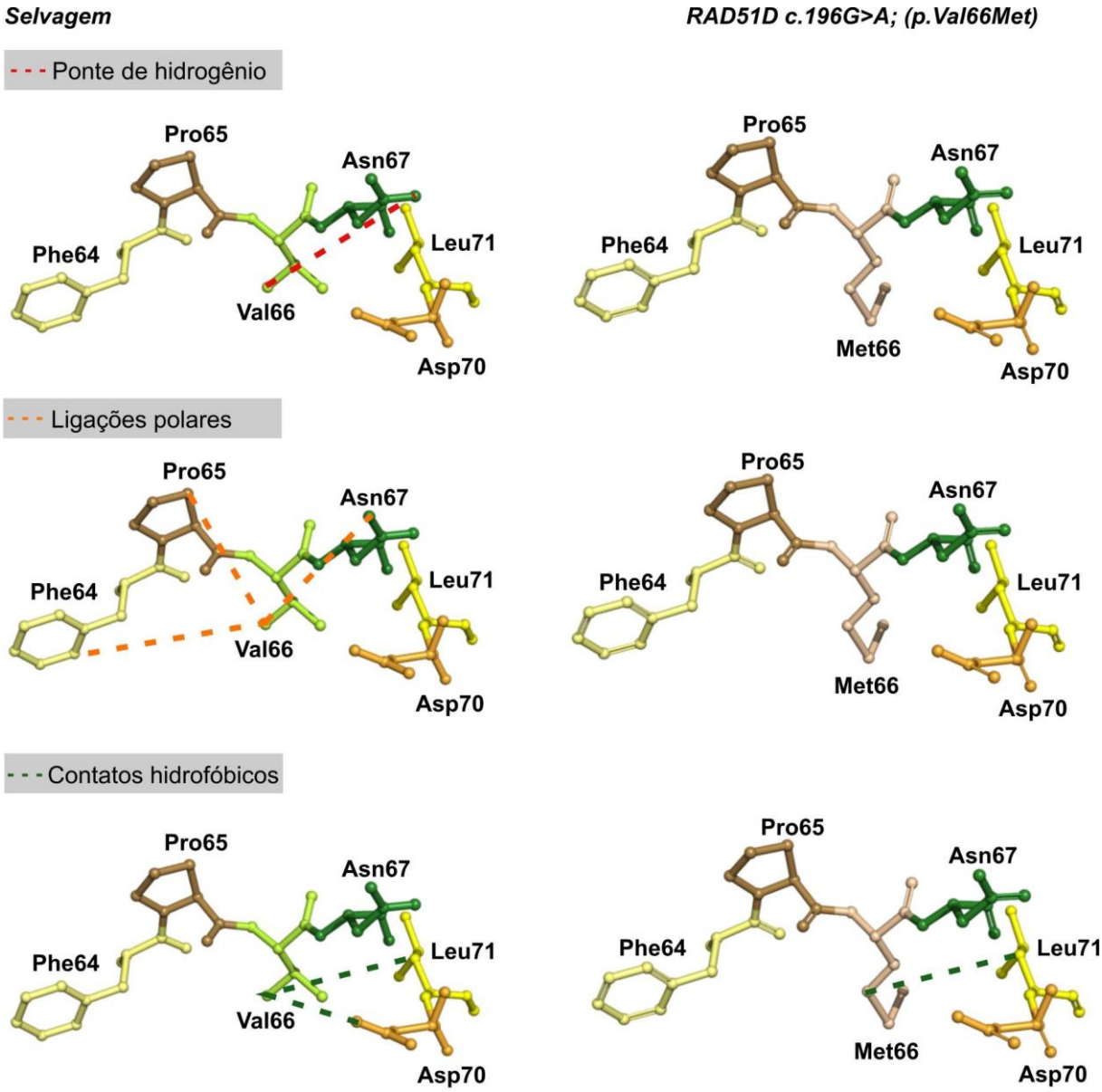
3.2.2.5 *RAD51D* c.196G>A; (p.Val66Met)

Localizada no éxon 4 do gene e no éxon 2 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), resulta de uma substituição de G para A na posição nucleotídica 196. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadoras).

A valina na posição 66 é substituída por metionina, um aminoácido com as propriedades físico-químicas similares, ambos são apolares. Os preditores *in silico* Mutation Taster, LIST-

02, FATH-MM, Provean, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Foi previsto que a Val66 faça ponte de hidrogênio (Val66 → Asn67), contatos hidrofóbicos (Val66 → Asp70 e Val66 → Leu7), interações polares (Val66 → Phe64, Val66 → Pro65, Val66 → Asn67) (Fig. 34). Após a troca de aminoácidos, a maioria das interações intermoleculares foram perdidas como a ponte de hidrogênio (Val66 → Asn67), contato hidrofóbico (Val66 → Asp70), interações polares (Val66 → Phe64, Val66 → Pro65, Val66 → Asn67) (Fig. 34). A perda da Val66 desestabiliza a estrutura da proteína, possivelmente afetando a interação com o DNA e outras proteínas (Quadro 6, na seção 3.2.6).

Figura 34 - Dinâmica molecular da variante *RAD51D* c.196G>A; (p.Val66Met)



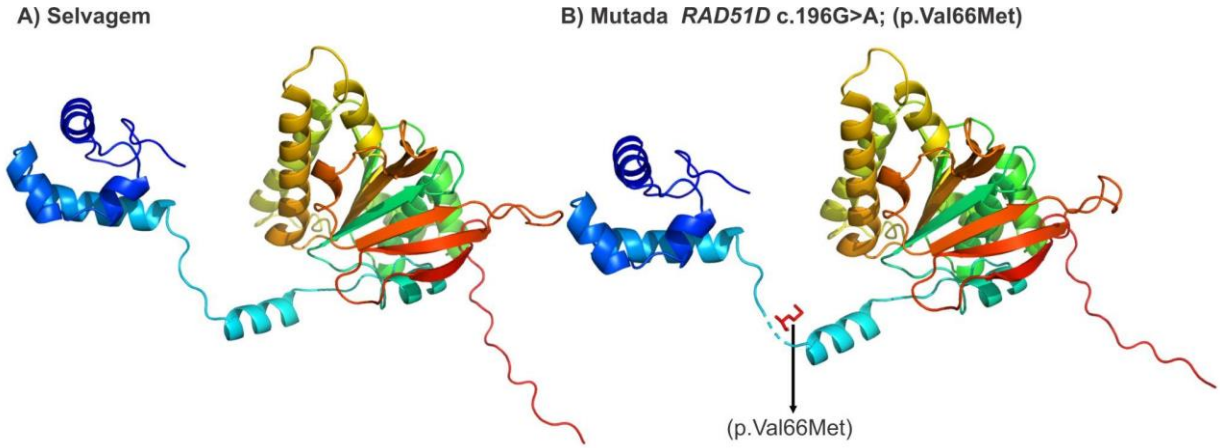
Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Phe64, Pro65, Val66, Asn67, Asp70 e Leu71. Na estrutura mutada, é mostrada a metionina substituindo a valina na posição 66.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Val66 está localizada em um sítio de ligação preferencialmente ao ssDNA (resíduos 1-83) e região de ligação (54-77) (Quadro 6, na seção 3.2.6). Considerando a

estrutura 3D da proteína, previu-se que a Val66 estaria localizada em uma alça, na qual nenhum impacto foi previsto na estrutura 3D (Fig. 35).

Figura 35 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.196G>A; (p.Val66Met))



Nota: Em A: estrutura 3D da proteína selvagem; em B: estrutura 3D da proteína mutada. A *RAD51D* c.196G>A; (p.Val66Met) está localizada em uma alça, indicada por uma seta. A estrutura 3D não mostra impactos significativos.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.5.1. Reclassificação da variante *RAD51D* c.196G>A; (p.Val66Met)

Considerando as informações acima sobre o impacto funcional da variante *RAD51D* c.196G>A; (p.Val66Met), propomos manter sua classificação como VUS, segundo os critérios ACMG/AMP: PM1, PP2 (ver na Seção 3.2.2) e BP4.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em *hotspot* de mutações, mas afeta sítio de interação transitória um "sítio de ligação preferencialmente ao ssDNA (resíduos 1-83)" e uma "região de ligação (54-77)". Conforme avaliado no banco de dados ClinVar, esta região da variante não contém variantes missense benignas. Para essa região (resíduos 1-83), as variantes benignas ou provavelmente benignas são sinônimas.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- Val66 não é conservada entre espécies de mamíferos e de aves, mas em anfíbios, répteis e peixes é, não sendo substituída por outros aminoácidos nessa posição;

- Os preditores *in silico* utilizados sugerem efeitos benignos para esta substituição;

- Foram previstas inúmeras alterações nas interações moleculares, em que a maioria das ligações (pontes de hidrogênio, ligações polares etc) foram perdidas. Essas mudanças podem contribuir para diferentes conformações da estrutura 3D da RAD51D. Isto, entretanto, não é evidência suficiente de patogenicidade.

3.2.2.5.2 Conclusão sobre a reclassificação da variante *RAD51D* c.196G>A; (p.Val66Met)

Em síntese, os algoritmos utilizados preveem a variante como provavelmente benigna, mesmo estando localizada em regiões funcionais responsáveis pela interação proteína-proteína. Sucede-se que, além de considerar as regiões funcionais, os programas *in silico* são baseados na conservação evolutiva. Entretanto, Val66 é altamente conservada apenas em peixes e anfíbios e répteis, e não é conservada nos mamíferos. Em mamíferos, a valina apresenta-se substituída ora por fenilalanina, isoleucina, leucina, cisteína ou serina. Nas aves, esse resíduo é substituído principalmente por alanina. A frequência populacional não é baixa (0.0003 (gnomAD), 0.0006 (1000 Genomes Project) e não encontrada no Exome Project Genome, impedindo a atribuição do critério PM2. Nas bases de dados públicas ClinVar, VarSome e ABraOM, a variante é conflitante, com três entradas para VUS, duas para benigna e seis para provavelmente benigna. Não foram encontrados na literatura estudos funcionais, que demonstrem evidências de impacto na função proteica, bem como a associação da variante com a Síndrome Hereditária do Câncer de Mama e Ovário (Yurgelun *et al.,* 2017; Hauke *et al,.* 2018, Tomioka *et al.,* 2021). A exceção é o estudo de Wei e colaboradores (2019), que mostra a co-ocorrência da variante *RAD51D* c.196G>A; (p.Val66Met) a variante *BRCA2* c.571_572insAT; (p.Asp191fs), classificada como patogênica, em indivíduos diagnosticados com câncer de próstata. Contudo, considerando o relatório do *Women's Health and Genetics - Laboratory Corporation of America (LabCorp),* disponível no ClinVar, essa informação não é suficientemente forte para sustentar a reclassificação da variante como patogênica, uma vez que, sua ocorrência com uma outra variante de alta penetrância, como genes da família *BRCA*, por si só, podem ser responsáveis pelo fenótipo em questão. Da mesma forma, a ocorrência simultânea da variante *RAD51D* c.196G>A; (p.Val66Met) **e** a variante *BRCA2* c.571_572insAT; (p.Asp191fs), descrita por Wei e colaboradores (2019), por si só, também

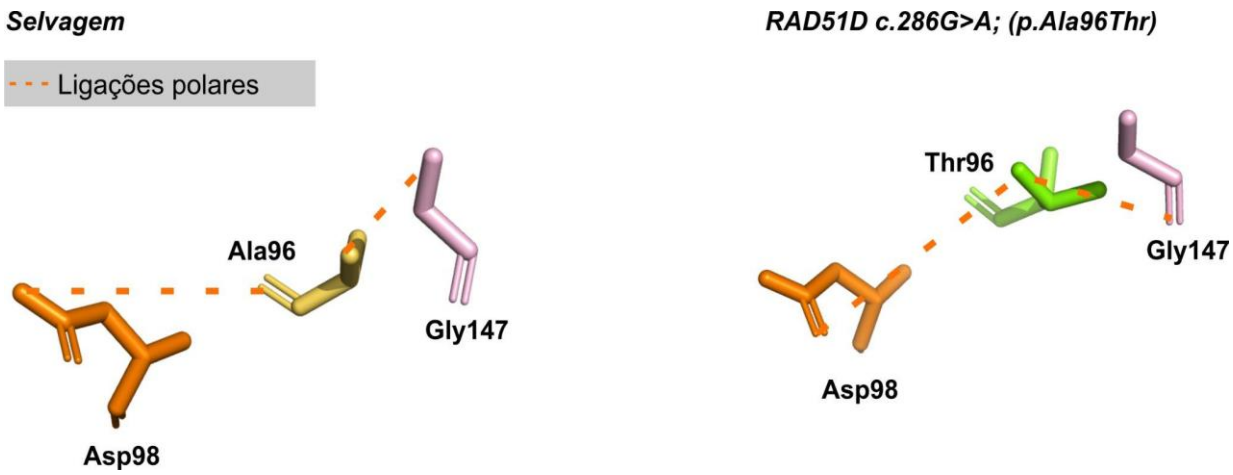
não forneceram evidências de suporte para a reclassificação da variante *RAD51D* c.196G>A; (p.Val66Met), ainda que o câncer seja sabidamente uma doença multifatorial e que os genes podem ter diferentes tipos de interação entre si. Também não foram encontradas variantes com a mesma alteração nesta posição de outros tipos de aminoácidos devidamente classificadas como patogênicas. Outras variantes encontradas nesta posição são classificadas como VUS (*RAD51D* c.196G>T (p.Val66Leu)) e benigna (*RAD51D* c.198G>T (p.Val66=)). Em síntese, a falta de dados clínicos como história familiar e dados de segregação limitam uma classificação assertiva da variante, mantendo sua classificação VUS.

3.2.2.6 *RAD51D* c.286G>A; (p.Ala96Thr)

Em função do *splicin*g alternativo, esta variante afeta apenas uma isoforma não canônica (isoforma 8). Ela está localizada no éxon 5 do gene e no éxon 2 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), resulta de uma substituição de G para A na posição nucleotídica 286. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores).

A alanina no códon 96 é substituída por treonina, um aminoácido com propriedades físico-químicas diferentes, visto que a alanina é apolar e a treonina é polar não carregado. Os preditores *in silico* Mutation Taster, LIST-02, FATH-MM, Provean, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Foi previsto que a Ala96 fizesse interações polares (Ala96 → Gly147), (Ala96 → Asp98) e essas interações permaneceram mesmo após a troca de aminoácidos (Fig. 36). A perda da Ala96 não leva a desestabilização da estrutura da proteína, bem como a habilidade de interação com o DNA e proteínas (Quadro 6, na seção 3.2.6).

Figura 36- Dinâmica molecular da variante *RAD51D* c.286G>A; (p.Ala96Thr)

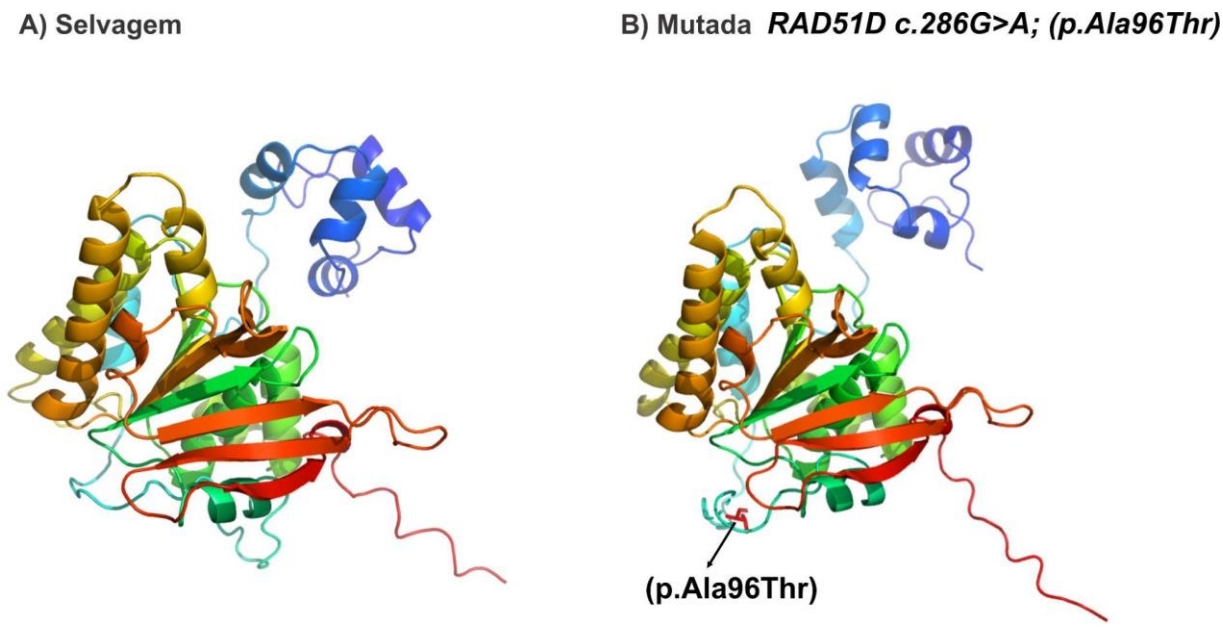


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço laranja representa as interações polares. Na estrutura selvagem, as interações polares acontecem entre os resíduos: Ala96, Asp98 e Gly147. Na estrutura mutada, estas interações não são abolidas.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios funcionais e características moleculares, a isoforma não canônica (isoforma 8) possui dois sítios, que medeiam a ligação do ATP conhecidos como *Walker* A (127-134) e *Walker* B (239-243). Ala96 não está localizado nesses sítios e nem nos domínios funcionais da proteína DNA recombination/repair_Rad51_C-terminal (107-303), AAA+ATPase (119-294) e ReacA_ATP-biding (98-271) (a numeração dos resíduos se refere à isoforma 8). Considerando a estrutura 3D da proteína, previu-se que a Ala96 está localizada em uma alça e, embora seja substituída por uma treonina, não foi previsto impacto na estrutura 3D (Fig. 37).

Figura 37- Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.286G>A; (p.Ala96Thr))



Nota: Em A está sendo mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada que é indicada na figura B. A *RAD51D* c.286G>A; (p.Ala96Thr) está localizada em uma alça, indicada por uma seta**.**

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.6.1 Reclassificação da variante *RAD51D* c.286G>A; (p.Ala96Thr)

Considerando as informações acima, a *RAD51D* c.286G>A; (p.Ala96Thr) está classificada nas diferentes bases de dados como benigna ou provavelmente benigna. As evidências encontradas no presente estudo reforçam a classificação como provavelmente benigna, segundo os critérios ACMG/AMP: PP2 (ver na Seção 3.2.2), BP4 e BP6.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A *RAD51D* c.286G>A; (p.Ala96Thr) acontece em uma isoforma não canônica, que é conservada apenas em humanos e primatas, limitando uma análise aprofundada da conservação evolutiva desse resíduo. A alanina 96 é conservada nas quatro espécies para as quais se tem informação sobre esta isoforma;
- Além disso, os preditores *in silico* utilizados sugerem efeitos benignos para esta substituição;

- Nenhuma alteração potencial nas interações intermoleculares (exceto alterações menores nas ligações polares, provavelmente desprovidas de impacto funcional) foram previstas.

BP6 - Fonte respeitável relata recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente

A variante é classificada como benigna em banco de dados públicos como ClinVar, VarSome e ABraOM.

3.2.2.6.2 Conclusão sobre a reclassificação da variante *RAD51D* c.286G>A; (p.Ala96Thr)

Em suma, semelhante à variante *RAD51D* c.190C>T; (p.Pro64Ser), a *RAD51D* c.286G>A; (p.Ala96Thr) afeta a isoforma 8. Devido ao *splicing* alternativo, a isoforma 8 contém o éxon 5. Na avaliação da conservação evolutiva, a isoforma 8 foi encontrada em três espécies, além de humanos. Dentre essas espécies, Ala96 é altamente conservada. Aqui, é importante salientar que se a análise da conservação evolutiva for feita sem separar as isoformas, este resíduo seria considerado não conservado, pesando possivelmente na classificação desta variante como benigna.

A variante é classificada como benigna pelos algoritmos utilizados, considerando que não foram encontrados efeitos significativos sobre os motivos, domínios, sítios de *splicing*, elementos regulatórios de *splicing* ou interações interatômicas. No ClinVar e VarSome, a variante é reportada como benigna/provavelmente benigna, considerando a frequência alélica, como, como no gnomAD no gnomAD (0.0003), no 1000 Genomes Project (0.0002) e no Exome Project Genome (0.00019), e ausência de relatos na literatura, que associam esta variante a casos de câncer de mama e/ou ovário. Diante desses resultados, propomos manter esta variante como provavelmente benigna.

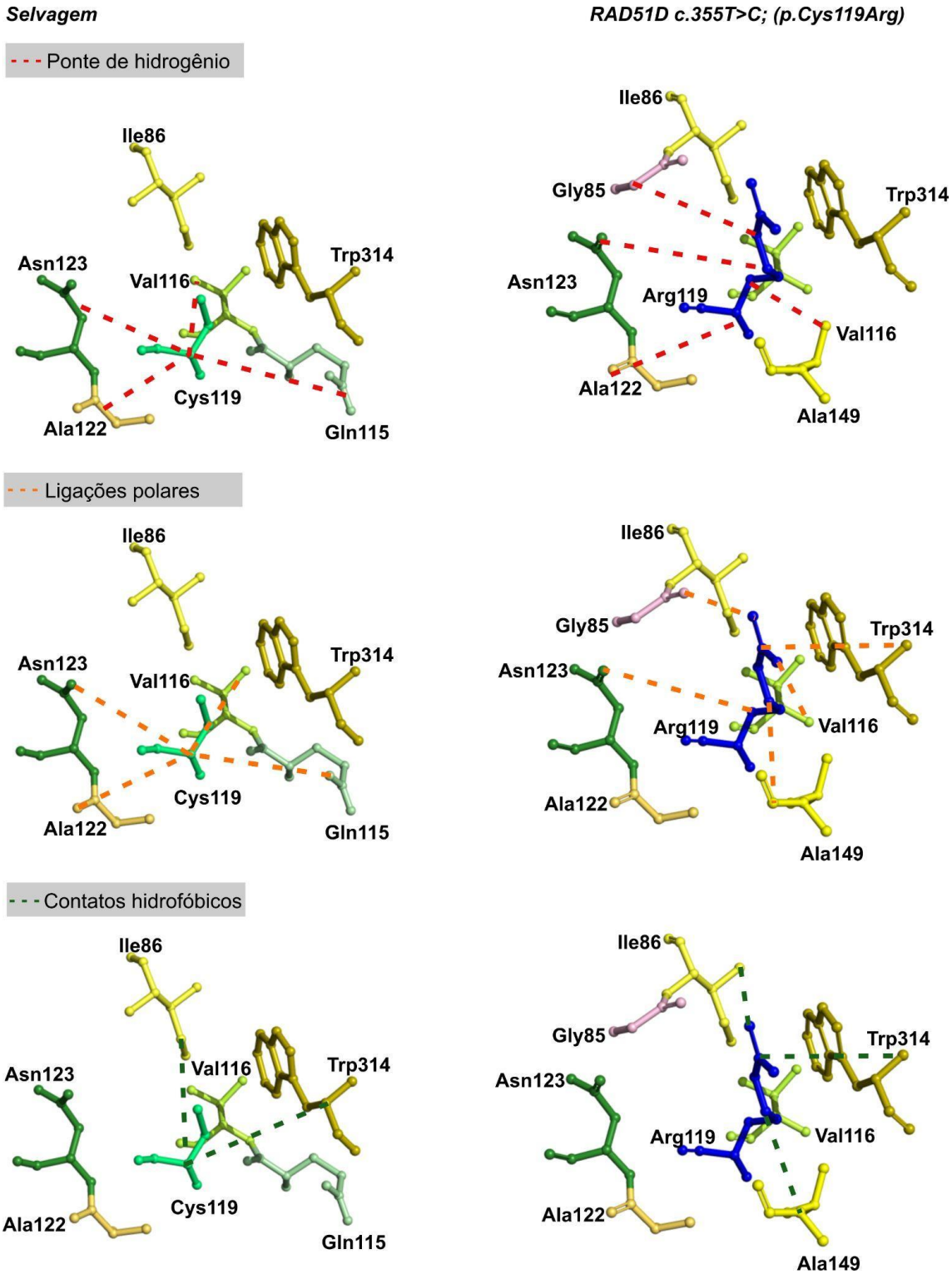
3.2.2.7 *RAD51D* c.355T>C; (p.Cys119Arg)

Localizada no éxon 7 do gene e no éxon 3 do transcrito canônico ENST00000345365.11 (Fig. 12 e seção 3.2), resulta de uma substituição de T para C na posição nucleotídica 355. Para essa substituição, não foram previstos efeitos significativos nas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores).

A cisteína no códon 119 é substituída por arginina, um aminoácido com propriedades diferentes, visto que a cisteína é polar não carregada e a arginina é básica. Os preditores *in silico* Mutation Taster, LIST-02, FATH-MM, Provean, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6).

No presente estudo, usando-se Dynamute, foi previsto que a Cys119 faça pontes de hidrogênio (Cys119 → Gln115, Cys119 → Val116, Cys119 → Ala122 e Cys119 → Asn123), contatos hidrofóbicos (Cys119 → Ile86, Cys119 → Trp314), interações polares (Cys119 → Gln115, Cys119 → Val116, Cys119 → Ala122, Cys119 → Asn123) (Fig. 38). Após a entrada da Arg119, algumas interações intramoleculares foram perdidas, como pontes de hidrogênio (Cys1199 → Gln115), polares (Cys119 → Gln115 e Cys119 → Asn123). Outras foram criadas, tal como pontes de hidrogênio (Arg119 → Gly85), contato hidrofóbico (Arg119 → Leu149), polar (Arg119 → Gly85) (Fig. 38). Usando-se Dianna, previu-se também que a Cys119 poderia estabelecer uma ponte dissulfeto com a Cys117. A perda de Cys119 levaria a perda dessa ligação covalente (Fig. 39). A perda da Cys119 também levaria à desestabilização da estrutura proteica, afetando a habilidade de interação com o ssDNA e com outras proteínas (Quadro 6, na Seção 3.2.6).

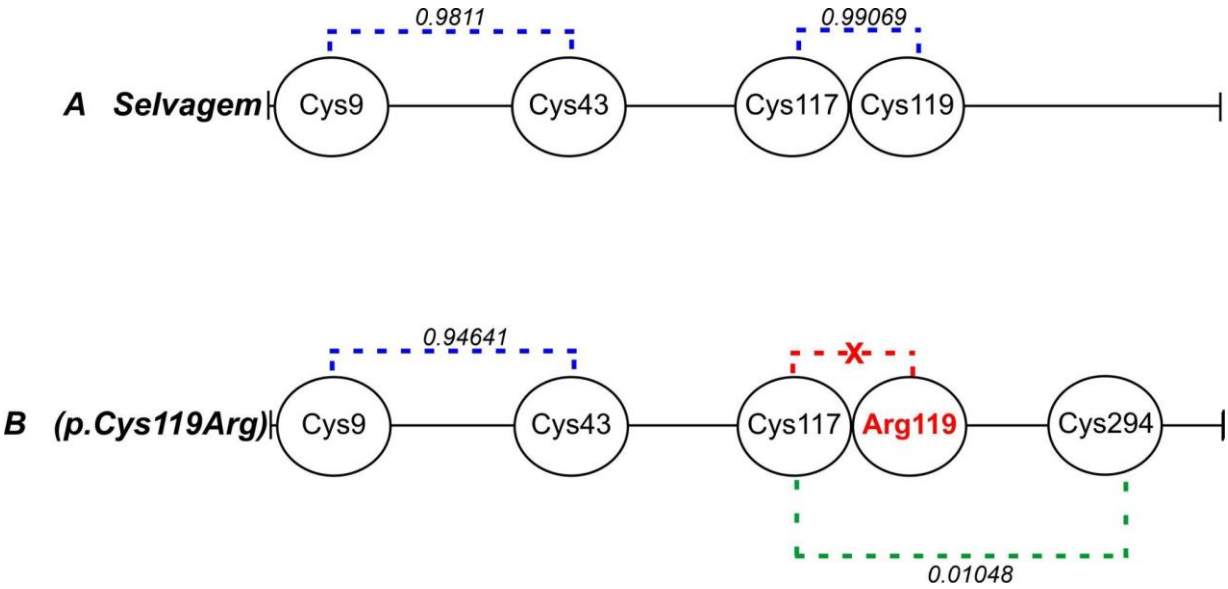
Figura 38 - Dinâmica molecular da variante *RAD51D* c.355T>C; (p.Cys119Arg)



Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Ile86, Gln115, Val116, Cys119, Ala122, Asn123, Ala149 e Trp314. Na estrutura mutada, estas interações se mantêm, mas surgem novas, além da arginina substituindo a cisteína na posição 119, uma Ala149 e uma Gly85. É importante salientar que os programas usados nesta predição não usam a informação relativa às pontes de dissulfeto, preditas entre a Cys117 e a Cys119.

Fonte: Elaborada pela autora, usando Dynamute.

Figura 39 - Impacto da variante *RAD51D* c.355T>C; (p.Cys119Arg) sobre pontes de dissulfeto da RAD51D usando Dianna

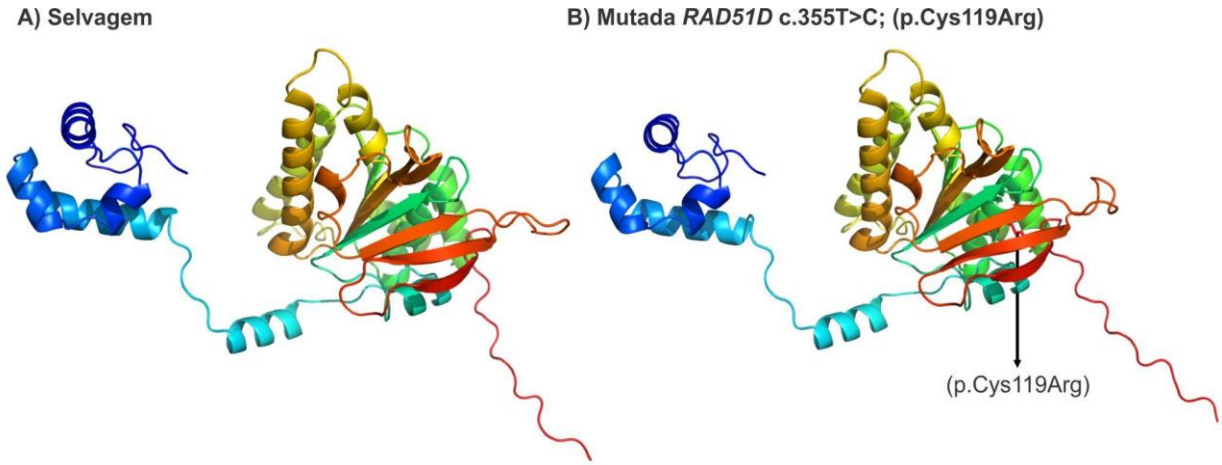


Nota: Em A, são mostrados os aminoácidos Cys9, Cys43, Cys117, Cys119 e Cys294, resíduos potenciais nas interações de pontes de dissulfeto. As pontes de dissulfeto são representadas pelas linhas pontilhadas azuis, por exemplo, a Cys9 interage com a Cys43 e assim por diante. Em B, é mostrada a perda da ponte de dissulfeto prevista entre a Cys117 e a Cys119, na presença da variante *RAD51D* c.355T>C; (p.Cys119Arg). O **X** indica a perda da ligação covalente. Os números junto às linhas tracejadas indicam o escore de probabilidade atribuído pelo Dianna para cada ponte de dissulfeto predita.

Fonte: Elaborada pela autora.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Cys119 está localizada nos três domínios funcionais RecA_ATP-biding (75-283), DNA recombination and repair protein (78-251) e AAA+ ATPase (99-274) (Quadro 6, na seção 3.2.6). Considerando-se a estrutura 3D da proteína, previu-se que a Cys119 está localizada em uma alça e, não foi previsto impacto na estrutura 3D devido a entrada de uma arginina. Entretanto, os efeitos da perda da ponte de dissulfeto não são avaliados pelo *software* usado (Figura 40).

Figura 40 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.355T>C; (p.Cys119Arg))



Nota: Em A, a estrutura da proteína 3D selvagem; em B, a da proteína mutada. A *RAD51D* c.355T>C; (p.Cys119Arg) está localizada em uma alça, indicada por uma seta.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.7.1 Reclassificação da variante *RAD51D* c.355T>C; (p.Cys119Arg)

Considerando as informações acima, a RAD51D c.355T>C; (p.Cys119Arg) manteria sua classificação VUS, segundo os critérios ACMG/AMP: PM2, PP2 (ver na Seção 3.2.2), BP4.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica em bancos de dados populacionais: gnomAD (0.00005), Exome Sequencing Project (0.00005) e não encontrada no 1000 Genomes Project. Além disso, não há nenhuma publicação associada a esta variante.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- Os preditores *in silic*o utilizados sugerem efeitos benignos para esta substituição, provavelmente por considerarem a conservação evolutiva e regiões funcionais;

- No presente estudo, observou-se que a Cys119 é conservada em espécies de mamíferos, exceto por espécies muito basais. Em aves, anfíbios e répteis e peixes, ela é substituída por outros aminoácidos nessa posição;

- No presente estudo, foram previstas inúmeras alterações nas interações moleculares, em que a maioria das ligações (pontes de hidrogênio, ligações polares etc) e um sítio potencial (escore de probabilidade = 0,99069) para estabelecimento de ponte de dissulfeto perdido; novas interações foram criadas.

- Entretanto, estes achados (conservação evolutiva em mamíferos e provável participação em ponte de dissulfeto) não são evidências suficientes de patogenicidade.

3.2.2.7.2 Conclusão sobre a reclassificação da *RAD51D* c.355T>C; (p.Cys119Arg)

Em suma, a perda de Cys119 implica na perda de um potencial ponte de dissulfeto, além de algumas interações interatômicas. Ambas as alterações poderiam causar mudanças conformacionais, que alterariam a dinâmica molecular "natural" da proteína, permitindo que novas interações ou mesmo novas funções ocorram. Contudo, os algoritmos usados neste estudo prevêem que a variante seja provavelmente benigna, pois consideraram que o resíduo seria pouco conservado, e que esta variante não levaria a perda de um domínio/motivo funcional específico.

Em relação a conservação evolutiva, a Cys119 é altamente conservada em mamíferos e torna-se menos conservada evolutivamente à medida que as espécies se tornam mais distantes, como aves, anfíbios e répteis e peixes, onde a Cys119 é frequentemente substituída por serina ou glicina. Em nove espécies de mamíferos esta cisteína foi substituída. Estas espécies estão agrupadas na base da radiação dos mamíferos, sugerindo que a Cys119 é conservada ao longo de 40 a 65 milhões de anos milhões de anos. Em cinco espécies de mamíferos (*Neophocaena asiaeorientalis* - cetáceo; *lagenorhynchus obliquidens* - golfinho de lado branco; *tursiops truncatus* - golfinho nariz de garrafa; *Orcinus orca* - Orca; *Globicephala melas* - Baleia-piloto-de-peitorais-longas) e uma das espécies de peixes, o *Xiphias gladius* (peixe-espada), há uma arginina nessa posição. Entretanto, considerando-se a distância evolutiva, seria importante avaliar o efeito desta substituição no conjunto da proteína.

Considerando-se as regiões funcionais da proteína, a Cys119 está localizada em grandes domínios funcionais, mas não impacta significativamente estas regiões. No ClinVar, a interpretação da patogenicidade dessa variante é conflitante, com oito entradas classificando-a como VUS e uma como provavelmente benigna. Além disso, não existem variantes com a mesma alteração de aminoácidos nesta posição previamente classificadas como patogênicas. De fato, são poucas as variantes afetando a Cys119, e elas são classificadas como VUS (*RAD51D* c.356G>C (p.Cys119Ser) e *RAD51D* c.356G>A (p.Cys119Tyr)), ou benigna (*RAD51D* c.357T>C (p.Cys119=)). A *RAD51D* c.355T>C; (p.Cys119Arg) tem baixa

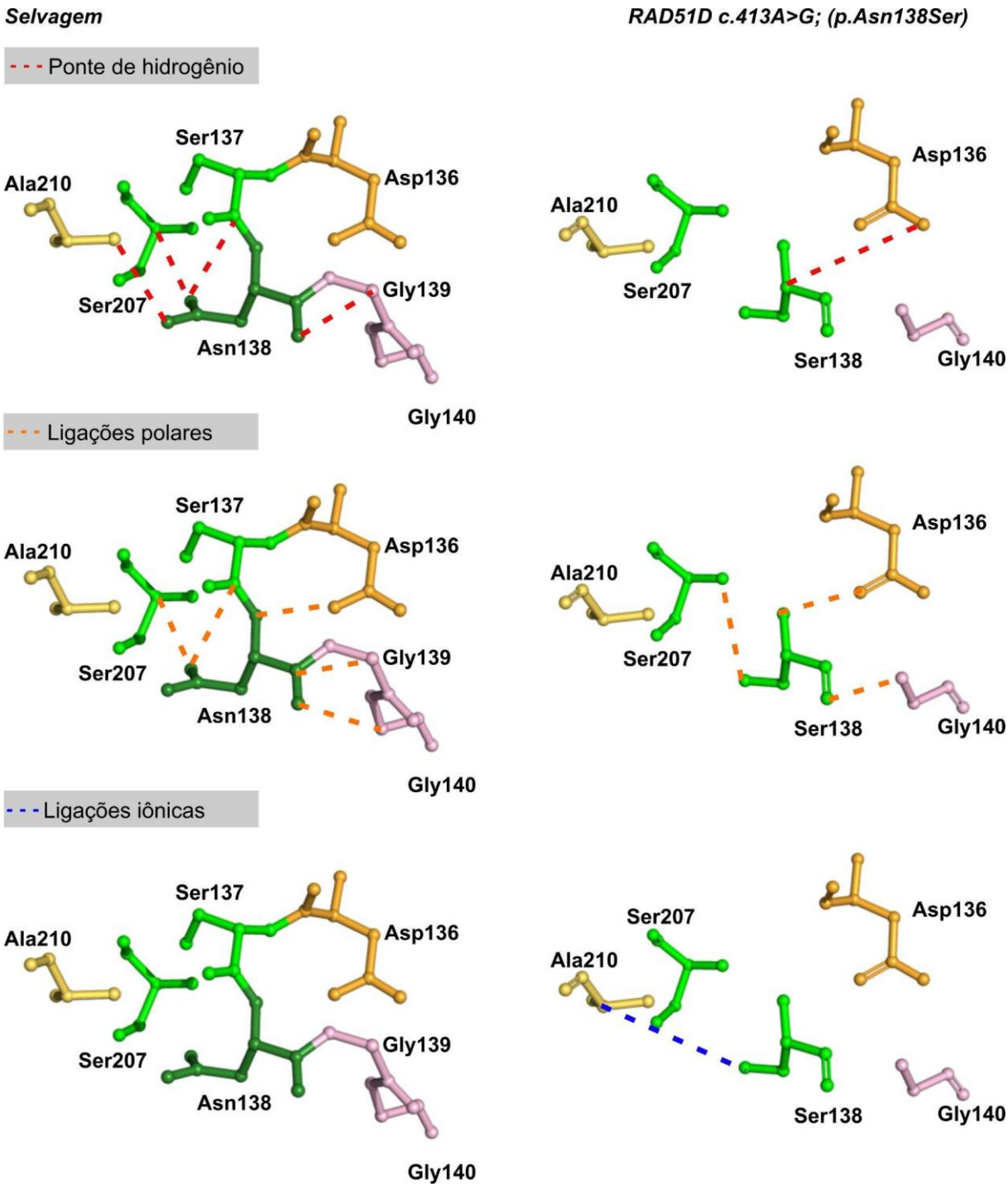
frequência em bancos de dados populacionais e está associada a casos de câncer de mama e/ou ovário na literatura (Osher *et al.,* 2012; Gutierrez-Enrıquez *et al.,* 2013; Song *et al.,* 2015; Penkert *et al.,* 2018). Não foram encontrados dados experimentais que avaliem o impacto desta variante na função da proteína. No presente estudo, essa variante foi encontrada em um paciente de 14 anos diagnosticado com linfoma de Hodgkin. Até o presente momento, esse fenótipo não está relacionado a variantes em *RAD51D*, consequentemente foi considerado um achado incidental. Considerando os dados obtidos para essa variante, sua classificação continuaria sendo VUS.

3.2.2.8 *RAD51D* c.413A>G; (p.Asn138Ser)

Localizada no éxon 7 do gene e no éxon 3 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), resulta de uma substituição de A para G na posição nucleotídica 413. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing*.

A asparagina no códon 138 é substituída por serina, um aminoácido com as mesmas propriedades físico-químicas, sendo ambos polares não carregados. Os preditores *in silico* Mutation Taster, LIST-02, FATH-MM, Provean, Polyphen, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação evolutiva do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Foi previsto que a Asn138 estabeleça pontes de hidrogênio (Asn138 → Asp136, Asn138 → Ser137, Asn138 → Gly139, Asn138 → Ser207 e Asn138 → Ala210), interações polares (Asn138 → Asp136, Asn138 → Ser137, Asn138 → Gly139, Asn138 → Gly140, Asn138 → Ser207 (Fig. 41). A maioria das interações foram perdidas após a troca dos aminoácidos, como as pontes de hidrogênio (Asn138 → Ser137, Asn138 → Gly139, Asn138 → Ser207 e Asn138 → Ala210) e as interações polares (Asn138 → Ser137, Asn138 → Gly139, Asn138 → Ser207). Uma nova ligação iônica foi criada (Ser138 → Ala210) (Fig. 41). No presente estudo, foi predito que a perda da Asn138 desestabiliza a estrutura da proteína, afetando sua habilidade de interação com o DNA e proteínas (Quadro 6, na Seção 3.2.6).

Figura 41 - Dinâmica molecular da variante *RAD51D* c.413A>G; (p.Asn138Ser)

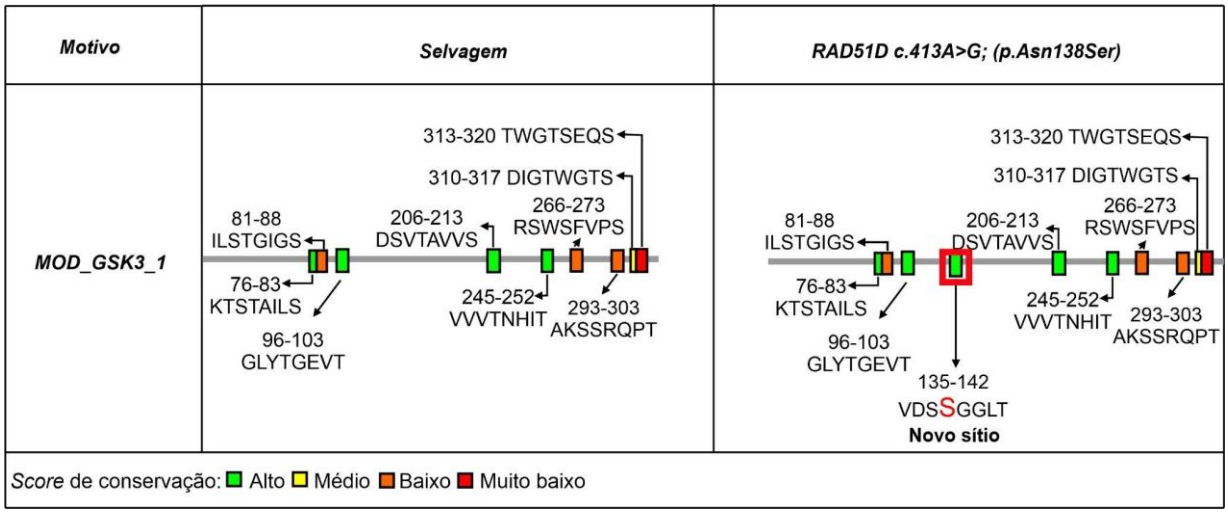


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em azul, ligações iônicas. Na estrutura selvagem, as interações acontecem entre os resíduos: Asp136, Ser137, Asn138, Gly139, Gly140, Ser207, Ala210. Na estrutura mutada, perdem-se as interações com os aminoácidos Ser137 e Gly139.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Asn138 está localizada nos três domínios funcionais *RecA_ATP-biding* (75-283), *DNA recombination and repair protein* (78-251) e *AAA+ ATPase* (99-274) (Tab. 3). A entrada da Ser138 criou um novo sítio de alta conservação, o do motivo *MOD_GSK3_1 (sítio de fosforilação GSK3)* na posição 135-142 [VDSSGGLT]. No entanto, não foram encontrados dados na literatura que confirmem essa predição (Fig. 42).

Figura 42 - Predição de impacto de *RAD51D* c.413A>G; (p.Asn138Ser) nos motivos proteicos

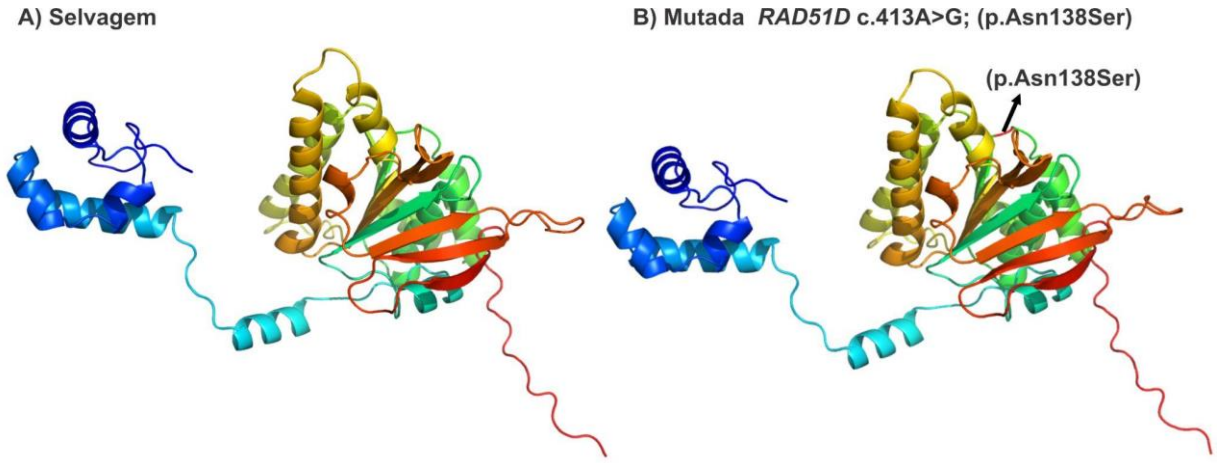


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo. Aqui, o novo sítio criado (VDS GGLT) é um motivo de fosforilação.

Fonte: Elaborada pela autora, usando ELM Motifs.

Considerando-se a estrutura 3D da proteína, previu-se que a Asn138 está localizada em uma alça e, embora seja substituída por uma serina, não foi previsto impacto a estrutura 3D (Fig. 43).

Figura 43- Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.413A>G; (p.Asn138Ser))



Nota: Em A, está sendo mostrada a estrutura 3D da proteína 3D selvagem; em B, na presença da variante *RAD51D* c.413A>G; (p.Asn138Ser), localizada em uma alça, indicada por uma seta. A estrutura 3D não mostra mudanças conformacionais significativas.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.8.1 Reclassificação da variante *RAD51D* c.413A>G; (p.Asn138Ser)

Considerando as informações acima, propomos manter a classificação da *RAD51D* c.413A>G; (p.Asn138Ser) VUS, segundo os critérios ACMG/AMP: PM1, PP2 (ver na Seção 3.2.2), BP4.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em *hotspot* de mutações, mas afeta domínios estruturais curtos, criando o sítio de alta conservação do motivo *MOD_GSK3_1* (sítio de fosforilação GSK3). Conforme avaliado no banco de dados ClinVar, esta região da variante não contém variantes *missense* benignas. Para esta região, as variantes benignas ou provavelmente benignas apresentam como consequência molecular substituições sinônimas.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A Asn138 não é conservada entre espécies de mamíferos, podendo aparecer três aminoácidos nesta posição, treonina, serina ou a própria asparagina. Nas espécies de aves, anfíbios e répteis, e peixes, o aminoácido presente nesta região foi sempre uma treonina;
- Os preditores *in silico* utilizados sugerem efeitos benignos para esta substituição;

- Foram previstas inúmeras alterações nas interações moleculares, em que a maioria das ligações (pontes de hidrogênio, ligações polares etc) foram perdidas. Entretanto, esse achado não é evidência suficiente de patogenicidade;
- Esta variante cria um sítio de alta conservação do motivo *MOD_GSK3_1* (sítio de fosforilação GSK3), mas nenhum dado referente a atividade desses motivos na função da RAD51D foi encontrado na literatura.

3.2.2.8.2 Conclusão sobre a reclassificação da variante *RAD51D* c.413A>G; (p.Asn138Ser)

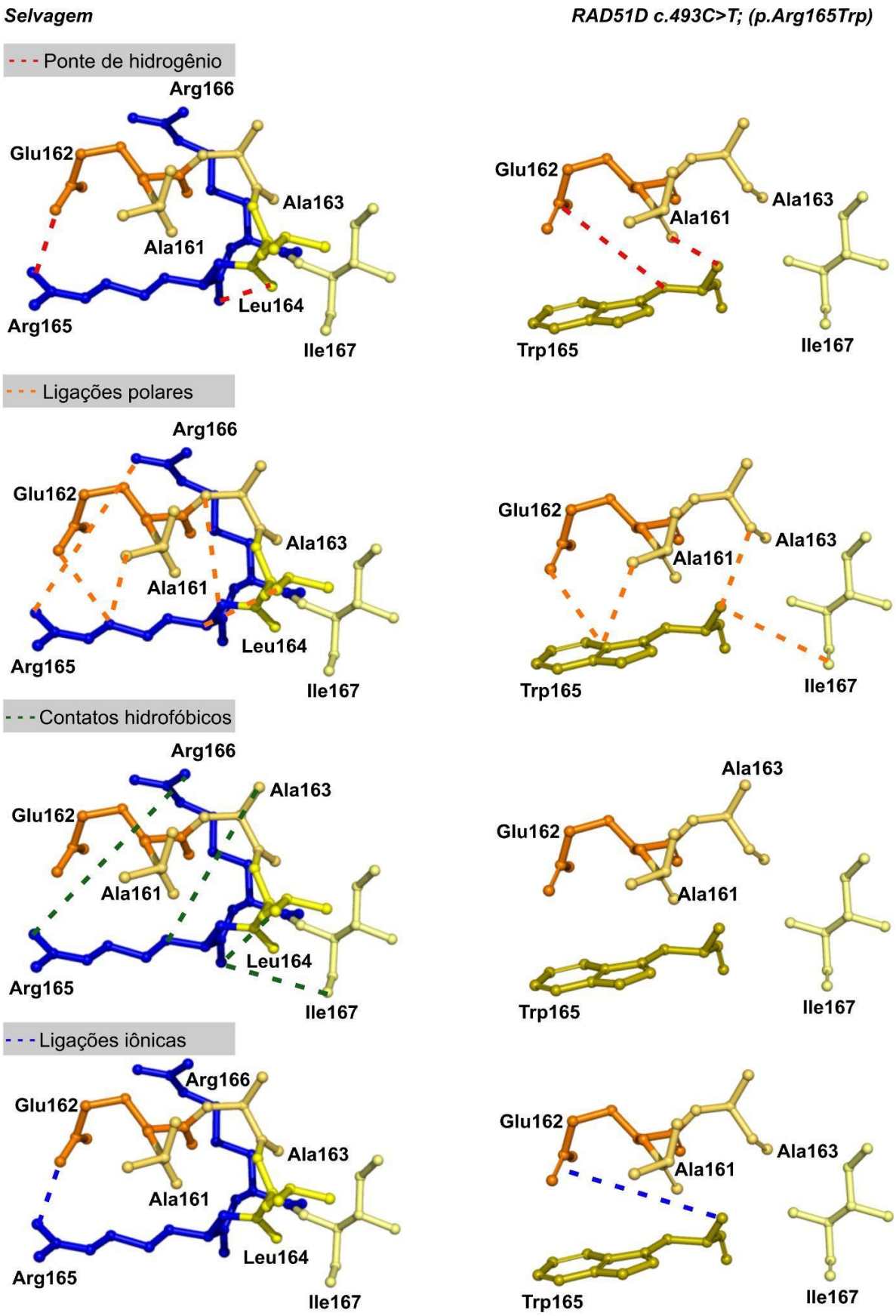
Em síntese, embora os preditores *in silico* prevejam efeitos benignos para esta variante, ela afeta interações intermoleculares e cria um novo sítio de fosforilação (*MOD_GSK3_1*). As interações intermoleculares moldam a dinâmica das proteínas e os domínios estruturais são abertos a influência de estímulos externos, que conduzem mudanças conformacionais "naturais" frequentes para estimular respostas a processos celulares (por exemplo, transdução de sinal). No ClinVar, a interpretação da patogenicidade dessa variante é conflitante, com quatro entradas para VUS e quatro para provavelmente benigna. Na literatura, a variante é associada a casos de câncer de mama e/ou ovário, mas não foi relatada associação com a síndrome hereditária do câncer de mama e ovário (Johnson *et al.,* 2014; Tung *et al.,* 2015; Ring *et al.,* 2016; Shindo *et al.,* 2017; Sánchez-Bermúdez *et al.,* 2018; Tsaousis *et al.,* 2019). Por exemplo, no estudo de Sánchez-Bermúdez *et al.,* 2018, esta variante foi identificada em uma família onde a mãe e a filha tiveram câncer de ovário. Além disso, neste mesmo estudo a variante foi encontrada em um controle saudável em heterozigose. Tung *et al.,* 2015, também identificaram esta variante em uma mulher com menos de 50 anos diagnosticada com câncer de mama. As frequências populacionais da variante foram: gnomAD, 0.0001; 1000 Genomes Project, 0.0002; e, Exome Project Genome, 0.0002. As outras variantes descritas em banco de dados, que envolvem esse mesmo resíduo, são classificadas como VUS (*RAD51D* c.412A>G (p.Asn138Asp) e *RAD51D* c.412A>C (p.Asn138His)) ou benigna (*RAD51D* c.414T>C (p.Asn138=)). A falta de dados clínicos como história familiar e dados de análise de segregação limita uma classificação mais assertiva dessa variante, mantendo sua classificação VUS.

3.2.2.9 *RAD51D* c.493C>T; (p.Arg165Trp)

Localizada no éxon 8 do gene e no éxon 5 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), esta variante resulta de uma substituição de C para T na posição nucleotídica 493. Para esta substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores ou silenciadores). A arginina no códon 165 é substituída por triptofano, um aminoácido com propriedades físico-químicas diferentes, visto que a arginina é um aminoácido básico e o triptofano é apolar. Os preditores *in silico* Mutation Taster, LIST-02, Provean, Polyphen e SIFT preveem **efeitos deletérios** para essa substituição, considerando a característica físico-química do resíduo e regiões funcionais na proteína (Quadro 6, na Seção 3.2.6).

No presente estudo, foi previsto que a Arg165 estabeleça ponte de hidrogênio (Arg165 → Ala161, Arg165 → Glu162, Arg165 → Leu164), contatos hidrofóbicos (Arg165 → Ala163, Arg165 → Leu164, Arg165 → Arg166, Arg165 → Ile167), ligações iônicas (Arg165 → Glu162), interações polares (R165 → Ala161, Arg165 → Glu162, Arg165 → Ala163, Arg165 → Leu164, Arg165 → Arg166, Arg165 → Ile167) (Fig. 44). Perderam-se todos os contatos hidrofóbicos, algumas pontes de hidrogênio (Arg165 → Leu164) e interações polares (Arg165 → Ala161, Arg165 → Ala163, Arg165 → Leu164, Arg165 → Arg166) (Fig. 44). A perda da Arg165 levaria a desestabilização da estrutura proteica, afetando a habilidade de interação com o DNA e proteínas (Quadro 6, na Seção 3.2.6), e levando à perda de desordem intrínseca em uma região, que não é significativamente desordenada (Fig. 45).

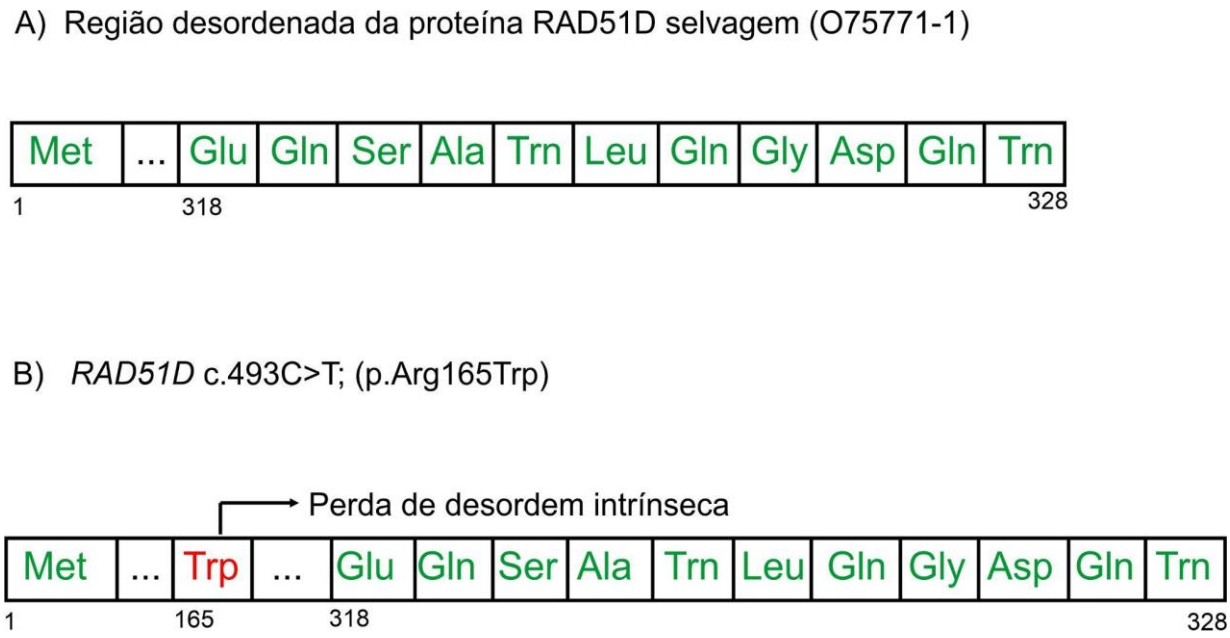
Figura 44 - Dinâmica molecular da variante *RAD51D* c.493C>T; (p.Arg165Trp)



Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico, e em azul, ligações iônicas. Na estrutura selvagem, as interações acontecem entre os resíduos: Glu162, Ala161, Ala163, Leu164, Arg165, Arg166, Ile167. Na estrutura mutada, perde-se interação com os aminoácidos Arg166 e Leu164.

Fonte: Elaborada pela autora, usando DynaMut.

Figura 45 - Impacto da variante RAD51D c.493C>T; (p.Arg165Trp) na desordem proteica

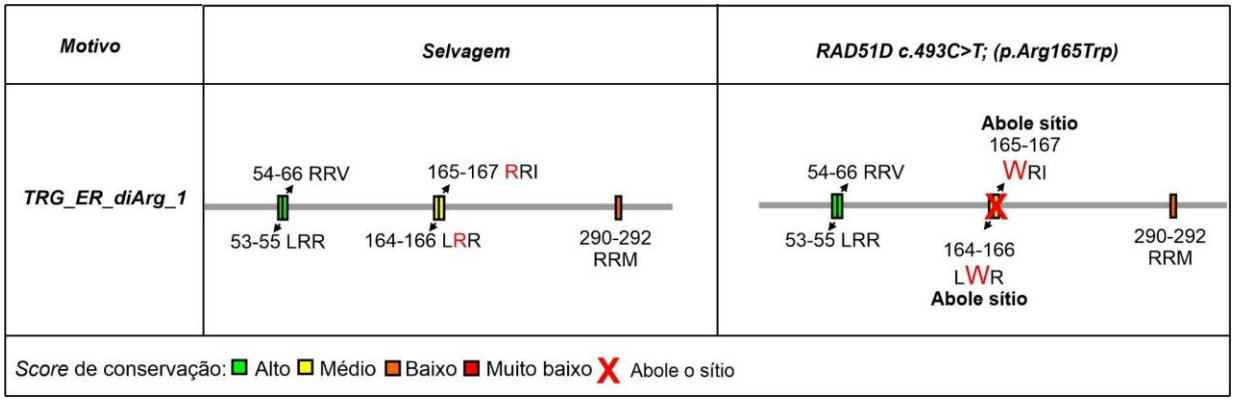


Nota: As informações sobre desordem intrínseca são mostradas com detalhes na seção sobre anotação funcional do gene e proteína. Aqui, está sendo mostrado o impacto da variante sobre a desordem. Em A, são os aminoácidos preditos como intrinsecamente desordenados na sequência selvagem. Em B, a entrada do Trp165 promove uma perda de desordem intrínseca (aminoácido vermelho na sequência) para este resíduo. É importante salientar, que este resíduo se encontra em uma região, onde a desordem intrínseca predita não havia atingido significância.

Fonte: Elaborada pela autora, usando PSIPred.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Arg165 está localizada nos domínios funcionais *RecA_ATP-biding* (75-283), *DNA recombination and repair protein* (78-251) e *AAA+ ATPase* (99-274), e em uma região funcional chamada *Coiled coil* (147-167) (Quadro 6, na Seção 3.2.6). A perda da Arg165 leva à perda de dois sítios moderadamente conservados do motivo *TRG_ER_diArg_1 (Motivo di-Arg)* nas posições 164-166 [LRR] e 165-167 [RRI] (Fig. 46). Não foram encontrados na literatura estudos descrevendo as repercussões funcionais desses motivos, bem como da sua perda.

Figura 46 - Predição de impacto de *RAD51D* c.493C>T; (p.Arg165Trp) nos motivos proteicos

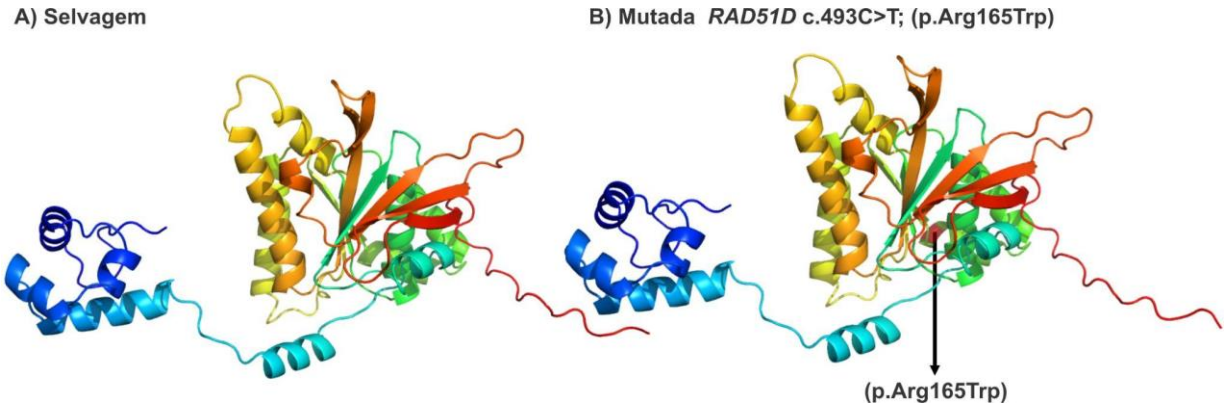


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo. Neste caso, a variante leva à perda do sítio.

Fonte: Elaborada pela autora, usando ELM Motifs.

Considerando-se a estrutura 3D da proteína, previu-se que a Arg165 está localizada em uma alfa-hélice e, embora seja substituída por um triptofano, não foi previsto impacto na estrutura 3D da proteína (Fig. 47).

Figura 47 - Estrutura 3D da RAD51D selvagem e da *RAD51D* c.493C>T; (p.Arg165Trp)



Nota: Em A, é mostrada a estrutura 3D da proteína selvagem, em B, da mutada. A *RAD51D* c.493C>T; (p.Arg165Trp) está localizada em uma alfa-hélice, indicada por uma seta. A estrutura 3D não mostra efeitos significativos sobre a dinâmica conformacional da RAD51D.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.9.1 Reclassificação da variante *RAD51D* c.493C>T; (p.Arg165Trp)

Considerando as informações acima, a *RAD51D* c.493C>T; (p.Arg165Trp) propomos manter a classificação como provavelmente patogênica, segundo os critérios ACMG/AMP: PM1, PM2, PP2 (ver Seção 3.2.2) e PP3.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em *hotspots* de mutações, mas afeta domínios estruturais curtos, levando à perda de domínios estruturais curtos (Motif di-Arg). Conforme avaliado no banco de dados ClinVar, esta região da variante contém uma variante *missense* benigna nesta posição *RAD51D* c.494G>A (p.Arg165Gln). No entanto, a glutamina é um resíduo fundador, uma vez que nesta posição apenas os resíduos de arginina ou glutamina são vistos, e a glutamina é o mais frequente em todas as espécies. Em vista disso, o critério PM1 é aplicado.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui frequência alélica baixa em bancos de dados populacionais, como gnomAD (0.00002), 1000 Genomes Project (0.0002) e não foi encontrada no Exome Sequencing Project).

PP3 – Múltiplas linhas de evidência computacional suportam um efeito deletério no gene ou produto gênico

- A Arg165 não é conservada entre espécies de mamíferos, aves, anfíbios e répteis e peixes. Até peixes, apenas dois aminoácidos são vistos na posição 165, arginina ou glutamina. A grande maioria das espécies apresenta uma glutamina na posição 165. Em aves e peixes, todas as espécies apresentam uma glutamina. Em anfíbios e répteis, todas as espécies amostradas nas bases de dados apresentam uma glutamina, exceto pela *Chelonia mydas*, que apresenta uma arginina. Em mamíferos, uma arginina aparece nesta posição em *Homo sapiens, Pan troglodytes, Gorilla gorilla, Artibeus jamaicensis, Loxodonta africana* e *Elephans maximus indicus*, sugerindo: 1. evolução convergente; 2. que a entrada de um triptofano não seja inócua, já que em nenhuma das espécies de vertebrados avaliadas um triptofano não ocupa esta posição;
- Os preditores *in silico* utilizados sugerem, que esta substituição seja deletéria;
- Foi previsto que a perda da Arg165 leve à perda de desordem intrínseca em uma região, que não é significativamente desordenada, achado este difícil de ponderar;
- Foi previsto que esta substituição afeta não só a função e a estabilidade da estrutura da proteína, mas também as diferentes características moleculares da RAD51D, que se prevê serem funcionais (sítios de interação e pequenos domínios de proteínas). Entretanto, não foram encontrados estudos descrevendo as repercussões funcionais destes motivos ou da sua perda.
- Foram previstas alterações nas interações moleculares, em que a maioria delas foi perdida. Essas mudanças podem contribuir para diferentes conformações da estrutura 3D da RAD51D.

3.2.2.9.2 Conclusão sobre a reclassificação da variante *RAD51D* c.493C>T; (p.Arg165Trp)

Em suma, observou-se na predição *in silico* que a Arg165 está localizada em regiões funcionais da proteína (domínios) e quando substituída por um triptofano conduz à perda de domínios estruturais curtos (Motivo di-Arg). O motivo di-Arg funciona como um domínio de recuperação e retenção no RE (retículo endoplasmático), cuja função é controlar o dobramento correto e a montagem do complexo proteico, que governa a saída do RE. Normalmente, este motivo é composto por dois resíduos de arginina situados lado a lado ou pode conter entre duas argininas ladeando um resíduo de aminoácido não conservado (RR ou

RXR), seguido por um resíduo hidrofóbico ou arginina adjacente que pode estar em ambos os lados do par RR. Este motivo é comumente encontrado em proteínas de membrana (algumas multiméricas) e, portanto, deve-se desconsiderar a correspondência de motivo di-Arg em proteínas não membranares (Michelsen *et al.*, 2005; Boulaflous *et al.*, 2009; Uemura *et al.*, 2009). No entanto, embora se preveja que a variante altere regiões funcionais da proteína, o critério PM1 não se aplicaria, devido à existência de uma variante de sentido trocado com a mesma posição (*RAD51D* c.494G>A;(p.Arg165Gln)) classificada como benigna no ClinVar, considerando principalmente a frequência populacional, que é próxima a 1% (Loveday *et al.*, 2011; Osher *et al.*, 2012; Jiraskova *et al.*, 2018). Entretanto, a variante p.Arg165Gln retorna o resíduo a sua sequência original, do ponto de vista evolutivo. Portanto, o fato de existir esta variante p.Arg165Gln, benigna, não exclui que a entrada de um triptofano nesta região seja patogênica. A propósito, a substituição de uma Arg por um Trp nesta posição é classificada como patogênica pelo SIFT. Consequentemente, o critério PM1 se aplicaria.

Na conservação evolutiva, duas argininas são vistas lado a lado, Arg165 e Arg166 apenas em *Homo sapien*s, *Pan troglodytes* e *Gorilla gorilla*, em *Loxodonta africana* e *Elephas maximus indicus* e em *Artibeus jamaicensis* (morcego da cara branca). Nas demais espécies de mamíferos, aves, anfíbios e répteis e peixes, a posição 165 é ocupada por uma glutamina, nenhum outro aminoácido aparece nesta posição. Estes achados sugerem, que a entrada de uma glutamina nesta posição pode ser bem tolerada. A Arg166, por outro lado, é muito conservada, com pouquíssimas espécies tendo uma lisina nesta posição, o que ocorre em tartarugas (*Mauremys mutica* e *Mauremys reevesii*) e mamíferos.

Os algoritmos utilizados preveem efeitos deletérios para a substituição p.Arg165Trp. Isso se deve às diferenças nas características físico-químicas entre os resíduos trocados, à localização em regiões funcionais (domínios estruturais curtos) e à baixa frequência populacional. Contudo, no ClinVar, a interpretação da patogenicidade dessa variante é conflitante, havendo nove submissões como VUS e uma como provavelmente patogênica (Janeiro, 2023). Nos bancos de dados populacionais, a variante é rara no gnomAD (0.00002) e no 1000 Genomes Project (0.0002), e não foi encontrada no Exome Project Genomes. Não foram encontrados na literatura dados experimentais, que expliquem as consequências moleculares desta variante na função proteica, porém alguns estudos têm associado esta variante a casos de câncer de mama e/ou ovário (Loveday *et al.*, 2011; Song *et al.*, 2015; Dominguez-Valentin *et al.*, 2018). Em nosso estudo, a variante foi relatada em um paciente com câncer de pâncreas, mas sem dados de história familiar ou segregação da variante com a doença. Mesmo com a falta de dados

clínicos sobre esta variante, sugerimos reclassificá-la para provavelmente patogênica, considerando seus impactos previstos na função da proteína citada acima.

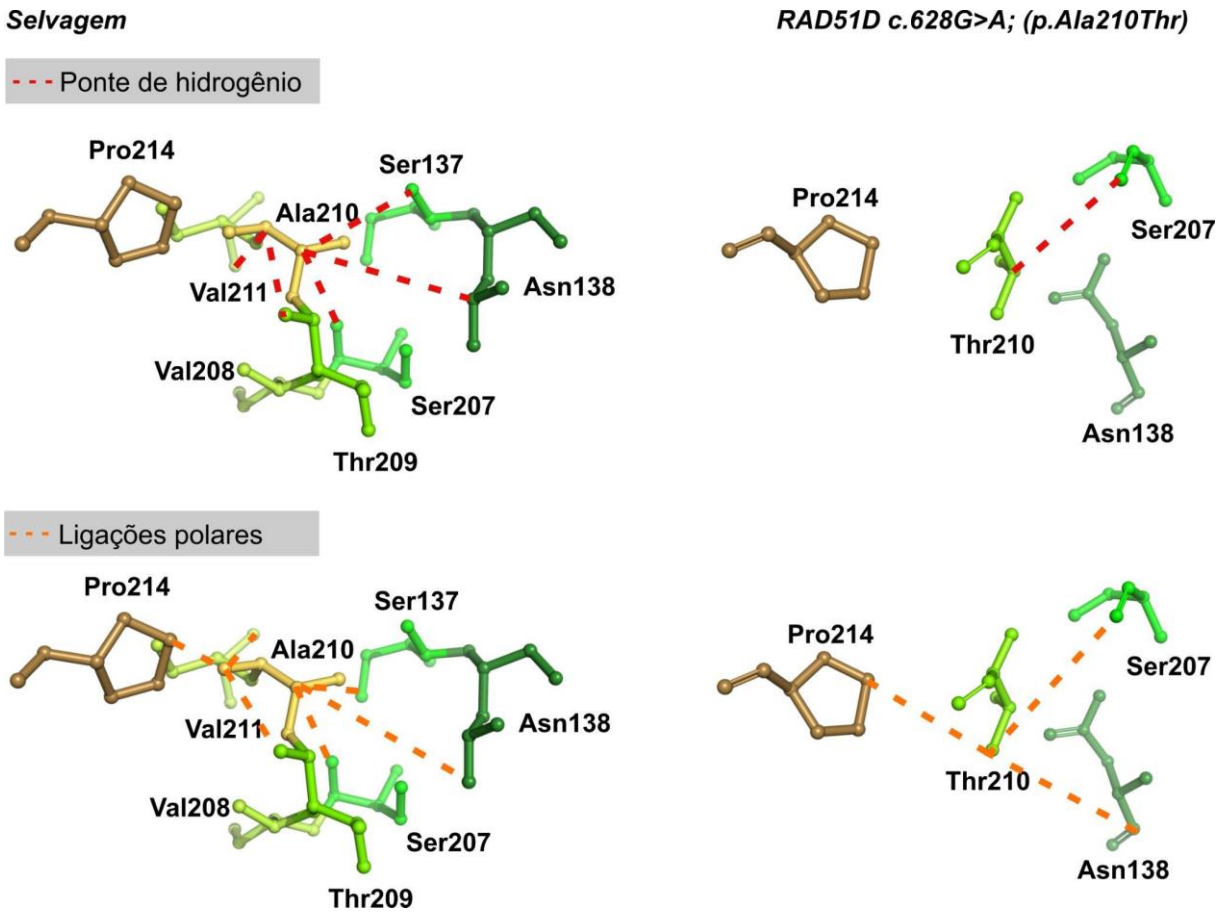
3.2.2.10 *RAD51D* c.628G>A;(p.Ala210Thr)

Localizada no éxon 10 do gene e no éxon 6 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), resulta de uma substituição de G para A na posição nucleotídica 628. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares regulatórias de *splicing* (acentuadores e silenciadores).

A alanina no códon 210 é substituída por treonina, um aminoácido com propriedades físico-químicas diferentes, visto que a alanina é um aminoácido apolar e a treonina um polar não carregado. A maioria dos preditores *in silico* utilizados, Mutation Taster, FATH-MM, Provean e SIFT preveem **efeitos benignos** para essa substituição (Quadro 6, na Seção 3.2.6).

No presente estudo, foi previsto que a Ala210 faça pontes de hidrogênio (Ala210 → Ser137, Ala210 → Asn138, Ala210 → Ser207, Ala210 → Thr209, Ala210 → Val211) e interações polares (Ala210 → Ser137, Ala210 → Asn138, Ala210 → Ser207, Ala210 → Val208, Ala210 → Thr209, Ala210 → Val211, Ala210 → Pro214) (Fig. 48). Perderam-se interações como pontes de hidrogênio (Ala210 → Ser137, Ala210 → Asn138, Ala210 → Thr209, Ala210 → Val211) e polares (Ala210 → Ser137, Ala210 → Asn138, Ala210-Ser207, Ala210 → Val208, Ala210 → Thr209, Ala210 → Val211) (Fig. 48). A perda da Ala210 leva a desestabilização da estrutura proteica, desestabilizando a habilidade de interação com o DNA e com outras proteínas (Quadro 6, na Seção 3.2.6).

Figura 48 - Dinâmica molecular da variante *RAD51D* c.628G>A; (p.Ala210Thr)



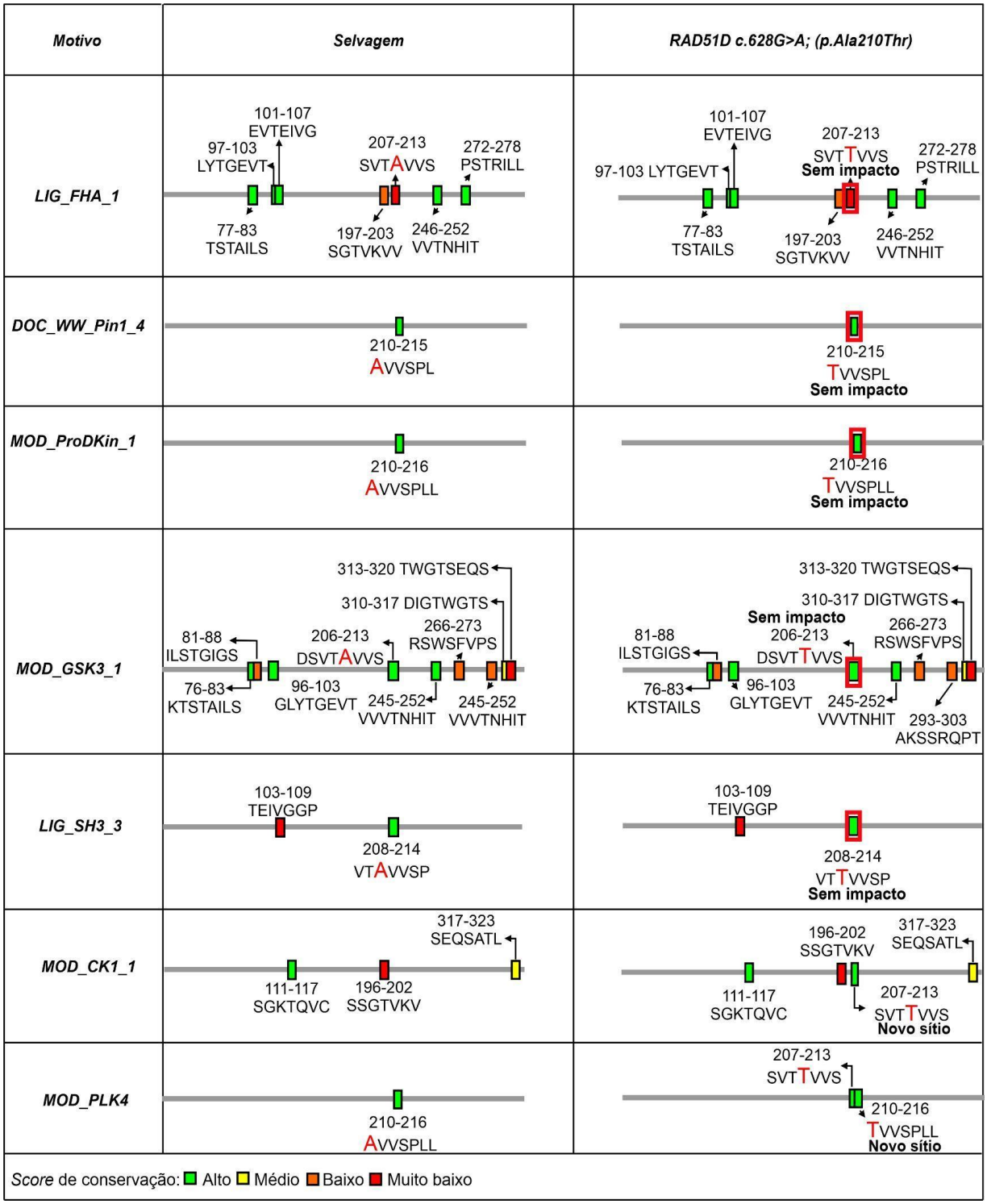
Nota: Interações interatômicas nas estruturas selvagem e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares. Na estrutura selvagem, as interações acontecem entre os resíduos: Ser137, Asn138, Ser207, Val208, Thr209, Ala210, Val211, Pro214. Na estrutura mutada, perdem-se interação com os aminoácidos Val208, Thr209, Ala210 e Val211.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Ala210 está localizada nos domínios funcionais *RecA_ATP-biding (75-283), DNA recombination and repair protein (78-251) e* AAA+ ATPase (99-274) (Quadro 6), bem como nos motivos: *LIG_FHA_1* (*Motivo fosfopeptídeo de ligação ao forkhead-associated - FHA*) na posição 207-213 [SVTAVVS] (Fig. 49); *MOD_GSK3_1* (*Sítio de fosforilação GSK3*) na posição 206-213 [DSVTAVVS]; *DOC_WW_Pin1_4 (Motivos de*

ligação ao domínio WW - medeiam a interação de proteínas sinalizadoras e regulatórias) na posição 210-215 [AVVSPL]*; MOD_ProDKin 1 (sítio de fosforilação de* MAPK) na posição 210-216 [AVVSPLL]; e *LIG_SH3_3* (*Motivo de ligação ao SH3)* na posição 208-214 [VTAVVSP] (Fig. 49). Embora a Ala210 esteja localizada nesses motivos, não foi predita a perda dos mesmos. A entrada da Thr210 criou um novo sítio de alta conservação do motivo *MOD_CK1_1* (*Sítio de fosforilação de Casein kinase 1 - CK1*) na posição 207-213 [SVTTVVS] e um novo sítio de alta conservação do motivo *MOD_PLK4 (Sítio de fosforilação Plk4)* na posição 207-213 [SVTTVVS] (Fig. 49). Não foram encontrados artigos descrevendo as repercussões funcionais desses motivos ou dos novos sítios criados.

Figura 49 - Predição de impacto de *RAD51D* c.628G>A; (p.Ala210Thr) nos motivos proteicos

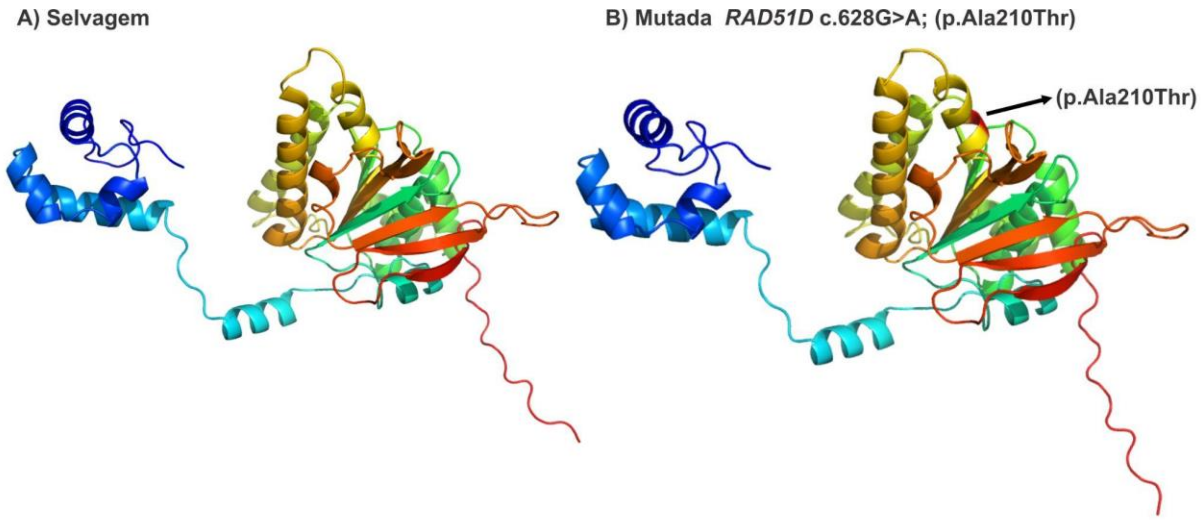


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM Motifs.

Considerando a estrutura 3D da proteína, a Ala210 está localizada em uma alfa-hélice e, embora seja substituída por uma treonina, não foi previsto impacto a estrutura 3D da proteína (Fig. 50).

Figura 50 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.628G>A; (p.Ala210Thr))



Nota: Estrutura 3D da RAD51D selvagem e mutada. Em A, está sendo mostrada a estrutura da proteína 3D selvagem. Em B, a estrutura 3D da proteína com a variante. A *RAD51D* c.628G>A; (p.Ala210Thr) está localizada em uma alfa-hélice, indicada por uma seta.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.10.1 Reclassificação da variante *RAD51D* c.628G>A; (p.Ala210Thr)

Considerando as informações acima, a *RAD51D* c.628G>A; (p.Ala210Thr) propomos manter a classificação como VUS, segundo os critérios ACMG/AMP: PM1, PP2 (ver na Seção 3.2.2), PP3.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante não está em *hotspot* de mutações, mas afeta domínios estruturais curtos, criando novos sítios do motivo *MOD_CK1_1* (*Sítio de fosforilação de Casein kinase 1 - CK1*) e *MOD_PLK4 (Sítio de fosforilação Plk4)*.

PP3 – Múltiplas linhas de evidência computacional suportam um efeito deletério no gene ou produto gênico

- A Ala210 é altamente conservada entre espécies de mamíferos, aves, anfíbios e répteis, e peixes, sem que nenhum outro aminoácido seja visto nessa posição;
- Foram previstas alterações nos motivos funcionais e nas interações moleculares, em que se perdem pontes de hidrogênio e ligações polares.

3.2.2.10.2 Conclusão sobre a reclassificação da variante *RAD51D* c.628G>A; (p.Ala210Thr)

A Ala210 é altamente conservada em mamíferos, anfíbios e répteis, e peixes e conservada em aves (raramente substituída por uma serina). Foram preditos efeitos significativos em domínios estruturais curtos, criando novos sítios de fosforilação. Prevê-se também que a variante altere as interações interatômicas, levando à perda de pontes de hidrogênio e ligações polares. As perdas de interações interatômicas podem modificar a dinâmica conformacional da proteína. No entanto, a repercussão funcional desses achados depende de estudos funcionais, sendo esta uma das principais limitações deste estudo, uma vez que ainda não foram encontrados dados experimentais para essa variante.

Alguns estudos publicados associam a variante a casos de câncer de mama ou ovário (Song *et al.*, 2015; Tung *et al.*, 2015; Yadav *et al.*, 2017; Ricker *et al.*, 2017). Por exemplo, nos estudos de Tung *et al.*, 2015, a variante é encontrada em três mulheres africanas e uma latino-americana com menos de 50 anos, diagnosticadas com câncer de mama. Além disto, no estudo de Ricker *et al.*, 2017, esta variante coocorre com uma variante patogênica (*MLH1* c.393_396del) em uma paciente do sexo feminino de 28 anos. Considerando o câncer como uma poligênica, pacientes com câncer podem ter mais de uma mutação afetando genes diferentes. Contudo, nenhuma variante patogênica foi encontrada com a mesma alteração de aminoácidos nesta posição. As variantes descritas nesta posição estão classificadas como VUS: *RAD51D* c.630G>A (p.Ala210=), *RAD51D* c.629C>A (p.Ala210Glu) e *RAD51D* c.629C>T (p.Ala210Val). No ClinVar, a interpretação desta variante é conflitante, com três entradas para VUS, cinco para benigna e quatro para provavelmente benigna. Contudo, devido a falta de dados clínicos como história familiar e dados de segregação, propomos manter a classificação VUS da variante. Entretanto, a alta conservação evolutiva do resíduo sugere que esta variante seja potencialmente patogênica.

3.2.2.11 *RAD51D* c.695G>A;(p.Arg232Gln)

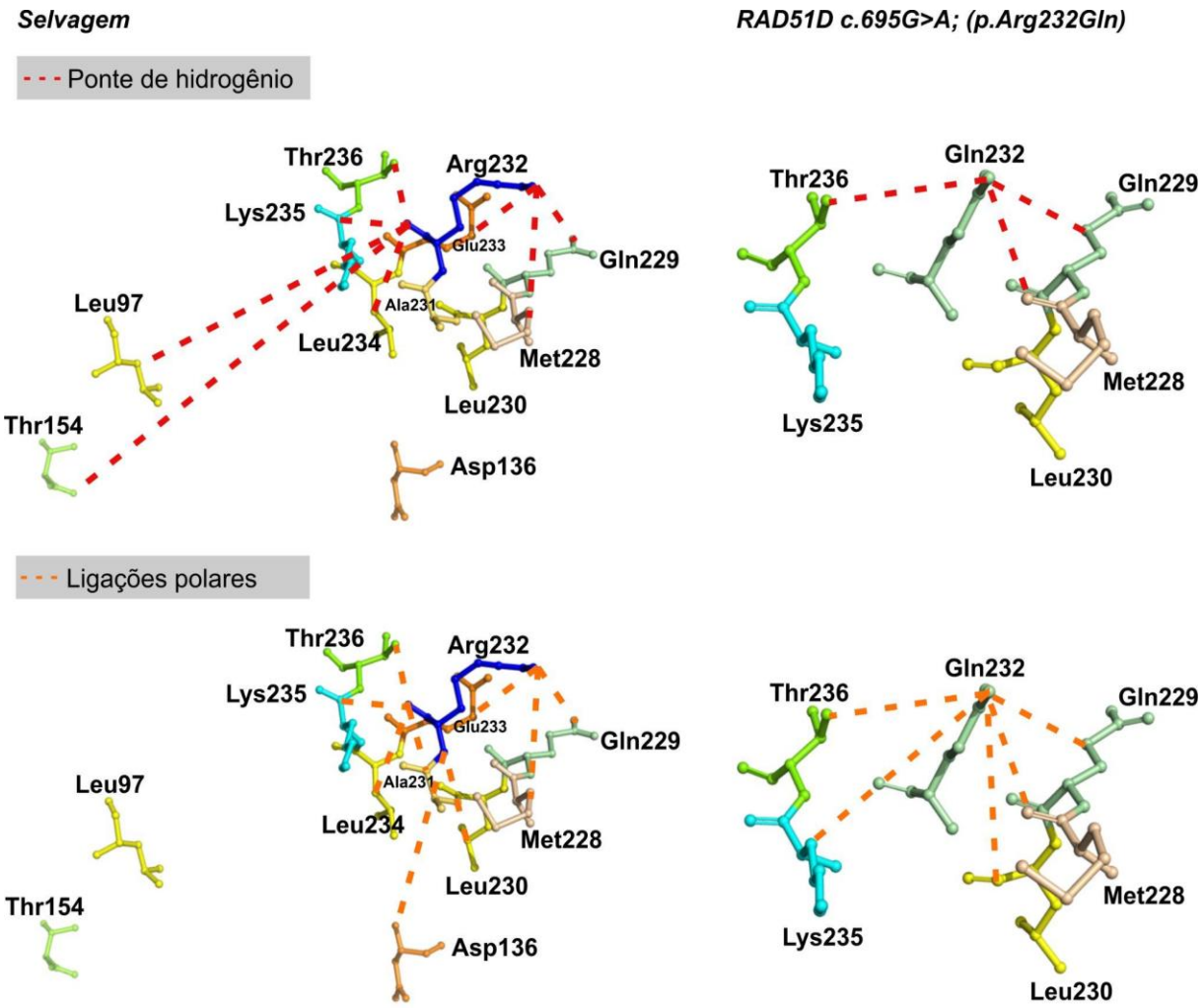
Localizada no éxon 11 do gene e no éxon 7 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), esta variante resulta de uma substituição de G para A na posição

nucleotídica 695. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores).

A arginina no códon 232 é substituída por glutamina, um aminoácido com propriedades físico-químicas diferentes, visto que a arginina é um aminoácido básico e a glutamina é polar não carregada. A maioria dos preditores *in silico*, Mutation Taster, FATH-MM, Provean e SIFT, aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na Seção 3.2.6).

No presente estudo, foi previsto que a Arg232 faça pontes de hidrogênio (Arg232 → Leu97, Arg232 → Met228, Arg232 → Gln229, Arg232 → Glu233, Arg232 → Leu234, Arg232 → Lys235, Arg232 → Thr236), interações polares (Arg232 → Asp136, Arg232 → Met228, Arg232 → Gln229, Arg232 → Leu230, Arg232 → Ala231, Arg232 → Glu233, Arg232 → Leu234, Arg232 → Lys235, Arg232 → Thr236) (Fig. 51). Perderam-se algumas interações após a troca de aminoácidos como pontes de hidrogênio (Arg232 → Leu97, Arg232 → Glu233, Arg232 → Leu234, Arg232 → Lys235) e interações polares (Arg232 → Asp136, Arg232 → Leu230, Arg232 → Ala231, Arg232 → Glu233, Arg232 → Leu234) (Fig. 51). A perda da Gln232 leva à desestabilização da estrutura proteica, afetando a habilidade de interação com o DNA e outras proteínas (Tab.6).

Figura 51 - Dinâmica molecular da variante *RAD51D* c.695G>A; (p.Arg232Gln)

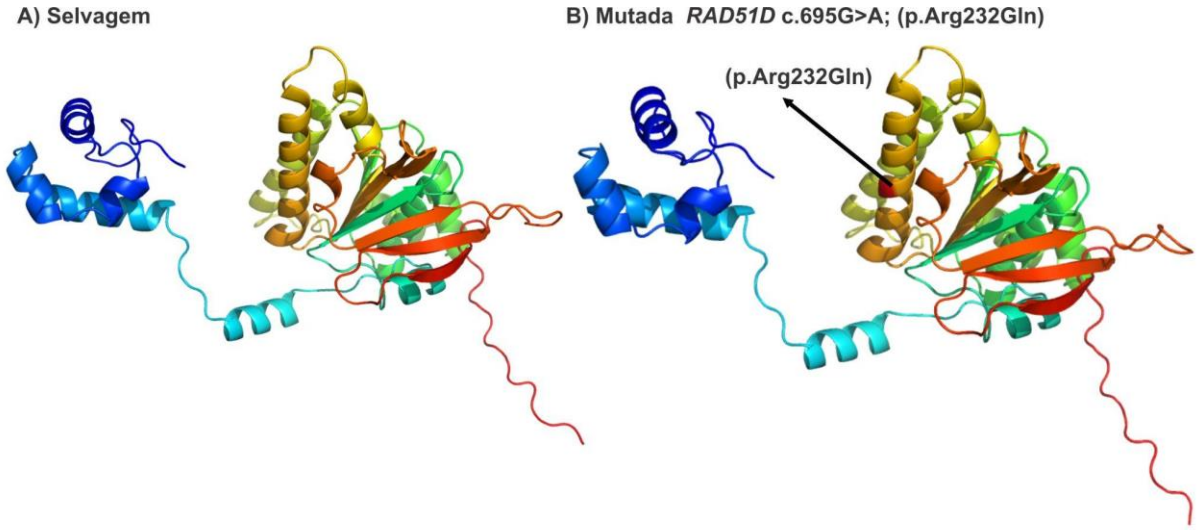


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares. Na estrutura selvagem, as interações acontecem entre os resíduos: Leu97, Asp136, Thr154, Met228, Gln229, Leu230, Ala231, Arg232, Glu233, Leu234, Lys235, Thr236. Na estrutura mutada, perde-se interação com os aminoácidos Leu97, Asp136, Thr154, Ala231, Arg232, Glu233 e Leu234.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Arg232 está localizada nos domínios funcionais *RecA_ATP-biding* (75-283), *DNA recombination and repair protein* (78-251) e *AAA+ ATPase* (99-274) (Quadro 6, na Seção 3.2.6). Considerando a estrutura 3D da proteína, previu-se que a Arg232 está localizada em uma alfa-hélice e, embora seja substituída por uma glutamina, não foi previsto impacto na estrutura 3D (Fig. 52).

Figura 52 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.695G>A; (p.Arg232Gln))



Nota: Em A, está mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada, que é indicada na figura B. A *RAD51D* c.695G>A; (p.Arg232Gln) está localizada em uma alfa-hélice e indicada por uma seta.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.11.1 Reclassificação da variante *RAD51D* c.695G>A; (p.Arg232Gln)

Considerando as informações acima, a *RAD51D* c.695G>A; (p.Arg232Gln) propomos manter a classificação como provavelmente benigna, segundo os critérios ACMG/AMP: PP2 (ver na seção 3.2.2), BP4, BP6.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A Arg232 é conservada entre espécies de mamíferos, mas algumas espécies apresentam leucina ou glutamina nesta posição; é altamente conservada em aves, sendo substituída por uma lisina em uma espécie; em anfíbios e répteis esta arginina é conservada; e é não conservada em peixes, sendo substituída por glicina nessa posição, em todas as espécies averiguadas;
- A maioria dos preditores *in silico* utilizados sugerem efeitos benignos para esta substituição. Isto pode refletir, pelo menos, em parte, a frequência relativamente alta

identificada para esta variante em diversas bases de dados (gnomAD 0.002; 1000 Genomes Project 0.005; Exome Project 0.005).

- Foram previstas diversas alterações nas interações moleculares, em que se perderam algumas pontes de hidrogênio e ligações polares. Entretanto, não foram preditos maiores impactos na estrutura tridimensional da proteína.

BP6 - Fonte respeitável relata recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente:

A variante é reportada como benigna/provavelmente benigna em bancos de dados públicos como ClinVar, VarSome e ABraOm.

3.2.2.11.2 Conclusão sobre a reclassificação da variante *RAD51D* c.695G>A; (p.Arg232Gln)

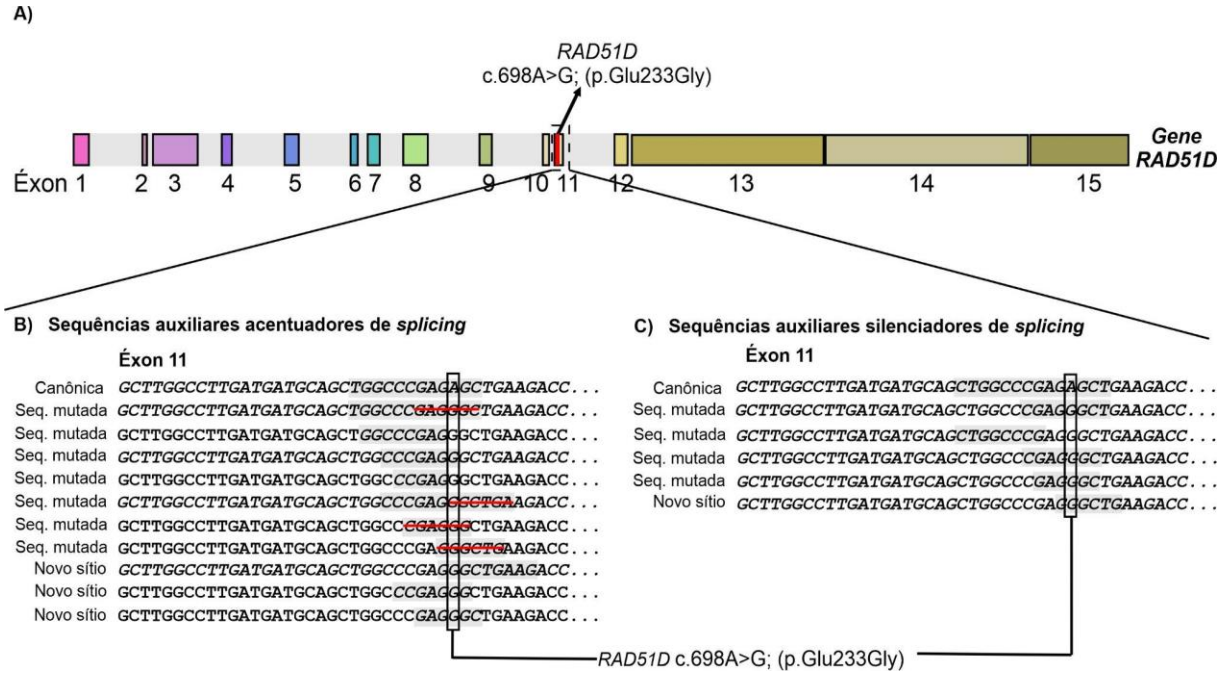
Em síntese, foi previsto que a troca de aminoácidos conduza à perda de interações interatômicas. As interações interatômicas modulam a dinâmica conformacional da proteína, e a perda dessas ligações pode alterar tanto a conformação quanto a função da proteína. Além disso, o resíduo Arg232 é conservado em mamíferos, aves e anfíbios e répteis (algumas vezes substituído por glutamina, leucina e lisina). No entanto, os algoritmos utilizados não sugerem impacto na função da proteína, considerando a conservação do resíduo e das regiões funcionais da proteína. Na literatura, essa variante foi associada a casos de câncer de mama ou ovário (Gutierrez-Enrıquez *et al.,* 2013; Pratz *et al.,* 2016), mas sem dados experimentais que investiguem o real impacto da variante na função da proteína. No estudo de Yurgelun e colaboradores (2014), a variante foi encontrada em quatro indivíduos com suspeita de síndrome de Lynch. No ClinVar, a variante foi classificada como benigna, devido principalmente a frequência populacional relativamente alta para ser uma variante patogênica. Este é um argumento que se aplica em doenças raras monogênicas. Entretanto, considerando-se que se trata de um gene penetrância moderada, as variantes podem ser mais frequentes, como tem sido visto nos escores de risco poligênico. Além disso, em uma das submissões do ClinVar arquivada pelo laboratório *Department of Pathology and Laboratory Medicine, Sinai Health System*, observou-se a ocorrência simultânea da variante *RAD51D* c.695G>A; (p.Arg232Gln) e a variante patogênica *BRCA2* c.6405_6409del em um indivíduo com HBOC. Segundo dados do relatório, essa evidência pode reduzir a possibilidade da mesma

desempenhar um papel patogênico. A hipótese alternativa é que se trata de uma variante associada à doença digênica dialélica. Diante desses fatos, propomos manter a sua classificação benigna/provavelmente benigna.

3.2.2.12 *RAD51D* c.698A>G; (p.Glu233Gly)

Localizada no éxon 11 do gene e no éxon 7 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), esta variante resulta de uma substituição de A para G na posição nucleotídica 698. Para essa substituição, foram previstos efeitos significativos nos sítios de *splicing* e nas sequências auxiliares reguladoras de *splicing* (Fig. 53). Novos sítios de reconhecimento, tanto acentuadores quanto silenciadores foram criados. Portanto, o real impacto destas alterações não pode ser avaliado apenas a partir destas evidências.

Figura 53 - Impacto da *RAD51D* c.698A>G; (p.Glu233Gly) nas sequências auxiliares regulatórias de sítios de *splicing*



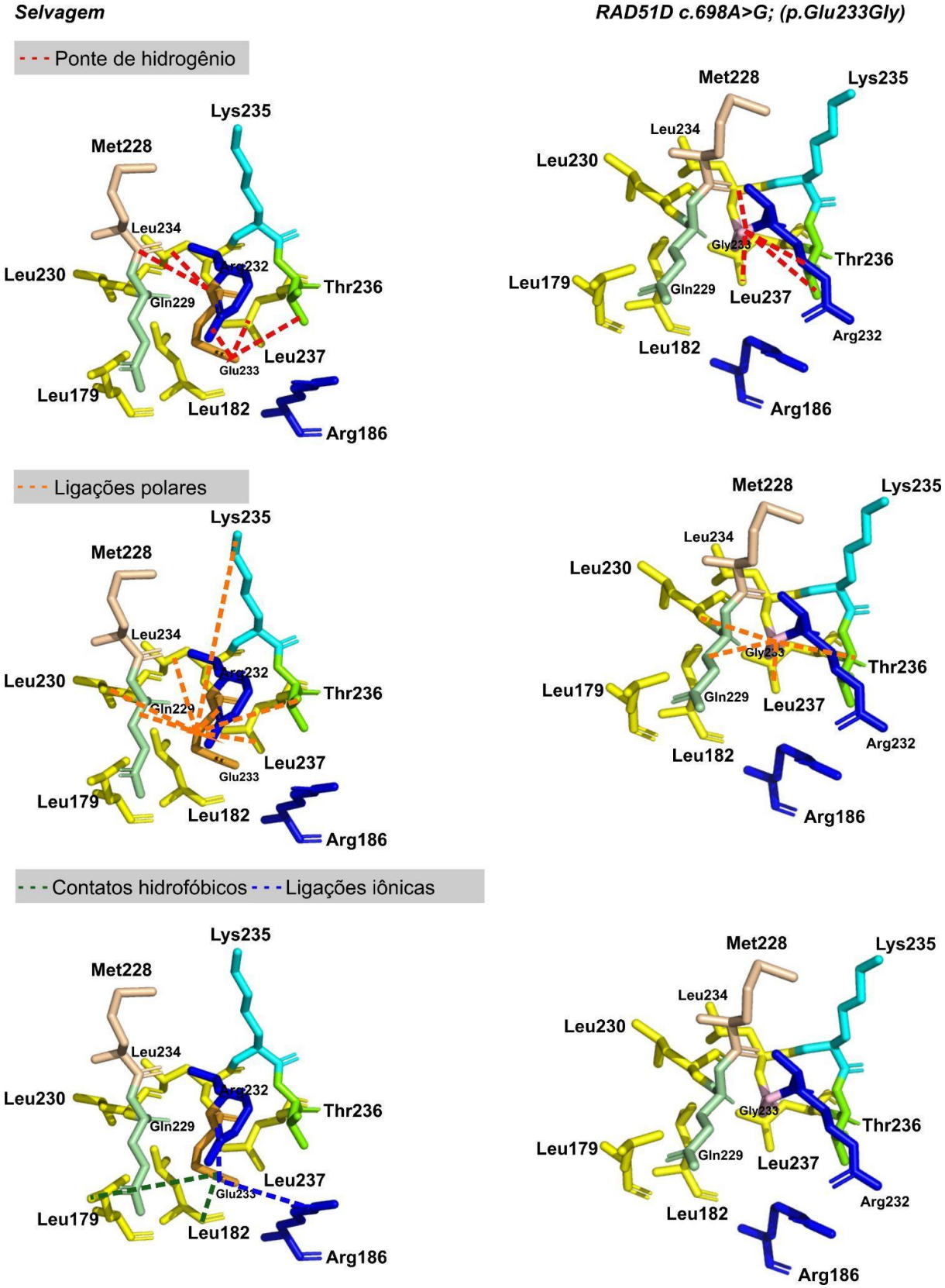
Nota: Em A, é mostrado o arranjo dos éxons ao longo do gene *RAD51D*. O *RAD51D* tem 15 éxons (cada um com uma cor atribuída). A variante *RAD51D* c.698A>G; (p.Glu233Gly) ocorre no éxon 11 do gene e sua localização pode ser vista por meio da seta indicando a variante e por uma linha vertical vermelha. Nas figuras B e C, são mostradas as sequências auxiliares, acentuadores ou silenciadores de *splicing*, que foram afetados após a troca de nucleotídeos. Em B, está sendo mostrado o impacto da variante em relação aos acentuadores de *splicing*. A sequência canônica contém o sítio acentuador de *splicing* de tipo selvagem sublinhado em cinza e abaixo dele estão as sequências mutadas. Alguns sítios não foram afetados mesmo após a troca de nucleotídeos (sequência intacta), outros foram abolidos (representados por uma linha vermelha sobre a sequência) e novos sítios foram gerados. Por outro lado, em C, está sendo mostrado o impacto da variante em relação aos silenciadores de *splicing*.

A sequência canônica contém o sítio silenciador de *splicing* de tipo selvagem e abaixo estão as sequências mutadas. Na imagem, é mostrado que sítios próximos à variante não foram afetados após a troca de nucleotídeos, nem mesmo o sítio que envolve a variante. Contudo, um novo sítio foi gerado no local.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

No nível da proteína, o ácido glutâmico no resíduo 233 foi substituído por glicina, um aminoácido com propriedades físico-químicas diferentes, visto que o ácido glutâmico é um aminoácido ácido e a glicina é apolar. A maioria dos preditores *in silico,* LIST-02, Provean, Polyphen e UMD-Predictor, aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos deletérios** para essa substituição, considerando a conservação evolutiva do resíduo e regiões funcionais na proteína (Quadro 6, na Seção 3.2.6). Foi previsto, que a Glu233 faça pontes de hidrogênio (Glu233 → Met228, Glu233 → Arg232, Glu233 → Leu234, Glu233 → Thr236, Glu233 → Leu237), contatos hidrofóbicos (Glu233 → Leu179, Glu233 → Leu182), ligações iônicas (Glu233 → Arg186, Glu233 → Arg232) e interações polares (Glu233 → Gln229, Glu233 → Leu230, Glu233 → Arg232, Glu233 → Leu234, Glu233 → Lys235, Glu233 → Thr236, Glu233 → Leu237) (Fig. 54). Perderam-se algumas interações após a troca de aminoácidos como ponte de hidrogênio (E233 → Met228), contatos hidrofóbicos (Glu233 → Leu179, Glu233 → Leu182), ligações iônicas (Glu233 → Arg186, Glu233 → Arg232) e polares (Glu233 → Arg232, Glu233 → Leu234, Glu233 → Lys235) (Fig. 54). A perda do Glu233 leva à desestabilização da estrutura proteica, com potencial efeito sobre a habilidade de interação com o DNA e outras proteínas (Quadro 6, na Seção 3.2.6).

Figura 54 - Dinâmica molecular da variante *RAD51D* c.698A>G; (p.Glu233Gly)

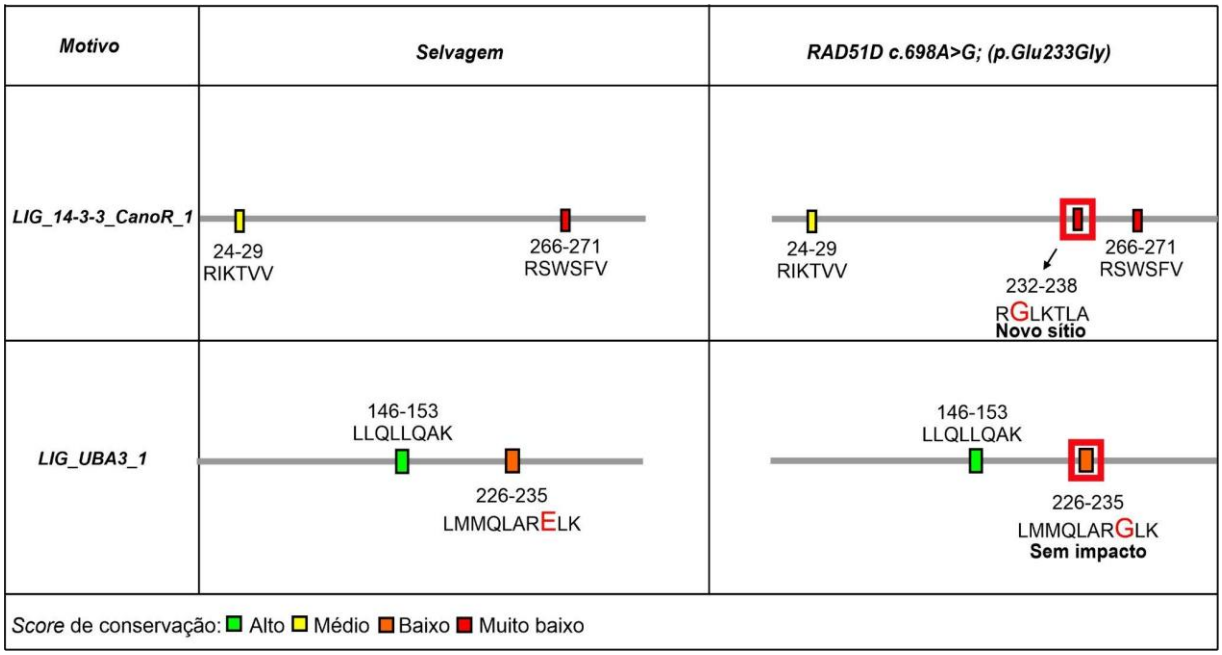


Nota: Interações interatômicas entre estruturas selvagem e mutada. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico; em azul, ligações iônicas. Na estrutura selvagem, as interações acontecem entre os resíduos: Leu179, Leu182, Arg186, Met228, Gln229, Leu230, Arg232, Glu233, Leu234, Lys235, Thr236, Leu237. Na estrutura mutada, perdem-se as interações com os aminoácidos Met228, Gln229, Arg232 e Leu234.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares, modificações pós-síntese, observou-se que o Glu233 está localizado nos domínios funcionais *RecA_ATP-biding* (75-283), *DNA recombination and repair protein* (78-251) e *AAA+ ATPase* (99-274) (Quadro 6, na Seção 3.2.6). Além disso, está localizado no motivo *LIG_UBA3_1 (Motivo de ligação para o domínio de adenilação UBA3)* na posição 226-235 [LMMQLARELK] (Fig. 55), mas não foi predita a perda deste motivo. Entretanto, a entrada da Gly233 cria um novo sítio de conservação muita baixa do motivo *LIG_14-3-3_CanoR_1 (Motivo de fosfopeptídeo de ligação 14-3-3)* na posição 232-238 [RGLKTLA] (Fig. 55).

Figura 55 - Predição de impacto de *RAD51D* c.698A>G; (p.Glu233Gly) nos motivos proteicos



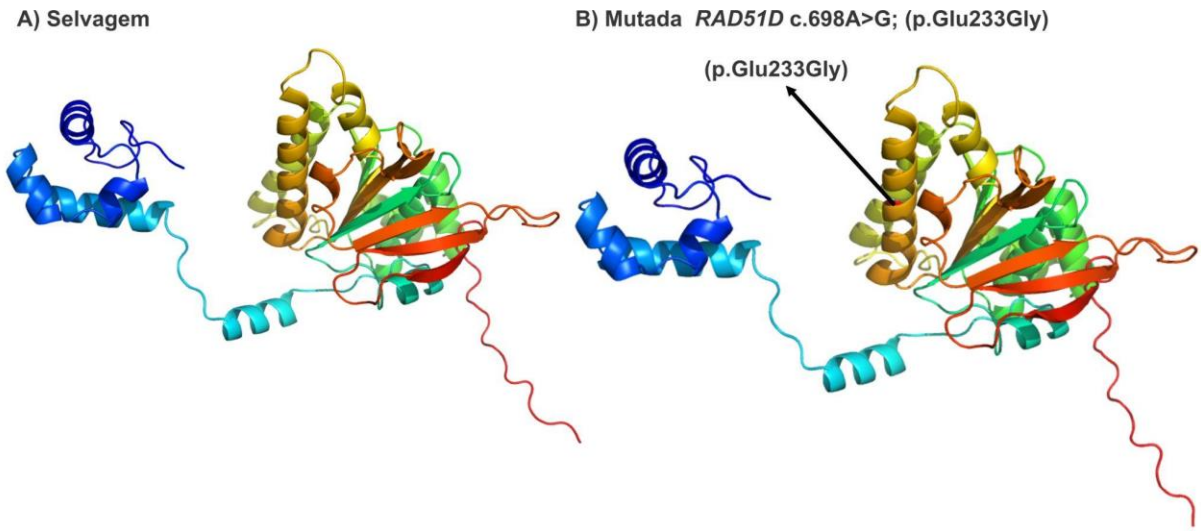
Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde

a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM Motif.

Considerando a estrutura 3D da proteína, previu-se que o Glu233 está localizado em uma alfa-hélice e, embora seja substituído por uma glicina, não foi previsto impacto na estrutura 3D da proteína (Fig. 56).

Figura 56 - Estrutura 3D da RAD51D selvagem e mutada *RAD51D* c.698A>G; (p.Glu233Gly)



Nota: Em A, está sendo mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada, que é indicada na figura B. A *RAD51D* c.698A>G; (p.Glu233Gly) está localizada em uma alça, indicada por uma seta.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.12.1 Reclassificação da variante *RAD51D* c.698A>G; (p.Glu233Gly)

A classificação desta variante é particularmente complexa. Ela aparece classificada como benigna ou provavelmente benigna. Entretanto, uma análise mais pormenorizada dá margem a dúvidas. A frequência alélica superior a 1% é geralmente assumida como evidência isolada de benignidade. Isto vale, certamente, para genes de alta penetrância. Mas existem evidências conflitantes.

Com base nas informações acima, a classificação provavelmente benigna da variante *RAD51D c.698A>G; (p.Glu233Gly)*, segundo os critérios ACMG/AMP: PP2 (ver na Seção 3.2.2), PP3, BS3 e BP6.

PP3 – Múltiplas linhas de evidência computacional suportam um efeito deletério no gene ou produto gênico:

- A Glu233 é altamente conservada entre espécies de mamíferos, aves e anfíbios e répteis. Em peixes não é conservada, sendo substituída por outros aminoácidos nesta posição (valina, alanina ou isoleucina);
- Os preditores *in silico* utilizados sugerem que esta substituição seja deletéria, afetando a estabilidade da estrutura da proteína, além de efeitos em pequenos domínios de proteínas.
- Foram previstas alterações nas interações moleculares, como perda de pontes de hidrogênio, contatos hidrofóbicos e ligações iônicas.
- Foi previsto que a troca de nucleotídeos ative novas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores).

BS3 - Estudos funcionais in vitro ou in vivo bem estabelecidos não mostram nenhum efeito prejudicial na função ou splicing da proteína

Nadkarni e colaboradores (2009) foi o primeiro grupo a estudar esta variante. Neste artigo, eles avaliaram o impacto molecular de RAD51D c.698A>G; (p.Glu233Gly) com um sistema de duplo híbrido em levedura, avaliando a percentagem de sobrevivência de células embrionárias tratadas com diversos agentes indutores de dano no DNA. A substituição de um ácido glutâmico por uma glicina na posição 233 levou ao *aumento da resistência celular* a agentes prejudiciais ao DNA, como mitomicina C, cisplatina, sulfonato de metilmetano, luz ultravioleta, taxol, quimioterápicos. Eles também determinaram efeitos protetores nos telômeros, pois a variante reduziu o índice de pontes anafásicas e aumentou a proliferação celular. Estes achados sugerem que esta variante seja benigna em relação à eficiência do reparo por recombinação. Entretanto, ela prejudica a resposta terapêutica aos quimioterápicos testados. Por outro lado, eles indicaram que a variante afeta negativamente a interação de RAD51D com RAD51C. Finalmente, estes autores observaram, através da modelagem molecular, que o Glu233 forma uma ponte salina com a Lys23 no domínio N-terminal de RAD51D e após a substituição essas interações ficaram comprometidas. Não é claro se estes mecanismos poderiam afetar negativamente a função da variante.

BP6 - Fonte respeitável relata recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente:

A variante é reportada como benigna/provavelmente benigna em bancos de dados públicos como ClinVar, VarSome e ABraOm.

3.2.2.12.2 Conclusão sobre a reclassificação da variante *RAD51D* c.698A>G; (p.Glu233Gly)

Em suma, foi previsto para esta variante impactos em elementos regulatórios de *splicing*, motivos e interações intermoleculares, que alteram a dinâmica conformacional do DNA e proteína e potencializam a incapacidade de interação transitória. Os algoritmos utilizados prevêem efeitos deletérios para esta variante, considerando a conservação evolutiva do resíduo e as regiões funcionais da proteína que foram previstas para serem afetadas pela variante (efeito sobre os domínios estruturais curtos). Domínios estruturais curtos alavancam interações transitórias fundamentais para processos celulares. Embora se preveja que a variante crie um novo sítio de baixa conservação do motivo de fosfopeptídeo de ligação 14-3-3, o critério PM1 não se aplicaria devido à presença de variantes benignas próximas ao resíduo Glu233. Os motivos de fosfopeptídeo de ligação 14-3-3 são motivos que contêm fosfoserina ou fosfotreonina. Esses fosfo-motivos permitem que as proteínas 14-3-3 formem homo- ou heterodímeros com seus alvos. As proteínas 14-3-3 são uma família de moléculas reguladoras que estão envolvidas em diversos processos celulares, incluindo transdução de sinal e regulação do ciclo celular. Existem fosfo-motivos canônicos, que contém arginina, e um grupo de motivos não canônicos, que são difíceis de classificar, mas utilizam interações hidrofóbicas adicionais (Muslin *et al.*, 1996; Obsil & Obsilova, 2011).

Em relação à conservação evolutiva, o Glu233 é altamente conservado, sendo substituído por ácido aspártico em uma espécie de mamíferos (*Dasypus novemcinctus*), sem variantes em aves, anfíbios e répteis. Em peixes, Glu233 não é conservado, sendo substituído por valina, isoleucina e alanina.

Na literatura, de acordo com Nadkarni e colaboradores (2009), esta variante é capaz de melhorar a atividade de reparo de DNA, uma vez que a resistência celular a drogas quimioterápicas como mitomicina C, cisplatina, sulfonato de metano de metila, luz ultravioleta e taxol é aumentada após a variante. No entanto, os autores identificaram efeitos nos processos celulares, cuja perda de Glu233 levou a uma diminuição da interação do mutante RAD51D com o RAD51C, mas sem afetar a formação do complexo (BCDX2). Eles também notaram que a proliferação celular se apresentou aumentada após a variante. Além

disso, essa variante foi identificada em pacientes com diagnóstico de câncer de mama. Segundo os autores, essa variante apresenta baixo ou nenhum risco de câncer de mama em casos de *BRCA*-negativo. Rodrígues-Lópes e colaboradores (2004), analisaram o papel da variante *RAD51D* c.698A>G; (p.Glu233Gly) na suscetibilidade ao câncer de mama e sugeriu um leve aumento ao risco de câncer de mama em famílias com história familiar positiva. Contrário de Dowtie e colaboradores (2008) e Jara e colaboradores (2010), que realizaram um estudo de caso-controle em um grupo de mulheres australianas e uma população sul americana, respectivamente. Em seus resultados, eles não identificaram associação entre a variante e o risco de câncer de mama em mulheres BRCA1/2-negativas com câncer de mama e história familiar positiva. Maxwell e colaboradores (2016), encontraram a variante ao estudar famílias afetadas pelo câncer de mama. Em seus resultados, embora a análise *in silico* tenha sido inconclusiva, eles classificam a variante como benigna, aplicando os seguintes critérios: BS1 - frequência do alelo é maior do que o esperado para o distúrbio. ; BS3 - estudo *in vitro* bem estabelecido da variante confirmam efeitos benignos (Nadkarni *et al.,* 2009); BS4 - variante não segrega com a doença em famílias; e BP6 - variante reportada como benigna/provavelmente benigna em uma fonte confiável. No estudo de Frimer e colaboradores (2016), a variante *RAD51D* c.698A>G; (p.Glu233Gly) foi identificada ocorrendo simultaneamente com as variantes *RAD51D* c.758 A>G;(p.Glu253Gly), *RAD51D* c. 362 A>G; (p.Glu121Gly) e *IGF1R* c.1532 G>A; (p.Arg511Gln), em uma paciente com diagnóstico de carcinoma seroso uterino estádio III, sem história familiar positiva para câncer de mama e ovário. Durante o tratamento, a paciente apresentou resistência à platina. Semelhante ao estudo de Frimer, no ClinVar, o *Women's Health and Genetics - Laboratory Corporation of America (LabCorp)* relata a variante *RAD51D* c.698A>G; (p.Glu233Gly) co-ocorrendo com outras variantes patogênicas ou provavelmente patogênicas, como *PALB2* c.196C>T (p.Gln66X), *MUTYH* c. 1187G>A (p.Gly396Asp), *BARD1* c.1690C>T (p.Gln564X). Foi proposto, que essa evidência possa enfraquecer a chance de um papel patogênico para essa variante. Entretanto, como discutido acima, as hipóteses alternativas seriam doença digênica ou que se trate de uma variante de baixa/moderada penetrância, componente do risco poligênico.

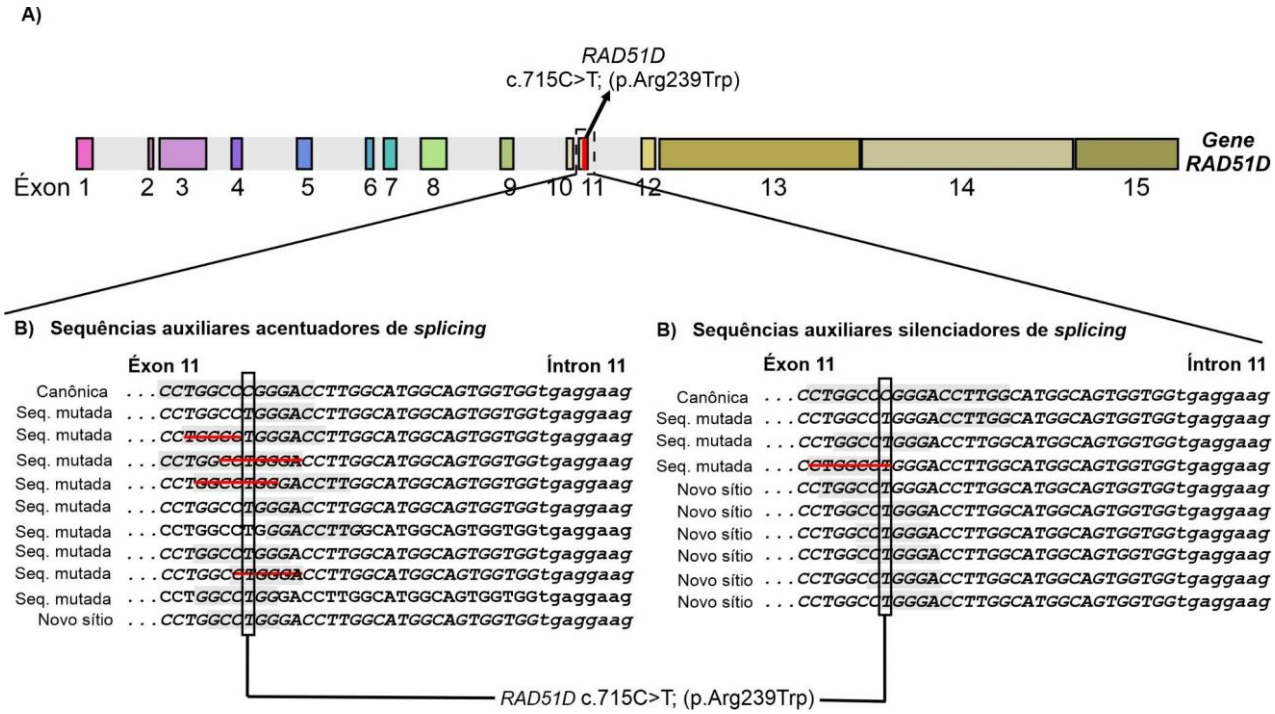
No ClinVar, essa variante é classificada como provavelmente benigna, devido à frequência populacional ser maior do que a esperada para uma variante patogênica *RAD51D* (acima de 1%), fornecendo evidências de suporte para um polimorfismo (Wickramanayake *et al*., 2012, Osher *et al*., 2012; Song *et al*., 2015, Nicolas *et al*., 2015). Não foram encontradas outras

variantes de sentido trocado classificadas como patogênicas envolvendo o resíduo Glu233. Existem tanto evidências de patogenicidade (conservação evolutiva) como de não patogenicidade para esta variante (estudo funcional, frequência relativamente alta e co-ocorrência com outras variantes patogênicas). Consequentemente, propomos manter a classificação desta variante provavelmente benigna. Entretanto, esta variante foi associada à resistência à quimioterápicos.

3.2.2.13 *RAD51D* c.715C>T (p.Arg239Trp)

Localizada no éxon 11 do gene e no éxon 7 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), resulta de uma substituição de C para T na posição nucleotídica 715. Para essa substituição, foram previstos efeitos significativos nas sequências auxiliares reguladoras de *splicing* (Fig. 57).

Figura 57 - Impacto da *RAD51D* c.715C>T (p.Arg239Trp) nas sequências auxiliares regulatórias de sítios de *splicing*



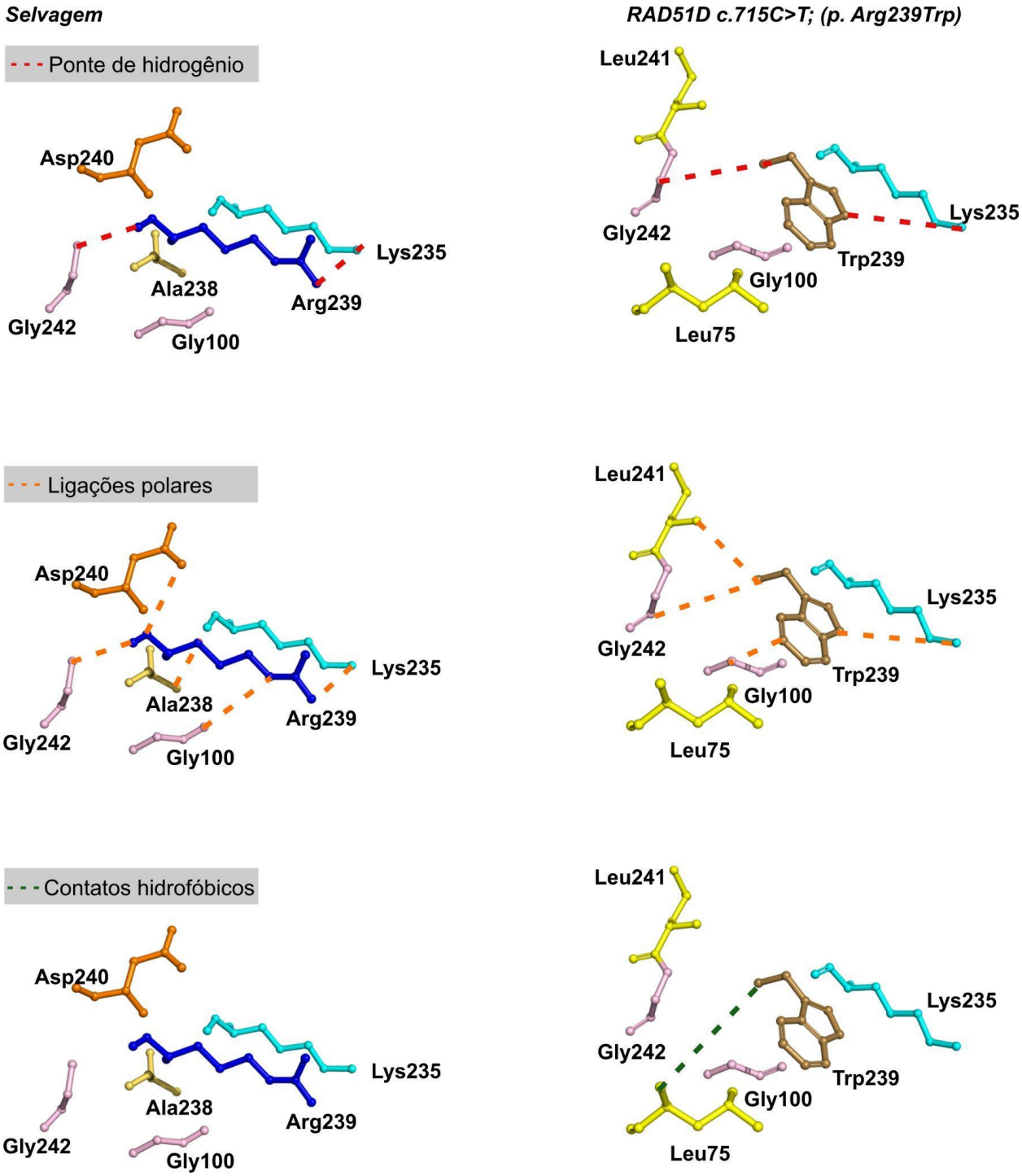
Nota: Em A, está sendo mostrado o arranjo dos éxons ao longo do gene RAD51D. O *RAD51D* tem 15 éxons (cada um com uma cor atribuída). A variante *RAD51D* c.715C>T (p.Arg239Trp) ocorre no éxon 11 do gene e sua localização pode ser vista por meio da seta indicando a variante e por uma linha vertical vermelha. Nas figuras B e C, são mostradas as sequências auxiliares, acentuadores ou silenciadores, que foram afetados após a troca de nucleotídeos. Em B, está sendo mostrado o impacto da variante em relação aos acentuadores de *splicing*. A sequência canônica contém o sítio acentuador de *splicing* de tipo selvagem em cinza e abaixo dele estão as sequências mutadas. Alguns sítios não foram afetados mesmo após a troca de nucleotídeos (sequência intacta), outros foram abolidos (representados por uma linha vermelha sobre a sequência) e novos sítios foram gerados. Por outro lado, em C, está sendo mostrado o impacto da variante em relação aos silenciadores de *splicing*. A

sequência canônica contém o sítio silenciador de *splicing* de tipo selvagem e abaixo estão as sequências mutadas. Na imagem, é mostrado que sítios próximos à variante não foram afetados após a troca de nucleotídeos, nem mesmo o sítio que envolve a variante. Contudo, novos sítios foram gerados no local.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

A arginina no resíduo 239 é substituída por triptofano, um aminoácido com propriedades diferentes, visto que a arginina é um aminoácido básico e o triptofano é apolar. Os preditores in silico Mutation Taster, FATH-MM, e SIFT aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, enquanto, LIST-02, Provean, Polyphen e UMD-Predictor, preveem **efeitos deletérios,** considerando a conservação evolutiva do resíduo e regiões funcionais na proteína (Quadro 6, na Seção 3.2.6). No presente estudo, foi previsto que a Arg239 faça pontes de hidrogênio (Arg239 → Lys235, Arg239 → Gly242) e interações polares (Arg239 → Gly100, Arg239 → Lys235, Arg239 → Ala238, Arg239 → Asp240) (Fig. 57). Perderam-se algumas interações polares após a troca de aminoácidos (Arg239 → Ala238, Arg239 → Asp240) e, outra foram geradas como contato hidrofóbico (Trp239 → Leu75), polar (Trp239 → Gly100, Trp239 → Leu241, Trp239 → Gly242) (Fig. 57). A perda Arg239 leva a desestabilização da estrutura proteica, desestabilizando a habilidade de interação com o DNA e proteínas (Quadro 6).

Figura 58 - Dinâmica molecular da variante *RAD51D* c.715C>T (p.Arg239Trp)

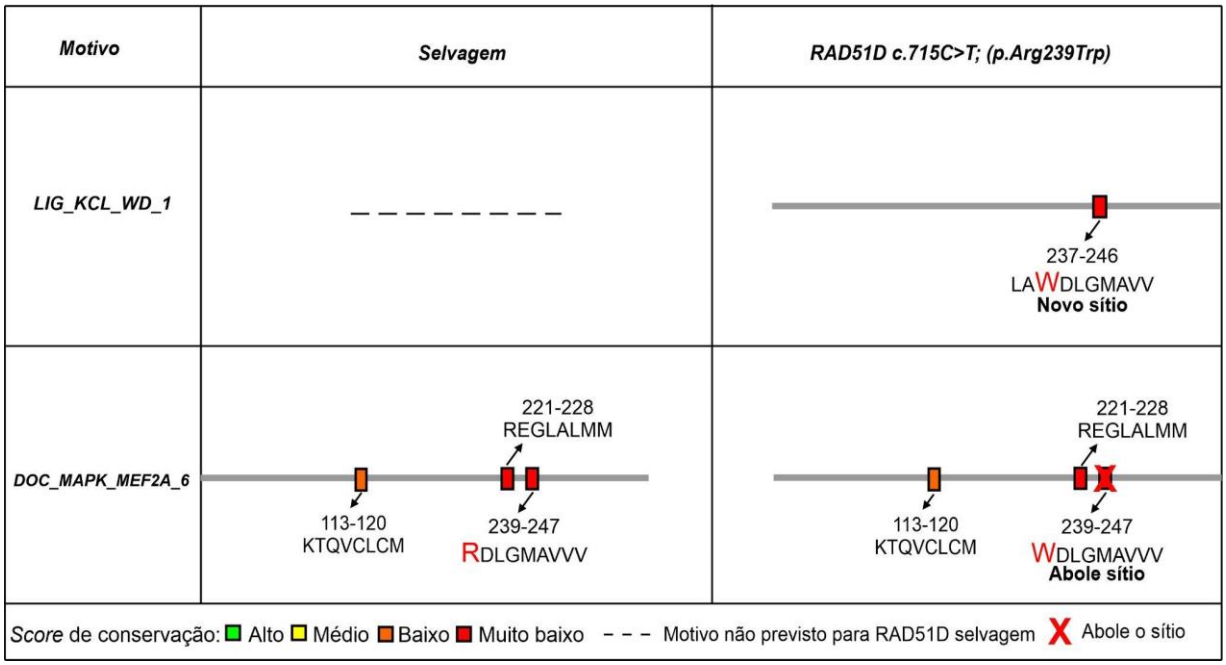


Nota: Interações interatômicas nas estruturas selvagem e mutada. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Ala161 e 163, Glu162, Leu164 , Ile167, Arg165 e Arg166. Na estrutura mutada, perde-se interação com os aminoácidos Leu164, Arg165 e Arg166.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Arg239 está localizada nos domínios funcionais *RecA_ATP-biding* (75-283), *DNA recombination and repair protein* (78-251) e *AAA+ ATPase* (99-274) (Tab. 3). A perda da Arg239 leva a perda de um sítio de baixíssima conservação do motivo *DOC_MAPK_MEF2A_6 (Motivo MAPK-docking)* na posição 239-247 [RDLGMAVVV] (Fig. 59). A entrada da Trp239 cria um sítio de baixa conservação do motivo *LIG_KCL_WD_1 (Motivos que ligam o domínio TPR da cadeia leve 1 de cinesina)* na posição 237-246 [LAWDLGMAVV] (Fig. 59).

Figura 59 - Predição de impacto de *RAD51D* c.715C>T (p.Arg239Trp) nos motivos proteicos

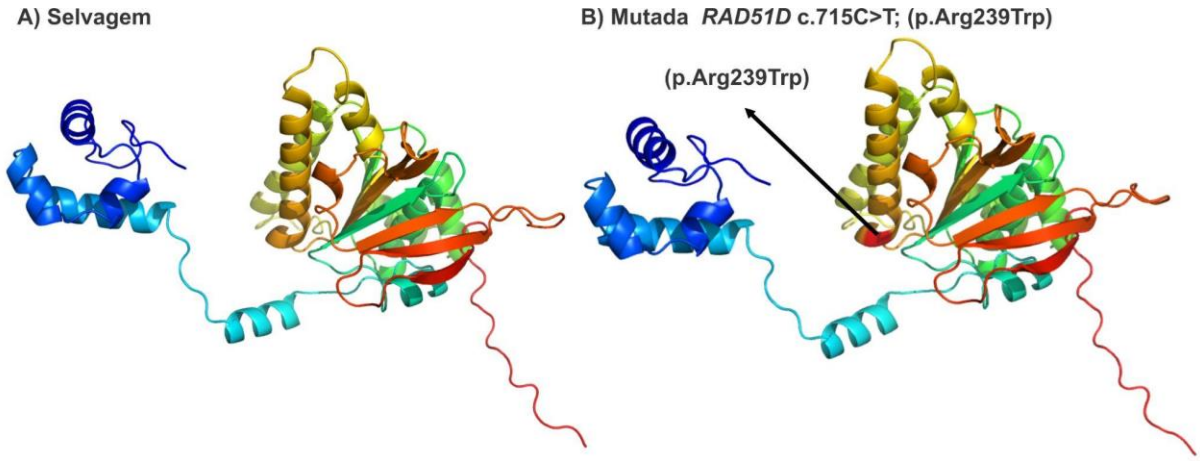


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM Motifs.

Considerando a estrutura 3D da proteína, previu-se que a Arg239 está localizada em uma alfa-hélice e, embora seja substituída por um triptofano, não foi previsto impacto na estrutura 3D (Fig. 60).

Figura 60 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.715C>T (p.Arg239Trp))



Nota: Em A, estrutura da proteína 3D selvagem; em B, estrutura mutada. A *RAD51D* c.715C>T (p.Arg239Trp) está localizada em uma alfa-hélice, indicada por uma seta.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.13.1 Reclassificação da variante *RAD51D* c.715C>T (p.Arg239Trp)

Com base nas informações acima, propomos manter a classificação VUS da variante *RAD51D c.715C>T (p.Arg239Trp)*, segundo os critérios ACMG/AMP: PM2, PP2 (ver Seção 3.2.2).

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica em bancos de dados populacionais (gnomAD (0.00004) e não foi encontrada no 1000 Genomes Project e no Exome Sequencing Project).

BP2 - Observado em trans com uma variante patogênica para um gene/distúrbio dominante totalmente penetrante ou observado em cis com uma variante patogênica em qualquer padrão de herança

No presente estudo, a *RAD51D* c.715C>T (p.Arg239Trp) foi encontrada em *cis* com uma variante patogênica *RAD51D* c.694C>T; (p.Arg232*) em uma paciente com câncer de mama. Isso está de acordo com o descrito na literatura, onde alguns autores encontraram ambas as

variantes segregando em membros afetados e não afetados, sugerindo que provavelmente as variantes estão em *cis*. Isso será discutido em mais detalhes na *Conclusão sobre a reclassificação da variante RAD51D c.715C>T (p.Arg239Trp)* (Seção 3.2.2.13.2).

3.2.2.13.2 Conclusão sobre a reclassificação da variante *RAD51D* c.715C>T (p.Arg239Trp)

No âmbito deste estudo, foi previsto que esta variante possa alterar elementos regulatórios de *splicing*, alterar as interações interatômicas e criar um sítio de baixa conservação do motivo *LIG_KCL_WD_1 (Motivos que ligam a regiões de repetição de tetratricopeptídeo)*, além de abolir um sítio de baixíssima conservação do motivo *DOC_MAPK_MEF2A_6 (Motivo MAPK-docking).* O motivo *LIG_KCL_WD_1* é encontrado em proteínas de carga, cujo motivo *WD* medeia a ligação da proteína às regiões de repetição do tetratricopeptídeo (RRT), que são encontradas na cinesina 1. As regiões RRT promovem interações transitórias entre as proteínas, reconhecendo os domínios WD nas proteínas de carga. Além dessas regiões RRT, as cinesinas contêm um domínio funcional que permite a hidrólise do ATP (Pernigo, 2013). Motivo *MAPK-docking* permite a interação de moléculas alvos com a MAP Quinase (*mitogen-activated protein kinase* - MAPK). MAPK é uma subfamília de serina/treonina quinase, que respondem a estímulos extracelulares e regulam várias atividades celulares, como expressão gênica, mitose, diferenciação, sobrevivência celular e apoptose (Biondi & Nebreda, 2003; Bardwell, 2006). Essas alterações em domínios curtos podem levar a mudanças na dinâmica da proteína. No entanto, os algoritmos usados no presente estudo fornecem previsões inconsistentes sobre o impacto na proteína.

A conservação evolutiva mostra que, em mamíferos, a Arg239 é conservada, exceto por sete espécies, onde é substituída por glutamina (*Rousettus aegyptiacus, Pteropus alecto, Pteropus giganteus*), triptofano (*Jaculus jaculus* e *Jaculus jaculus 2*) ou leucina (*Nomascus leucogenys* e *Hylobates moloch*); em aves, também é conservada, algumas vezes sendo substituída por lisina; em anfíbios e répteis, ocorre nesta posição apenas arginina; e, em peixes não é conservada, somente lisinas são vistas nessa posição. A variante tem baixa frequência populacional e foi encontrada em indivíduos com câncer de mama ou ovário (Song *et al.*, 2015; Gutierrez-Enrıquez *et al.*, 2013; Wickramanayake *et al.*, 2012; Sánchez-Bermúdez *et al.*, 2018; Hauke *et al*, 2018). Em nosso estudo, a variante foi encontrada em uma paciente com câncer de mama co-ocorrendo com uma variante patogênica *RAD51D* c.694C>T; (p.R232*). Outros relatos semelhantes foram encontrados, onde alguns estudos relataram as

mesmas variantes co-ocorrendo em indivíduos afetados com câncer de ovário ou não: 1) Gutiérrez-Enriquez e colaboradores (2013) identificaram as mesmas variantes ocorrendo simultaneamente em duas famílias. Na primeira família, a probanda tinha 44 anos quando foi diagnosticada com câncer de ovário, enquanto, na segunda família, a probanda tinha 42 anos e foi diagnosticada com câncer de ovário bilateral. Nesta família, a variante também foi encontrada em quatro de seus irmãos saudáveis. Esses autores também sugerem que ambas as variantes estão provavelmente em *cis* e a mudança do nucleotídeo selvagem, *RAD51D* c.694C>T; (p.R232*), pode ser a mutação fundadora; 2) Sanchez-Bermudez e colaboradores (2018) relataram essas variantes em outra família, onde estas variantes foram encontradas em duas irmãs com câncer de ovário e seus três filhos saudáveis. No ClinVar, há conflito na interpretação de patogenicidade da variante, com duas entradas para benigna e oito para VUS. Além disso, até o momento, não foi encontrado nenhum estudo funcional, que investigue o impacto da variante na função da proteína. Diante desses fatos, a variante mantém sua classificação VUS, embora esteja inclinada para não patogênica, pelo menos quando em *cis* com uma patogênica.

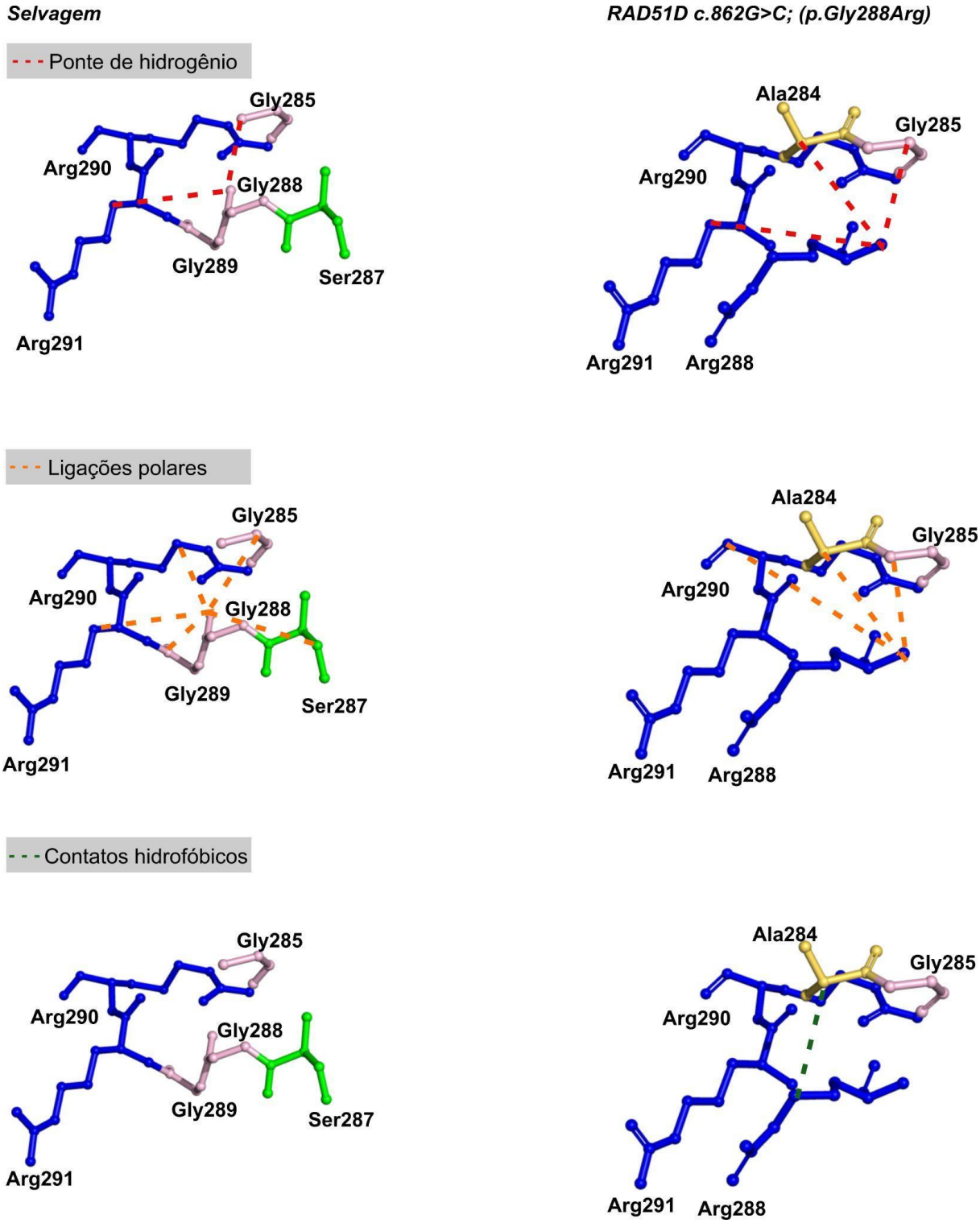
3.2.2.14 RAD51D c.862G>C (p.Gly288Arg)

Localizada no éxon 12 do gene e no éxon 8 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), esta variante resulta de uma substituição de G para C na posição nucleotídica 862. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores ou silenciadores).

A glicina no resíduo 288 é substituída por arginina, um aminoácido com propriedades físico-químicas diferentes, visto que a glicina é um aminoácido apolar e a arginina é básico. Os preditores *in silico* Mutation Taster, LIST-02, FATH-MM, Provean, Polyphen, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na Seção 3.2.6). Foi previsto que a Gly288 faça pontes de hidrogênio (Gly288 → Gly285, Gly288 → Arg291) e interações polares (Gly288 → Gly285, Gly288 → Ser287, Gly288 → Gly289, Gly288 → Arg290, Gly288 → Arg291) (Fig. 61). Perderam-se algumas interações polares após a troca de aminoácidos (Arg288 → Ala284, Gly288 → Ser287, Gly288 → Gly289, Gly288 → Arg291) e outras interações polares foram geradas (Trp288 → Ala284, Trp288 → Arg290) (Fig. 61).

No âmbito do presente estudo, foi previsto que a perda da Gly288 leve a desestabilização da estrutura proteica, afetando habilidade de interação com o DNA e outras proteínas (Quadro 6, na Seção 3.2.6).

Figura 61 - Dinâmica molecular da variante *RAD51D* c.862G>C (p.Gly288Arg)

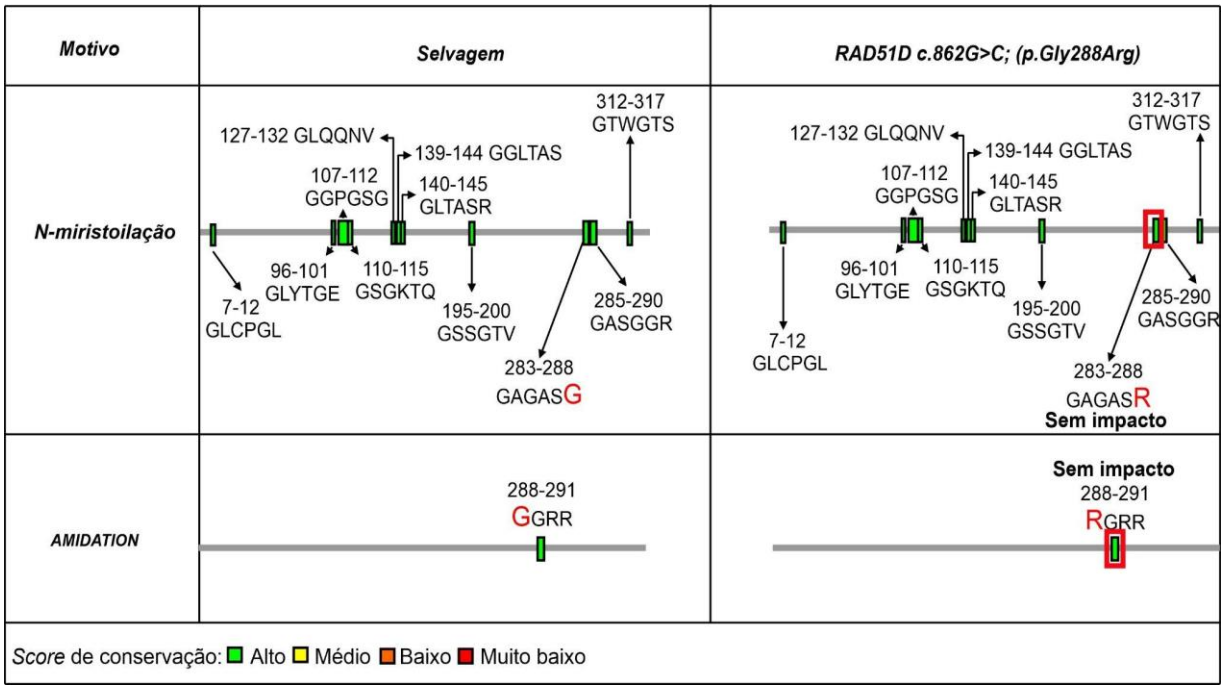


Nota: Interações interatômicas entre estruturas selvagem e mutada. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico. Na estrutura selvagem, as interações acontecem entre os resíduos: Gly285, Gly288 e Gly289, Arg 290 e Arg291 e Ser287. Na estrutura mutada, perde-se interação com os aminoácidos Ser287; e uma nova interação foi prevista com a Ala284.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a Gly288 está localizada no sítio de N-miristoilação na posição 283-288 [GAGASG] e Amidação na posição 288-291 [GGRR], mas sem efeitos deletérios sobre esses sítios (Fig. 62).

Figura 62 - Predição de impacto de *RAD51D* c.862G>C (p.Gly288Arg) nos motivos proteicos

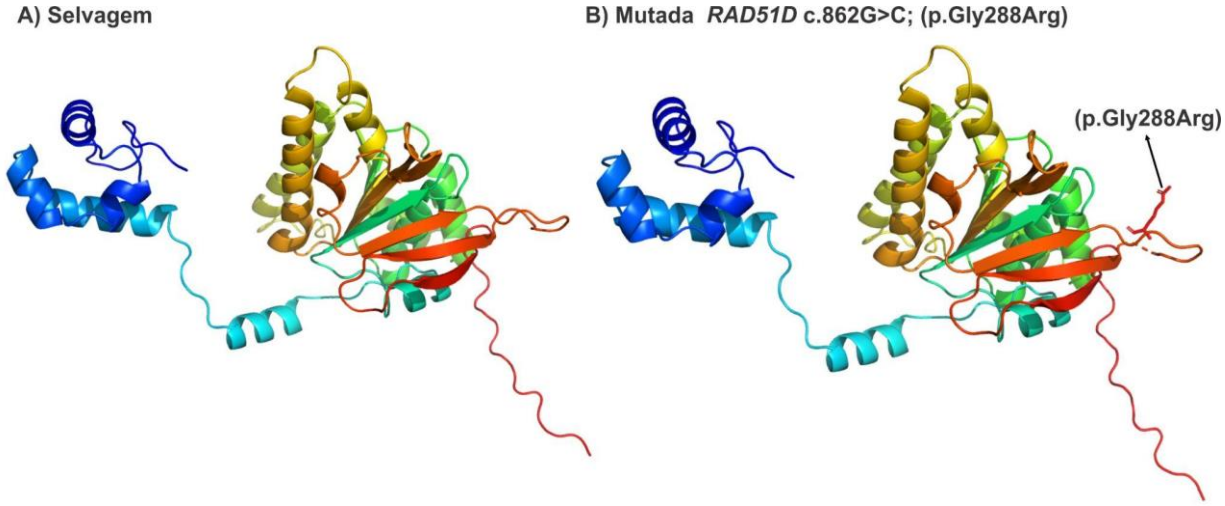


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM Motifs.

Considerando a estrutura 3D da proteína, previu-se que a Gly288 estaria localizada em uma alça e, embora seja substituída por uma arginina, não foi previsto impacto na estrutura 3D (Fig. 63).

Figura 63 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.862G>C (p.Gly288Arg))



Nota: Em A, está sendo mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada que é indicada na figura B. A *RAD51D* c.862G>C (p.Gly288Arg) está localizada em uma alça, indicada por uma seta e por um stick.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.14.1 Reclassificação da variante *RAD51D* c.862G>C (p.Gly288Arg)

Com base nas informações acima, propomos manter a classificação VUS da variante *RAD51D c.862G>C (p.Gly288Arg)*, segundo os critérios ACMG/AMP: PM2, PP2 (ver na Seção 3.2.2) e BP4.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica no gnomAD (0.00004) e não é encontrada no 1000 Genomes Project ou no Exome Sequencing Project.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A Gly288 não é conservada. Em mamíferos esta posição pode ser ocupada por glicina, ácido aspártico, alanina, histidina ou serina), em aves (cisteína, serina, treonina, metionina, ou glicina), anfíbios e répteis (apenas serina) e peixes (serina ou prolina). Aqui, é importante considerar que esta é uma região menos conservada da proteína;

- Os preditores *in silico* utilizados sugerem efeitos benignos para esta substituição;

- Foram previstas inúmeras alterações nas interações moleculares, em que se perderam algumas interações polares após a troca de aminoácido. No entanto, esse achado não é evidência suficiente de patogenicidade.

3.2.2.14.2 Conclusão sobre a reclassificação da variante RAD51D c.862G>C (p.Gly288Arg)

Em síntese, a Gly288 é pouco conservada em mamíferos, sendo substituído por alanina, histidina, ácido aspártico e serina; em aves, anfíbios e répteis e peixes não é conservada, substituída por treonina, serina, cisteína, prolina e metionina. Embora se preveja que a variante altere as interações interatômicas, os algoritmos utilizados preveem efeitos benignos para essa variante. As repercussões funcionais dessas predições só podem ser investigadas por meio de estudos funcionais, o que é uma limitação deste estudo. A variante foi encontrada em associação com câncer de ovário (Song *et al.*, 2015). Uma limitação importante é a falta de dados clínicos, como histórico familiar ou dados de segregação. Até o momento, as evidências encontradas para esta variante são insuficientes e limitadas, mantendo sua classificação como VUS, mesmo tendo potencial para ser uma variante benigna, levando-se em conta a baixa conservação evolutiva.

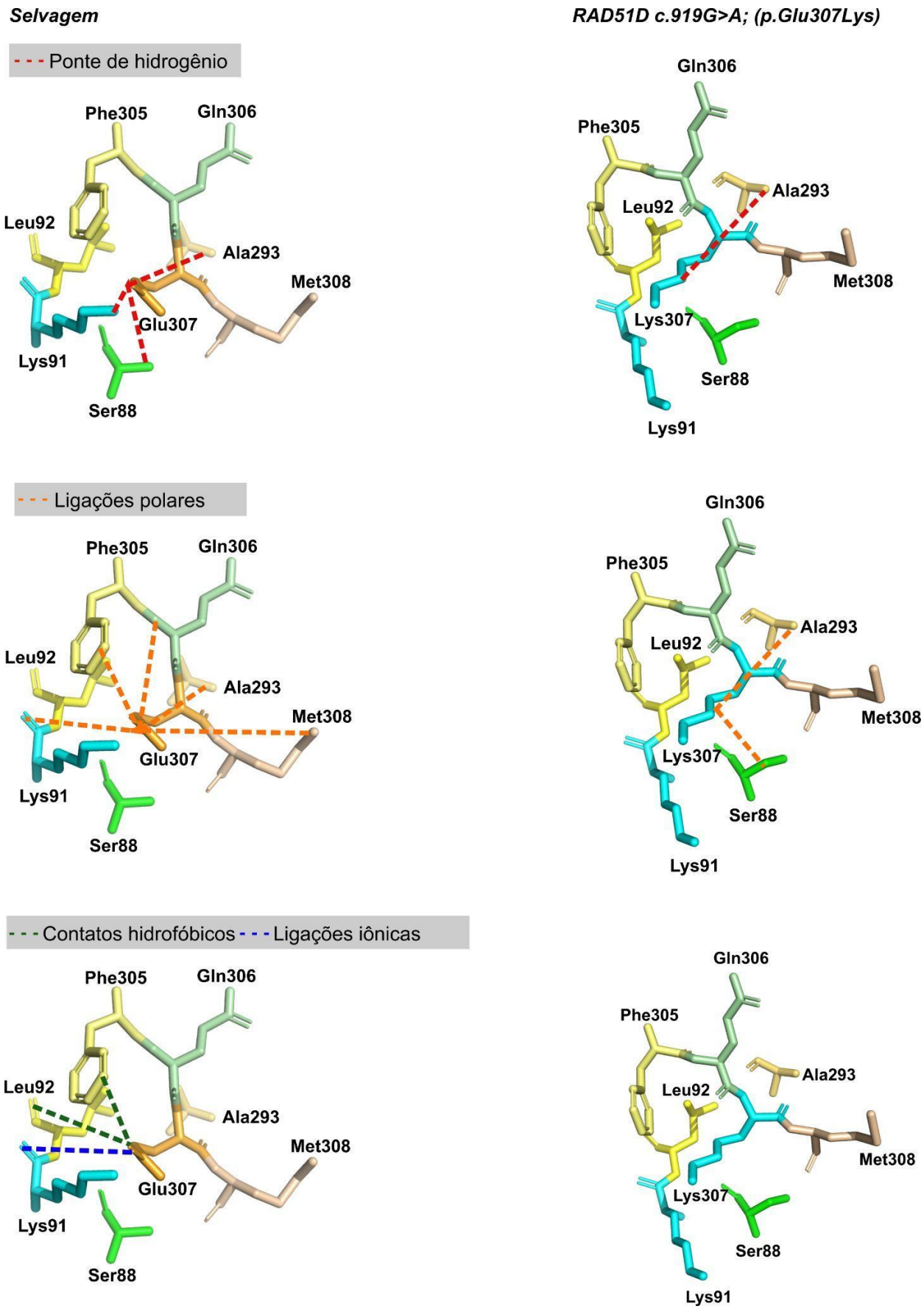
3.2.2.15 *RAD51D* c.919G>A; (p.Glu307Lys)

Localizada no éxon 12 do gene e no éxon 8 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), esta variante resulta de uma substituição de G para A na posição nucleotídica 919. Para essa substituição, não foram previstos efeitos significativos nos sítios de *splicing* ou nas sequências auxiliares reguladoras de *splicing* (acentuadores ou silenciadores).

Esta variante provoca troca de classe do resíduo, visto que o ácido glutâmico (ácido) no códon 307 é substituído por lisina (básico). Os preditores *in silico* Mutation Taster, LIST-02, FATH-MM, Provean, Polyphen, SIFT e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na Seção 3.2.6). Foi previsto que o Glu307 faça pontes de hidrogênio (Glu307 → Ser88, Glu307 → Lys91, Glu307 → Ala293), contatos hidrofóbicos (Glu307 → Leu92, Glu307 → Phe305), ligação iônica (Glu307 → Lys91), interações polares (Glu307 → Lys91, Glu307 →

Ala293, Glu307 → Phe305, Glu307 → Gln306, Glu307 → Met308) (Fig. 64). Perderam-se algumas interações após a troca de aminoácidos como pontes de hidrogênio (Glu307 → Ser88, Glu307 → Lys91), contatos hidrofóbicos (Glu307 → Leu92, Glu307 → Phe305), ligação iônica (Glu307 → Lys91) e interações polares (Glu307 → Lys91, Glu307 → Phe305, Glu307 → Gln306, Glu307 → Met308), enquanto outra interação polar foi gerada (Lys307 → Ser88) (Fig. 64). No âmbito do presente estudo, previu-se que a perda da Glu307 leve a desestabilização da estrutura proteica, afetando a habilidade de interação com outras proteínas (Quadro 6, na Seção 3.2.6).

Figura 64 - Dinâmica molecular da variante RAD51D c.919G>A; (p.Glu307Lys)

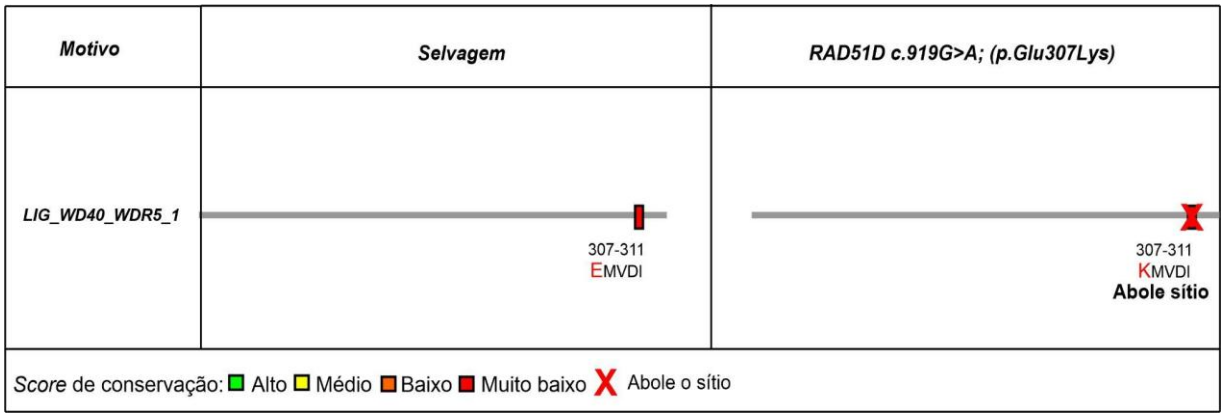


Nota: Interações interatômicas entre estruturas selvagem e mutada. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares; em verde, contato hidrofóbico; em azul ligações iônicas. Na estrutura selvagem, as interações acontecem entre os resíduos: Ser88, Lys91 e Leu92, Ala293, Phe305, Gln306, Glu307, Met308. Na estrutura mutada, são vistos os mesmos aminoácidos.

Fonte: Elaborada pela autora, usando DynaMut.

Em relação aos domínios, motivos, características moleculares e modificações pós-síntese, observou-se que a perda da Glu307 leva a perda de um sítio do motivo *LIG_WD40_WDR5_1* 307-311 [EMVDI] (Fig. 65).

Figura 65 - Predição de impacto de *RAD51D* c.919G>A; (p.Glu307Lys) nos motivos proteicos

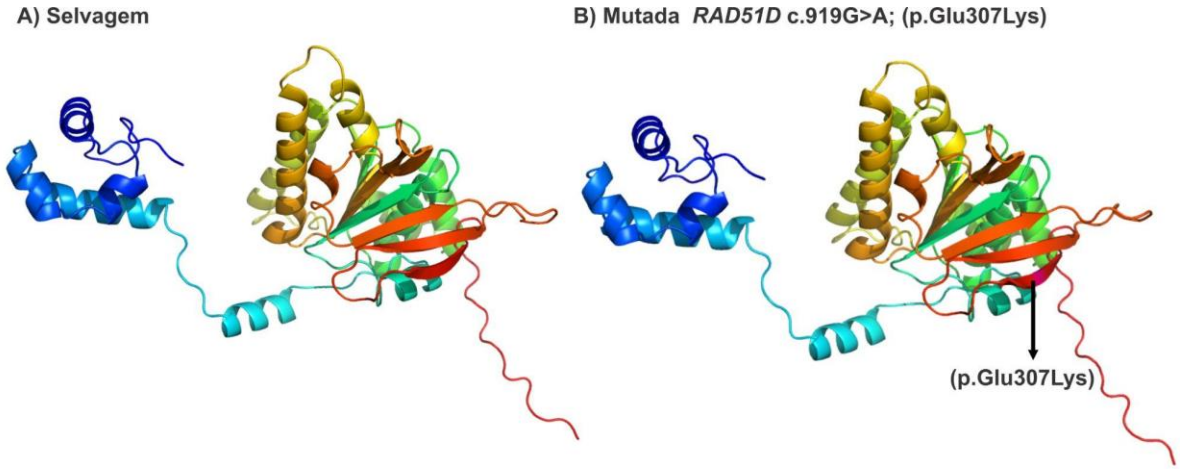


Nota: Cada barra horizontal em cinza representa a organização da proteína RAD51D. Os motivos são representados por retângulos coloridos em verde, amarelo, laranja ou vermelho. Retângulos verdes são sítios de motivos altamente conservados; os sítios amarelos são moderadamente conservados; laranja e vermelho são sítios de baixa ou nenhuma conservação. Os quadrados vermelhos indicam o local onde a variante está situada e a seta abaixo do quadrado aponta a consequência da variante, por exemplo, se abole o sítio, se ativa um novo sítio ou se não há impacto significativo.

Fonte: Elaborada pela autora, usando ELM Motifs.

Considerando a estrutura 3D da proteína, previu-se que a Glu307 está localizada em uma folha beta e, embora seja substituída por uma lisina, não foi previsto impacto na estrutura 3D (Fig. 66).

Figura 66 - Estrutura 3D da RAD51D selvagem e mutada (*RAD51D* c.919G>A; (p.Glu307Lys))



Nota: Em A está sendo mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada que é indicada na figura B. A *RAD51D* c.919G>A; (p.Glu307Lys) está localizada em uma folha beta, indicada por uma seta.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.15.1 Reclassificação da variante RAD51D c.919G>A; (p.Glu307Lys)

Com base nas informações acima, propomos manter a classificação VUS da variante *RAD51D c.919G>A; (p.Glu307Lys)*, segundo os critérios ACMG/AMP: PM1, PP2 (ver na Seção 3.2.2) e BP4.

PM1 - Variante em hotspot ou domínio funcional sem variante benigna frequente

A variante está localizada em um sítio de baixa conservação do motivo *LIG_WD40_WDR5_1*. Conforme avaliado na base de dados do ClinVar, essa região não contém variantes de sentido trocado benignas. Para essa região, as variantes benignas ou provavelmente benignas são sinônimas.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.):

- O Glu307 está presente em todas as espécies de mamíferos e na maioria das de peixes (onde em duas espécies é substituída por uma alanina). Torna-se não conservada em aves e anfíbios e répteis (todas as espécies têm uma valina nesta posição)
- Os preditores *in silico* utilizados sugerem efeitos benignos para esta substituição;
- Foram previstas inúmeras alterações nas interações moleculares, em que se perderam algumas interações como ponte de hidrogênio, contato hidrofóbico, ligação intrônica e polares

após a troca de aminoácidos. No entanto, esse achado não é evidência suficiente de patogenicidade.

3.2.2.15.2 Conclusão sobre a reclassificação da variante *RAD51D* c.919G>A; (p.Glu307Lys)

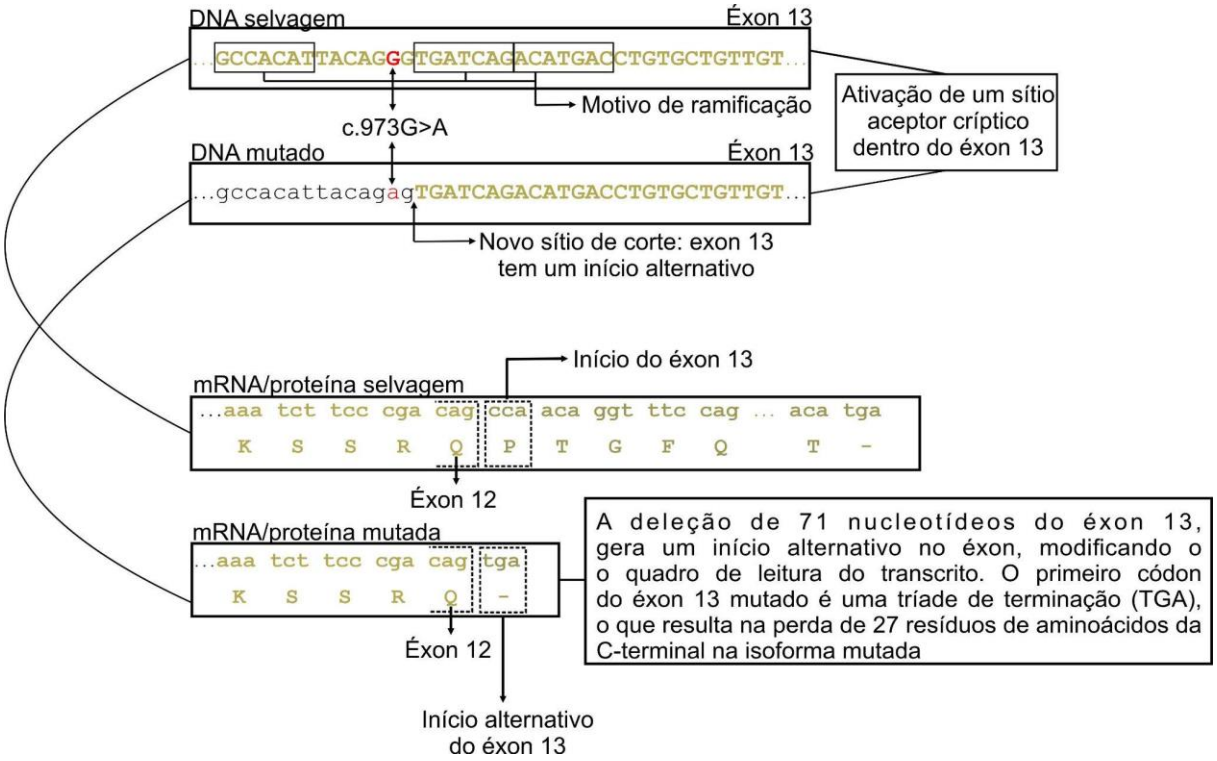
Os preditores *in silico* sugerem que a variante seja benigna, provavelmente considerando a conservação evolutiva do resíduo, ausência de efeitos em domínios funcionais grandes ou regiões de interação molecular. Entretanto, em relação à conservação evolutiva, o Glu307 é altamente conservado em mamíferos, tornando-se menos conservado em organismos mais distantes. Por outro lado, no presente estudo, foi previsto a substituição de um ácido glutâmico por uma lisina levou a perda de um sítio do motivo *LIG_WD40_WDR5_1,* cuja função é ligar-se aos domínios de repetição WD40 de WDR5. O domínio de repetição WD40 WDR5 é encontrado em proteínas, que desempenha um papel fundamental na metilação de H3K4 e acetilação de H4K16, atuando como uma proteína de suporte para a montagem do respectivo complexo de metilação e acetilação de histonas, que são conservados através da evolução (Odho *et al.,* 2010). Entretanto, não se sabe a função deste domínio na RAD51D.

A variante foi identificada em bancos de dados populacionais com uma frequência alélica de 0,0003 (gnomAD), 0,0006 (1000 Genomes Project) e 0,00092 (Exome Project Genome). Tung e colaboradores encontraram essa variante em duas mulheres com câncer de mama, que foram diagnosticadas com menos de 50 anos. Segundo o autor, considerando a história familiar das pacientes, o risco associado à presença dessa variante seria alto. Até o momento, não foram encontrados dados experimentais ou associação desta variante com a síndrome hereditária do câncer de mama e ovário. Esta variante causa uma troca de carga no resíduo; entretanto, a avaliação do efeito desta troca de carga fica dificultada porque o resíduo Glu307 está na cauda C-terminal de RAD51D, que por ser muito flexível e intrinsecamente desordenada, poderia tolerar maior variabilidade de resíduos na região. Considerando estes achados, é possível que a *RAD51D* c.919G>A;(p.Glu307Lys) seja uma variante benigna de baixa frequência. Entretanto, a conservação evolutiva do resíduo sugere que outro aminoácido não seja tolerado nesta posição em mamíferos. Consequentemente, esta variante permaneceria classificada como VUS.

3.2.2.16 *RAD51D* c.973G>A (p.Gly325Ser)

Localizada no éxon 13 do gene e no éxon 9 do transcrito canônico ENST00000345365.11 (Fig. 12, na seção 3.2), resulta de uma substituição de G para A na posição nucleotídica 973. Foi predito que essa substituição conduza a ativação do sítio aceptor críptico, que gera um início alternativo no éxon 13, levando à perda de 71 nucleotídeos. Com essa deleção, a sequência C-terminal da proteína é modificada, gerando um novo códon de parada, levando à perda de 27 resíduos de aminoácidos (Fig. 67).

Figura 67 - Impacto da *RAD51D* c.973G>A (p.Gly325Ser) nas sequências auxiliares regulatórias de sítios de *splicing*



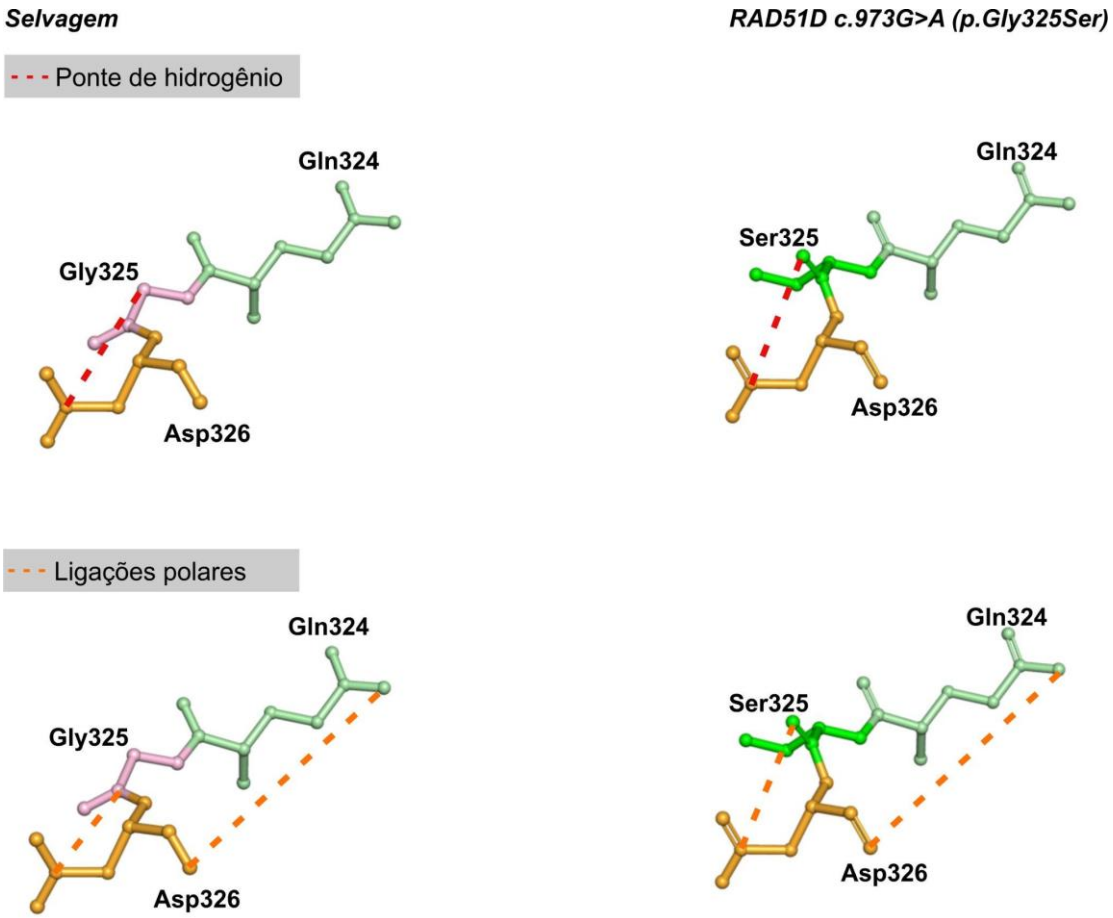
Nota: Na imagem, a troca de guanina por adenina (nucleotídeos em vermelho) leva à ativação de um sítio aceptor críptico dentro do éxon 13. Três pontos de ramificação são vistos muito próximos à posição da variante, o que pode levar a não ativação desse novo sítio. Porém, se houver a ativação desse aceptor, 71 nucleotídeos do éxon 13 serão perdidos, passando a fazer parte do íntron 12. Essa alteração muda o quadro de leitura da proteína e um códon de parada prematuro é gerado no primeiro códon do éxon 13, resultando em uma isoforma menor do que a isoforma de tipo selvagem, com a perda dos 27 aminoácidos da C-terminal, que são codificados pelo exon 13.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

Para que este aceptor seja efetivamente usado, outras características são necessárias, como a proximidade de um sítio de ramificação, permitindo o novo corte do RNA, e sequências acentuadores, permitindo a entrada deste novo éxon.

Caso este novo sítio aceptor não seja ativado, serão observados os efeitos de substituição de aminoácidos. A glicina no códon 325 é substituída por serina, um aminoácido com propriedades físico-químicas diferentes, visto que a glicina é um aminoácido apolar e a serina é polar não carregado. Os preditores in silico Mutation Taster, LIST-02, FATH-MM, Provean, Polyphen e UMD-Predictor aplicados para a determinação do potencial efeito das variantes em relação a função proteica preveem **efeitos benignos** para essa substituição, considerando a conservação do resíduo e regiões funcionais na proteína (Quadro 6, na seção 3.2.6). Foi previsto que a Gly325 estabeleça ponte de hidrogênio (Gly325 → Asp326) e interações polares (Gly325 → Gln324, Gly325 → Asp326) (Fig. 68). Não houve perda de interação polar após a troca de aminoácidos.

Figura 68 - Dinâmica molecular da variante *RAD51D* c.973G>A (p.Gly325Ser)

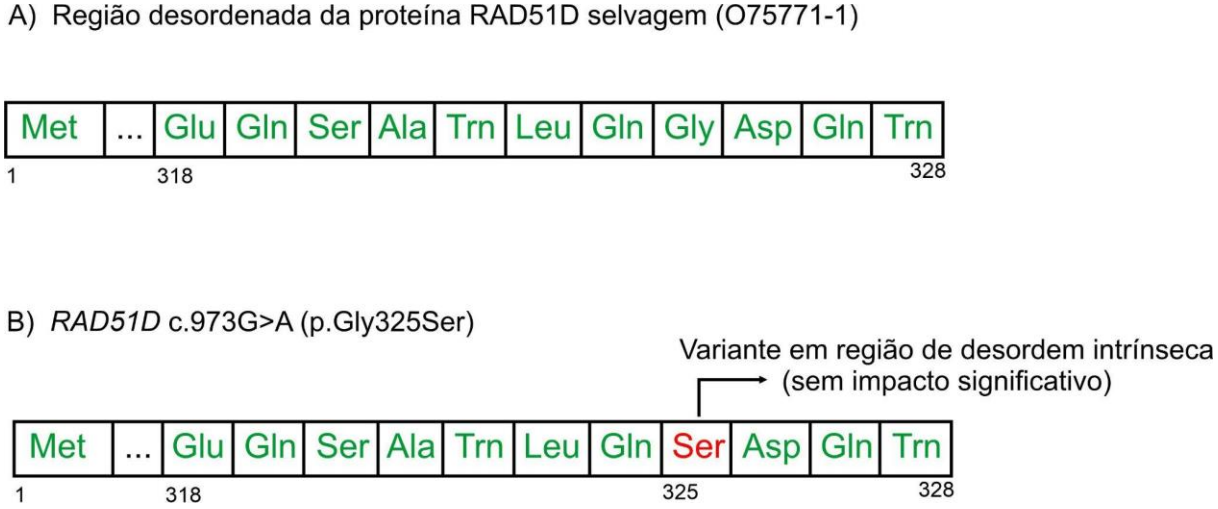


Nota: Interações interatômicas entre estruturas selvagens e mutadas. O traço vermelho representa a ponte de hidrogênio; em laranja, as interações polares. Na estrutura selvagem, as interações acontecem entre os resíduos: Gln234, Gly325 e Asp326. Na estrutura mutada, a entrada da Ser325 não altera as interações.

Fonte: Elaborada pela autora, usando DynaMut.

A perda da Gly325 leva a desestabilização da estrutura proteica, afetando a habilidade de interação com o DNA (Quadro 6, na Seção 3.2.6).

Figura 69 - Impacto da variante *RAD51D* c.973G>A (p.Gly325Ser) na desordem proteica

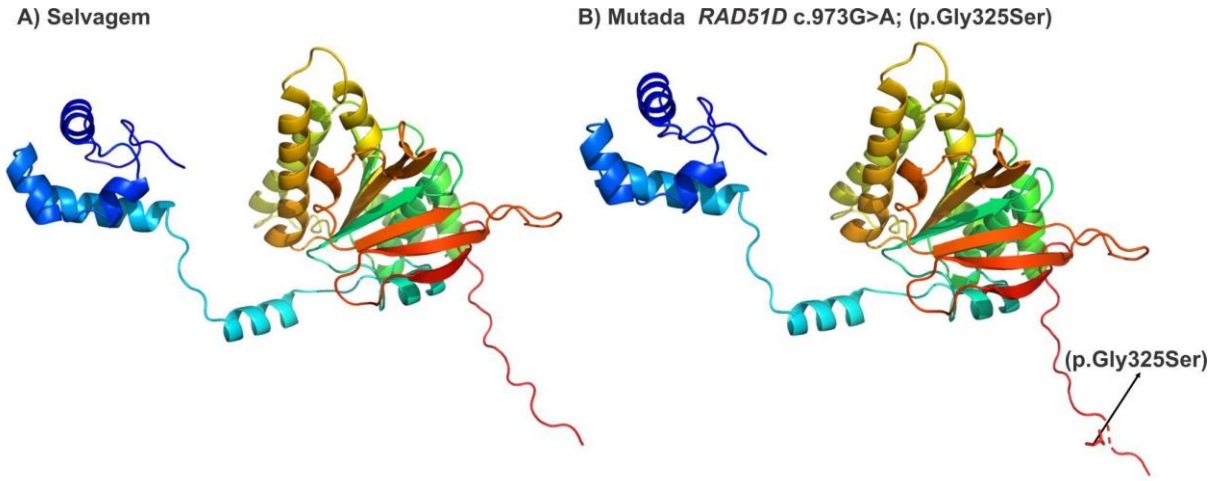


Nota: As informações sobre desordem intrínseca são mostradas com detalhes na parte de anotação funcional do gene e proteína. Aqui, está sendo mostrado o impacto da variante sobre a desordem, onde em A estão alguns dos aminoácidos preditos como intrinsecamente desordenados na sequência selvagem. Em B, está sendo mostrado que a Gly325 (aminoácido em vermelho) está localizada em uma região de desordem e mesmo sendo substituída por uma serina não apresenta impactos significativos para essa região.

Fonte: Elaborada pela autora, usando PSIPred.

Considerando a estrutura 3D da proteína, previu-se que a Gly325 está localizada em uma alça e, embora seja substituída por uma serina, não foi previsto impacto a estrutura 3D (Fig. 69).

Figura 70 - Estrutura 3D da RAD51D selvagem e mutada *RAD51D* c.973G>A (p.Gly325Ser)



Nota: Em A está sendo mostrada a estrutura da proteína 3D selvagem em relação a estrutura mutada que é indicada na figura B. A *RAD51D* c.973G>A (p.Gly325Ser) está localizada em uma alça, indicada por uma seta. A estrutura 3D não mostra efeitos significativos causados pela variante.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.2.16.1 Reclassificação da variante *RAD51D* c.973G>A (p.Gly325Ser)

Com base nas informações acima, propomos manter a classificação VUS da variante RAD51D c.973G>A (p.Gly325Ser), segundo os critérios ACMG/AMP: PM2, PP2 (ver na Seção 3.2.2) e BP4.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica em bancos de dados populacionais, como o gnomAD (0.000008) e não foi encontrada no 1000 Genomes Project ou no Exome Sequencing Project.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A Gly325 é conservada entre espécies de mamíferos (sendo substituída em seis espécies por valina, ácido aspártico, ácido glutâmico ou arginina) e anfíbios e répteis (substituída por uma alanina em uma espécie). Aves apresentam nesta posição uma treonina (valina em dois organismos); Peixes apresentam uma arginina nesta posição.

- A cauda C-terminal da RAD51D é predita como desordenada. Contudo, embora a Gly325 esteja localizada entre esses resíduos, a troca de aminoácidos não apresenta efeitos significativos a essa região (ganho ou diminuição da desordem intrínseca);
- Os preditores *in silico* utilizados sugerem efeitos benignos para esta substituição;
- Foi previsto que a troca de nucleotídeos ative um sítio aceptor críptico, mas os pontos de ramificação estão muito próximos, o que pode inativar esse aceptor. No entanto, se o aceptor for ativado, ocorre a perda 71 nucleotídeos do éxon 13 (o éxon selvagem tem 84 nucleotídeos). Considerando que o primeiro códon do éxon 13 é uma tríade de terminação, a proteína não apresentaria o éxon 13 e resultaria na perda de 27 aminoácidos da C-terminal.
- Não foram previstas alterações nas interações intermoleculares;
- Não foram encontrados efeitos sobre os motivos e domínios funcionais.

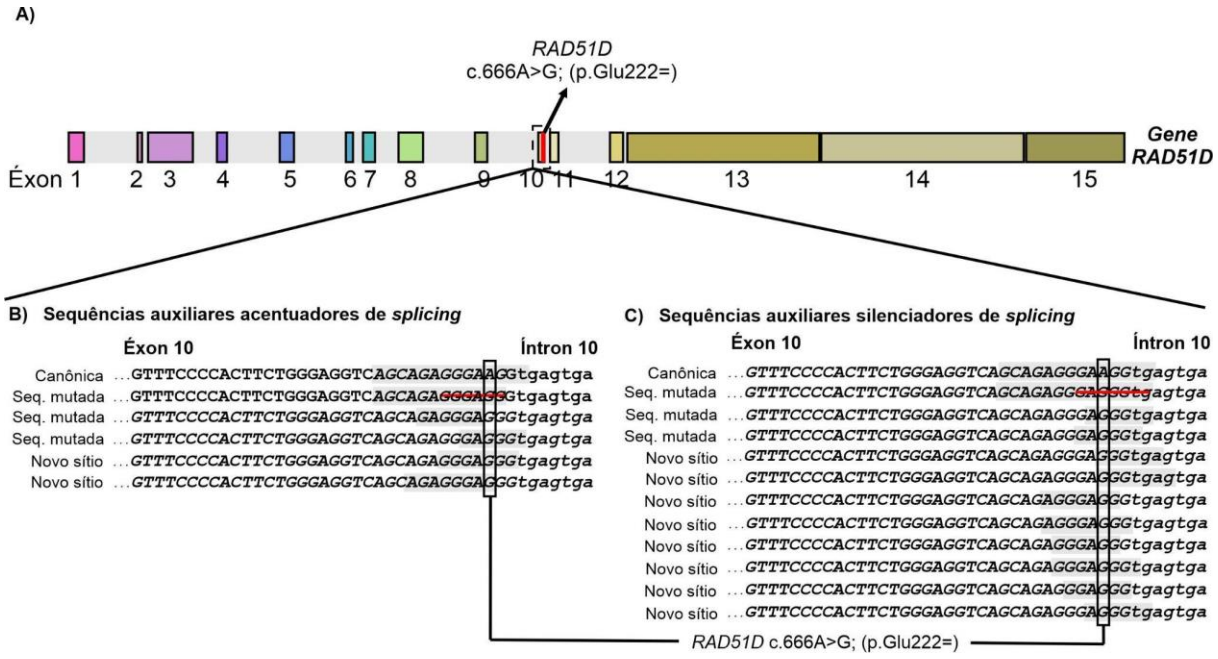
3.2.2.16.2 Conclusão sobre a reclassificação da variante *RAD51D* c.973G>A (p.Gly325Ser)

A variante está localizada em uma região desordenada, cuja sequência varia em organismos mais distantes, limitando a avaliação da conservação evolutiva do resíduo. No entanto, Gly325 é conservada entre mamíferos e anfíbios e répteis, às vezes sendo substituída por Ala, Gln, Arg, Val, Glu e Asp. Devido à posição do resíduo, a variante não afeta motivos, domínios ou regiões moleculares funcionais, portanto, os preditores utilizados não prevêem efeitos deletérios para esta variante. Contudo, Tung e colaboradores (2015) relatam a variante em uma paciente latino-americana diagnosticada com câncer de mama com menos de 50 anos. Considerando a história familiar o risco para câncer de mama é alto. A variante também foi encontrada nos estudos de Yurgelun e colaboradores (2015), que realizaram testes genéticos clínicos para a síndrome de Lynch em 1.260 indivíduos com histórico de câncer e/ou pólipos associados à síndrome de Lynch. Mas, não foi encontrada associação com a síndrome hereditária do câncer de mama e ovário e não foram encontrados dados experimentais, que expliquem a consequência molecular dessa variante sobre a função da proteína. A frequência alélica é muito baixa no gnomAD (0.000008) e não foi encontrada no 1000 Genomes Project ou no Exome Sequencing Project. No entanto, esses achados são insuficientes para determinar a patogenicidade desta variante, mantendo sua classificação VUS.

3.2.3 Estudo da variante germinativa sinônima *RAD51D* c.666A>G; (p.Glu222=)

Localizada no éxon 10 do gene e no éxon 6 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2), resulta de uma substituição de A para G na posição nucleotídica 666. Para essa substituição, foram previstos efeitos significativos nas sequências auxiliares reguladoras de *splicing* (Fig. 71).

Figura 71- Impacto da *RAD51D* c.666A>G; (p.Glu222=) nas sequências auxiliares regulatórias de sítios de *splicing*



Nota: Em A, é mostrado o arranjo dos éxons ao longo do gene RAD51D. O *RAD51D* tem 15 éxons (cada um com uma cor atribuída). A variante *RAD51D* c.666A>G; (p.Glu222=) está indicada por uma seta e por uma linha vertical vermelha. Nas figuras B e C, são mostradas as sequências auxiliares, acentuadores ou silenciadores, que foram afetados após a troca de nucleotídeos. Em B, é mostrado o impacto da variante em relação aos acentuadores de *splicing*. A sequência canônica contém o sítio acentuador de *splicing* de tipo selvagem sublinhado em cinza e abaixo dele estão as sequências mutadas. A maioria dos sítios não foi afetada pela troca de nucleotídeos (sequência intacta), dois foram abolidos (representados por uma linha vermelha sobre a sequência) ou criados (novo sítio). Em C, está sendo mostrado o impacto da variante em relação aos silenciadores de *splicing*. A sequência canônica contém o sítio silenciador de *splicing* de tipo selvagem e abaixo estão as sequências mutadas. Na imagem, é mostrado que sítios próximos à variante não foram afetados após a troca de nucleotídeos, mas um sítio envolvendo a variante foi abolido e novos sítios foram gerados no local.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

Foi investigado o efeito da variante na conformação do RNA; não foram previstos efeitos deletérios (Quadro 5). A variante também não apresentou efeitos significativos nos sítios CpG ou viés de uso de códons (Anexo 1).

Quadro 5 - Predição de impacto molecular da *RAD51D* c.666A>G; (p.Glu222=) sobre a estrutura de RNA, usando ViennaRNA

Estrutura selvagem	**Estrutura mutada *RAD51D* c.666A>G; (p.Glu222=)**
Energia livre mínima da estrutura secundária: -379,60 kcal/mol	Energia livre mínima da estrutura secundária mutada: -378.30 kcal/mol
Energia livre do conjunto termodinâmico: -396,57 kcal/mol	Energia livre do conjunto termodinâmico: -395.71 kcal/mol
Energia livre mínima da frequência da estrutura no conjunto é 0,00 %	Energia livre mínima da frequência da estrutura no conjunto é 0,00 %
A diversidade do conjunto: 228,88	A diversidade do conjunto: 267.70.

Fonte: Elaborado pela autora

3.2.3.1 Reclassificação da variante *RAD51D* c.666A>G; (p.Glu222=)

Com base nas informações acima, propomos manter a classificação VUS da variante *RAD51D c.666A>G; (p.Glu222=)*, segundo os critérios ACMG/AMP: PP2 (ver na Seção 3.2.2), BP4 e BP6.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

- A Glu222 é altamente conservada entre as espécies de mamíferos, aves e anfíbios e répteis. Em peixes é conservada, sendo substituída por ácido aspártico.
- Foi previsto, que a troca de nucleotídeos ative novas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadores). Entretanto, a avaliação do impacto destes achados depende de estudos *in vitro* ou *in vivo*;
- Não foram previstos efeitos deletérios em relação a conformação de RNA, sítios CpG e viés de uso de códons.

BP6 - Fonte respeitável relata recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente:

A variante é reportada como provavelmente benigna em bancos de dados públicos como ClinVar, VarSome e ABraOm.

3.2.3.2 Conclusão sobrea a reclassificação da variante *RAD51D* c.666A>G; (p.Glu222=)

Durante a investigação da patogenicidade de variantes genéticas, variantes sinônimas assumem automaticamente um papel benigno, devido ao seu baixo ou nenhum potencial de impacto sobre a proteína. No entanto, existem estudos que desmistificam essa ideia, sugerindo mecanismos alternativos de patogenicidade atuando em nível de RNA. Por exemplo, o impacto pode ser sobre: 1. os mecanismos de *splicing* e a regulação do *splicing* alternativo; 2. Mudanças na dinâmica conformacional do mRNA, causando aumento ou diminuição na energia de dobramento do mRNA, podendo afetar sua estabilidade e conformação; 3. Viés no uso de códon, ou seja, a substituição faz com que um códon mais frequente seja substituído por outro, menos frequente. Como a frequência de uso do códon corresponde, em termos gerais, à abundância do tRNA na célula, a entrada de um códon raro pode ocasionar a parada da tradução, por escassez de substrato; 4. a substituição pode afetar elementos regulatórios que atuam em nível do DNA, como dinucleotídeos de CpG sujeitos a metilação variável.

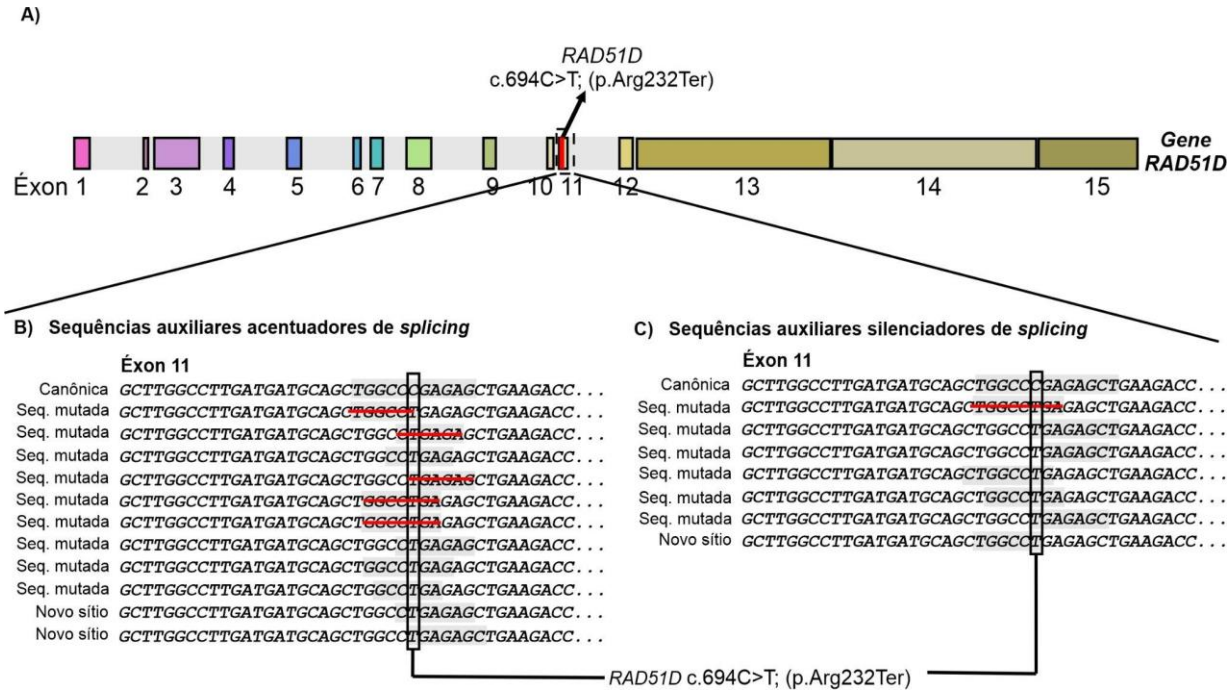
Contudo, em nossos resultados, não foram previstos impactos significativos na conformação do RNA, sítios CpG ou mesmo viés de uso de códons, fornecendo evidências de um papel benigno para a variante *RAD51D* c.666A>G; (p.Glu222=). Além disso, prevê-se que a variante ative novas sequências auxiliares (acentuadores e silenciadores) de *splicing*, mas o impacto funcional deste achado só pode ser avaliado por meio de estudos funcionais, o que é uma das limitações deste estudo. Nas bases de dados ClinVar e VarSome, a variante é classificada como benigna/provavelmente benigna, considerando principalmente a consequência molecular (variante sinônima) e a frequência alélica, relativamente alta: gnomAD (0,0004); 1000 Genome Project (0,001); e Exome Project Genome (0,002). Até o momento, não foi descrita associação dessa variante com casos de câncer de mama e/ou ovário. Diante desses achados, propomos manter a classificação benigna para essa variante.

3.2.4 Estudo da variante germinativa perda de sentido *RAD51D* c.694C>T; (p.Arg232Ter)

A variante *RAD51D* c.694C>T; (p.Arg232Ter) está localizada no éxon 11 do gene e no éxon 7 do transcrito canônico ENST00000345365.11 (Fig. 12, na Seção 3.2) e resulta de uma substituição de C para T na posição nucleotídica 694. A substituição leva a uma mudança no quadro de leitura, onde a arginina 232 é substituída por um códon de parada dentro do éxon. De acordo com o Mutation Taster, a alteração resulta em perda de função devido ao códon de parada prematura, estando sujeita à degradação do mRNA mediada por variantes sem sentido (do Inglês, *nonsense mediated decay)*.

Além disso, foram previstos efeitos significativos sobre os sítios de *splicing*, bem como nas sequências auxiliares reguladoras de *splicing* (acentuadores e silenciadoras) (Fig. 72). Entretanto, na presença de NMD, estes achados provavelmente perdem relevância.

Figura 72 - Impacto da *RAD51D* c.694C>T; (p.Arg232Ter) nas sequências auxiliares regulatórias de sítios de *splicing*



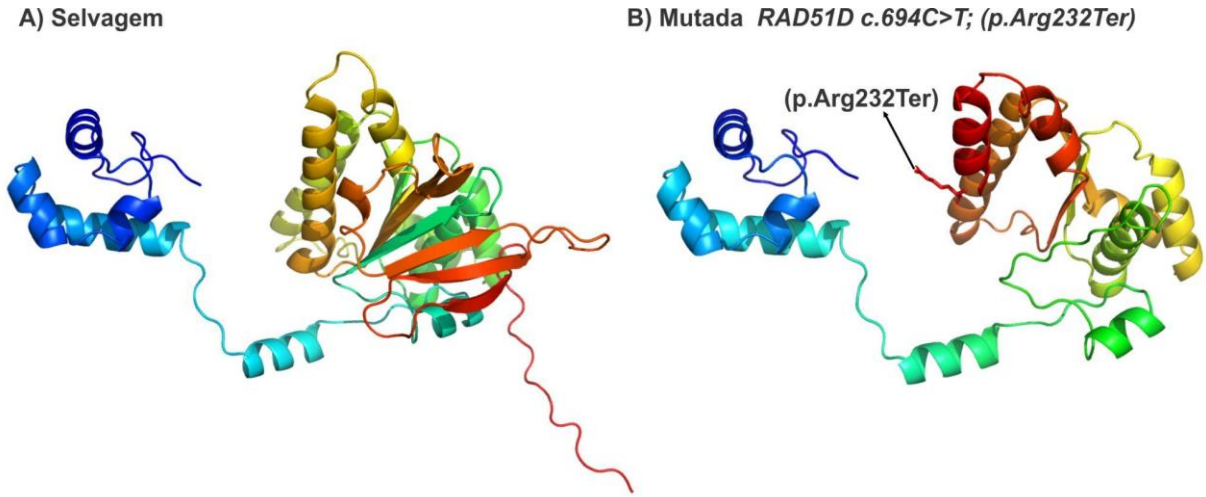
Nota: Em A, está sendo mostrado o arranjo dos éxons ao longo do gene RAD51D. O *RAD51D* tem 15 éxons (cada um com uma cor atribuída). A variante *RAD51D* c.694C>T; (p.Arg232Ter) ocorre no éxon 11 do gene e sua localização pode ser vista por meio da seta indicando a variante e por uma linha vertical vermelha. Nas figuras B e C, são mostradas as sequências auxiliares, acentuadores ou silenciadores, que foram afetados após a troca de nucleotídeos. Em B, está sendo mostrado o impacto da variante em relação aos acentuadores de *splicing*. A sequência canônica contém o sítio acentuador

de *splicing* de tipo selvagem destacados em cinza e abaixo dele estão as sequências mutadas. Alguns sítios não foram afetados mesmo após a troca de nucleotídeos (sequência intacta), outros foram abolidos (representados por uma linha vermelha sobre a sequência) ou criados (novo sítio). Por outro lado, em C, está sendo mostrado o impacto da variante em relação aos silenciadores de *splicing*. A sequência canônica contém o sítio silenciador de *splicing* de tipo selvagem e abaixo estão as sequências mutadas. Na imagem é mostrado que sítios próximos à variante não foram afetados após a troca de nucleotídeos, mas um sítio envolvendo a variante foi abolido e um novo sítio foi gerado no local.

Fonte: Elaborada pela autora, usando Human Splicing Finder.

Caso não ocorra o NMD, conforme previsto acima, poderia haver impactos no nível da proteína. Considerando a estrutura 3D da proteína, previu-se que a Arg232Ter estaria localizada em uma alça e levaria a perda de 96 aminoácidos do C-terminal (Fig. 73).

Figura 73 - Estrutura 3D da RAD51D selvagem e mutada *RAD51D* c.694C>T; (p.Arg232Ter)



Nota: Em A, está sendo mostrada a estrutura da proteína 3D selvagem; em B, da proteína mutada. A *RAD51D c.694C>T; (p.Arg232Ter)* está localizada em uma alça, indicada por uma seta. A estrutura 3D mostra que a perda de 96 aminoácidos da C-terminal modifica a estrutura 3D de RAD51D.

Fonte: Elaborada pela autora, usando AlphaFold.

3.2.4.1 Reclassificação da variante RAD51D c.694C>T; (p.Arg232Ter)

Esta variante é classificada como patogênica, por criar um códon de terminação prematura. No presente estudo, buscamos evidências sobre os mecanismos que possam justificar seu efeito.

PVS1 - Variante nula (nonsense, frameshift, canônico ±1 ou 2 sítios de splicing, códon de iniciação, deleção única ou multiexonica) em um gene onde LOF é um mecanismo conhecido da doença

Essa mudança de sequência cria um códon de parada prematura, que resulta em uma proteína mais curta. Essa alteração está sujeita à NMD.

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica no gnomAD: (0,000003), e não foi encontrada no 1000 Genomes Project ou no Exome Sequencing Project.

PP5 - Uma fonte respeitável relatou recentemente uma variante como patogênica, mas a evidência não está disponível para o laboratório realizar uma avaliação

A variante é reportada como patogênica em bancos de dados públicos como ClinVar, VarSome e ABraOm.

3.2.4.2 Conclusão sobre a reclassificação da variante *RAD51D* c.694C>T; (p.Arg232Ter)

A variante resulta na perda de função da proteína devido à perda de 96 resíduos C-terminais causada pelo ganho prematuro do códon de parada no códon 232, além da possibilidade de NMD. Dados experimentais que investigam os efeitos da variante na função da proteína ainda não foram encontrados na literatura, embora ela tenha sido associada a casos de câncer de mama e/ou ovário (Wickramanayake *et al.,* 2012;Gutierrez-Enrıquez *et al.,* 2013; Pennington *et al.,* 2014; Baker *et al.,* 2015; Susswein *et al.,* 2016; Tedaldi *et al.,* 2017; Sánchez-Bermúdez *et al.,* 2018; Suszynska *et al.,* 2020). Gutierrez-Enrıquez e colaboradores (2013) identificaram esta variante em uma paciente diagnosticada com câncer de ovário aos 44 anos; Janatova e colaboradores (2015) encontraram a variante em duas pacientes jovens com diagnóstico de câncer de ovário sem história familiar; González-Rivera e colaboradores (2016) rastreou esta variante em uma paciente com câncer de ovário bilateral com menos de

50 anos; Tedaldi e colaboradores (2017) relataram a variante em uma paciente com 47 anos diagnosticada com câncer de ovário; Susswein e colaboradores (2015) encontram em uma paciente com câncer de mama; Todos estes estudos sugeriram que a variante seja patogênica. No estudo de Suszynska e colaboradores (2020), a variante esteve presente em 11 de 22584 casos (frequência = 0,0487), resultando em frequência ~16 vezes maior do que a observada nos controles (4/131873 = 0,003). Curiosamente, esta variante foi rastreada em pacientes que também eram portadores da variante *RAD51D* c.715C>T (p.Arg239Trp), sugerindo que provavelmente ambas as variantes estejam em *cis* (Wickramanayake *et al.,* 2012; Gutierrez-Enrıquez *et al.,* 2013; Sánchez-Bermúdez *et al.,* 2018). No presente estudo, a variante também foi encontrada co-ocorrendo com a variante *RAD51D* c.715C>T (p.Arg239Trp), em paciente com diagnóstico de câncer mama. Diante desses resultados, a variante mantém sua classificação patogênica. Nos estudos anteriores, os mecanismos de perda de função não haviam sido analisados. No presente estudo, identificamos dois possíveis mecanismos: degradação do mRNA por NMD e degradação da proteína em função da falta do C-terminal.

3.2.5 Estudo das variantes germinativas intrônicas *RAD51D* c.345+5A>G e a *RAD51D* c.481-7G>A

Para as variantes intrônicas (*RAD51D* c.345+5A>G e a *RAD51D* c.481-7G>A) não foram previstos efeitos significativos nos sítios de *splicing* ou sequências auxiliares reguladoras de *splicing*.

3.2.5.1 Reclassificação da variante *RAD51D* c.345+5A>G

PM2 - Ausente de controles (ou em frequência extremamente baixa se recessivo) no Exome Sequencing Project, 1000 Genomes Project ou Exome Aggregation Consortium

A variante possui uma baixa frequência alélica no gnomAD (0,000004) e não foi encontrada no 1000 Genomes Project e no Exome Sequencing Project.

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.):

- Não foram previstos efeitos nos sítios de *splicing* ou sequências auxiliares reguladores de *splicing*.

3.2.5.1.2 Conclusão sobre a reclassificação da variante *RAD51D* c.345+5A>G

Em nossos resultados, nenhum impacto significativo foi previsto nos sítios de *splicing*, ativando ou abolindo sítios aceptores ou doadores crípticos e aceptores ou doadores selvagens, respectivamente. No entanto, até o momento, não existem dados experimentais que confirmem essa previsão. Além disso, nenhum estudo que associa essa variante a casos de câncer de mama e/ou ovário foi encontrado na literatura. A frequência do alelo nos bancos de população é baixa no gnomAD (0,000004) e ela não foi encontrada no Projeto 1000 Genomas ou no Projeto Genoma Exome. No ClinVar, a interpretação da patogenicidade dessa variante é conflitante, havendo três entradas para VUS e uma para provavelmente benigna. Devido a falta de evidências de suporte, dados de segregação em famílias ou mesmo estudos de caso-controle, propomos manter a classificação dessa variante como VUS.

3.2.5.2 Reclassificação da variante *RAD51D* c.481-7G>A

BP4 – Múltiplas linhas de evidência computacional sugerem nenhum impacto no gene ou produto gênico (conservação, evolução, impacto de splicing, etc.)

Não foram previstos efeitos nos sítios de *splicing* ou sequências auxiliares reguladoras do splicing.

BP6 - Fonte respeitável relata recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente

A variante é reportada como provavelmente benigna em bancos de dados públicos como ClinVar, VarSome e ABraOm.

3.2.5.2.1 Conclusão sobre a reclassificação da variante *RAD51D* c.481-7G>A

Semelhante a c.345+5A>G, não foram previstos impactos significativos nos sítios de *splicing* possivelmente causados pela variante c. 481-7G>A. No entanto, a repercussão funcional desses resultados só pode ser investigada por meio de estudos funcionais, que estão ausentes para essa variante. Tung e colaboradores (2014) encontraram essa variante em três três mulheres africanas e uma mulher latino-americana todas diagnosticadas com câncer de mama antes dos 50 anos. Considerando a história familiar, as quatro mulheres apresentavam baixo risco. No estudo de Yurgelun e colaboradores (2015), a variante foi encontrada em um dos indivíduos submetidos ao teste genético com suspeita de síndrome de Lynch. No ClinVar, a

variante é classificada como benigna/provavelmente benigna, principalmente considerando a frequência populacional que é relativamente alta: gnomaAD (0,0005), 1000 Genomes Project (0,0008) e Exome Project Genome (0,001). Diante desses achados, a variante mantém sua classificação benigna.

3.2.6 Síntese dos resultados

Todos os resultados sobre a investigação da patogenicidade das variantes estão sintetizados no Quadro 6, com exceção das variantes intrônicas e sinônimas que foram estudadas nos níveis de DNA, RNA e proteína. Para variantes de sentido trocado, foram utilizados algoritmos de repercussão funcional que avaliam os efeitos benignos ou deletérios na estabilidade da proteína, domínios, motivos funcionais e sítios de modificação pós-síntese, efeitos em sítios de *splicing* e conservação evolutiva, bem como para a variante de *frameshift* (Quadro 6, na Seção 3.2.6). Procurou-se investigar mudanças conformacionais na estrutura de DNA/RNA e viés de uso de códons para a variante sinônima. E, por fim, para as variantes intrônicas, foram consideradas alterações quanto aos sítios de *splicing* e elementos regulatórios de *splicing*.

Quadro 6 - Descrição dos resultados obtidos na análise in silico, evidenciando os efeitos no DNA/RNA e proteína RAD51D

Variantes RAD51D	**Conservação evolutiva**	**Predição de impactos nos sítios de splicing**	**Predição de impacto na proteína**							**Predição de impacto na estabilidade da estrutura proteica**			**Predição de impactos moleculares em proteínas**		**Classificação prévia/ proposta de classificação**
	BLASTP/Clustal Omega	**Human Splicing Finder**	**MUTATION**	**LIST-2**	**FATH-MMM**	**PROVEAN**	**POLYPHEN**	**SIFT**	**UMD-PRED**	**mCSM predição**	**mCSM proteína**	**mCSM DNA**	**DynaMut** **Predição de efeitos nas interações intermoleculares**	**UniProt, InterPro, MutPred2, PhosphoSitePlus** **Impacto em domínios de proteínas, motivos estruturais, modificações pós-síntese**	**ACMG/AMP**
c.26G>C; (p.Cys9Ser) rs140825795	Altamente conservada em mamíferos, aves, anfíbios e répteis e peixes	Alteração de sequências auxiliares (acentuadores e silenciadores)	B	P	B	P	P	P	P	↓ Des.	↓ Des.	↓ Des.	Foi previsto que o Cys9 estabeleça ponte de dissulfeto com a Cys43; ligações de hidrogênio, contatos hidrofóbicos e ligações polares. A entrada de Ser9, uma nova ligação polar é criada e o contato hidrofóbico é perdido.	Localiza-se em uma região de ligação ao ssDNA (1-83 aa). Cria novos motivos de DOC_WW_Pin1_4 (motivos de ligação de domínio WW), MOD_ProDKin_1 (sítio de fosforilação MAPK) e MOD_GSK3_1 (sítio de fosforilação GSK3)	VUS/ Provavelmente patogênica

												A perda de Cys9 também promove o ganho de desordem intrínseca da proteína.		
c.137C>G; (p.Ser46Cys) rs587780102	Altamente conservada em mamíferos, aves, anfíbios e répteis e peixes	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	?	P	B	P	P	B	P	↓ Des.	↓ Des.	↓ Des.	Previu-se que Ser46 fizesse ligações de hidrogênio e ligações polares. Após a entrada de Cys46, as pontes de hidrogênio e a interação polar são perdidas e um novo contato hidrofóbico. Prevê-se que Cys46 (aminoácido substituto) estabeleça ponte dissulfeto com Cys294 de baixa probabilidade	Está localizada na região de ligação ao ssDNA (1-83 aa). A substituição também leva à perda do sítio de acetilação K48. Causa perda do motivo PKC_PHOSPHO_SITE (sítio de fosforilação da proteína quinase C)	VUS/ Provavelmente patogênica
c.140A>G; (p.Tyr47Cys) rs776528342	Altamente conservada em mamíferos, aves, anfíbios e répteis e peixes	Alteração de sequências auxiliares (acentuadores e silenciadores)	P	P	B	P	P	B	P	↓ Des.	↓ Des.	↑ Est.	Tyr47 foi previsto para fazer ligações de hidrogênio, contatos hidrofóbicos e ligações polares. Após a entrada de Cys47, ligações como pontes de hidrogênio, contatos hidrofóbicos e ligações polares foram perdidas. Prevê-se que Cys47 (aminoácido substituto) possa estabelecer ponte dissulfeto de baixo	Localiza-se na região de ligação ao ssDNA (1-83 aa) e a região HELIX (47-61 aa). A substituição leva a perda do sítio de acetilação K48. Abole LIG_TYR_ITYM (motivo de inibidor baseado em tirosina do receptor imunológico - ITIM), TRG_ENDOCITYC_2 (motivo de sinal de classificação endocítica baseado em Y). A entrada Cys47 cria um novo sítio de motivo MOD_NEK_2_2 (sítio de fosforilação NEK2)	VUS/ Provavelmente patogênica

												score, interagindo com Cys294		
c.190C>T; (p.Pro64Ser) rs201506572	Conservada em mamíferos*	Cria um novo site aceptor de *splicing*, levando à ativação de exon críptico	B	B	B	B	B	B	B	↓ Des.	↓ Des.	↓ Des.	Não foram previstas interações intermoleculares da prolina com outros resíduos, porém, após a troca de aminoácidos, uma nova ligação polar surgiu entre a serina e a valina na posição 66.	Pro64 não está localizado nos domínios funcionais da proteína DNA recombination/repair_Rad51_C-terminal (107-303), AAA+ATPase (119-294) e ReacA_ATP-biding	Benigna/ Provavelmente benigna
c.196G>A; (p.Val66Met) rs56026142	Não conservado em espécies de mamíferos e aves. No entanto, altamente conservada em anfíbios e répteis e peixes	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	B	B	B	B	?	B	B	↓ Des.	↓ Des.	↑ Est.	previsto que Val66 estabeleça pontes de hidrogênio, contatos hidrofóbicos e ligações polares. Após a troca de aminoácidos, a maioria das interações intermoleculares foram perdidas, como pontes de hidrogênio, contato hidrofóbico e interações polares	Localiza-se em uma região que se liga preferencialmente ao ssDNA (1-83 aa) e em uma região de ligação (54-77 aa)	VUS/VUS
c.286G>A; (p.Ala96Thr) rs147933658	Conservada em mamíferos*	Sem impacto significativo nos sítios de *splicing* e elementos regulatórios	B	B	B	B	B	B	B	↓ Des.	↓ Des.	↑ Est.	Foi previsto que a Ala96 faça interações polares e essas interações permaneceram mesmo após a troca de aminoácidos	Ala96 não está localizada nos domínios funcionais da proteína DNA recombination/repair_Rad51_C-terminal (107-303), AAA+ATPase (119-294) e ReacA_ATP-biding	Benigna/ Provavelmente benigna

c.345+5A>G rs878854562	Não se aplica	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	NÃO SE APLICA											VUS/VUS
c.355T>C; (p.Cys119Arg) rs201313861	Conservada em mamíferos. No entanto, não conservada em aves, anfíbios e répteis e peixes	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	B	B	B	B	?	B	B	↓ Des.	↓ Des.	↓ Des.	Foi previsto que Cys 119 estabeleça pontes de hidrogênio, contatos hidrofóbicos e ligações polares. Prevê-se que Cys119 estabeleça pontes dissulfeto com Cys117. Após a entrada de Arg119, esta ponte seria perdida. Algumas interações, como pontes de hidrogênio, foram perdidas e novas interações de pontes de hidrogênio, contato hidrofóbico e ligações polares foram criadas.	Está localizado nos domínios RecA_ATP-biding (75-283 aa), recombinação de DNA e proteína de reparo (78-251 aa) e AAA+ ATPase (99-274 aa)	VUS/VUS
c.413A>G; (p.Asn138Ser) rs201676898	Pouco conservada em mamíferos. Não conservada em aves, anfíbios e répteis e peixes	Sem impacto significativo nos sítios de *splicing* e elementos regulatórios	B	B	B	B	B	B	B	↓ Des.	↑ Est.	↓ Des.	Foi previsto que a Asn138 estabeleça pontes de hidrogênio e ligações polares. A entrada de Ser138 levou à perda da maioria das interações e uma nova ligação iônica foi	Está localizado nos domínios RecA_ATP-biding (75-283 aa), recombinação de DNA e proteína de reparo (78-251 aa) e AAA+ ATPase (99-274 aa). A entrada Ser138 criou um novo sítio do motivo MOD_GSK3_1 (sítio de fosforilação GSK3)	VUS/VUS

												criada.		
c.481-7G>A rs145832514	Não se aplica	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	NÃO SE APLICA											Benigna/ Provavelmente benigna
c.493C>T; (p.Arg165Trp) rs544654228	Aminoácido típico nesta posição é a glutamina. Quando substituída, entra a arginina	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	P	P	B	P	P	P	B	↓ Des.	↓ Des.	↓ Des.	Foi previsto que a Arg165 faça pontes de hidrogênio, contatos hidrofóbicos, ligações iônicas e ligações polares. Todos os contatos hidrofóbicos, pontes de hidrogênio e interações polares foram perdidos	Está localizado nos domínios RecA_ATP-biding (75-283 aa), recombinação de DNA e proteína de reparo (78-251 aa), AAA+ ATPase (99-274 aa) e *Coiled coil* (147-167 aa). A perda de Arg165 leva à perda de dois sítios do motivo TRG_ER_diArg_1 (motivo di-Arg)	VUS/VUS
c.628G>A; (p.Ala210Thr) rs80116829	Altamente conservada em mamíferos, aves, anfíbios e répteis e peixes	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	B	P	B	B	P	B	P	↓ Des.	↓ Des.	↑ Est.	Foi previsto que a Ala210 estabeleça pontes de hidrogênio e ligações polares. Algumas dessas interações foram perdidas	Está localizada em domínios RecA_ATP-biding (75-283 aa), recombinação de DNA e proteína de reparo (78-251 aa) e AAA+ ATPase (99-274 aa). A entrada Thr210 criou um novo motivo MOD_CK1_1 (caseína quinase 1 - sítio de fosforilação CK1) e MOD_PLK4 (sítio de fosforilação Plk4)	VUS/VUS
c.666A>G; (p.Glu222=) rs114012742	**Não se aplica**	Alteração de sequências auxiliares (acentuadores e	**Estes critérios de avaliação da patogenicidade não se aplicam à avaliação da variante sinônima. Então, foram considerados possíveis impactos na conformação do RNA, nas ilhotas CpG ou efeitos decorrentes do viés de uso de códons. Nenhuma evidência de patogenicidade significativa foi encontrada nestas análises**											Benigna/ Provavelmente benigna

		silenciadores)												
c.694C>T; (p.Arg232Ter) rs587780104	Não se aplica	Alteração de sequências auxiliares (acentuadores e silenciadores)	**NÃO SE APLICA**									Perda de parte dos domínios RecA_ATP-biding (75-283 aa), recombinação de DNA e proteína de reparo (78-251 aa) e AAA+ ATPase (99-274 aa)		Patogênica/ Patogênica
c.695G>A; (p.Arg232Gln) rs28363283	Conservada em mamíferos. Altamente conservado em aves e anfíbios e répteis. Não conservada em peixes	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	B	P	B	B	P	B	P	↓ Des.	↑ Est.	↓ Des.	Foi previsto que a Arg232 estabeleça pontes de hidrogênio e ligações polares. Após a entrada do Gln232, algumas dessas interações foram perdidas	Está localizado nos domínios RecA_ATP-biding (75-283 aa), recombinação de DNA e proteína de reparo (78-251 aa) e AAA+ ATPase (99-274 aa).	Benigna/ Provavelmente benigna
c.698A>G; (p.Glu233Gly) rs28363284	Altamente conservada em mamíferos, aves e anfíbios e répteis. Não conservada em peixes	Alteração de sequências auxiliares (acentuadores e silenciadores)	B	P	B	P	P	B	P	↓ Des.	↓ Des.	↓ Des.	Foi previsto que a Glu233 estabeleça pontes de hidrogênio, contatos hidrofóbicos, ligações iônicas e ligações polares. Após a entrada algumas interações foram perdidas após a troca de aminoácidos, como pontes de hidrogênio, contatos hidrofóbicos, ligações iônicas e polares	Está localizada nos domínios funcionais RecA_ATP-biding (75-283 aa), DNA recombination and repair protein (78-251 aa) e AAA+ ATPase (99-274 aa). A entrada Gly233 cria um novo sítio do motivo LIG_14-3-3_CanoR_1 (motivo de ligação de fosfopeptídeo 14-3-3)	Benigna/ Provavelmente benigna
c.715C>T; (p.Arg239Trp)	Conservada em mamíferos e aves. Altamente conservado em	Alteração de sequências auxiliares (acentuadores e silenciadores)	B	?	B	P	P	B	P	↓ Des.	↓ Des.	↓ Des.	Foi previsto que a Arg239 estabeleça pontes de hidrogênio e ligações polares. A entrada de Trp239	Está localizada nos domínios funcionais RecA_ATP-biding (75-283 aa), DNA recombination and repair protein (78-251 aa) e AAA+ ATPase (99-274 aa). A entrada da	VUS/VUS

	anfíbios e répteis. Não conservada em peixes											leva à perda de algumas ligações polares e gera uma nova interação de contato hidrofóbica	Trp239 cria um sítio de baixa conservação do motivo LIG_KCL_WD_1 na posição (237-246 aa)	
c.862G>C; (p.Gly288Arg) rs1391505912	Conservada em mamíferos. Não conservado em aves, anfíbios e répteis e peixes	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	B	B	B	B	B	B	B	↓ Des.	↓ Des.	↓ Des.	Foi previsto que a Gly288 estabeleça ponte de hidrogênio e ligações polares. Algumas ligações polares foram perdidas após a entrada de Arg288 e novas interações foram geradas	Está localizada sobre os motivos N-miristoilação (283-288 aa) e Amidação (288-291 aa), mas sem efeitos deletérios a esses sítios	VUS/VUS
c.919G>A; (p.Glu307Lys) rs115031549	Altamente conservada em mamíferos e peixes. Não conservado em aves e anfíbios e répteis	Sem impacto significativo nos sítios de *splicing* e sequências auxiliares (acentuadores e silenciadores)	B	B	B	B	B	B	B	↓ Des.	↓ Des.	↑ Est.	Foi previsto que a Glu307 estabeleça ponte de hidrogênio, contato hidrofóbico, ligação iônica e polares. Algumas dessas interações foram perdidas após a troca de aminoácidos, e uma nova ligação polar foi gerada	Perda de um sítio do motivo LIG_WD40_WDR5_1 (ligar-se aos domínios de repetição WD40 de WDR5 - 307-311 aa)	VUS/VUS
c.973G>A; (p.Gly325Ser) rs587780106	Conservada em mamíferos e anfíbios e répteis. Não conservado em aves e peixes	Cria um novo site de *splicing* Aceptor (ativação de éxon críptico)	B	B	B	B	B	P	B	↓ Des.	↑ Est.	↓ Des.	Foi previsto que a Gly325 estabeleça ponte de hidrogênio, ligação polar. Não houve perda de interação após a troca de aminoácidos	Não está localizada em domínios funcionais ou motivos	VUS/VUS

Nota: Uma descrição detalhada das substituições nucleotídica e aminoacídica das variantes são apresentadas na primeira coluna da tabela, seguida do efeito molecular de cada variante (de sentido trocado, sinônima e intrônica). A predição de impacto foi desempenhada no Human Splicing Finder & Alternative Splice Site Predictor, que investigou alterações quanto aos sítios de *splicing* e elementos regulatórios de *splicing.* Do resultado, foi predito que apenas duas variantes de sentido trocado criaram novos sítio de *Splicing Acceptor;* uma variante intrônica criou um novo sítio *donor;* Seis variantes apresentaram potencial para alterações quanto aos elementos regulatórios de *splicing.* Para a avaliação da conservação evolutiva foi utilizado o *software* BLASTP, Clustal Omega. O resultado é empregado em: **pouco conservado,** quando os resíduos são muito variáveis em todas as sequências avaliadas; **conservados**, quando a variação de um resíduo para outro é baixa; e **altamente conservada** quando para os resíduos raramente são substituídos por outros aminoácidos. A avaliação do impacto funcional das variantes também foi desempenhada através dos *softwares* Mutation Taster, LIST-2, FATH-MM, Provean, PolyPhen, SIFT e UMD-predictor, que classificam as variantes quanto ao seu impacto na proteína, sendo sem efeitos significativos (Benigna - B) e com efeitos significativos (Patogênica - P). A estabilidade da estrutura da proteína RAD51D foi avaliada no mCSM, que demonstrou que a estrutura proteica se desestabiliza na maioria das variantes (↓ Des.); a estabilidade da afinidade de interação da proteína RAD51D com outras proteínas e com o DNA foram avaliadas no mCSMPPACUM e mCSMPDACUM, que variou entre estabilizante (↑ Est.) e desestabilizante (↓ Des.). Através do DynaMut e DIANNA, foram preditos os efeitos das variantes sobre as interações entre os resíduos de aminoácidos tais como pontes de hidrogênio, ligações iônicas, ligações hidrofóbicas e ligações de dissulfeto. Por fim, foi predito o efeito molecular de cada uma das variantes em relação aos domínios, motivos, sítios de interações e modificações pós-síntese. ACMG: American College of Medical Genetics and Genomics – ACMG; AMP: Association for Molecular Pathology dos EUA – AMP.

4. CONSIDERAÇÕES FINAIS

A triagem de mutações baseada no sequenciamento de painéis multigênicos ou de exoma é frequentemente recomendada para pacientes afetados e familiares ou com predisposição a câncer de mama e/ou ovário e outras neoplasias (Li *et al*., 2020). Este tipo de metodologias leva a uma grande quantidade de variantes, que precisam ser classificadas, quanto a seu potencial patogenicidade, com vistas a orientar as condutas médicas. Na prática clínica, as variantes patogênicas ou provavelmente patogênicas fomentam ações automáticas como recomendações de cirurgias profiláticas, quimioterápicos, entre outras. Entretanto, nem todas as variantes patogênicas deveriam ser tratadas da mesma forma, considerando a penetrância dos genes (alta, moderada e baixa) e o impacto funcional específico de cada variante no nível molecular (Katsoulakis *et al*., 2020, Graffeo *et al*., 2022).

Em função disto, são usados internacionalmente os critérios de classificação propostos pelo ACMG/AMP. Estes critérios não são específicos para câncer. O exoma é usado no apoio ao diagnóstico, fazendo parte da investigação, por exemplo, da criança polimalformada, sem hipótese clínica. Além disso, entre as doenças raras, existe um grande número de doenças órfãs, cuja base genética é desconhecida. Por outro lado, há um grande número de genes órfãos. Em função disto, os critérios de classificação de variantes genéticas precisam ser muito rigorosos.

A classificação das variantes é baseada em critérios, que consideram dados clínicos e história familiar, frequência populacional da variante, conservação evolutiva do resíduo, tipo de variante e impacto funcional predito por Bioinformática e artigos, descrevendo estudos funcionais ou a presença da variante em outros afetados. Entretanto, o potencial de patogenicidade ainda é influenciado pela qualidade da informação associando o gene, onde ocorreu a variante, à doença.

No câncer, tem-se uma quantidade relativamente grande de informações sobre genes que contribuem para o desenvolvimento da doença, assim como a sua penetrância em relação a diferentes fenótipos. Ainda assim, o número de variantes genéticas classificadas como VUS é grande.

A quantidade de VUS gerada é alta mesmo em genes de alta penetrância, e.g., cerca de 30% para *BRCA1* e 50% para *BRCA2* (Chiang *et al*., 2021; Gervas e*t al*., 2022). O *RAD51D* é um gene de penetrância moderada, o que acrescenta uma dificuldade adicional à classificação das

variantes. Não surpreendentemente, a maior parte das variantes em *RAD51D* são classificadas como VUS.

Um dos critérios importantes na classificação de uma variante é o PP2 (PP2 - Variante *missense* em um GENE, que tem uma baixa taxa de variação *missense* benigna e na qual variantes *missense* são um mecanismo comum de doença). O *RAD51D* possui 1307 variantes descritas no ClinVar, incluindo sinônimas, não sinônimas, intrônicas e truncantes (acesso em novembro 2022). Destas 1307, 749 são VUS, das quais 553 são variantes não sinônimas. São descritas 388 variantes benignas ou provavelmente benignas para *RAD51D*, mas a maioria delas são intrônicas ou sinônimas e apenas 12 são não-sinônimas são classificadas como benignas. Apenas 170 variantes foram classificadas como patogênicas/provavelmente patogênicas, refletindo em parte as dificuldades inerentes ao processo de classificação. Em função disto, no presente estudo, consideramos que este critério se aplica ao gene *RAD51D*, porque ele está claramente relacionado ao fenótipo. Entretanto, essa proposta não está descrita na literatura.

Muitas variantes classificadas como VUS coloca um problema adicional. Às vezes não fornecem informações acionáveis. Por exemplo, intervenções cirúrgicas redutoras de risco não podem ser recomendadas em função da presença de VUS porque, após reanálise, elas poderiam ser consideradas benignas, tornando a intervenção clínica inadequada e irreversível (Loong *et al*., 2022).

Geralmente, estudos aprofundados, como a avaliação da co-segregação da variante com a expressão da doença em famílias, também não são recomendados para pacientes portadores de VUS. Após um esforço de classificação aprofundado, usando-se ferramentas de Bioinformática uma VUS por ter baixo nível de suspeição, ou seja, não se encontra qualquer evidência de patogenicidade. Por outro lado, algumas VUS são suspeitas, porque são encontradas evidências de patogenicidade na análise Bioinformática, mas não há outras evidências de suporte. Estas evidências de suporte podem vir da história familiar ou de estudos funcionais, embora, formalmente, a presença de uma VUS não embase estes procedimentos. Isto ressalta como é importante a coleta de uma boa história familiar, porque a presença de recorrência e, se positiva, a análise de segregação da variante realizada antes da sua classificação pode mudar esta classificação. Contudo, uma história familiar com diversos afetados não é tão frequente e, consequentemente, este critério na maioria das vezes não estará disponível. Neste contexto, é importante salientar no laudo que a variante não justifica

procedimentos redutores de risco ou mesmo triagem em cascata. Entretanto, nos casos em que a variante seja uma VUS “suspeita”, seria importante avaliar se a variante é expressa em nível de RNA ou avaliar seu efeito em estudos funcionais, que normalmente não são realizados em função dos custos e da laboriosidade. Neste contexto, é importante que a interpretação dos testes genéticos seja meticulosa e que a classificação seja precisa e oportuna, pois os médicos contam com esse resultado para modular o tratamento, o prognóstico, aconselhamento genético e acompanhamento das famílias (Mersch *et al*., 2018; Graffeo *et al*., 2022).

Cientistas clínicos e geneticistas médicos, como Nykamp e colaboradores (2017) e o consórcio Clinical Genome Resource (ClinGen), estão constantemente revisando os critérios de classificação para variantes problemáticas, especialmente em genes de suscetibilidade ao câncer, estabelecendo novas diretrizes e ferramentas de classificação, que podem reduzir a superestimação da patogenicidade de variantes e vice-versa. Contudo, o refinamento dos critérios ainda é uma necessidade na prática clínica para o desenvolvimento de orientações sobre como avaliar variantes genéticas em genes de penetrância moderada e baixa, bem como orientações específicas da doença, para aumentar a precisão e a consistência da interpretação das variantes genéticas (Amendola *et al*., 2016; Loong *et al.*, 2022).

Com base nos resultados aqui apresentados, poucas variantes seriam reclassificadas. Isso porque, apesar da investigação *in silico* dos efeitos deletérios ou toleráveis ser fundamental no estudo de cada variante, é preciso considerar dados fenotípicos, populacionais e estudos funcionais (se disponíveis) para cada variante para tirar conclusões válidas e reclassificar. Contudo, estudos funcionais *in vivo* ou *in vitro* ainda são muito limitados. Por exemplo, no presente estudo, das 20 variantes, apenas duas (p.Ser46Cys e p.Glu233Gly) tiveram estudos funcionais bem estabelecidos, avaliando a eficiência do reparo da via de recombinação homóloga. Células expressando a variante (p.Ser46Cys) apresentaram instabilidade, prejudicando a recombinação homóloga em resposta ao aumento das quebras de fita dupla do DNA. Isto afeta a sensibilidade celular ao olaparibe e ao talazoparibe, conforme previsto nas análises *in silico*, permitindo-nos aplicar o critério PS3 como suporte de patogenicidade para esta variante (Alenezi *et al.,* 2022). Por outro lado, o estudo funcional da variante (p.Glu233Gly) relatou efeitos benéficos com melhora significativa no funcionamento da via de reparo por recombinação homóloga. Entretanto, os testes *in vitro* mostraram aumento da resistência a quimioterápicos, que atuam causando aumento da frequência de quebras no DNA. A variante é um exemplo de como a classificação de variantes pode ser complexa. No nível molecular, ela acarretou melhora funcional, no nível individual, causando resistência à

quimioterapia. Atualmente, esta variante é classificada como benigna, entretanto, ela deveria ser informada (Nadkarni *et al*., 2009).

Além disso, os dados clínicos como história familiar, dados fenotípicos e dados de segregação da variante com a doença, em grande parte ausentes, são limitações recorrentes na prática clínica. Isto afetou a classificação das variantes na presente amostra, mas também em bases de dados. Isso provoca um aumento do número de VUS, pois a maioria dos critérios de classificação do ACMG/AMP depende dos dados clínicos para sustentar as evidências de patogenicidade ou benignidade encontradas nos estudos *in silico* (Richards *et al*., 2015).

O estudo *in silico* aprofundado deve incluir uma predição das evidências de patogenicidade ao nível de DNA, RNA e proteína. Mesmo com análises *in silico* detalhadas, como as desenvolvidos no presente estudo, o uso da informação obtida sobre a potencial patogenicidade ou benignidade de uma variante genética fica limitado aos critérios ACMG/AMP (BP4 ou PP3), que só podem ser aplicados quando todos os algoritmos classificatórios apresentarem resultados semelhantes, ou seja, sugerirem benignidade ou patogenicidade, respectivamente. Contudo, há oposição a essa regra, pois a maioria dos algoritmos *in silico,* que classificam a substituição de aminoácidos como tolerável ou deletéria, consideram basicamente a conservação evolutiva do resíduo, a frequência populacional da variante e a localização em regiões funcionais já descritas em bancos de dados. Além destes, é fundamental considerar outros aspectos, como: 1. os efeitos nos sítios de *splicing* (criando ou abolindo sítios aceptores e doadores) e nas sequências auxiliares reguladoras do *splicing* (acentuadores e silenciadores); 2. interações inter- ou intramoleculares das quais resíduos de aminoácido participam; 3. ativação ou perda de novos domínios curtos críticos para a função da proteína, entre outros. Contudo, alterações em sítios de *splicing* (em posições menos conservadas dos sítios doador, aceptor e de ramificação) ou sequências auxiliares do *splicing* podem ser pouco informativas, quando não há estudo funcional. Da mesma forma a predição de localização em regiões funcionais (domínios e *motifs*), que denota possível impacto da variante na função da proteína, entretanto, na maioria das vezes não é possível afirmar-se que este efeito ocorra *in vivo*. De maneira alternativa, alguns laboratórios estão associando o sequenciamento do RNA, para confirmar se o alelo é de fato expresso. Neste contexto, a predição de um efeito sobre *splicing* ou regulação de *splicing* poderia ser

usado como mecanismo de triagem para indicar para quais variantes o sequenciamento do RNA deveria ser feito.

Em nossos resultados, a maioria das variantes foram preditas como criando ou abolindo sítios de motivos funcionais (domínios curtos), que são importantes para a interação transitória entre RAD51D e outras moléculas, sugerindo a possibilidade de impactos na formação de complexos e/ou na atividade da RAD51D em si. Tipicamente, esse achado não foi avaliado pelos algoritmos *in silico* que consideraram a substituição de aminoácidos como não deletéria, embora sugestiva de patogenicidade, como foi o caso da variante (p.Ala210Thr). Pouco se sabe sobre as repercussões funcionais desses motivos, pois não foram encontrados estudos funcionais na literatura, descrevendo a função destes motivos na RAD51D.

Semelhantemente, nenhuma das variantes apresentou mudanças significativas em relação a sua estrutura 3D, embora efeitos em interações intermoleculares e ligações covalentes tenham sido previstos. Este achado reflete algumas limitações que afetam este estudo: 1. ausência da estrutura tridimensional completa da proteína de RAD51D e outras parálogas por ressonância nuclear magnética e cristalografia estão ausentes nos bancos de dados. As estruturas tridimensionais obtidas por métodos *in silico*, foram geradas em estudos investigando um grande número de proteínas simultaneamente, podendo não capturar informações relativas à pontes de dissulfeto, associação com grupos prostéticos, íons, e modificações pós-síntese em geral; 2. conhece-se alguns dos complexos que a RAD51D íntegra, mas não se sabe o efeito da modificação de regiões específicas sobre a entrada da RAD51D nestes complexos. 3. Preditores de impacto de variantes não usam informação relativa às características da estrutura 3D da proteína, como a formação de estruturas específicas, em função da presença de pontes de dissulfeto, por exemplo. Além de tudo, a conservação evolutiva deve ser avaliada adequadamente. A existência de isoformas de uma proteína em bancos de dados coloca uma dificuldade adicional na avaliação da conservação evolutiva de cada resíduo. Em nossa revisão da literatura e das bases de dados, encontramos que a RAD51D tem 8 isoformas produzidas por *splicing* alternativo. A mistura delas no alinhamento múltiplo, pode fazer com que determinados resíduos sejam classificados como pouco conservados, quando na verdade são conservados. Para resolver este problema, coletamos nas bases de dados as sequências disponíveis para a RAD51D, separamos aquelas que correspondem à isoforma canônica ou a isoforma 8, e conduzimos a avaliação da conservação de cada resíduo, separadamente, conforme a isoforma.

Outra questão é a forma como a conservação evolutiva é avaliada por preditores *in silico,* que consideram todas as sequências disponíveis, independentemente da distância evolutiva e da conservação evolutiva no entorno do resíduo. Um exemplo claro deste problema apareceu na avaliação da conservação evolutiva do resíduo Cys119. Este resíduo é conservado em mamíferos, sendo substituído por outro resíduo apenas em espécies basais de mamíferos e em cetáceos. Ou seja, uma cisteína ocupa esta posição há cerca de 40 a 60 milhões de anos.

Da mesma forma, durante a análise da conservação evolutiva do resíduo Arg165, que aparentemente possui uma glutamina como resíduo fundador (a maioria das espécies possui este resíduo nesta posição) e, quando substituída, possui uma arginina nesta posição, sugerindo uma evolução convergente. Por isso, a Arg165 foi considerada em outros estudos como não conservada.

Em síntese, se esses detalhes não fossem considerados, os resíduos Cys119 e Arg165 seriam classificados como "não conservados", o que implicaria na aplicação de um critério de benignidade (por exemplo, BP4). Diante disso, a conservação evolutiva do resíduo é uma das análises fundamentais no estudo das variantes e, por si só, implica no nível de patogenicidade estimado. Por exemplo, resíduos altamente conservados costumam ter atividade bioquímica específica para a função da proteína (e.g., interações inter- e intramoleculares), onde poucos ou nenhum outro aminoácido caberiam na posição dada. Ainda que não se consiga estabelecer claramente porque um resíduo é fundamental para a atividade da proteína, a profundidade da sua conservação evolutiva, assim como do seu entorno, são indicativos muito fortes de relevância funcional. Considerando a consequência bioquímica da substituição de aminoácidos, a conservação evolutiva é uma das principais evidências na investigação da patogenicidade e, portanto, deveria ser um critério à parte para classificação de variantes.

Finalmente, conclui-se que a ausência das evidências de suporte, como os dados clínicos, resultou na manutenção da classificação de VUS para as variantes: c.196G>A; (p.Val66Met), c.355T>C; (p.Cys119Arg), c.413A>G; (p.Asn138Ser), c.628G>A; (p.Ala210Thr), c.695G>A; (p.Arg232Gln), c.698A>G; (p.Glu233Gly), c.715C>T; (p.Arg239Trp), c.862G>C; p.Gly288Arg), c.919G>A; (p.Glu307Lys) e c.973G>A; (p.Gly325Ser). As análises *in silico* sugerem que a variante (p.Ala210Thr) tenha um potencial patogênico. Entretanto, as evidências de patogenicidade isoladas não são suficientes para se reclassificar uma variante de VUS para provavelmente patogênica.

Baseando-se em evidências de suporte como estudos caso-controle, estudos funcionais ou dados relevantes disponíveis na literatura ou em bases de dados públicas, este estudo propõe a reclassificação das variantes *RAD51D* c.26G>C; (p.Cys9Ser), *RAD51D* c.137C>G; (p.Ser46Cys), *RAD51D* c.140A>G; (p.Tyr47Cys) e *RAD51D* c.493C>T; (p.Arg165Trp), como provavelmente patogênica, patogênica, provavelmente patogênica e provavelmente patogênica, respectivamente. Isso reforça a necessidade de estabelecer colaboração clínico-laboratorial para produzir estruturas de consenso na interpretação de variantes por meio da análise simultânea de dados fenotípicos e genotípicos, reduzindo a incerteza, ampliando a aplicabilidade do teste genético e garantindo que as informações obtidas sejam úteis tanto para os médicos assistentes quanto para os indivíduos portadores dessas variantes.

As análises *in silico* são ferramentas competentes para serem aplicadas no estudo inicial de variantes genéticas, principalmente pelo menor custo. Entretanto, os resultados *in silico* devem ser avaliados de forma combinada ao interpretar-se uma variante, e não isolada. Apesar da investigação *in silico* dos efeitos deletérios ou toleráveis ser fundamental no estudo de variantes, essas evidências não descartam a necessidade de estudos funcionais, mas auxiliam na escolha de quais variantes deveriam ser encaminhadas para estudos *in vitro*.

Os critérios do ACMG/AMP são utilizados para interpretar a patogenicidade de variantes genéticas e buscam minimizar a possibilidade de uma classificação incorreta, que poderia levar a condutas definitivas inadequadas. Contudo, a interpretação final e a compreensão dos critérios de classificação ficam a cargo de cada laboratório. Isto pode aumentar as inconsistências na classificação de variantes, pois a maioria dos laboratórios não têm acesso a dados fundamentais para a classificação como a recorrência familiar, análise de segregação nas famílias ou evidência de tumores (número, tipo, recorrência). Portanto, são poucos os laboratórios, que usam os critérios exatamente como descritos pelo ACMG/AMP (Amendola *et al*., 2016; Niehaus *et al*., 2019). A análise desenvolvida no presente estudo, mais detalhada do que a usual, trouxe uma melhor compreensão das limitações encontradas no processo de classificação de variantes assim como dos aspectos que poderiam ser melhorados, particularmente na interação com a clínica e com laboratórios de pesquisa básica.

6. REFERÊNCIAS

Adzhubei, I.A., Schmidt, S., Peshkin, L. et al. A method and server for predicting damaging missense mutations. **Nat Methods**, v.7, p. 248-249, 2010. https://doi.org/10.1038/nmeth0410-248.

Ahmad, K.A. et al. Protein kinase CK2--a key suppressor of apoptosis. **Adv Enzyme Regul**, v.48, p.179-87, 2008. https://doi.org/10.1016/j.advenzreg.2008.04.002.

Aiad, H.A., Samaka, R.M., Asaad, N.Y. et al. Relationship of CK8/18 expression pattern to breast cancer immunohistochemical subtyping in Egyptian patients. **Ecancermedicalscience**, v.8, p.404, 2014. https://doi.org/10.3332/ecancer.2014.404.

Alenezi, W.M., Milano, L., Fierheller, C.T. et al. The Genetic and Molecular Analyses of RAD51C and RAD51D Identifies Rare Variants Implicated in Hereditary Ovarian Cancer from a Genetically Unique Population. **Cancers (Basel)**, v.14, p.2251, 2022. https://doi.org/10.3390/cancers14092251.

Altschul, S.F. et al. Basic local alignment search tool. **J Mol Biol**, n. 215, p. 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Amendola, L.M., Jarvik, G.P., Leo, M.C. et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. **Am J Hum Genet**. v. 2, n. 98, p. 1067-1076, 2016. https://doi.org/10.1016/j.ajhg.2016.03.024.

American Cancer Society**. Ovarian Cancer Risk Factors**. <https://www.cancer.org/cancer/ovarian-cancer/causes-risks-prevention/risk-factors.html#:~:text=The%20lifetime%20ovarian%20cancer%20risk,and%2030%25%20by%20age%2070>. Acesso em: 10/2022

Amerik, A.Y., Hochstrasser, M. Mechanism and function of deubiquitinating enzymes. **Biochim Biophys Acta**. v.1695, p.189-207, 2004. https://doi.org/10.1016/j.bbamcr.2004.10.003.

Andres,S.N., Vergnes, A., Ristic, D. et al. A human XRCC4-XLF complex bridges DNA. **Nucleic Acids Res**, v.40, p.1868-78, 2012. https://doi.org/10.1093/nar/gks022.

Apweiler, R., Attwood, T. K., Bairoch, A. et al. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. **Nucleic Acids Res**, n. 29, p. 37-40, 2001. https://doi.org/10.1093/nar/29.1.37.

Bains, S.B.S., Dotzler, S.M.B.A., Krijger, C.B.S. et al. A phenotype-enhanced variant classification framework to decrease the burden of missense variants of uncertain significance in type 1 long QT syndrome. **Heart Rhythm**, v.19, p.435-442, 2022. https://doi.org/10.1016/j.hrthm.2021.11.017.

Baker, J.L. et al. Breast cancer in a RAD51D mutation carrier: case report and review of the literature. **Clin Breast Cancer**, n. 15, p. e71-5, 2015. https://doi.org/10.1016/j.clbc.2014.08.005.

Bardwell, L. Mechanisms of MAPK signalling specificity. **Biochem Soc Trans**, v.34, p.837-41, 2006. https://doi.org/10.1042/BST0340837.

Berliner, J.L. et al. Risk assessment and genetic counseling for hereditary breast and ovarian cancer syndromes—Practice resource of the National Society of Genetic Counselors. **J Genet Couns**, 2021. https://doi.org/10.1002/jgc4.1374.

Berti, M., Teloni, F., Mijic, S. et al. Sequential role of RAD51 paralog complexes in replication fork remodeling and restart. **Nat Commun**, n. 11, p. 3531, 2020. https://doi.org/10.1038/s41467-020-17324-z.

Biondi, R.M., Nebreda, AR. Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions**. Biochem J**, v.15, p.1-13, 2003. https://doi.org/10.1042/BJ20021641.

Boasquevisque, P.C.R. Classificação do grau histológico de malignidade em câncer de mama utilizando programas de análise automatizada de imagens e aprendizado de máquina. **COLLAB**, Ciências Farmacêuticas, 2020. https://repositorio.uvv.br//handle/123456789/543.

Bodelon, C., Killian, J.K., Sampson, J.N. et al. Molecular Classification of Epithelial Ovarian Cancer Based on Methylation Profiling: Evidence for Survival Heterogeneity. **Clin Cancer Res**, v.25, p.5937-5946, 2019. https://doi.org/10.1158/1078-0432.CCR-18-3720.

Boll, W., Rapoport, I., Brunner, C. et al. The mu2 subunit of the clathrin adaptor AP-2 binds to FDNPVY and YppØ sorting signals at distinct sites. **Traffic**, v.3, p.590-600, 2002. https://doi.org/10.1034/j.1600-0854.2002.30808.x.

Boulaflous, A., Saint-Jore-Dupas, C., Herranz-Gordo, M.C. et al. Cytosolic N-terminal arginine-based signals together with a luminal signal target a type II membrane protein to the plant ER. **BMC Plant Biol**, v. 8, p.144, 2009. https://doi.org/10.1186/1471-2229-9-144.

Bray, F., Ferlay, J., Soerjomataram, I. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA Cancer J Clin**, n. 68, p. 394-424, 2018. https://doi.org/10.3322/caac.21492.

Brown, J.S., Jackson, S.P. Ubiquitylation, neddylation and the DNA damage response. **Open Biol,** v. 5, p. 150018, 2015. https://doi.org/10.1098/rsob.150018.

Buchan, D. W. A., Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. **Nucleic Acids Res**, 2019. https://doi.org/10.1093/nar/gkz297.

Caracciolo, D., Riillo, C., Di Martino M.T. et al. Alternative Non-Homologous End-Joining: Error-Prone DNA Repair as Cancer's Achilles' Heel. **Cancers (Basel)**, v. 13, p. 1392, 2021. https://doi.org/10.3390/cancers13061392.

Carmo, G.B. Estudos de genes e variantes genéticas associadas ao câncer de mama familial: impactos no aconselhamento genético. **Universidade de São Paulo**, 2018.

Cartwright, R. et al Isolation of novel human and mouse genes of the recA/RAD51 recombination-repair gene family. **Nucleic Acids Res**, v.1, p.1653-9, 1998. https://doi.org/10.1093/nar/26.7.1653.

Castro, A. et al. The anaphase-promoting complex: a key factor in the regulation of cell cycle. **Oncogene**, v. 24, p. 314-25, 2005. https://doi.org/10.1038/sj.onc.1207973.

Castro, B.G.R., Souza, C. P., Andrade, C.E.M.C. et al. Tumores de células de Sertoli-Leydig ovarianos: fatores epidemiológicos, clínicos e prognósticos. **Rev Bras Ginecol Obstet**, n. 41, 2019. https://doi.org/10.1055/s-0039-1693056.

Catalano, F., Borea, R., Puglisi, S. et al. Targeting the DNA Damage Response Pathway as a Novel Therapeutic Strategy in Colorectal Cancer. **Cancers**, v.14, p.1388, 2022. https://doi.org/10.3390/cancers14061388

Cheon, K.W., Baek, K.H. HAUSP as a therapeutic target for hematopoietic tumors (review). **Int J Oncol**, v. 28, p. 1209-15, 2006.

Chiang, J., Chia, T.H., Yuen, J. et al. Impact of Variant Reclassification in Cancer Predisposition Genes on Clinical Care. **JCO Precis Oncol**, n. 5, p. 577-584, 2021. https://doi.org/10.1200/PO.20.00399.

Choi, Y. et al. Predicting the Functional Effect of Amino Acid Substitutions and Indels. **PLoS ONE**, v. 7, p. e46688, 2012. https://doi.org/10.1371/journal.pone.0046688.

Clapperton, J.A., Manke, I.A., Lowery, D.M. et al. Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. **Nat Struct Mol Biol**, n. 11, p. 512-8, 2004. https://doi.org/10.1038/nsmb775.

Consortium U (2019) UniProt: a worldwide hub of protein knowledge. **Nucleic acids Res,** n. 47(D1), p. D506-D515, 2022.

Dai, X., Li, T., Bai, Z. et al. Breast cancer intrinsic subtype classification, clinical use and future trends. **Am J Cancer Res**, v.5(10), p.2929-43, 2015.

David, A. et al. The AlphaFold Database of Protein Structures: A Biologist's Guide. **Journal of Molecular Biology**, v.434, p.167336, 2022. https://doi.org/10.1016/j.jmb.2021.167336.

De Castro, E., Sigrist, C.J.A., Gattiker, A. et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. **Nucleic Acids Res**, v. 1, p. W362-5, 2006.

Desmet, F. O. et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. **Nucleic Acids Res**, n. 37, p. e67, 2009. https://doi.org/10.1093/nar/gkp215.

Dieci, M.V. et al. Rare breast cancer subtypes: histological, molecular, and clinical peculiarities. **Oncologist**, n. 19, p. 805-13, 2014. https://doi.org/10.1634/theoncologist.2014-0108.

Dominguez-Valentin, M., Evans, D.G.R., Nakken, S. et al. Genetic variants of prospectively demonstrated phenocopies in BRCA1/2 kindreds. **Hered Cancer Clin Pract**, v. 15, n. 16, p. 4, 2018. https://doi.org/10.1186/s13053-018-0086-0.

Dowty, J.G., Lose, F., Jenkins, M.A. et al. The RAD51D E233G variant and breast cancer risk: population-based and clinic-based family studies of Australian women. **Breast Cancer Res Treat**, v.112, p.35-9, 2008. https://doi.org/10.1007/s10549-007-9832-0.

Durocher, D., Jackson, S.P. The FHA domain. **FEBS Lett**, v. 20, n. 513, p. 58-66 2002. https://doi.org/10.1016/s0014-5793(01)03294-x.

Elshami, M., Tuffaha, A., Yaseen, A. et al. Awareness of ovarian cancer risk and protective factors: A national cross-sectional study from Palestine. **PLOS ONE**, v.17, p.e0265452, 2022. https://doi.org/10.1371/journal.pone.0265452.

ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. **Nature**, v. 6, n. 489 (7414), p. 57-74, 2012. https://doi.org/10.1038/nature11247.

Esteves, F.P., Alves, T.A.F.A., Amorim, B.J. et al. Câncer de ovário recidivado. **Sociedade brasileira de medicina nuclear**, 2018. Disponível <https://amb.org.br/wp-content/uploads/2021/09/CANCER-DE-OVARIO-RECIDIVADO-PET_CT-COM-FDG-18F-FINAL-2018.pdf>. Acesso em: 10/2022.

Exome Variant Server. NHLBI GO Exome Sequencing Project (ESP), **Seattle**, WA, 2021. (http://evs.gs.washington.edu/EVS/).

Fairley, S. et al. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. **Nucleic Acids Res**, n. 48D1, p. D941–D947, 2021. https://doi.org/10.1093/nar/gkz836.

Falck, J. et al. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. **Nature**, v. 31, n. 434, p. 605-11, 2005. https://doi.org/10.1038/nature03442.

Fantz, D.A. et al. Docking sites on substrate proteins direct extracellular signal-regulated kinase to phosphorylate specific residues. **J Biol Chem**, v. 20, n. 276, p. 27256-65, 2001. https://doi.org/10.1074/jbc.M102512200.

Fell, V.L., Schild-Poulter, C. Ku regulates signaling to DNA damage response pathways through the Ku70 von Willebrand A domain. **Mol Cell Biol**, n. 32, p. 76-87, 2012. https://doi.org/10.1128/MCB.05661-11.

Ferlay, J., Colombet, M., Soerjomataram, I. et al. Cancer statistics for the year 2020: An overview. **Int J Cancer**, 2021. https://doi.org/10.1002/ijc.33588

Figueiredo, M.C.P. Câncer de mama hereditário: rastreamento de mutações nos genes BRCA1 e BRCA2 e busca de novos genes de susceptibilidade. **Fundação Antônio Prudente**, 2014. Disponível <https://accamargo.phlnet.com.br/Doutorado/2014/MarciaCPFigueiredo/MarciaCPFigueiredo.pdf>. Acesso em: 10/2022

Fry, A.M. The Nek2 protein kinase: a novel regulator of centrosome structure. **Oncogene**, v. 9, n. 21, p. 6184-94, 2002. https://doi.org/10.1038/sj.onc.1205711.

Fridman, A.L., Raz, A., Timmermans, S. et al. Fragmented responsibility: views of Israeli HCPs regarding patient recontact following variant reclassification. **J Community Genet**, v. 13, p. 13–18, 2022. https://doi.org/10.1007/s12687-021-00556-x.

Friedenreich, C.M. et al. Physical activity, obesity and sedentary behavior in cancer etiology: epidemiologic evidence and biologic mechanisms. **Mol Oncol**, n. 15, p. 790-800, 2021. https://doi.org/10.1002/1878-0261.12772.

Frimer, M., Levano, K.S., Rodriguez-Gabin, A. et al. Germline mutations of the DNA repair pathways in uterine serous carcinoma. **Gynecol Oncol**, n. 141, p. 101-7, 2016. https://doi.org/10.1016/j.ygyno.2015.12.034.

Gandhi, T., Bhatt. Cancer Antigen 125. **StatPearls [Internet]**, 2022. Available from: https://www.ncbi.nlm.nih.gov/books/NBK562245/

Garcin EB, Gon S, Sullivan MR. et al. Differential Requirements for the RAD51 Paralogs in Genome Repair and Maintenance in Human Cells. **PLoS Genet**, n. 4, v. 15, p. e1008355, 2019. https://doi.org/0.1371/journal.pgen.1008355.

Garrido-Castro, A.C. et al. Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. **Cancer Discov**, n. 9, p. 176-198, 2019. https://doi.org/10.1158/2159-8290.CD-18-1177.

Gervas, P., Molokov, A., Babyshkina, N. et al. Pathogenicity Reclassification of Genetic Variants Related to Early-Onset Breast Cancer among Women of Mongoloid Origin. **Asian Pac J Cancer Prev,** v. 1, n. 23, p. 2027-2033, 2022. https://doi.org/10.31557/APJCP.2022.23.6.2027.

GLOBOCAN**. International Agency for Research on Cance**r, 2020. Disponível <https://gco.iarc.fr/today>. Acesso em: 20/10/2022.

Glover, J.N. et al. Interactions between BRCT repeats and phosphoproteins: tangled up in two. **Trends Biochem Sci**, n. 29, p. 579-85, 2004. https://doi.org/10.1016/j.tibs.2004.09.010.

González-Rivera, M., Lobo, M., López-Tarruella, S. et al. Frequency of germline DNA genetic findings in an unselected prospective cohort of triple-negative breast cancer patients participating in a platinum-based neoadjuvant chemotherapy trial. **Breast Cancer Res Treat**, n. 156, p. 507-515, 2016. https://doi.org/10.1007/s10549-016-3792-1.

Graffeo, R., Rana, H.Q., Conforti, F. et al. Moderate penetrance genes complicate genetic testing for breast cancer diagnosis: ATM, CHEK2, BARD1 and RAD51D. **The Breast**, v.65, p.32-40, 2022. https://doi.org/10.1016/j.breast.2022.06.003.

GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans**. Science**, n. 348, p. 648-660, 2015. https://doi.org/10.1126/science.1262110.

Guerrero, V.G. et al. Modifiable risk factors for breast cancer: an obligation for health professionals. **Rev Panam Salud Publica**, n. 41, p. e80, 2017. https://iris.paho.org/handle/10665.2/34054?locale-attribute=pt.

Gutiérrez-Enríquez, S., Bonache, S,. Garibay, G.R. et al. About 1% of the breast and ovarian Spanish families testing negative for BRCA1 and BRCA2 are carriers of RAD51D pathogenic variants. **Cancer Genetics**, v. 134, p. 2088-2097, 2013. https://doi.org/10.1002/ijc.28540.

Hanahan, D., Weinberg, R.A. The hallmarks of cancer. **Cell**, v.100, p.57-70, 2000. https://doi.org/10.1016/s0092-8674(00)81683-9.

Hanahan, D., Weinberg, R.A. Hallmarks of cancer: the next generation. **Cell**, v.144, p.646-74, 2011. https://doi.org/10.1016/j.cell.2011.02.013.

Hashmi, A.A., Aijaz, S., Khan, S.M. et al. Prognostic parameters of luminal A and luminal B intrinsic breast cancer subtypes of Pakistani patients. **World J Surg Onc**, v. 16, p. 1, 2018. https://doi.org/10.1186/s12957-017-1299-9

Hauke, J., Horvath, J., Groß E. et al. Gene panel testing of 5589 BRCA1/2-negative index patients with breast cancer in a routine diagnostic setting: results of the German Consortium for Hereditary Breast and Ovarian Cancer. **Cancer Med**, v. 7, p. 1349-1358, 2018. https://doi.org/10.1002/cam4.1376.

Hecker, C.M. et al. Specification of SUMO1- and SUMO2-interacting motifs. **J Biol Chem**, v. 9, n. 281, p. 16117-27, 2006. https://doi.org/10.1074/jbc.M512757200.

Heery, D.M. et al. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. **Nature,** v. 387, p. 733-6, 1997. https://doi.org/10.1038/42750.

Hertz, E.P.T., Kruse, T., Davey, N.E. et al. A Conserved Motif Provides Binding Specificity to the PP2A-B56 Phosphatase. **Mol Cell**, v. 63, p. 686-695, 2018. https://doi.org/10.1016/j.molcel.2016.06.024.

Hoffman-Andrews, L. The known unknown: the challenges of genetic variants of uncertain significance in clinical practice. **J Law Biosci**, v. 4, p. 648-657, 2018. https://doi.org/10.1093/jlb/lsx038.

Holowaty, M.N. et al. Protein interaction domains of the ubiquitin-specific protease, USP7/HAUSP. **J Biol Chem**, v. 28, p. 47753-61, 2003. https://doi.org/10.1074/jbc.M307200200.

Hornbeck, P.V., Kornhauser, J.M., Tkachev, S. et al. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. **Nucleic Acids Res**, v. 40, p. 261-70, 2012. https://doi.org/10.1093/nar/gkr1122.

Howe, K. L. et al. Ensembl. **Nucleic Acids Res,** n. 49(D1), p.D884-D891, 2021. https://doi.org/10.1093/nar/gkaa942.

Huang, D.T., Miller, D.W., Mathew, R. . A unique E1-E2 interaction required for optimal conjugation of the ubiquitin-like protein NEDD8. **Nat Struct Mol Biol**, n. 11, p. 927-35, 2004. https://doi.org/10.1038/nsmb826.

Huo, J., Chen, X., Zhang, H. et al. Bcl-3 promotes proliferation and chemosensitivity in BL1 subtype of TNBC cells. **Acta Biochim Biophys Sin**, v.50, p.1141–1149, 2018. https://doi.org/10.1093/abbs/gmy117

Inic, Z., Zegarac, M., Inic, M. et al. Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information. **Clin Med Insights Oncol,** v. 11, n. 8, p. 107-11, 2014. https://doi.org/10.4137/CMO.S18006.

Instituto Nacional Câncer. Estimativa. INCA, 2014. Disponível <http://www.inca.gov.br/estimativa/2014/>. Acesso em: 10/2022.

Instituto Nacional Câncer. Estimativa. **INCA,** 2020. Disponível <https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2020-incidencia-de-cancer-no-brasil.pdf>. Acesso em: 10/2022.

Instituto Nacional Câncer. Estimativas de câncer no Brasil. **INCA**, 2020. Disponível <http://www.oncoguia.org.br/conteudo/estimativas-no-brasil/1705/1/#:~:text=O%20Brasil%20dever%C3%A1%20registrar%20625,da%20doen%C3%A7a%20para%20cada%20ano.>. Acesso em: 10/2022.

Instituto Nacional Câncer. Estimativa. **INCA**, 2022. Disponível <https://www.inca.gov.br/estimativa/introducao>. Acesso em: 10/2022.

Janatova, M., Soukupova, J., Stribrna, J. et al. Mutation Analysis of the RAD51C and RAD51D Genes in High-Risk Ovarian Cancer Patients and Families from the Czech Republic. **PLoS One**, v. 9, n. 10, p. e0127711, 2015. https://doi.org/10.1371/journal.pone.0127711.

Jara, L., Dubois, K., Gaete, D. et al. Variants in DNA double-strand break repair genes and risk of familial breast cancer in a South American population. **Breast Cancer Res Treat**, v.122, p.813-22, 2010. https://doi.org/10.1007/s10549-009-0709-2.

Jeong, Y. et al. Characterization of NIP2/centrobin, a novel substrate of Nek2, and its potential role in microtubule stabilization. **J Cell Sci**, v. 15, n. 120, p. 2106-16, 2007. https://doi.org/10.1242/jcs.03458.

Jiraskova, K., Hughes, D.J., Brezina, S. et al. Functional Polymorphisms in DNA Repair Genes Are Associated with Sporadic Colorectal Cancer Susceptibility and Clinical Outcome. **Int J Mol Sci**, v.27, n. 20, p. 97, 2018. https://doi.org/10.3390/ijms20010097.

Johnson, A.M., Zuhlke, K.A., Plotts, C. et al. Mutational landscape of candidate genes in familial prostate cancer. **Prostate**. n. 74, p. 1371-8, 2014. https://doi.org/10.1002/pros.22849.

JosahKian, J.A. Avaliação da influência da expressão de STAT1 na respostas ao tratamento quimioterápico no câncer de ovário sseroso de alto grau. **Faculdade de Medicina de Ribeirão Preto**, 2016. Disponível

<https://www.teses.usp.br/teses/disponiveis/17/17135/tde-21072016-151516/publico/JULIANAALVESJOSAHKIANOrig.pdf>.

Kaku, T., Ogawa, S., Kawano, Y. et al. Histological classification of ovarian cancer. **Med Electron Microsc**, n. 36, p. 9-17, 2003. https://doi.org/10.1007/s007950300002.

Kaneko, T. et al. Phosphotyrosine recognition domains: the typical, the atypical and the versatile. **Cell Commun Signal**, v. 7, n. 10, p. 32, 2012. https://doi.org/10.1186/1478-811X-10-32. PMID: 23134684; PMCID: PMC3507883.

Karczewski, K. J., Francioli, L. C., Tiao, G. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. **Nature**, n. 581, p. 434–443, 2020. https://doi.org/10.1038/s41586-020-2308-7.

Kathawala, R.J. et al. The Chemoprevention of Ovarian Cancer: the Need and the Options. **Curr Pharmacol Rep**, n. 4, p. 250-260, 2018. https://doi.org/10.1007/s40495-018-0133-6.

Katsoulakis E. et al. Comparison of Annotation Services for Next-Generation Sequencing in a Large-Scale Precision Oncology Program**. JCO Precision Oncology**, n. 4, p. 212-221, 2020. https://doi.org/ 10.1200/PO.19.00118.

Kawabata, M., Saeki, K. Multiple Alternative Transcripts of the Human Homologue of the MouseTRAD/R51H3/RAD51DGene, a Member of therec A/RAD51Gene Family. **Biochem Biophys Res Commun**, v.257, p. 156-162, 1999. https://doi.org/10.1006/bbrc.1999.0413

Kchouk, M. et al. Generations of Sequencing Technologies: From First to Next Generation. Biol Med, n. 9, p. 3, 2017. https://doi.org/10.4172/0974-8369.1000395.

Kettenbach, A.N., Wang, T., Faherty, B.K. et al. Rapid determination of multiple linear kinase substrate motifs by mass spectrometry. **Chem Biol**, v. 25, n. 19, p. 608-18, 2012. https://doi.org/10.1016/j.chembiol.2012.04.011.

Kopanos, C., Tsiolkas, V., Kouris, A. et al. VarSome: the human genomic variant search engine. **Bioinformatics**, n. 35, p. 1978, 2019. https://doi.org/10.1093/bioinformatics/bty897.

Kraemer, D., Azzarello-Burri, S., Steindl, K. et al. Prevalence of genetic susceptibility for breast and ovarian cancer in a non-cancer related study population: secondary germline findings from a Swiss single centre cohort. **Swiss Med Wkly**, v. 18, n. 149, p. w20092, 2019. https://doi.org/10.4414/smw.2019.20092.

Kumar, K.R. et al. Next-Generation Sequencing and Emerging Technologies. **Thieme**, n. 45, p. 661-673, 2019. https://doi.org/10.1055/s-0039-1688446.

Kumar, M., Michael, S., Alvarado-Valverde, J. et al. The Eukaryotic Linear Motif resource: 2022 release. **Nucleic Acids Res,** v. 50, p. D497–D508, 2022. https://doi.org/10.1093/nar/gkab975.

Kurman, R.J., Shih, I.M. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. **Am J Surg Pathol**, n. 34, p. 433-43, 2010. https://doi.org/10.1097/PAS.0b013e3181cf3d79.

Landrum M. J., Chitipiralla S., Brown G. R. et al. ClinVar: improvements to accessing data. **Nucleic Acids Res**, n. 48(D1), p. D835-D844, 2020. https://doi.org/10.1093/nar/gkx1153.

Laskowski R.A. et al. PROCHECK - a program to check the stereochemical quality of protein structures. **J. App. Cryst**., v. 26, . 283-291, 1993.

Laskowski R. A., Thornton J M. PDBsum extras: SARS-CoV-2 and AlphaFold models. **Prot. Sci**., 31, 283-289, 2021.

Leo, C., Chen, J.D. The SRC family of nuclear receptor coactivators. **Gene**, v. 7, n. 245, p. 1-11, 2000. https://doi.org/10.1016/s0378-1119(00)00024-x.

Leung, G.C. et al. Determination of the Plk4/Sak consensus phosphorylation motif using peptide spots arrays. **FEBS Lett**, v. 9, n. 581, p. 77-83, 2007. https://doi.org/10.1016/j.febslet.2006.11.080.

Li, H., LaDuca, H., Pesaran, T.M.A. et al. Classification of variants of uncertain significance in BRCA1 and BRCA2 using personal and family history of cancer from individuals in a large hereditary cancer multigene panel testing cohort. **Genetics in Medicine**, v.22, p. 701-708, 2020. https://doi.org/10.1038/s41436-019-0729-1.

Li, J. et al. The FHA domain mediates phosphoprotein interactions. **J Cell Sci**, v. 113, n. 23, p. 143-9, 2000. https://doi.org/10.1242/jcs.113.23.4143. PMID: 11069759.

Lima, M.S., Siqueira, H.F.F., Moura, A.R., et al. Temporal trend of cancer mortality in a Brazilian state with a medium Human Development Index (1980–2018). **Sci Rep**, v.10, p.21384, 2020. https://doi.org/10.1038/s41598-020-78381-4.

Long, E.O. Regulation of immune responses through inhibitory receptors. **Annu Rev Immunol**, n. 17, p. 875-904, 1999. https://doi.org/10.1146/annurev.immunol.17.1.875.

Loong, L., Garrett, A., Allen, S. et al. Reclassification of clinically-detected sequence variants: Framework for genetic clinicians and clinical scientists by CanVIG-UK (Cancer Variant Interpretation Group UK). **Genetics in Medicine**, v.24, P.1867-1877, 2022. https://doi.org/10.1016/j.gim.2022.05.002

Loveday, C., Turnbull, C., Ramsay, E. et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. **Nat Genet,** v. 7, p. 879-882, 2011. https://doi.org/10.1038/ng.893.

Lu, P.J., Zhou, X.Z., Shen, M., Lu, K.P. Function of WW domains as phosphoserine- or phosphothreonine-binding modules. **Science**, v. 26, n. 283, p. 1325-8, 1999. doi: 10.1126/science.283.5406.1325.

Lucci-Cordisco, E., Amenta, S., Panfili, A. et al. Variants of uncertain significance (VUS) in cancer predisposing genes: What are we learning from multigene panels? **Eur J Hum Genet**, v.65, p.104400, 2022. https://doi.org/10.1016/j.ejmg.2021.104400.

Luz, A.A. da, et al. Alterações histopatológicas do câncer de mama. **Revista Multidisciplinar Em Saúde**, n. 2, p. 14, 2021. https://doi.org/10.51161/rems/2722.

Madeddu, C., Kotsonis, P., Lavra, F. et al. Next generation sequencing driven successful combined treatment with laparoscopic surgery and immunotherapy for relapsed stage IVB cervical and synchronous stage IV lung cancer. **Oncotarget**, v. 12, p. 2012-2021, 2019. https://doi.org/10.18632/oncotarget.26769.

Madeira, F., Pearce, M., Tivey, A.R.N. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. **Nucleic Acids Res**, v. 240, 2022. https://doi.org/10.1093/nar/gkac240.

Makki, J. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. **Clin Med Insights Pathol**, v.8, p.23-31, 2015. https://doi.org/10.4137/CPath.S31563.

Malhis, N. et al. . LIST-S2: taxonomy based sorting of deleterious missense mutations across species. **Nucleic Acids Res**, 2020. https://doi.org/10.1093/nar/gkaa288.

Mattera, R et al.. Conservation and diversification of dileucine signal recognition by adaptor protein (AP) complex variants. **J Biol Chem,** v. 21, n. 286, p. 2022-30, 2011. https://doi.org/ 10.1074/jbc.M110.197178.

Maxwell, K.N., Hart, S.N., Vijai, J. et al. Evaluation of ACMG-Guideline-Based Variant Classification of Cancer Susceptibility and Non-Cancer-Associated Genes in Families Affected by Breast Cancer. **Am J Hum Genet**, v. 98, p. 801-817, 2016. https://doi.org/10.1016/j.ajhg.2016.02.024.

McCright, B. et al. The B56 family of protein phosphatase 2A (PP2A) regulatory subunits encodes differentiation-induced phosphoproteins that target PP2A to both nucleus and cytoplasm. **J Biol Chem**, v. 6, n. 271, p. 22081-9, 1996. https://doi.org/10.1074/jbc.271.36.22081.

Meggio, F., Pinna, L.A. One-thousand-and-one substrates of protein kinase CK2? **FASEB J**, n. 17, p. 349-68, 2003. https://doi.org/10.1096/fj.02-0473rev.

Mersch, J., Brown, N., Pirzadeh-Miller, S. et al. Prevalence of Variant Reclassification Following Hereditary Cancer Genetic Testing. **JAMA**, n. 320, p. 1266–1274, 2018. https://doi.org/10.1001/jama.2018.13152.

Meyer, K., Kirchner, M., Uyar, B. et al. Mutations in Disordered Regions Can Cause Disease by Creating Dileucine Motifs. **Cell**, v. 20, n. 175, p. 239-253.e17, 2018. https://doi.org/10.1016/j.cell.2018.08.019.

Michelsen, K. et al. Hide and run. Arginine-based endoplasmic-reticulum-sorting motifs in the assembly of heteromultimeric membrane proteins. **EMBO Rep**, n. 6, p. 717-22, 2005. https://doi.org/10.1038/sj.embor.7400480.

Ministério da Saúde. A situação do câncer no Brasil. 2007. Disponível <https://bvsms.saude.gov.br/bvs/publicacoes/situacao_cancer_brasil.pdf>. Acesso em: 10/2022

Ministério da Saúde. Dia Nacional de Luta contra o Câncer de Mama. 2022. Disponível <https://bvsms.saude.gov.br/27-dia-nacional-de-luta-contra-o-cancer-de-mama/#:~:text=O%20c%C3%A2ncer%20de%20mama%20tamb%C3%A9m,ocorrem%20ap%C3%B3s%20os%2050%20anos>. Acesso em: 10/2022

Mirdita, M., Schütze, K., Moriwaki, Y. et al. ColabFold: Making protein folding accessible to all. **Nature Methods,** 2022

Moshous, D., Callebaut, I., de Chasseval, R. et al. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. **Cell**, v. 20, n. 105, p. 177-86, 2001. https://doi.org/10.1016/s0092-8674(01)00309-9.

Moslehi, R., Tsao, HS., Zeinomar, N. et al. Integrative genomic analysis implicates ERCC6 and its interaction with ERCC8 in susceptibility to breast cancer. **Sci Rep,** v. 10, p. 21276, 2020. https://doi.org/10.1038/s41598-020-77037-7

Muslin, A.J. et al. Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. **Cell**, v. 22, n. 84, p. 889-97, 1996. https://doi.org/10.1016/s0092-8674(00)81067-3.

NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. **Nucleic Acids Res**, n. 46(D1), D8-D13, 2018. https://doi.org/10.1093/nar/gkx1095.

Nadkarni, A., Furda, A., Rajesh, C. et al. Functional characterization of the RAD51D E233G genetic variant. **Pharmacogenet Genomics**, n. 19, p. 153-60, 2009. https://doi.org/10.1097/FPC.0b013e32831db2fd.

Naslavsky, M.S., Yamamoto, G.L., de Almeida, T.F., et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. **Human Mutation**, v. 38, p. 751– 763, 2017. https://doi.org/10.1002/humu.23220

Nathans, J.F. Cell cycle inertia underlies a bifurcation in cell fates after DNA damage. Science Advances, v.7, n. 3. https://doi.org/10.1126/sciadv.abe3882.

National Cancer Institute. What Is Cancer? Reuse of NCI Information, 2022. Disponível <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>.

Nicolas, E., Arora, S., Zhou, Y. et al. Systematic evaluation of underlying defects in DNA repair as an approach to case-only assessment of familial prostate cancer. **Oncotarget**, v. 24, n. 6, p. 39614-33, 2015. https://doi.org/10.18632/oncotarget.5554.

Niefind, K. et al. Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme. **EMBO J**, v. 1, n. 20, p. 5320-31, 2001. https://doi.org/10.1093/emboj/20.19.5320.

Niehaus, A., Azzariti, D.R., Harrison, S.M. et al. A survey assessing adoption of the ACMG-AMP guidelines for interpreting sequence variants and identification of areas for continued improvement. **Genet Med**, n. 21, p. 1699-1701, 2019. https://doi.org/10.1038/s41436-018-0432-7.

Nolte, R.T., Wisely, G.B., Westin, S. et al. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. **Nature,** v. 10, n. 395, p. 137-43, 1998. https://doi.org/10.1038/25931.

Nykamp, K., Anderson, M., Powers, M. et al. Sherloc: a comprehensive refinement of the ACMG–AMP variant classification criteria. Genet Med, n. 19, p. 1105–1117, 2017. https://doi.org/10.1038/gim.2017.37

Obsil, T, Obsilova, V. Structural basis of 14-3-3 protein functions. **Semin Cell Dev Biol**, n. 22, p. 663-72, 2011. https://doi.org/10.1016/j.semcdb.2011.09.001.

Odho, Z. et al. Characterization of a novel WDR5-binding site that recruits RbBP5 through a conserved motif to enhance methylation of histone H3 lysine 4 by mixed lineage leukemia protein-1. **J Biol Chem**, v. 22, n. 285, p. 32967-32976, 2010. https://doi.org/10.1074/jbc.M110.159921.

Oliveira, A.C. Influência do gene PTEN na expressão de RAD51 e suas parálogas RAD51C e RAD51B em linhagens de glioblastoma tratadas com etoposídeo. RIbeirão Preto, 2016.

Örd, M. et al. Cyclin-Specific Docking Mechanisms Reveal the Complexity of M-CDK Function in the Cell Cycle. **Mol Cell**, v. 11, n. 75, p. 76-89.e3, 2019. https://doi.org/10.1016/j.molcel.2019.04.026.

Organização Mundial da Saúde. Câncer. 2020. Disponível <https://www.paho.org/pt/topicos/cancer>, Acesso em: 10/2022.

Osher, D.J., De Leeneer, K., Michils, G. et al. Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families. **Br J Cancer**, v.106, p.1460-3, 2012. https://doi.org/10.1038/bjc.2012.87.

Park, C, Park, K, Kim, J. et al. Prognostic values of negative estrogen or progesterone receptor expression in patients with luminal B HER2-negative breast cancer. **World J Surg Onc,** v.14, p.244, 2016. https://doi.org/10.1186/s12957-016-0999-x.

Paunola, E. et al. WH2 domain: a small, versatile adapter for actin monomers. **FEBS Lett**. v. 20, n. 513, p. 92-7, 2002. https://doi.org/10.1016/s0014-5793(01)03242-2.

Pawson, T., Gish, G.D., Nash, P. SH2 domains, interaction modules and cellular wiring. **Trends Cell Biol**, n. 11, p. 504-11, 2001. https://doi.org/10.1016/s0962-8924(01)02154-7.

Pejaver, V., Urresti, J., Lugo-Martinez, J. et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. **Nat Commun**, v. 11, p. 5918, 2020. https://doi.org/10.1038/s41467-020-19669-x.

Penkert, J., Schmidt, G., Hofmann, W. et al. Breast cancer patients suggestive of Li-Fraumeni syndrome: mutational spectrum, candidate genes, and unexplained heredity. **Breast Cancer Res.** v. 7, p. 87, 2018. https://doi.org/10.1186/s13058-018-1011-1.

Pennington, K.P., Walsh, T., Harrell, M.I. et al. Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas. **Clin Cancer Res,** v. 1, n. 20(3), p. 764-75, 2014. https://doi.org/10.1158/1078-0432.CCR-13-2287.

Pernigo, S. et al. Structural basis for kinesin-1:cargo recognition. **Science**, v. 19, n. 340, p. 356-9. https://doi.org/10.1126/science.1234264..

Peters J.M. The anaphase promoting complex/cyclosome: a machine designed to destroy. **Nat Rev Mol Cell Biol**, n. 7, p. -56, 2006. https://doi.org/10.1038/nrm1988.

Petersen B.O., Wagener C., Marinoni F. et al. Cell cycle- and cell growth-regulated proteolysis of mammalian CDC6 is dependent on APC-CDH1. **Genes Dev**, v. 15, n. 14, p. 2330-43, 2000. https://doi.org/10.1101/gad.832500.

Pires, D. E. V. et al. mCSM: predicting the effects of mutations in proteins using graph-based signatures. **Bioinformatics**, n. 30, p. 335-342, 2013. https://doi.org/10.1093/bioinformatics/btt691

Pittman, D.L. et al. Identification, characterization, and genetic mapping of Rad51d, a new mouse and human RAD51/RecA-related gene. **Genomics**, v. 1, n. 49, p. 103-11, 1998. https://doi.org/10.1006/geno.1998.5226.

Potter, S.C. et al. HMMER web server. **Nucleic Acids Res**, n. 46, p. W200-W204, 2018. https://doi.org/10.1016/j.jbc.2021.101210.

Pratz, K.W., Koh, B.D., Patel, A.G. et al. Poly (ADP-Ribose) Polymerase Inhibitor Hypersensitivity in Aggressive Myeloproliferative Neoplasms. **Clin Cancer Res**, v. 1, n. 22, p. 3894-902, 2016. https://doi.org/10.1158/1078-0432.CCR-15-2351.

Rachidi, N., Taly, J.F., Durieu, E. et al. Pharmacological assessment defines Leishmania donovani casein kinase 1 as a drug target and reveals important functions in parasite viability and intracellular infection. **Antimicrob Agents Chemother**, n. 58, p. 1501-15, 2014. https://doi.org/10.1128/AAC.02022-13.

Rámon, A.M.C. Diseño y caracterizacíon de proteínas quiméricas para el estudio del reconecimiento molecular entre ligandos peptidícos y dominios SH3. Universidade de Granada, 2008. Disponível <https://digibug.ugr.es/bitstream/handle/10481/1778/17328196.pdf?sequence=1>.

Rehder, C, Bean, LJH, Bick, D. et al. Next-generation sequencing for constitutional variants in the clinical laboratory, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG) . **Genetics in Medicine**, 2021. https://doi.org/10.1038/s41436-021-01139-4.

Reis, F.J.C. Rastreamento e diagnóstico das neoplasias de ovário – papel dos marcadores tumorais. Rev Bras Ginecol Obstet, n. 27, p. 222-7, 2005. Disponível < https://www.scielo.br/j/rbgo/a/H4nRNRbjPXLpGDyZDXf8JJS/?format=pdf&lang=pt>.

Richards, S., Aziz N., Bale S. 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**, n. 17, p. 405-24, 2015. https://doi.org/10.1038/gim.2015.30.

Ricker, C.N., Hanna, D.L., Peng, C. et al. DNA mismatch repair deficiency and hereditary syndromes in Latino patients with colorectal cancer. **Cancer**, v. 1, n. 123, p. 3732-3743, 2017. https://doi.org/10.1002/cncr.30790.

Ring, K.L., Brueg,l A.S., Allen, B.A. et al. Germline multi-gene hereditary cancer panel testing in an unselected endometrial cancer cohort. **Mod Pathol**. n. 29, p. 1381-1389, 2016. https://doi.org/10.1038/modpathol.2016.135.

Rodrigues, C. H. M. et al. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. **Nucleic Acids Res**, n. 46W1, p. W350–W355, 2018. https://doi.org/10.1093/nar/gky300.

Rodríguez-López, R, Osorio, A, Ribas, G. et al. The variant E233G of the RAD51D gene could be a low-penetrance allele in high-risk breast cancer families without BRCA1/2 mutations. **Cancer Genetics**, v. 110, p.845-849, 2004. https://doi.org/10.1002/ijc.20169.

Rosa, H.T. Anemia de fanconi: observações recentes em células linfoblastóides imortalizadas de pacientes como modelo de estudo. **Universidade Federal de Ciências da Saúde de Porto Alegre**, 2016. Disponível <https://repositorio.ufcspa.edu.br/jspui/bitstream/123456789/474/1/%5BDISSERTA%C3%87%C3%83O%5D%20Silva%2C%20Helen%20Tais%20da%20Rosa>.

Rogers M. F. et al. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features, **Bioinformatics**, v. 34, p. 511–513, 2018. https://doi.org/10.1093/bioinformatics/btx536.

Romagna, D.V., Oliveira, M.M. de, Abreu, L.G. et al. Incidence and mortality rates of lip, oral cavity, and pharynx cancers in Brazil: time-trend and age-period-cohort analysis from the last 30 years, Global Burden of Disease Study. **Rev Soc Bras Med Trop**, 2022. https://doi.org/10.1590/0037-8682-0286-2021.

Ruff, K.M. et al. AlphaFold and Implications for Intrinsically Disordered Proteins. **J Mol Biol**, v.433, p.167208, 2021. https://doi.org/10.1016/j.jmb.2021.167208

Salgado, D., Desvignes, J. P., Rai, G. et al. UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution. **Hum Mutat**, n. 37(5), p. 439-46, 2016. https://doi.org/ 10.1002/humu.22965.

Sánchez-Bermúdez, A.I., Sarabia-Meseguer, M.D., García-Aliaga, Á. et al. Mutational analysis of RAD51C and RAD51D genes in hereditary breast and ovarian cancer families from Murcia (southeastern Spain). **Eur J Med Genet**, v. 61, p. 355-361, 2018. https://doi.org/10.1016/j.ejmg.2018.01.015.

Senter, N.C. et al. Identification of Recurrent Chromosome Breaks Underlying Structural Rearrangements in Mammary Cancer Cell Lines. **Genes**, n. 13, p. 1228, 2022. https://doi.org/10.3390/genes13071228.

Schaab, C. et al. Analysis of high accuracy, quantitative proteomics data in the MaxQB database. **Mol Cell Proteomics**, n. 11, p. M111.014068, 2012. https://doi.org/10.1074/mcp.M111.014068.

Schittek, B., Sinnberg, T. Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. **Mol Cancer**, v. 11, n. 13, p. 231, 2014. https://doi.org/10.1186/1476-4598-13-231.

Sherry, S.T., Ward, M.H., Kholodov, M. et al. dbSNP: the NCBI database of genetic variation. **Nucleic Acids Res,** n. 29, p. 308-11, 2001. https://doi.org/10.1093/nar/29.1.308.

Scheid, M. R. Análise de ferramentas in silico para variantes sinônimas em genes de predisposição ao câncer. Universidade Federal do Rio Grande do Sul, 2018. http://hdl.handle.net/10183/199014.

Shih, I-M. et al. The Origin of Ovarian Cancer Species and Precancerous Landscape. **Am J Clin Pathol**, v.191, p.26-39, 2021. https://doi.org/10.1016/j.ajpath.2020.09.006.

Shindo, K., Yu, J., Suenaga, M. et al. Deleterious Germline Mutations in Patients With Apparently Sporadic Pancreatic Adenocarcinoma. **J Clin Oncol**, v. 20, n. 35, p. 3382-3390, 2017. https://doi.org/10.1200/JCO.2017.72.3502.

Slatko, B.E., Gardner, A.F., Ausubel, F.M. Overview of next-generation sequencing technologies. **Curr Protoc Mol Biol**, n. 122, p. e59, 2018. https://doi.org/10.1002/cpmb.59.

Somyajit, K. et al. RAD51C: a novel cancer susceptibility gene is linked to Fanconi anemia and breast cancer. **Carcinogenesis**, v. 31, p. 2031-8, 2010. https://doi.org/10.1093/carcin/bgq210.

Song, H., Dicks, E., Ramus, S.J. et al. Contribution of Germline Mutations in the RAD51B, RAD51C, and RAD51D Genes to Ovarian Cancer in the Population. **J Clin Oncol**. v.10, p. 2901-7, 2015. https://doi.org/10.1200/JCO.2015.61.2408.

Song, J. et al. Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. **Proc Natl Acad Sci U S A**. v. 5, p. 14373-8, 2004. https://doi.org/10.1073/pnas.0403498101.

Spitale, A. et al. Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. **Ann Oncol**, v.20, p.628-35, 2009. https://doi.org/10.1093/annonc/mdn675.

Steinhaus, R., Proft, S., Schuelke M. et al. MutationTaster. **Nucleic Acids Res**, n. 49(W1), p. W446-W451, 2021. https://doi.org/10.1093/nar/gkab266.

Sudol, M., Sliwa, K., Russo, T. Functions of WW domains in the nucleus. **FEBS Lett**, v. 16, p. 90-5, 2001. doi: https://doi.org/10.1016/s0014-5793(01)02122-6.

Sullivan, M.R., Bernstein, K.A. RAD-ical New Insights into RAD51 Regulation. **Genes (Basel)**. v. 13, p. 629, 2018. https://doi.org/10.3390/genes9120629.

Susswein, L.R., Marshall, M.L., Nusbaum, R. et al. Pathogenic and likely pathogenic variant prevalence among the first 10,000 patients referred for next-generation cancer panel testing. **Genet Med**, n. 18, p. 823-32, 2016. https://doi.org/10.1038/gim.2015.166.

Suszynska, M. et al. BRIP1, RAD51C, and RAD51D mutations are associated with high susceptibility to ovarian cancer: mutation prevalence and precise risk estimates based on a pooled analysis of ~30,000 cases. **J Ovarian Res,** v. 2, n. 13, p. 50, 2020. https://doi.org/10.1186/s13048-020-00654-3.

Taghizadeh, H., Mader, R.M., Müllauer, L. et al. Outcome of Targeted Therapy Recommendations for Metastatic and Recurrent Head and Neck Cancers. **Cancers (Basel)**, v. 15, p. 3381, 2020. https://doi.org/10.3390/cancers12113381.

Tang, Z. et al. Protein interactions in the sumoylation cascade: lessons from X-ray structures. **FEBS J**, n. 275, p. 3003-15, 2008. https://doi.org/10.1111/j.1742-4658.2008.06459.x.

Tedaldi, G., Tebaldi, M., Zampiga, V. et al. Multiple-gene panel analysis in a case series of 255 women with hereditary breast and ovarian cancer. **Oncotarget,** v. 18, n. 8, p. 47064-47075, 2017. https://doi.org/10.18632/oncotarget.16791.

Tomioka, K., Miyamoto, T., Akutsu, S.N. et al. NBS1 I171V variant underlies individual differences in chromosomal radiosensitivity within human populations. **Sci Rep**, v. 4; n. 11, p. 19661, 2021. https://doi.org/10.1038/s41598-021-98673-7.

Toss, A., Tomasello, C., Razzaboni, E. et al. Hereditary ovarian cancer: not only BRCA 1 and 2 genes. **Biomed Res Int**, p. 341723, 2015. https://doi.org/10.1155/2015/341723.

Tsaousis, G.N., Papadopoulou, E., Apessos, A. et al. Analysis of hereditary cancer syndromes by using a panel of genes: novel and multiple pathogenic mutations. **BMC Cancer** v. 19, p. 535, 2019. https://doi.org/10.1186/s12885-019-5756-4.

Tung, N., Battelli, C., Allen, B. et al. Frequency of mutations in individuals with breast cancer referred for BRCA1 and BRCA2 testing using next-generation sequencing with a 25-gene panel. **Cancer**, v. 1, n. 121, p. 25-33, 2015. https://doi.org/10.1002/cncr.29010.

Uemura, S. et al. The cytoplasmic tail of GM3 synthase defines its subcellular localization, stability, and in vivo activity. **Mol Biol Cell**, n. 20, p. 3088-100, 2009. https://doi.org/10.1091/mbc.e08-12-1219.

Unkeless, J.C., Jin, J. Inhibitory receptors, ITIM sequences and phosphatases. **Curr Opin Immunol**, v. 9, p. 338-43, 1997. https://doi.org/10.1016/s0952-7915(97)80079-9.

US Preventive Services Task Force. Risk Assessment, Genetic Counseling, and Genetic Testing for BRCA-Related Cancer: US Preventive Services Task Force Recommendation Statement. **JAMA,** n. 322, p. 652–665, 2019. https://doi.org/10.1001/jama.2019.10987.

Vaser, R. et al. SIFT missense predictions for genomes. **Nat Protocols**, n. 11, p. 1-9, 2016. https://doi.org/10.1038/nprot.2015.123.

Vogelstein, B., Kinzler, K.W. Cancer genes and the pathways they control. **Nat Med,** v. 10, p. 789-99, 2004. https://doi.org/10.1038/nm1087. PMID: 15286780.

Wang, C. et al. The Association between Endometriosis, Tubal Ligation, Hysterectomy and Epithelial Ovarian Cancer: Meta-Analyses. **Int J Environ Res Public Health**, n. 13, p. 1138, 2016. https://doi.org/10.3390/ijerph13111138.

Wang, M., Marín, A. Characterization and Prediction of Alternative Splice Sites. **Gene**, n. 366, p. 219-227, 2006.

Webb. B., Sali. A. Comparative Protein Structure Modeling Using Modeller. **Curr Protoc Bioinformatics**, v. 54, p. 5.6.1-5.6.37, 2016.

Wei, Y., Wu, J., Gu, W., et al. Germline DNA Repair Gene Mutation Landscape in Chinese Prostate Cancer **Patients**. v.76, p.280-283, 2019. https://doi.org/10.1016/j.eururo.2019.06.004.

Weinberg, R.A. Mechanisms of malignant progression. **Carcinogenesis**, n. 29, p. 1092-5, 2008. https://doi.org/10.1093/carcin/bgn104.

Wesch, N. et al. Atg8-Family Proteins-Structural Features and Molecular Interactions in Autophagy and Beyond. **Cells**, v. 1, n. 9, p. 2008, 2020. https://doi.org/10.3390/cells9092008.

Wickramanayake, A., Bernier, G., Pennil, C. et al. Loss of function germline mutations in RAD51D in women with ovarian carcinoma. **Gynecol Oncol**, n. 127, p. 552-5, 2012. https://doi.org/10.1016/j.ygyno.2012.09.009.

Wilson III, D.M., Deacon, A.M., Duncton, M.A.J. et al. Fragment- and structure-based drug discovery for developing therapeutic agents targeting the DNA Damage Response. **Progress in Biophysics and Molecular Biology**, 163, 130-142, 2021. https://doi.org/10.1016/j.pbiomolbio.2020.10.005.

Yadav, S. et al. Outcomes of retesting BRCA negative patients using multigene panels. **Fam Cancer**, n. 16, p. 319-328, 2017. https://doi.org/10.1007/s10689-016-9956-7.

Yang X., Song H., Leslie G. et al. Ovarian and Breast Cancer Risks Associated With Pathogenic Variants in RAD51C and RAD51D. **J Natl Cancer Inst**, n. 14, v. 112, p. 1242-1250, 2020. https://doi.org/10.1093/jnci/djaa030.

Yin, L., Duan, J.J., Bian, X.W. et al. Triple-negative breast cancer molecular subtyping and treatment progress. **Breast Cancer Res**, v. 22, p. 61, 2020. https://doi.org/10.1186/s13058-020-01296-5

Yongkiettrakul, S. et al. The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues. **Biochemistry**, v. 6, n. 43, p. 3862-9, 2004 https://doi.org/10.1021/bi036195f.

Yoshida, R. Hereditary breast and ovarian cancer (HBOC): review of its molecular characteristics, screening, treatment, and prognosis. **Breast Cancer,** n. 28, p. 167-1180, 2021. https://doi.org/10.1007/s12282-020-01148-2.

Yurgelun, M.B., Allen, B., Kaldate, R.R. et al.. Identification of a Variety of Mutations in Cancer Predisposition Genes in Patients With Suspected Lynch Syndrome. **Gastroenterology**, n. 149, p.604-13.e20, 2015. https://doi.org/10.1053/j.gastro.2015.05.006.

Yurgelun, M.B., Kulke, M.H., Fuchs, C.S. et al. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. **J Clin Oncol**, v. 1, p. 1086-1095. https://doi.org/10.1200/JCO.2016.71.0012.

Zheng, H, Zhang, G, Zhang, L, et al. Comprehensive Review of Web Servers and Bioinformatics Tools for Cancer Prognosis Analysis. **Front Oncol**, v.10, 2020. https://doi.org/10.3389/fonc.2020.00068.

Zhu, G., Fujii, K., Liu, Y. et al. A single pair of acidic residues in the kinase major groove mediates strong substrate preference for P-2 or P-5 arginine in the AGC, CAMK, and STE kinase families. **J Biol Chem**, v. 28, n. 280, p. 36372-9, 2005. https://doi.org/10.1074/jbc.M505031200.

Zhu, K. et al. Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. **Proteins**, n. 82, p. 1646-1655, 2014. https://doi.org/10.1002/prot.24551.

Zielińska, M. et al. HER2-positive breast cancer – Available anti-HER2 therapies and new agents under investigation. **J Pre Clin Clin Res**, n. 14, p. 44-48, 2020. https://doi.org/10.26444/jpccr/122068.

Zore, T. et al. Polycystic Ovarian Syndrome: Long-Term Health Consequences. **Semin Reprod Med**, n. 35, p. 271-281. 2017. https://doi.org/10.1055/s-0037-1603096

ANEXO 1 - COEP UFMG

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Patogênica ou não? Avaliação do impacto de mutações descobertas em pessoas com câncer

Pesquisador: Maria Raquel Santos Carvalho

Área Temática: Genética Humana:
(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP;);

Versão: 1

CAAE: 48770621.6.0000.5149

Instituição Proponente: UNIVERSIDADE FEDERAL DE MINAS GERAIS

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.865.051

Apresentação do Projeto:

Estudo prospectivo/retrospectivo, visando avaliar o impacto funcional de mutações, descobertas em pacientes com diversos tipos de cânceres. As mutações estão sendo detectadas em exames realizados pelos pacientes, por solicitação de seus médicos. Estes exames são baseados no sequenciamento do DNA. Podem ser baseados em painéis compostos por diversos genes associados a tipos específicos de cânceres familiares ou pelo sequenciamento completo do exoma. Como existe muita variação genética nas populações humanas, é comum que se encontrem mutações. As mutações são classificadas em benignas, provavelmente benignas, variantes de significado incerto (VUS), potencialmente patogênicas ou patogênicas. Existem protocolos de cuidados para os pacientes/famílias que possuem patogênicas ou potencialmente patogênicas. Quando a mutação é uma VUS, o médico não tem um suporte para decidir o que fazer (investiga outros membros da família? monitora a recorrência de quanto em quanto tempo? usa um esquema terapêutico específico para mutações mais agressivas?). Após os exames genéticos, muitos pacientes recebem laudos descrevendo VUS. Estes laudos são baseados em programas de computador que avaliam poucas características das mutações. Avaliações mais pormenorizadas dos impactos nos níveis de DNA, RNA e proteína podem permitir a reclassificação das VUS. Estas análises

são baseadas em ferramentas de Bioinformática relativamente sofisticadas, mais acessíveis nos serviços universitários. Neste projeto, preve avaliar mutações encaminhadas pelos médicos assistentes, particularmente VUS, e devolver para o médico um laudo, reclassificando ou não a(s) mutação(ões) encaminhada(s).

Objetivo da Pesquisa:

1. Avaliar o impacto na predição de patogenicidade de uma mutação, do uso de uma avaliação computacional detalhada, baseada em ferramentas da Bioinformática, considerando os impactos nos níveis da proteína, do RNA e do DNA;
2. Identificar mutações novas, e investigar sua associação com cânceres hereditários, presentes na população de Minas Gerais ou na brasileira;

Avaliação dos Riscos e Benefícios:

risco de gerar ansiedade no participante, por se tratar da ampliação de um exame realizado. Risco de exposição de dados.

benefícios individuais como a possibilidade de reclassificação de riscos e benefícios geral de identificação de banco de informações de mutações genéticas

Comentários e Considerações sobre a Pesquisa:

Trata-se de importante trabalho de avaliação comparada de riscos através das análises das mutações.

No projeto não está claro qual o número de indivíduos da pesquisa, quem , quando e onde serão convidados a participarem. Não estabelecendo período de pesquisa e não sendo possibilitado a coleta de autorização no TCLE, como se chegará aos laudos dos pacientes= arquivos de consultório?de hospitais?quais? Não há há anuência de hospitais ou laboratórios.

Não está anexado o pedido de dispensa do TCLE

Considerações sobre os Termos de apresentação obrigatória:

A pesquisadora anexa

projeto de pesquisa, ficha clínica

aprovação do Depto. Genética, Ecologia e Evolução, ICB e parecer departamental

TCLE, TCUs,

Recomendações:

Recomenda-se aprovação do protocolo de pesquisa.

Conclusões ou Pendências e Lista de Inadequações:

SMJ,somos pela aprovação do projeto de pesquisa.

Considerações Finais a critério do CEP:

Tendo em vista a legislação vigente (Resolução CNS 466/12), o CEP-UFMG recomenda aos Pesquisadores: comunicar toda e qualquer alteração do projeto e do termo de consentimento via emenda na Plataforma Brasil, informar imediatamente qualquer evento adverso ocorrido durante o desenvolvimento da pesquisa (via documental encaminhada em papel), apresentar na forma de notificação relatórios parciais do andamento do mesmo a cada 06 (seis) meses e ao término da pesquisa encaminhar a este Comitê um sumário dos resultados do projeto (relatório final).

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1745433.pdf	25/06/2021 18:19:23		Aceito
Projeto Detalhado / Brochura Investigador	Projeto.pdf	25/06/2021 17:49:57	Maria Raquel Santos Carvalho	Aceito
Folha de Rosto	Folha_de_Rosto_Assinada.pdf	25/06/2021 17:22:53	Maria Raquel Santos Carvalho	Aceito
Outros	Ficha_Clinica.pdf	01/05/2021 17:57:10	Maria Raquel Santos Carvalho	Aceito
Outros	Aprovacao_DGEE.pdf	01/05/2021 17:56:44	Maria Raquel Santos Carvalho	Aceito
Outros	Parecer_do_Departamento.pdf	01/05/2021 17:55:47	Maria Raquel Santos Carvalho	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_Cancer_no_ICB.pdf	01/05/2021 17:55:17	Maria Raquel Santos Carvalho	Aceito
Outros	TCUD.pdf	01/05/2021 17:29:57	Maria Raquel Santos Carvalho	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

BELO HORIZONTE, 24 de Julho de 2021

Assinado por:
Críssia Carem Paiva Fontainha
(Coordenador(a))

Endereço: Av. Presidente Antônio Carlos,6627 2º Ad SI 2005
Bairro: Unidade Administrativa II **CEP:** 31.270-901
UF: MG **Município:** BELO HORIZONTE
Telefone: (31)3409-4592 **E-mail:** coep@prpq.ufmg.br

ANEXO 2 - COEP FACULDADE DE MEDICINA DA UFMG

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Perfil de mutações germinativas em pacientes submetidas ao aconselhamento genético para câncer hereditário de mama, ovário e endométrio, em Minas Gerais, Brasil

Pesquisador: Agnaldo Lopes da Silva Filho

Área Temática:

Versão: 2

CAAE: 01758418.0.0000.5149

Instituição Proponente: Faculdade de Medicina da UFMG

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.134.990

Apresentação do Projeto:

Trata-se de versão do projeto de pesquisa que responde diligências elencadas no parecer de número 3.078.927, emitido por este CEP.

Desenho inalterado em relação àquele parecer:

Conforme informado no formulário de informações básicas:

"Avaliar o perfil de mutações germinativas em pacientes com câncer de mama, ovário e endométrio submetidas a aconselhamento genético para câncer hereditário, bem como associar a presença dessas mutações com as características clínicas das pacientes e o tipo de tumor, caracterizar as variantes genéticas detectadas nos genes entre as pacientes com câncer ginecológico hereditário e comparar os dados obtidos com dados já

reportados na literatura da população brasileira e 'The Cancer Genome Atlas'.

[...]

Metodologia Proposta:

Serão analisados os prontuários de todos pacientes que realizaram o aconselhamento genético, devido à suspeita de câncer hereditário. Para a composição da amostra deste estudo serão selecionados 300 pacientes que apresentaram história pessoal ou familiar de câncer de mama, ovário, endométrio e outros tipos de câncer associado a síndromes hereditárias".

Objetivo da Pesquisa:

Conforme formulário de informações básicas atual:

Objetivo Primário:

Avaliar o perfil de mutações germinativas em pacientes com câncer de mama, ovário e endométrio submetidas a aconselhamento genético para câncer hereditário.

Avaliação dos Riscos e Benefícios:

Conforme formulário de informações básicas atual, descrito pelos pesquisadores:

"Riscos:

Os riscos decorrentes da participação nesta pesquisa sao minimos, pois podem decorrer de possível constrangimento dos participantes com a divulgação de dados confidenciais, invasão de privacidade e segurança dos prontuários. Para mimizar estes riscos, o pesquisador irá tratar a identidade dos participantes com padrões profissionais de sigilo, divulgação do nome ou qualquer material que indique o participante sem sua

permissão.

Benefícios:

Não haverá benefício direto para as pacientes, exceto nos casos em que forem identificadas mutações genéticas que já estão associadas, na literatura, a síndromes neoplásicas, pois muda o manejo do risco da doença e permite que os pacientes e seus familiares em risco se beneficiem de medidas voltadas para o diagnóstico precoce e intervenção, condutas profiláticas voltadas para redução considerável do risco para o câncer e o prognóstico seja melhor para esses indivíduos, com diminuição da morbidade e mortalidade".

Comentários e Considerações sobre a Pesquisa:

Pendências emitidas no parecer de número 3.078.927 foram atendidas, a saber:

1) inclusão do Hermes Pardini como centro coparticipante;

2) inclusão de novo TCLE exclusivo para a pesquisa (independente do documento de TCLE de aconselhamento genético). Este novo TCLE esclarece que serão utilizados dados de prontuários e que tais dados serão tratados sob sigilo, com uso exclusivo para pesquisa.

Considerações sobre os Termos de apresentação obrigatória:

Termos adequados.

Recomendações:

Afirmar, no TCLE, que os dados serão utilizados de maneira agregada, sem possibilitar identificação do participante.

Conclusões ou Pendências e Lista de Inadequações:

Tendo sido atendidas as diligências do parecer de número 3.078.927, aprova-se o projeto de pesquisa.

Considerações Finais a critério do CEP:

Tendo em vista a legislação vigente (Resolução CNS 466/12), o CEP-UFMG recomenda aos Pesquisadores: comunicar toda e qualquer alteração do projeto e do termo de consentimento via emenda na Plataforma Brasil, informar imediatamente qualquer evento adverso ocorrido durante o desenvolvimento da pesquisa (via documental encaminhada em papel), apresentar na forma de notificação relatórios parciais do andamento do mesmo a cada 06 (seis) meses e ao término da pesquisa encaminhar a este Comitê um sumário dos resultados do projeto (relatório final).

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1190397.pdf	06/01/2019 19:40:32		Aceito
Projeto Detalhado / Brochura Investigador	projeto_detalhado.docx	06/01/2019 19:40:18	CAMILA MARTINS DE CARVALHO	Aceito
Outros	Carta_resposta_ao_parecer.docx	05/01/2019 21:04:23	CAMILA MARTINS DE CARVALHO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_para_pesquisa.doc	05/01/2019 21:03:11	CAMILA MARTINS DE CARVALHO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx	15/10/2018 15:12:47	CAMILA MARTINS DE CARVALHO	Aceito
Outros	parecer_departamento.pdf	13/10/2018 17:38:22	CAMILA MARTINS DE CARVALHO	Aceito
Outros	carta_anuencia_hermes_pardini.pdf	13/10/2018 17:37:15	CAMILA MARTINS DE CARVALHO	Aceito
Folha de Rosto	FOLHA_DE_ROSTO.pdf	13/10/2018 17:35:09	CAMILA MARTINS DE CARVALHO	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

BELO HORIZONTE, 06 de Fevereiro de 2019

Assinado por:
Eliane Cristina de Freitas Rocha
(Coordenador(a))

Endereço: Av. Presidente Antônio Carlos,6627 2º Ad SI 2005
Bairro: Unidade Administrativa II **CEP:** 31.270-901
UF: MG **Município:** BELO HORIZONTE
Telefone: (31)3409-4592 **E-mail:** coep@prpq.ufmg.br

ANEXO 3 - Resultado da *RAD51D* c.666A>G; (p.Glu222=)

Análises complementares foram realizadas para a variante *RAD51D* c.666A>G; (p.Glu222=), para prever os efeitos na sequência de DNA. Utilizando o software SMS CpG Islands, foram avaliados os possíveis impactos nas ilhotas CpG (Fig. 1).

Figura 1. Possíveis impacto nas ilhosta CpG, usando SMS CpG Islands

SELVAGEM

	Primer	Start	Size	Tm	GC%	'C's	Sequence
1	Left primer	137	25	57.93	60.00	8	TTTATAAGGTTTTGGTTGTTTTGAG
	Right primer	326	25	58.62	48.00	4	AAACCTCCTACAATTTCAATCACTT
	Product size: 190, Tm: 71.6, CpGs in product: 5						
2	Left primer	460	26	59.06	57.69	5	ATTTAGGATGAGGAGGAATAGGTAGA
	Right primer	657	25	59.61	60.00	7	CTAACCTCCCAAAAATAAAAAAACC
	Product size: 198, Tm: 71.8, CpGs in product: 4						
3	Left primer	460	26	59.06	57.69	5	ATTTAGGATGAGGAGGAATAGGTAGA
	Right primer	658	25	57.50	60.00	8	ACTAACCTCCCAAAAATAAAAAAAC
	Product size: 199, Tm: 71.7, CpGs in product: 4						
4	Left primer	129	25	59.22	52.00	6	TGGTTTGTTTTATAAGGTTTTGGTT
	Right primer	326	25	58.62	48.00	4	AAACCTCCTACAATTTCAATCACTT
	Product size: 198, Tm: 71.8, CpGs in product: 5						
5	Left primer	461	26	59.90	57.69	5	TTTAGGATGAGGAGGAATAGGTAGAA
	Right primer	657	25	59.61	60.00	7	CTAACCTCCCAAAAATAAAAAAACC
	Product size: 197, Tm: 71.6, CpGs in product: 4						

MUTADA

	Primer	Start	Size	Tm	GC%	'C's	Sequence
1	Left primer	137	25	57.93	60.00	8	TTTATAAGGTTTTGGTTGTTTTGAG
	Right primer	326	25	58.62	48.00	4	AAACCTCCTACAATTTCAATCACTT
	Product size: 190, Tm: 71.6, CpGs in product: 5						
2	Left primer	460	26	59.06	57.69	5	ATTTAGGATGAGGAGGAATAGGTAGA
	Right primer	657	25	59.61	60.00	7	CTAACCTCCCAAAAATAAAAAAACC
	Product size: 198, Tm: 71.8, CpGs in product: 4						
3	Left primer	460	26	59.06	57.69	5	ATTTAGGATGAGGAGGAATAGGTAGA
	Right primer	658	25	57.50	60.00	8	ACTAACCTCCCAAAAATAAAAAAAC
	Product size: 199, Tm: 71.7, CpGs in product: 4						
4	Left primer	129	25	59.22	52.00	6	TGGTTTGTTTTATAAGGTTTTGGTT
	Right primer	326	25	58.62	48.00	4	AAACCTCCTACAATTTCAATCACTT
	Product size: 198, Tm: 71.8, CpGs in product: 5						
5	Left primer	461	26	59.90	57.69	5	TTTAGGATGAGGAGGAATAGGTAGAA
	Right primer	657	25	59.61	60.00	7	CTAACCTCCCAAAAATAAAAAAACC
	Product size: 197, Tm: 71.6, CpGs in product: 4						

Nota: Elaborado pela autora, usando SMS CpG Islands. Stothard P (2000) The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102-1104

Para predizer os efeitos sobre a preferência por códons sinônimos específicos, foi usado o software SMS codon usage, no qual o resultado pode ser observado na figura 2.

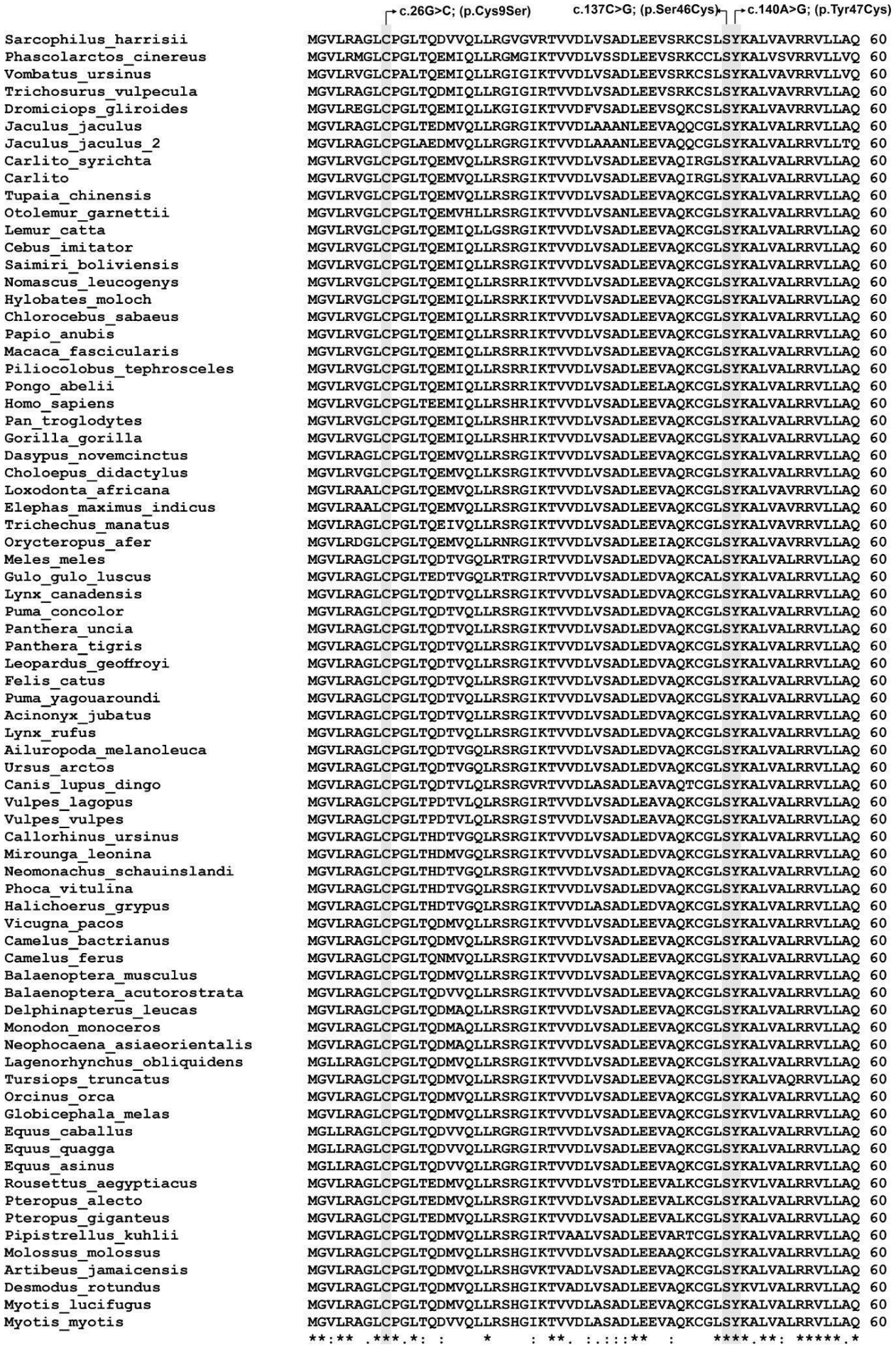
Figura 2. Efeitos da variante quanto ao viés n uso de códon

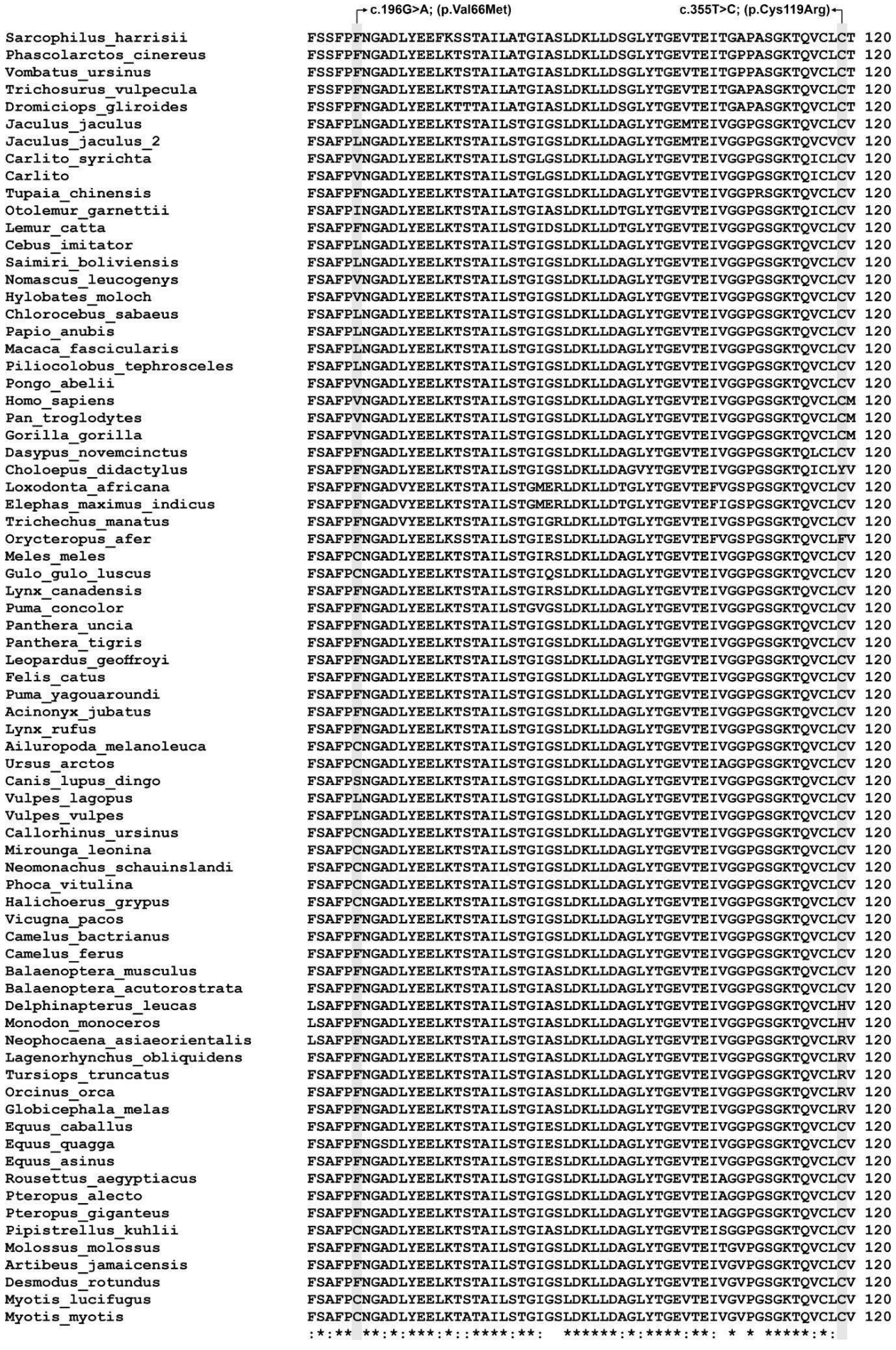
SELVAGEM					MUTADA				
aa	Códon	Num.	/1000	Fração	AmAcid	Codon	Number	/1000	Fração
Ala	GCG	2.00	6.08	0.07	Ala	GCG	2.00	6.08	0.07
Ala	GCA	8.00	24.32	0.28	Ala	GCA	8.00	24.32	0.28
Ala	GCT	8.00	24.32	0.28	Ala	GCT	8.00	24.32	0.28
Ala	GCC	11.00	33.43	0.38	Ala	GCC	11.00	33.43	0.38
Cys	TGT	4.00	12.16	0.80	Cys	TGT	4.00	12.16	0.80
Cys	TGC	1.00	3.04	0.20	Cys	TGC	1.00	3.04	0.20
Asp	GAT	7.00	21.28	0.44	Asp	GAT	7.00	21.28	0.44
Asp	GAC	9.00	27.36	0.56	Asp	GAC	9.00	27.36	0.56
Glu	GAG	11.00	33.43	0.61	Glu	GAG	12.00	36.47	0.67
Glu	GAA	7.00	21.28	0.39	Glu	GAA	6.00	18.24	0.33
Phe	TTT	2.00	6.08	0.33	Phe	TTT	2.00	6.08	0.33
Phe	TTC	4.00	12.16	0.67	Phe	TTC	4.00	12.16	0.67
Gly	GGG	4.00	12.16	0.12	Gly	GGG	4.00	12.16	0.12
Gly	GGA	9.00	27.36	0.27	Gly	GGA	9.00	27.36	0.27
Gly	GGT	6.00	18.24	0.18	Gly	GGT	6.00	18.24	0.18
Gly	GGC	14.00	42.55	0.42	Gly	GGC	14.00	42.55	0.42
His	CAT	2.00	6.08	0.50	His	CAT	2.00	6.08	0.50
His	CAC	2.00	6.08	0.50	His	CAC	2.00	6.08	0.50
Ile	ATA	1.00	3.04	0.09	Ile	ATA	1.00	3.04	0.09
Ile	ATT	4.00	12.16	0.36	Ile	ATT	4.00	12.16	0.36
Ile	ATC	6.00	18.24	0.55	Ile	ATC	6.00	18.24	0.55
Lys	AAG	5.00	15.20	0.45	Lys	AAG	5.00	15.20	0.45
Lys	AAA	6.00	18.24	0.55	Lys	AAA	6.00	18.24	0.55
Leu	TTG	3.00	9.12	0.07	Leu	TTG	3.00	9.12	0.07
Leu	TTA	1.00	3.04	0.02	Leu	TTA	1.00	3.04	0.02
Leu	CTG	21.00	63.83	0.47	Leu	CTG	21.00	63.83	0.47
Leu	CTA	1.00	3.04	0.02	Leu	CTA	1.00	3.04	0.02
Leu	CTT	7.00	21.28	0.16	Leu	CTT	7.00	21.28	0.16
Leu	CTC	12.00	36.47	0.27	Leu	CTC	12.00	36.47	0.27
Met	ATG	9.00	27.36	1.00	Met	ATG	9.00	27.36	1.00
Asn	AAT	3.00	9.12	0.60	Asn	AAT	3.00	9.12	0.60
Asn	AAC	2.00	6.08	0.40	Asn	AAC	2.00	6.08	0.40
Pro	CCG	0.00	0.00	0.00	Pro	CCG	0.00	0.00	0.00
Pro	CCA	3.00	9.12	0.43	Pro	CCA	3.00	9.12	0.43
Pro	CCT	2.00	6.08	0.29	Pro	CCT	2.00	6.08	0.29
Pro	CCC	2.00	6.08	0.29	Pro	CCC	2.00	6.08	0.29
Gln	CAG	22.00	66.87	0.96	Gln	CAG	22.00	66.87	0.96
Gln	CAA	1.00	3.04	0.04	Gln	CAA	1.00	3.04	0.04
Arg	AGG	8.00	24.32	0.40	Arg	AGG	8.00	24.32	0.40
Arg	AGA	0.00	0.00	0.00	Arg	AGA	0.00	0.00	0.00
Arg	CGG	5.00	15.20	0.25	Arg	CGG	5.00	15.20	0.25
Arg	CGA	4.00	12.16	0.20	Arg	CGA	4.00	12.16	0.20
Arg	CGT	0.00	0.00	0.00	Arg	CGT	0.00	0.00	0.00
Arg	CGC	3.00	9.12	0.15	Arg	CGC	3.00	9.12	0.15
Ser	AGT	2.00	6.08	0.09	Ser	AGT	2.00	6.08	0.09
Ser	AGC	5.00	15.20	0.22	Ser	AGC	5.00	15.20	0.22
Ser	TCG	2.00	6.08	0.09	Ser	TCG	2.00	6.08	0.09
Ser	TCA	3.00	9.12	0.13	Ser	TCA	3.00	9.12	0.13
Ser	TCT	4.00	12.16	0.17	Ser	TCT	4.00	12.16	0.17
Ser	TCC	7.00	21.28	0.30	Ser	TCC	7.00	21.28	0.30
Thr	ACG	0.00	0.00	0.00	Thr	ACG	0.00	0.00	0.00
Thr	ACA	5.00	15.20	0.21	Thr	ACA	5.00	15.20	0.21
Thr	ACT	11.00	33.43	0.46	Thr	ACT	11.00	33.43	0.46
Thr	ACC	8.00	24.32	0.33	Thr	ACC	8.00	24.32	0.33
Val	GTG	21.00	63.83	0.64	Val	GTG	21.00	63.83	0.64
Val	GTA	5.00	15.20	0.15	Val	GTA	5.00	15.20	0.15
Val	GTT	4.00	12.16	0.12	Val	GTT	4.00	12.16	0.12
Val	GTC	3.00	9.12	0.09	Val	GTC	3.00	9.12	0.09
Trp	TGG	2.00	6.08	1.00	Trp	TGG	2.00	6.08	1.00
Tyr	TAT	2.00	6.08	0.50	Tyr	TAT	2.00	6.08	0.50
Tyr	TAC	2.00	6.08	0.50	Tyr	TAC	2.00	6.08	0.50
End	TGA	1.00	3.04	1.00	End	TGA	1.00	3.04	1.00
End	TAG	0.00	0.00	0.00	End	TAG	0.00	0.00	0.00
End	TAA	0.00	0.00	0.00	End	TAA	0.00	0.00	0.00

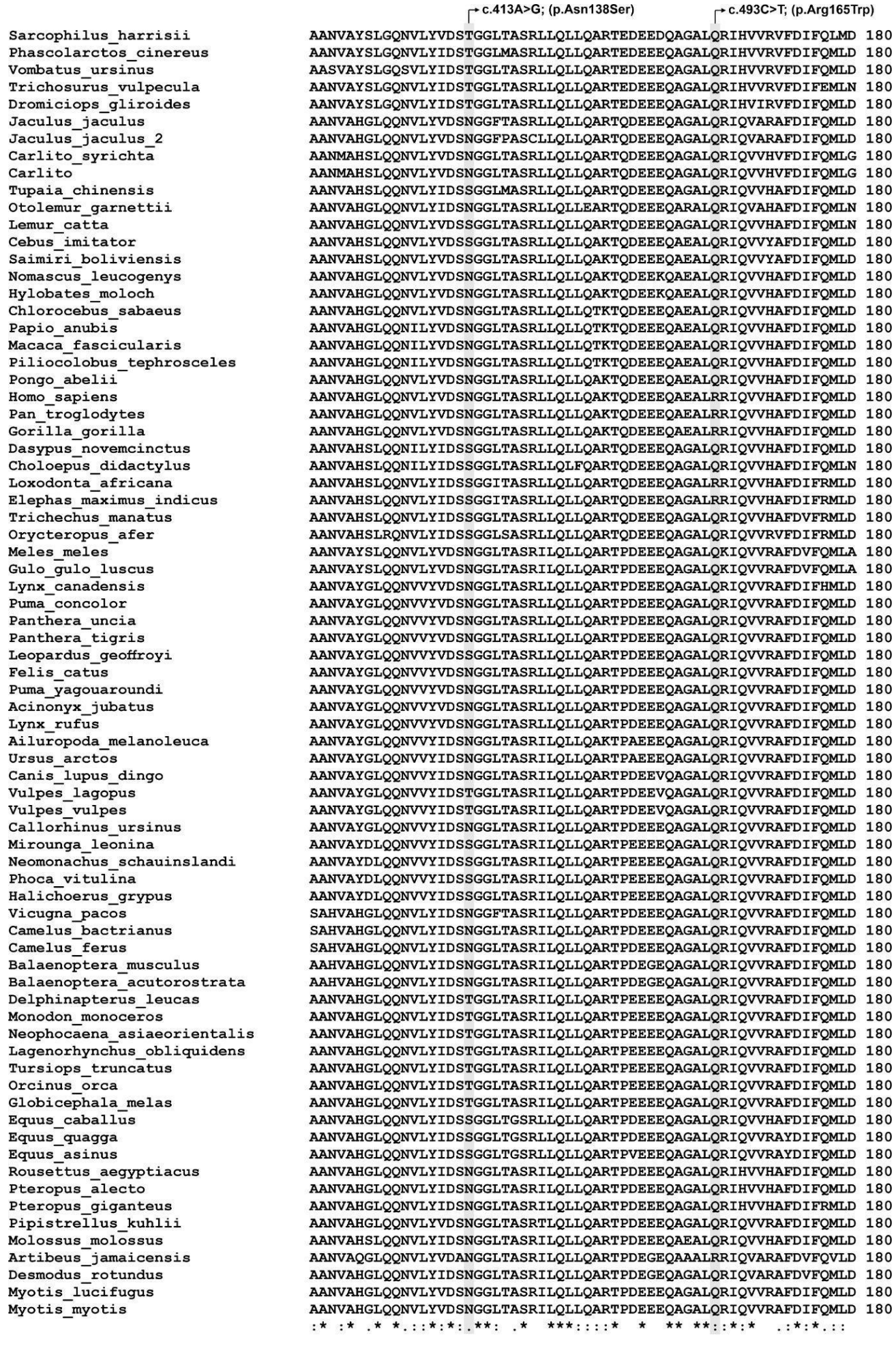
Nota: Elaborado pela autora, usando SMS codon usage Stothard P (2000) The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102-1104.

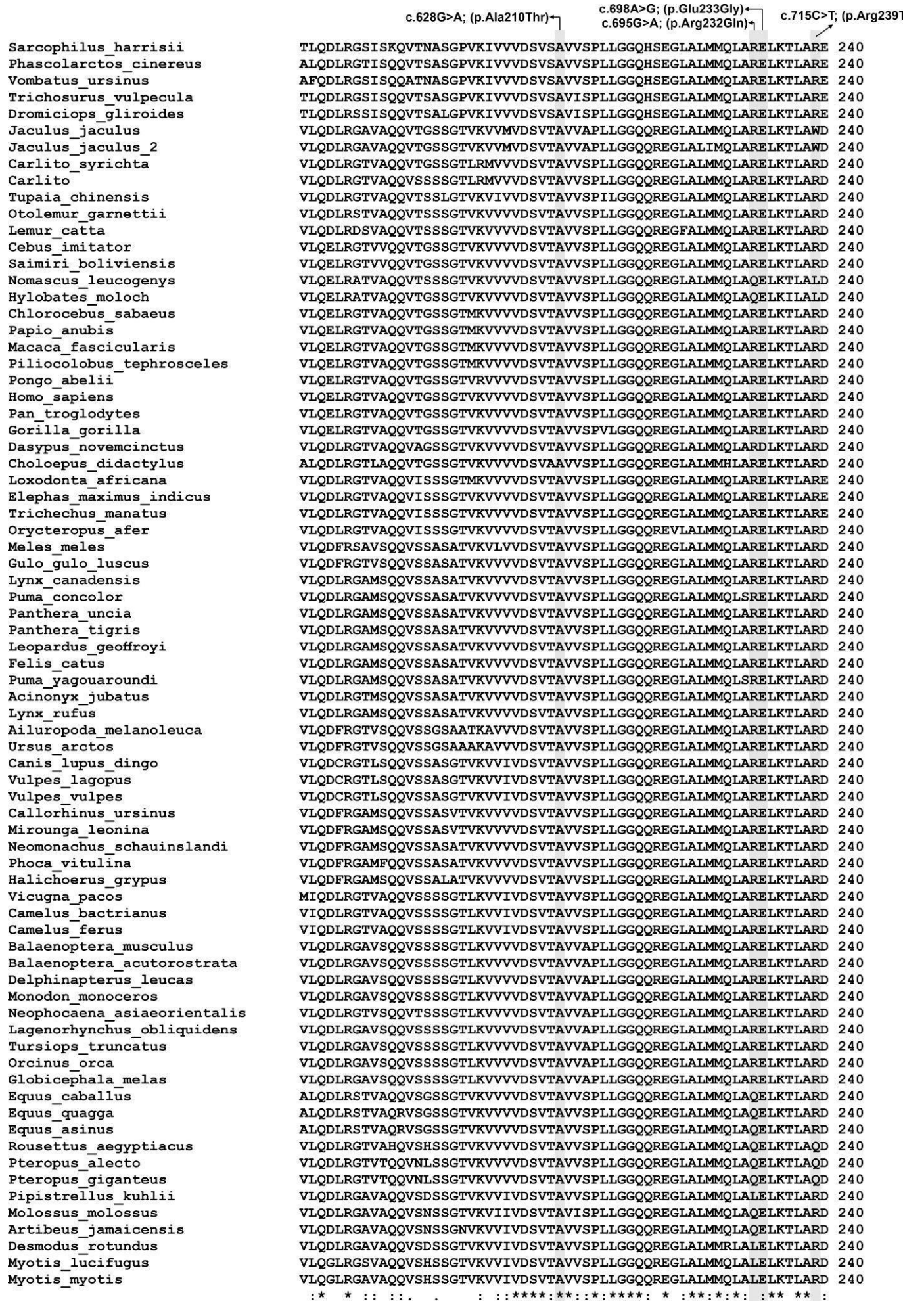
ANEXO 4 - Alinhamento múltiplo da isoforma canônica (O75771-1)

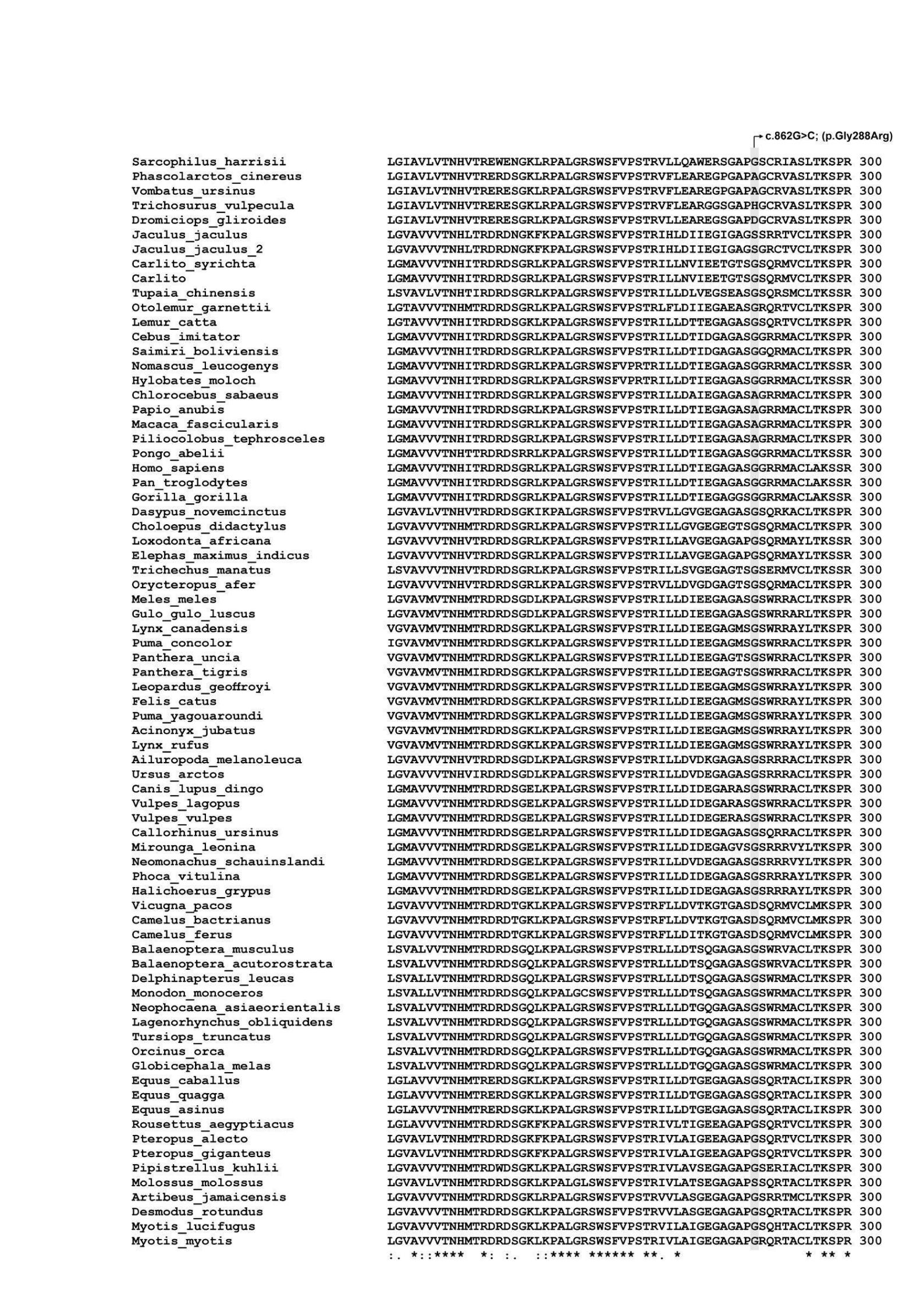
Alinhamento com mamíferos

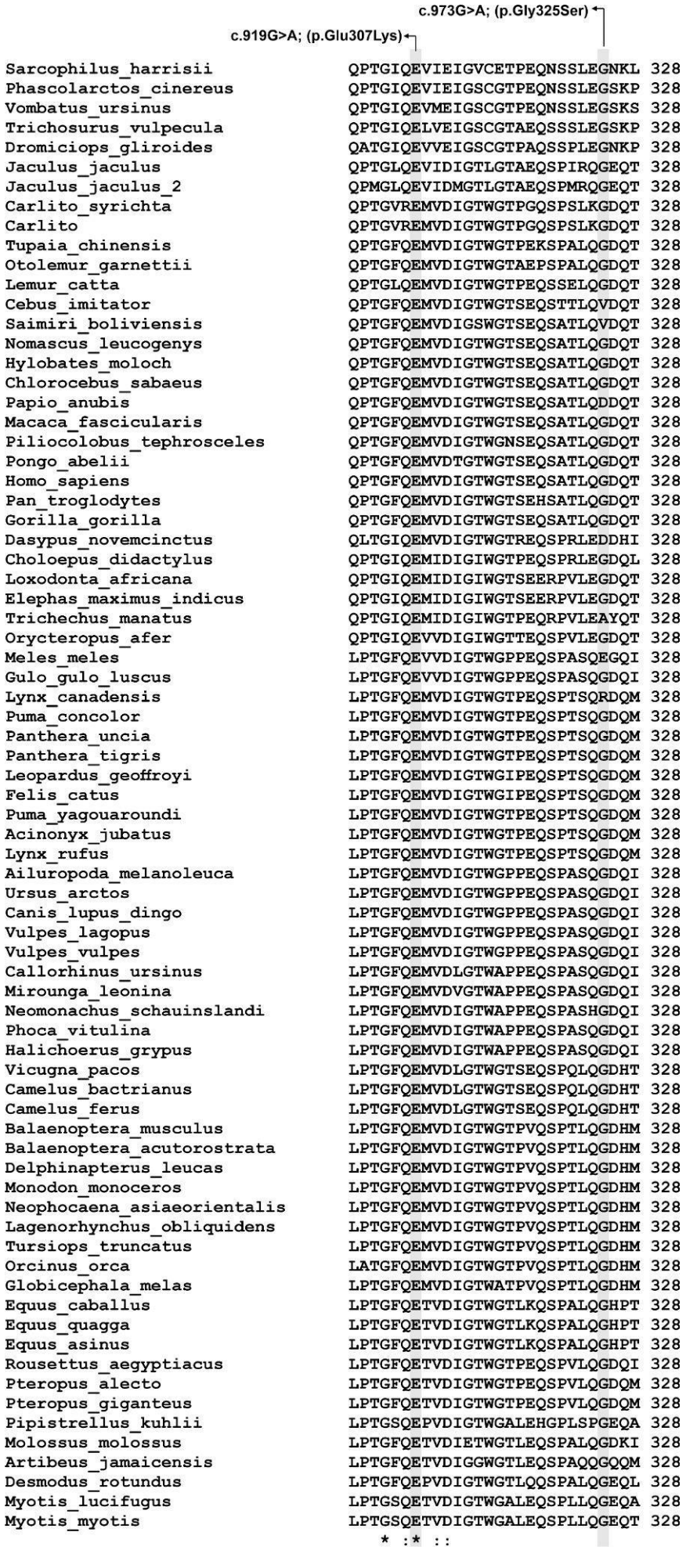

c.26G>C; (p.Cys9Ser)
c.137C>G; (p.Ser46Cys)
c.140A>G; (p.Tyr47Cys)
Sarcophilus_harrisii MGVLRAGLCPGLTQDVVQLLRGVGVRTVVDLVSADLEEVSRKCSLSYKALVAVRRVLLAQ 60
Phascolarctos_cinereus MGVLRMGLCPGLTQEMIQLLRGMGIKTVVDLVSSDLEEVSRKCCLSYKALVSVRRVLLVQ 60
Vombatus_ursinus MGVLRVGLCPALTQEMIQLLRGIGIKTVVDLVSADLEEVSRKCSLSYKALVAVRRVLLVQ 60
Trichosurus_vulpecula MGVLRAGLCPGLTQDMIQLLRGIGIRTVVDLVSADLEEVSRKCSLSYKALVAVRRVLLAQ 60
Dromiciops_gliroides MGVLREGLCPGLTQEMIQLLKGIGIKTVVDFVSADLEEVSQKCSLSYKALVAVRRVLLAQ 60
Jaculus_jaculus MGVLRAGLCPGLTEDMVQLLRGRGIKTVVDLAAANLEEVAQQCGLSYKALVALRRVLLAQ 60
Jaculus_jaculus_2 MGVLRAGLCPGLAEDMVQLLRGRGIKTVVDLAAANLEEVAQQCGLSYKALVALRRVLLTQ 60
Carlito_syrichta MGVLRVGLCPGLTQEMVQLLRSRGIKTVVDLVSADLEEVAQIRGLSYKALVALRRVLLAQ 60
Carlito MGVLRVGLCPGLTQEMVQLLRSRGIKTVVDLVSADLEEVAQIRGLSYKALVALRRVLLAQ 60
Tupaia_chinensis MGVLRVGLCPGLTQEMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Otolemur_garnettii MGVLRVGLCPGLTQEMVHLLRSRGIKTVVDLVSANLEEVAQKCGLSYKALVALRRVLLAQ 60
Lemur_catta MGVLRVGLCPGLTQEMIQLLGSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Cebus_imitator MGVLRVGLCPGLTQEMIQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Saimiri_boliviensis MGVLRVGLCPGLTQEMIQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Nomascus_leucogenys MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Hylobates_moloch MGVLRVGLCPGLTQEMIQLLRSRKIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Chlorocebus_sabaeus MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Papio_anubis MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Macaca_fascicularis MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Piliocolobus_tephrosceles MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Pongo_abelii MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEELAQKCGLSYKALVALRRVLLAQ 60
Homo_sapiens MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Pan_troglodytes MGVLRVGLCPGLTQEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Gorilla_gorilla MGVLRVGLCPGLTQEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Dasypus_novemcinctus MGVLRAGLCPGLTQEMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Choloepus_didactylus MGVLRVGLCPGLTQEMVQLLKSRGIKTVVDLVSADLEEVAQRCGLSYKALVALRRVLLAQ 60
Loxodonta_africana MGVLRAALCPGLTQEMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVAVRRVLLAQ 60
Elephas_maximus_indicus MGVLRAALCPGLTQEMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVAVRRVLLAQ 60
Trichechus_manatus MGVLRAGLCPGLTQEIVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVAVRRVLLAQ 60
Orycteropus_afer MGVLRDGLCPGLTQEMVQLLRNRGIKTVVDLVSADLEEIAQKCGLSYKALVAVRRVLLAQ 60
Meles_meles MGVLRAGLCPGLTQDTVGQLRTRGIRTVVDLVSADLEDVAQKCALSYKALVALRRVLLAQ 60
Gulo_gulo_luscus MGVLRAGLCPGLTEDTVGQLRTRGIRTVVDLVSADLEDVAQKCALSYKALVALRRVLLAQ 60
Lynx_canadensis MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Puma_concolor MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Panthera_uncia MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Panthera_tigris MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Leopardus_geoffroyi MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Felis_catus MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Puma_yagouaroundi MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Acinonyx_jubatus MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Lynx_rufus MGVLRAGLCPGLTQDTVQLLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Ailuropoda_melanoleuca MGVLRAGLCPGLTQDTVGQLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Ursus_arctos MGVLRAGLCPGLTQDTVGQLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Canis_lupus_dingo MGVLRAGLCPGLTQDTVLQLRSRGVRTVVDLASADLEAVAQTCGLSYKALVALRRVLLAQ 60
Vulpes_lagopus MGVLRAGLCPGLTPDTVLQLRSRGIRTVVDLVSADLEAVAQKCGLSYKALVALRRVLLAQ 60
Vulpes_vulpes MGVLRAGLCPGLTPDTVLQLRSRGISTVVDLVSADLEAVAQKCGLSYKALVALRRVLLAQ 60
Callorhinus_ursinus MGVLRAGLCPGLTHDTVGQLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Mirounga_leonina MGVLRAGLCPGLTHDMVGQLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Neomonachus_schauinslandi MGVLRAGLCPGLTHDTVGQLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Phoca_vitulina MGVLRAGLCPGLTHDTVGQLRSRGIKTVVDLVSADLEDVAQKCGLSYKALVALRRVLLAQ 60
Halichoerus_grypus MGVLRAGLCPGLTHDTVGQLRSRGIKTVVDLASADLEDVAQKCGLSYKALVALRRVLLAQ 60
Vicugna_pacos MGVLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Camelus_bactrianus MGVLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Camelus_ferus MGVLRAGLCPGLTQNMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Balaenoptera_musculus MGVLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Balaenoptera_acutorostrata MGVLRAGLCPGLTQDVVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Delphinapterus_leucas MGVLRAGLCPGLTQDMAQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Monodon_monoceros MGVLRAGLCPGLTQDMAQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Neophocaena_asiaeorientalis MGVLRAGLCPGLTQDMAQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Lagenorhynchus_obliquidens MGLLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Tursiops_truncatus MGVLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVAQRRVLLAQ 60
Orcinus_orca MGVLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Globicephala_melas MGVLRAGLCPGLTQDMVQLLRSRGIKTVVDLVSADLEEVAQKCGLSYKVLVALRRVLLAQ 60
Equus_caballus MGLLRAGLCPGLTQDVVQLLRGRGIRTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Equus_quagga MGLLRAGLCPGLTQDVVQLLRGRGIRTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Equus_asinus MGLLRAGLCPGLTQDVVQLLRGRGIRTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Rousettus_aegyptiacus MGVLRAGLCPGLTEDMVQLLRSRGIKTVVDLVSTDLEEVALKCGLSYKVLVALRRVLLAQ 60
Pteropus_alecto MGVLRAGLCPGLTEDMVQLLRSRGIKTVVDLVSADLEEVALKCGLSYKALVALRRVLLAQ 60
Pteropus_giganteus MGVLRAGLCPGLTEDMVQLLRSRGIKTVVDLVSADLEEVALKCGLSYKALVALRRVLLAQ 60
Pipistrellus_kuhlii MGVLRAGLCPGLTQDMVQLLRSRGIRTVAALVSADLEEVARTCGLSYKALVALRRVLLAQ 60
Molossus_molossus MGVLRAGLCPGLTQDMVQLLRSHGIKTVVDLVSADLEEAAQKCGLSYKALVALRRVLLAQ 60
Artibeus_jamaicensis MGVLRAGLCPGLTQDMVQLLRSHGVKTVADLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Desmodus_rotundus MGVLRAGLCPGLTQDMVQLLRSHGIKTVADLVSADLEEVAQKCGLSYKVLVALRRVLLAQ 60
Myotis_lucifugus MGVLRAGLCPGLTQDMVQLLRSHGIKTVVDLASADLEEVAQKCGLSYKALVALRRVLLAQ 60
Myotis_myotis MGVLRAGLCPGLTQDVVQLLRSHGIKTVVDLASADLEEVAQKCGLSYKALVALRRVLLAQ 60


c.196G>A; (p.Val66Met)
c.355T>C; (p.Cys119Arg)
Sarcophilus_harrisii FSSFPFNGADLYEEFKSSTAILATGIASLDKLLDSGLYTGEVTEITGAPASGKTQVCLCT 120
Phascolarctos_cinereus FSSFPFNGADLYEELKTSTAILATGIASLDKLLDSGLYTGEVTEITGPPASGKTQVCLCT 120
Vombatus_ursinus FSSFPFNGADLYEELKTSTAILATGIASLDKLLDSGLYTGEVTEITGPPASGKTQVCLCT 120
Trichosurus_vulpecula FSSFPFNGADLYEELKTSTAILATGIASLDKLLDSGLYTGEVTEITGAPASGKTQVCLCT 120
Dromiciops_gliroides FSSFPFNGADLYEELKTTTAILATGIASLDKLLDSGLYTGEVTEITGAPASGKTQVCLCT 120
Jaculus_jaculus FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEMTEIVGGPGSGKTQVCLCV 120
Jaculus_jaculus_2 FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEMTEIVGGPGSGKTQVCVCV 120
Carlito_syrichta FSAFPVNGADLYEELKTSTAILSTGLGSLDKLLDAGLYTGEVTEIVGGPGSGKTQICLCV 120
Carlito FSAFPVNGADLYEELKTSTAILSTGLGSLDKLLDAGLYTGEVTEIVGGPGSGKTQICLCV 120
Tupaia_chinensis FSAFPFNGADLYEELKTSTAILATGIGSLDKLLDAGLYTGEVTEIVGGPRSGKTQVCLCV 120
Otolemur_garnettii FSAFPINGADLYEELKTSTAILSTGIASLDKLLDTGLYTGEVTEIVGGPGSGKTQICLCV 120
Lemur_catta FSAFPFNGADLYEELKTSTAILSTGIDSLDKLLDTGLYTGEVTEIVGGPGSGKTQVCLCV 120
Cebus_imitator FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Saimiri_boliviensis FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Nomascus_leucogenys FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Hylobates_moloch FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Chlorocebus_sabaeus FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Papio_anubis FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Macaca_fascicularis FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Piliocolobus_tephrosceles FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Pongo_abelii FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Homo_sapiens FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM 120
Pan_troglodytes FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM 120
Gorilla_gorilla FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM 120
Dasypus_novemcinctus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQLCLCV 120
Choloepus_didactylus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGVYTGEVTEIVGGPGSGKTQICLYV 120
Loxodonta_africana FSAFPFNGADVYEELKTSTAILSTGMERLDKLLDTGLYTGEVTEFVGSPGSGKTQVCLCV 120
Elephas_maximus_indicus FSAFPFNGADVYEELKTSTAILSTGMERLDKLLDTGLYTGEVTEFIGSPGSGKTQVCLCV 120
Trichechus_manatus FSAFPFNGADVYEELKTSTAILSTGIGRLDKLLDTGLYTGEVTEIVGSPGSGKTQVCLCV 120
Orycteropus_afer FSAFPFNGADLYEELKSSTAILSTGIESLDKLLDAGLYTGEVTEFVGSPGSGKTQVCLFV 120
Meles_meles FSAFPCNGADLYEELKTSTAILSTGIRSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Gulo_gulo_luscus FSAFPCNGADLYEELKTSTAILSTGIQSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Lynx_canadensis FSAFPFNGADLYEELKTSTAILSTGIRSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Puma_concolor FSAFPFNGADLYEELKTSTAILSTGVGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Panthera_uncia FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Panthera_tigris FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Leopardus_geoffroyi FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Felis_catus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Puma_yagouaroundi FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Acinonyx_jubatus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Lynx_rufus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Ailuropoda_melanoleuca FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Ursus_arctos FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIAGGPGSGKTQVCLCV 120
Canis_lupus_dingo FSAFPSNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Vulpes_lagopus FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Vulpes_vulpes FSAFPLNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Callorhinus_ursinus FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Mirounga_leonina FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Neomonachus_schauinslandi FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Phoca_vitulina FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Halichoerus_grypus FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Vicugna_pacos FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Camelus_bactrianus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Camelus_ferus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Balaenoptera_musculus FSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Balaenoptera_acutorostrata FSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Delphinapterus_leucas LSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLHV 120
Monodon_monoceros LSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLHV 120
Neophocaena_asiaeorientalis LSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLRV 120
Lagenorhynchus_obliquidens FSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLRV 120
Tursiops_truncatus FSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLRV 120
Orcinus_orca FSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLRV 120
Globicephala_melas FSAFPFNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLRV 120
Equus_caballus FSAFPFNGADLYEELKTSTAILSTGIESLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Equus_quagga FSAFPFNGSDLYEELKTSTAILSTGIESLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Equus_asinus FSAFPFNGADLYEELKTSTAILSTGIESLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCV 120
Rousettus_aegyptiacus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIAGGPGSGKTQVCLCV 120
Pteropus_alecto FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIAGGPGSGKTQVCLCV 120
Pteropus_giganteus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIAGGPGSGKTQVCLCV 120
Pipistrellus_kuhlii FSAFPCNGADLYEELKTSTAILSTGIASLDKLLDAGLYTGEVTEISGGPGSGKTQVCLCV 120
Molossus_molossus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEITGVPGSGKTQVCLCV 120
Artibeus_jamaicensis FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGVPGSGKTQVCLCV 120
Desmodus_rotundus FSAFPFNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGVPGSGKTQVCLCV 120
Myotis_lucifugus FSAFPCNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGVPGSGKTQVCLCV 120
Myotis_myotis FSAFPCNGADLYEELKTATAILSTGIGSLDKLLDAGLYTGEVTEIVGVPGSGKTQVCLCV 120
:*:** **:*:***:*::****:**: ******:*:****:**: * * *****:*:


c.413A>G; (p.Asn138Ser)
c.493C>T; (p.Arg165Trp)
Sarcophilus_harrisii AANVAYSLGQNVLYVDSTGGLTASRLLQLLQARTEDEEDQAGALQRIHVVRVFDIFQLMD 180
Phascolarctos_cinereus AANVAYSLGQNVLYIDSTGGLMASRLLQLLQARTEDEEEQAGALQRIHVVRVFDIFQMLD 180
Vombatus_ursinus AASVAYSLGQSVLYIDSTGGLTASRLLQLLQARTEDEEEQAGALQRIHVVRVFDIFQMLD 180
Trichosurus_vulpecula AANVAYSLGQNVLYIDSTGGLTASRLLQLLQARTEDEEEQAGALQRIHVVRVFDIFEMLN 180
Dromiciops_gliroides AANVAYSLGQNVLYIDSTGGLTASRLLQLLQARTEDEEEQAGALQRIHVIRVFDIFQMLD 180
Jaculus_jaculus AANVAHGLQQNVLYVDSNGGFTASRLLQLLQARTQDEEEQAGALQRIQVARAFDIFQMLD 180
Jaculus_jaculus_2 AANVAHGLQQNVLYVDSNGGFPASCLLQLLQARTQDEEEQAGALQRIQVARAFDIFQMLD 180
Carlito_syrichta AANMAHSLQQNVLYVDSNGGLTASRLLQLLQARTQDEEEQAGALQRIQVVHVFDIFQMLG 180
Carlito AANMAHSLQQNVLYVDSNGGLTASRLLQLLQARTQDEEEQAGALQRIQVVHVFDIFQMLG 180
Tupaia_chinensis AANVAHSLQQNVLYIDSSGGLMASRLLQLLQARTQDEEEQAGALQRIQVVHAFDIFQMLD 180
Otolemur_garnettii AANVAHGLQQNVLYIDSNGGLTASRLLQLLEARTQDEEEQARALQRIQVAHAFDIFQMLN 180
Lemur_catta AANVAHGLQQNVLYVDSSGGLTASRLLQLLQARTQDEEEQAGALQRIQVVHAFDIFQMLN 180
Cebus_imitator AANVAHSLQQNVLYVDSSGGLTASRLLQLLQAKTQDEEEQAEALQRIQVVYAFDIFQMLD 180
Saimiri_boliviensis AANVAHSLQQNVLYVDSSGGLTASRLLQLLQAKTQDEEEQAEALQRIQVVYAFDIFQMLD 180
Nomascus_leucogenys AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEKQAEALQRIQVVHAFDIFQMLD 180
Hylobates_moloch AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEKQAEALQRIQVVHAFDIFQMLD 180
Chlorocebus_sabaeus AANVAHGLQQNVLYVDSNGGLTASRLLQLLQTKTQDEEEQAEALQRIQVVHAFDIFQMLD 180
Papio_anubis AANVAHGLQQNILYVDSNGGLTASRLLQLLQTKTQDEEEQAEALQRIQVVHAFDIFQMLD 180
Macaca_fascicularis AANVAHGLQQNILYVDSNGGLTASRLLQLLQTKTQDEEEQAEALQRIQVVHAFDIFQMLD 180
Piliocolobus_tephrosceles AANVAHGLQQNILYVDSNGGLTASRLLQLLQTKTQDEEEQAEALQRIQVVHAFDIFQMLD 180
Pongo_abelii AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALQRIQVVHAFDIFQMLD 180
Homo_sapiens AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD 180
Pan_troglodytes AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD 180
Gorilla_gorilla AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD 180
Dasypus_novemcinctus AANVAHSLQQNILYIDSSGGLTASRLLQLLQARTQDEEEQAGALQRIQVVHAFDIFQMLD 180
Choloepus_didactylus AANVAHSLQQNILYIDSSGGLTASRLLQLFQARTQDEEEQAGALQRIQVVHAFDIFQMLN 180
Loxodonta_africana AANVAHSLQQNVLYIDSSGGITASRLLQLLQARTQDEEEQAGALRRIQVVHAFDIFRMLD 180
Elephas_maximus_indicus AANVAHSLQQNVLYIDSSGGITASRLLQLLQARTQDEEEQAGALRRIQVVHAFDIFRMLD 180
Trichechus_manatus AANVAHSLQQNVLYIDSSGGLTASRLLQLLQARTQDEEEQAGALQRIQVVHAFDVFRMLD 180
Orycteropus_afer AANVAHSLRQNVLYIDSSGGLSASRLLQLLQARTQDEEEQAGALQRIQVVRVFDIFRMLD 180
Meles_meles AANVAYSLQQNVLYVDSNGGLTASRILQLLQARTPDEEEQAGALQKIQVVRAFDVFQMLA 180
Gulo_gulo_luscus AANVAYSLQQNVLYVDSNGGLTASRILQLLQARTPDEEEQAGALQKIQVVRAFDVFQMLA 180
Lynx_canadensis AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFHMLD 180
Puma_concolor AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Panthera_uncia AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Panthera_tigris AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Leopardus_geoffroyi AANVAYGLQQNVVYVDSSGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Felis_catus AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Puma_yagouaroundi AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Acinonyx_jubatus AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Lynx_rufus AANVAYGLQQNVVYVDSNGGLTASRLLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Ailuropoda_melanoleuca AANVAYGLQQNVVYIDSNGGLTASRILQLLQAKTPAEEEQAGALQRIQVVRAFDIFQMLD 180
Ursus_arctos AANVAYGLQQNVVYIDSNGGLTASRILQLLQARTPAEEEQAGALQRIQVVRAFDIFQMLD 180
Canis_lupus_dingo AANVAYGLQQNVVYIDSNGGLTASRILQLLQARTPDEEVQAGALQRIQVVRAFDIFQMLD 180
Vulpes_lagopus AANVAYGLQQNVVYIDSTGGLTASRILQLLQARTPDEEVQAGALQRIQVVRAFDIFQMLD 180
Vulpes_vulpes AANVAYGLQQNVVYIDSTGGLTASRILQLLQARTPDEEVQAGALQRIQVVRAFDIFQMLD 180
Callorhinus_ursinus AANVAYGLQQNVVYIDSNGGLTASRILQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Mirounga_leonina AANVAYDLQQNVVYIDSSGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Neomonachus_schauinslandi AANVAYDLQQNVVYIDSSGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Phoca_vitulina AANVAYDLQQNVVYIDSSGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Halichoerus_grypus AANVAYDLQQNVVYIDSSGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Vicugna_pacos SAHVAHGLQQNVLYIDSNGGFTASRILQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Camelus_bactrianus SAHVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Camelus_ferus SAHVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Balaenoptera_musculus AAHVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEGEQAGALQRIQVVRAFDIFQMLD 180
Balaenoptera_acutorostrata AAHVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEGEQAGALQRIQVVRAFDIFQMLD 180
Delphinapterus_leucas AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Monodon_monoceros AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Neophocaena_asiaeorientalis AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Lagenorhynchus_obliquidens AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Tursiops_truncatus AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Orcinus_orca AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Globicephala_melas AANVAHGLQQNVLYIDSTGGLTASRILQLLQARTPEEEEQAGALQRIQVVRAFDIFQMLD 180
Equus_caballus AANVAHGLQQNVLYIDSSGGLTGSRLLQLLQARTPDEEEQAGALQRIQVVHAFDIFQMLD 180
Equus_quagga AANVAHGLQQNVLYIDSSGGLTGSRLLQLLQARTPDEEEQAGALQRIQVVRAYDIFQMLD 180
Equus_asinus AANVAHGLQQNVLYIDSSGGLTGSRLLQLLQARTPVEEEQAGALQRIQVVRAYDIFQMLD 180
Rousettus_aegyptiacus AANVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEEEQAGALQRIHVVHAFDIFQMLD 180
Pteropus_alecto AANVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEEEQAGALQRIHVVHAFDIFQMLD 180
Pteropus_giganteus AANVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEEEQAGALQRIHVVHAFDIFRMLD 180
Pipistrellus_kuhlii AANVAHGLQQNVLYVDSNGGLTASRTLQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Molossus_molossus AANVAHSLQQNVLYIDSNGGLTASRILQLLQARTPDEEEQAEALQRIQVVHAFDIFQMLD 180
Artibeus_jamaicensis AANVAQGLQQNVLYVDANGGLTASRILQLLQARTPDEGEQAAALRRIQVARAFDVFQVLD 180
Desmodus_rotundus AANVAHGLQQNVLYIDSNGGLTASRILQLLQARTPDEGEQAGALQRIQVARAFDVFQMLD 180
Myotis_lucifugus AANVAHGLQQNVLYVDSNGGLTASRILQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180
Myotis_myotis AANVAHGLQQNVLYVDSNGGLTASRILQLLQARTPDEEEQAGALQRIQVVRAFDIFQMLD 180


c.628G>A; (p.Ala210Thr)
c.698A>G; (p.Glu233Gly)
c.695G>A; (p.Arg232Gln)
c.715C>T; (p.Arg239T
Sarcophilus_harrisii TLQDLRGSISKQVTNASGPVKIVVVDSVSAVVSPLLGGQHSEGLALMMQLARELKTLARE 240
Phascolarctos_cinereus ALQDLRGTISQQVTSASGPVKIVVVDSVSAVVSPLLGGQHSEGLALMMQLARELKTLARE 240
Vombatus_ursinus AFQDLRGSISQQATNASGPVKIVVVDSVSAVVSPLLGGQHSEGLALMMQLARELKTLARE 240
Trichosurus_vulpecula TLQDLRGSISQQVTSASGPVKIVVVDSVSAVISPLLGGQHSEGLALMMQLARELKTLARE 240
Dromiciops_gliroides TLQDLRSSISQQVTSALGPVKIVVVDSVSAVISPLLGGQHSEGLALMMQLARELKTLARE 240
Jaculus_jaculus VLQDLRGAVAQQVTGSSGTVKVVMVDSVTAVVAPLLGGQQREGLALMMQLARELKTLAWD 240
Jaculus_jaculus_2 VLQDLRGAVAQQVTGSSGTVKVVMVDSVTAVVAPLLGGQQREGLALIMQLARELKTLAWD 240
Carlito_syrichta VLQDLRGTVAQQVTGSSGTLRMVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Carlito VLQDLRGTVAQQVSSSSGTLRMVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Tupaia_chinensis VLQDLRGTVAQQVTSSLGTVKVIVVDSVTAVVSPILGGQQREGLALMMQLARELKTLARD 240
Otolemur_garnettii VLQDLRSTVAQQVTSSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Lemur_catta VLQDLRDSVAQQVTSSSGTVKVVVVDSVTAVVSPLLGGQQREGFALMMQLARELKTLARD 240
Cebus_imitator VLQELRGTVVQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Saimiri_boliviensis VLQELRGTVVQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Nomascus_leucogenys VLQELRATVAQQVTSSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLALD 240
Hylobates_moloch VLQELRATVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKILALD 240
Chlorocebus_sabaeus VLQELRGTVAQQVTGSSGTMKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Papio_anubis VLQELRGTVAQQVTGSSGTMKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Macaca_fascicularis VLQELRGTVAQQVTGSSGTMKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Piliocolobus_tephrosceles VLQELRGTVAQQVTGSSGTMKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Pongo_abelii VLQELRGTVAQQVTGSSGTVRVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Homo_sapiens VLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Pan_troglodytes VLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Gorilla_gorilla VLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPVLGGQQREGLALMMQLARELKTLARD 240
Dasypus_novemcinctus VLQDLRGTVAQQVAGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARDLKTLARD 240
Choloepus_didactylus ALQDLRGTLAQQVTGSSGTVKVVVVDSVAAVVSPLLGGQQREGLALMMHLARELKTLARD 240
Loxodonta_africana VLQDLRGTVAQQVISSSGTMKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARE 240
Elephas_maximus_indicus VLQDLRGTVAQQVISSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARE 240
Trichechus_manatus VLQDLRGTVAQQVISSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARE 240
Orycteropus_afer VLQDLRGTVAQQVISSSGTVKVVVVDSVTAVVSPLLGGQQREVLALMMQLARELKTLARD 240
Meles_meles VLQDFRSAVSQQVSSASATVKVLVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Gulo_gulo_luscus VLQDFRGTVSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Lynx_canadensis VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Puma_concolor VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLSRELKTLARD 240
Panthera_uncia VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Panthera_tigris VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Leopardus_geoffroyi VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Felis_catus VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Puma_yagouaroundi VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLSRELKTLARD 240
Acinonyx_jubatus VLQDLRGTMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Lynx_rufus VLQDLRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Ailuropoda_melanoleuca VLQDFRGTVSQQVSSGSAATKAVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Ursus_arctos VLQDFRGTVSQQVSSGSAAAKAVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Canis_lupus_dingo VLQDCRGTLSQQVSSASGTVKVVIVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Vulpes_lagopus VLQDCRGTLSQQVSSASGTVKVVIVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Vulpes_vulpes VLQDCRGTLSQQVSSASGTVKVVIVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Callorhinus_ursinus VLQDFRGAMSQQVSSASVTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Mirounga_leonina VLQDFRGAMSQQVSSASVTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Neomonachus_schauinslandi VLQDFRGAMSQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Phoca_vitulina VLQDFRGAMFQQVSSASATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Halichoerus_grypus VLQDFRGAMSQQVSSALATVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Vicugna_pacos MIQDLRGTVAQQVSSSSGTLKVVIVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Camelus_bactrianus VIQDLRGTVAQQVSSSSGTLKVVIVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Camelus_ferus VIQDLRGTVAQQVSSSSGTLKVVIVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Balaenoptera_musculus VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Balaenoptera_acutorostrata VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Delphinapterus_leucas VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Monodon_monoceros VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Neophocaena_asiaeorientalis VLQDLRGTVSQQVTSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Lagenorhynchus_obliquidens VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Tursiops_truncatus VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Orcinus_orca VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Globicephala_melas VLQDLRGAVSQQVSSSSGTLKVVVVDSVTAVVAPLLGGQQREGLALMMQLARELKTLARD 240
Equus_caballus ALQDLRSTVAQQVSGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLARD 240
Equus_quagga ALQDLRSTVAQRVSGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLARD 240
Equus_asinus ALQDLRSTVAQRVSGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLARD 240
Rousettus_aegyptiacus VLQDLRGTVAHQVSHSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLAQD 240
Pteropus_alecto VLQDLRGTVTQQVNLSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLAQD 240
Pteropus_giganteus VLQDLRGTVTQQVNLSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLAQD 240
Pipistrellus_kuhlii VLQDLRGAVAQQVSDSSGTVKVVIVDSVTAVVSPLLGGQQREGLALMMQLALELKTLARD 240
Molossus_molossus VLQDLRGAVAQQVSNSSGTVKVIIVDSVTAVISPLLGGQQREGLALMMQLAQELKTLARD 240
Artibeus_jamaicensis VLQDLRGAVAQQVSNSSGNVKVVIVDSVTAVVSPLLGGQQREGLALMMQLAQELKTLARD 240
Desmodus_rotundus VLQDLRGAVAQQVSDSSGTVKVVIVDSVTAVVSPLLGGQQREGLALMMRLALELKTLARD 240
Myotis_lucifugus VLQGLRGSVAQQVSHSSGTVKVVIVDSVTAVVSPLLGGQQREGLALMMQLALELKTLARD 240
Myotis_myotis VLQGLRGAVAQQVSHSSGTVKVVIVDSVTAVVSPLLGGQQREGLALMMQLALELKTLARD 240

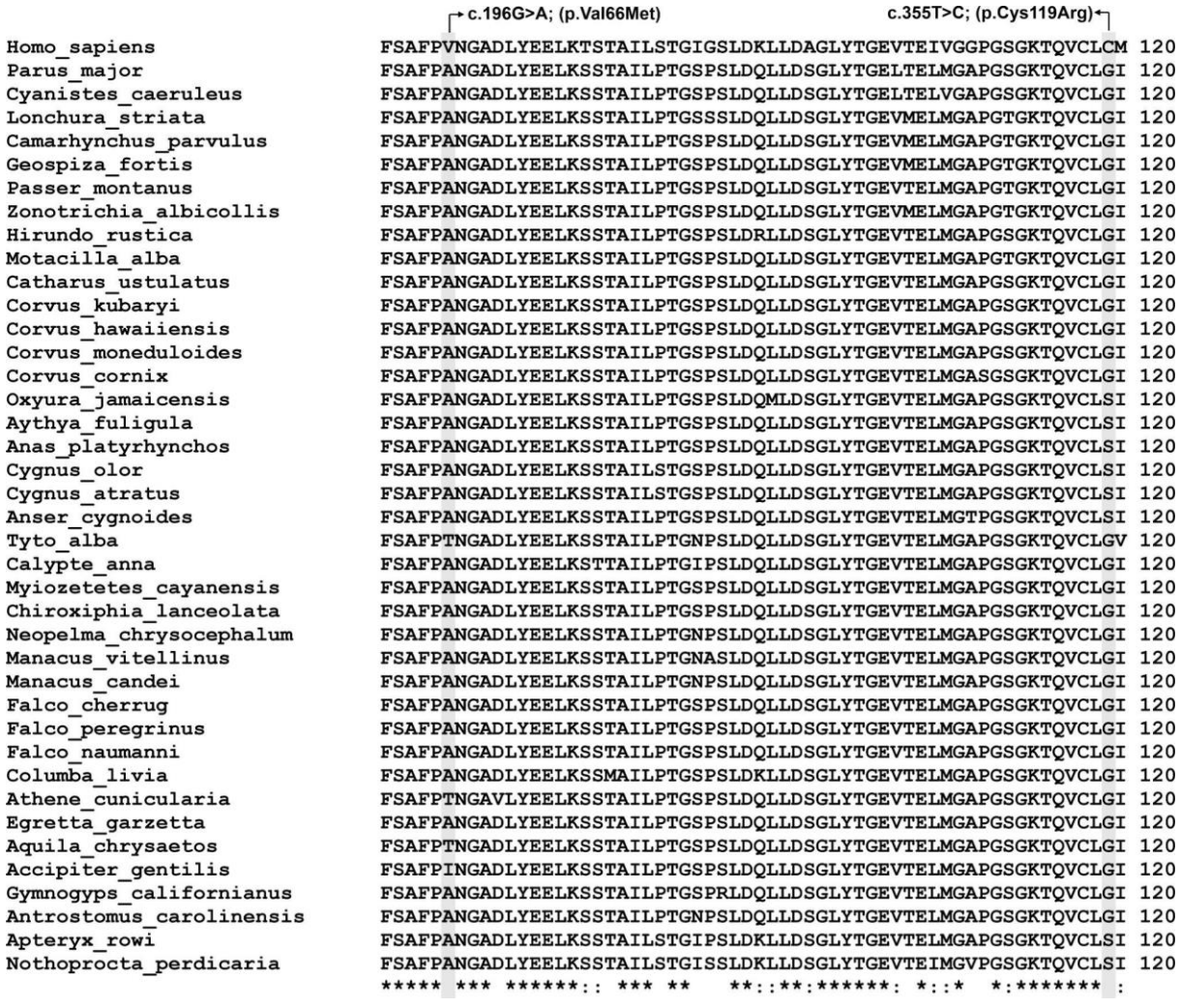

c.862G>C; (p.Gly288Arg)
Sarcophilus_harrisii LGIAVLVTNHVTREWENGKLRPALGRSWSFVPSTRVLLQAWERSGAPGSCRIASLTKSPR 300
Phascolarctos_cinereus LGIAVLVTNHVTRERDSGKLRPALGRSWSFVPSTRVFLEAREGPGAPAGCRVASLTKSPR 300
Vombatus_ursinus LGIAVLVTNHVTRERESGKLRPALGRSWSFVPSTRVFLEAREGPGAPAGCRVASLTKSPR 300
Trichosurus_vulpecula LGIAVLVTNHVTRERESGKLRPALGRSWSFVPSTRVFLEARGGSGAPHGCRVASLTKSPR 300
Dromiciops_gliroides LGIAVLVTNHVTRERESGRLKPALGRSWSFVPSTRVLLEAREGSGAPDGCRVASLTKSPR 300
Jaculus_jaculus LGVAVVVTNHLTRDRDNGKFKPALGRSWSFVPSTRIHLDIIEGIGAGSSRRTVCLTKSPR 300
Jaculus_jaculus_2 LGVAVVVTNHLTRDRDNGKFKPALGRSWSFVPSTRIHLDIIEGIGAGSGRCTVCLTKSPR 300
Carlito_syrichta LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLNVIEETGTSGSQRMVCLTKSPR 300
Carlito LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLNVIEETGTSGSQRMVCLTKSPR 300
Tupaia_chinensis LSVAVLVTNHTIRDRDSGRLKPALGRSWSFVPSTRILLDLVEGSEASGSQRSMCLTKSSR 300
Otolemur_garnettii LGTAVVVTNHMTRDRDSGRLKPALGRSWSFVPSTRLFLDIIEGAEASGRQRTVCLTKSPR 300
Lemur_catta LGTAVVVTNHITRDRDSGKLKPALGRSWSFVPSTRILLDTTEGAGASGSQRTVCLTKSPR 300
Cebus_imitator LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIDGAGASGGRRMACLTKSPR 300
Saimiri_boliviensis LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIDGAGASGGQRMACLTKSPR 300
Nomascus_leucogenys LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPRTRILLDTIEGAGASGGRRMACLTKSSR 300
Hylobates_moloch LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPRTRILLDTIEGAGASGGRRMACLTKSSR 300
Chlorocebus_sabaeus LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDAIEGAGASAGRRMACLTKSPR 300
Papio_anubis LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASAGRRMACLTKSPR 300
Macaca_fascicularis LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASAGRRMACLTKSPR 300
Piliocolobus_tephrosceles LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASAGRRMACLTKSPR 300
Pongo_abelii LGMAVVVTNHTTRDRDSRRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLTKSSR 300
Homo_sapiens LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSSR 300
Pan_troglodytes LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGASGGRRMACLAKSSR 300
Gorilla_gorilla LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGGSGGRRMACLAKSSR 300
Dasypus_novemcinctus LGVAVLVTNHVTRDRDSGKIKPALGRSWSFVPSTRVLLGVGEGAGASGSQRKACLTKSPR 300
Choloepus_didactylus LGVAVVVTNHMTRDRDSGRLKPALGRSWSFVPSTRILLGVGEGEGTSGSQRMACLTKSPR 300
Loxodonta_africana LGVAVVVTNHVTRDRDSGRLKPALGRSWSFVPSTRILLAVGEGAGAPGSQRMAYLTKSSR 300
Elephas_maximus_indicus LGVAVVVTNHVTRDRDSGRLKPALGRSWSFVPSTRILLAVGEGAGAPGSQRMAYLTKSSR 300
Trichechus_manatus LSVAVVVTNHVTRDRDSGRLKPALGRSWSFVPSTRILLSVGEGAGTSGSERMVCLTKSSR 300
Orycteropus_afer LGVAVVVTNHVTRDRDSGRLKPALGRSWSFVPSTRVLLDVGDGAGTSGSQRMACLTKSPR 300
Meles_meles LGVAVMVTNHMTRDRDSGDLKPALGRSWSFVPSTRILLDIEEGAGASGSWRRACLTKSPR 300
Gulo_gulo_luscus LGVAVMVTNHMTRDRDSGDLKPALGRSWSFVPSTRILLDIEEGAGASGSWRRARLTKSPR 300
Lynx_canadensis VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRAYLTKSPR 300
Puma_concolor IGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRACLTKSPR 300
Panthera_uncia VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGTSGSWRRACLTKSPR 300
Panthera_tigris VGVAVMVTNHMIRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGTSGSWRRACLTKSPR 300
Leopardus_geoffroyi VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRAYLTKSPR 300
Felis_catus VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRAYLTKSPR 300
Puma_yagouaroundi VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRAYLTKSPR 300
Acinonyx_jubatus VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRAYLTKSPR 300
Lynx_rufus VGVAVMVTNHMTRDRDSGKLKPALGRSWSFVPSTRILLDIEEGAGMSGSWRRAYLTKSPR 300
Ailuropoda_melanoleuca LGVAVVVTNHVTRDRDSGDLKPALGRSWSFVPSTRILLDVDKGAGASGSRRRACLTKSPR 300
Ursus_arctos LGVAVVVTNHVIRDRDSGDLKPALGRSWSFVPSTRILLDVDEGAGASGSRRRACLTKSPR 300
Canis_lupus_dingo LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDIDEGARASGSWRRACLTKSPR 300
Vulpes_lagopus LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDIDEGARASGSWRRACLTKSPR 300
Vulpes_vulpes LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDIDEGERASGSWRRACLTKSPR 300
Callorhinus_ursinus LGMAVVVTNHMTRDRDSGELRPALGRSWSFVPSTRILLDIDEGAGASGSQRRACLTKSPR 300
Mirounga_leonina LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDIDEGAGVSGSRRRVYLTKSPR 300
Neomonachus_schauinslandi LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDVDEGAGASGSRRRVYLTKSPR 300
Phoca_vitulina LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDIDEGAGASGSRRRAYLTKSPR 300
Halichoerus_grypus LGMAVVVTNHMTRDRDSGELKPALGRSWSFVPSTRILLDIDEGAGASGSRRRAYLTKSPR 300
Vicugna_pacos LGVAVVVTNHMTRDRDTGKLKPALGRSWSFVPSTRFLLDVTKGTGASDSQRMVCLMKSPR 300
Camelus_bactrianus LGVAVVVTNHMTRDRDTGKLKPALGRSWSFVPSTRFLLDVTKGTGASDSQRMVCLMKSPR 300
Camelus_ferus LGVAVVVTNHMTRDRDTGKLKPALGRSWSFVPSTRFLLDITKGTGASDSQRMVCLMKSPR 300
Balaenoptera_musculus LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTSQGAGASGSWRVACLTKSPR 300
Balaenoptera_acutorostrata LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTSQGAGASGSWRVACLTKSPR 300
Delphinapterus_leucas LSVALLVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTSQGAGASGSWRMACLTKSPR 300
Monodon_monoceros LSVALLVTNHMTRDRDSGQLKPALGCSWSFVPSTRLLLDTSQGAGASGSWRMACLTKSPR 300
Neophocaena_asiaeorientalis LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTGQGAGASGSWRMACLTKSPR 300
Lagenorhynchus_obliquidens LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTGQGAGASGSWRMACLTKSPR 300
Tursiops_truncatus LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTGQGAGASGSWRMACLTKSPR 300
Orcinus_orca LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTGQGAGASGSWRMACLTKSPR 300
Globicephala_melas LSVALVVTNHMTRDRDSGQLKPALGRSWSFVPSTRLLLDTGQGAGASGSWRMACLTKSPR 300
Equus_caballus LGLAVVVTNHMTRERDSGKLKPALGRSWSFVPSTRILLDTGEGAGASGSQRTACLIKSPR 300
Equus_quagga LGLAVVVTNHMTRERDSGKLKPALGRSWSFVPSTRILLDTGEGAGASGSQRTACLIKSPR 300
Equus_asinus LGLAVVVTNHMTRERDSGKLKPALGRSWSFVPSTRILLDTGEGAGASGSQRTACLIKSPR 300
Rousettus_aegyptiacus LGLAVVVTNHMTRDRDSGKFKPALGRSWSFVPSTRIVLTIGEEAGAPGSQRTVCLTKSPR 300
Pteropus_alecto LGVAVLVTNHMTRDRDSGKFKPALGRSWSFVPSTRIVLAIGEEAGAPGSQRTVCLTKSPR 300
Pteropus_giganteus LGVAVLVTNHMTRDRDSGKFKPALGRSWSFVPSTRIVLAIGEEAGAPGSQRTVCLTKSPR 300
Pipistrellus_kuhlii LGVAVVVTNHMTRDWDSGKLKPALGRSWSFVPSTRIVLAVSEGAGAPGSERIACLTKSPR 300
Molossus_molossus LGVAVLVTNHMTRDRDSGKLKPALGLSWSFVPSTRIVLATSEGAGAPSSQRTACLTKSPR 300
Artibeus_jamaicensis LGVAVVVTNHMTRDRDSGKLRPALGRSWSFVPSTRVVLASGEGAGAPGSRRTMCLTKSPR 300
Desmodus_rotundus LGVAVVVTNHMTRDRDSGKLKPALGRSWSFVPSTRVVLASGEGAGAPGSQRTACLTKSPR 300
Myotis_lucifugus LGVAVVVTNHMTRDRDSGKLKPALGRSWSFVPSTRVILAIGEGAGAPGSQHTACLTKSPR 300
Myotis_myotis LGVAVVVTNHMTRDRDSGKLKPALGRSWSFVPSTRIVLAIGEGAGAPGRQRTACLTKSPR 300

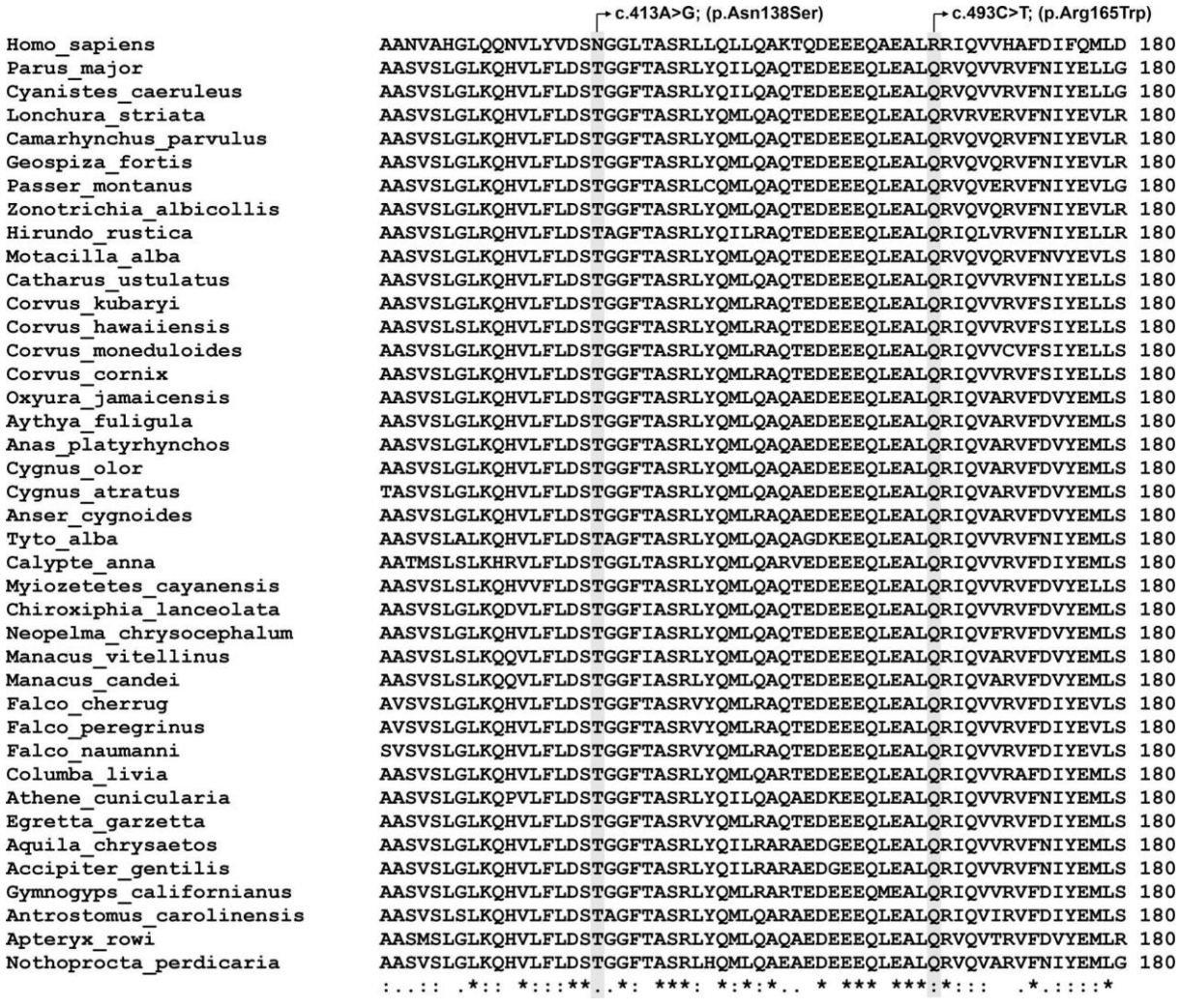

c.973G>A; (p.Gly325Ser)
c.919G>A; (p.Glu307Lys)
Sarcophilus_harrisii QPTGIQEVIEIGVCETPEQNSSLEGNKL 328
Phascolarctos_cinereus QPTGIQEVIEIGSCGTPEQNSSLEGSKP 328
Vombatus_ursinus QPTGIQEVMEIGSCGTPEQNSSLEGSKS 328
Trichosurus_vulpecula QPTGIQELVEIGSCGTAEQSSSLEGSKP 328
Dromiciops_gliroides QATGIQEVVEIGSCGTPAQSSPLEGNKP 328
Jaculus_jaculus QPTGLQEVIDIGTLGTAEQSPIRQGEQT 328
Jaculus_jaculus_2 QPMGLQEVIDMGTLGTAEQSPMRQGEQT 328
Carlito_syrichta QPTGVREMVDIGTWGTPGQSPSLKGDQT 328
Carlito QPTGVREMVDIGTWGTPGQSPSLKGDQT 328
Tupaia_chinensis QPTGFQEMVDIGTWGTPEKSPALQGDQT 328
Otolemur_garnettii QPTGFQEMVDIGTWGTAEPSPALQGDQT 328
Lemur_catta QPTGLQEMVDIGTWGTPEQSSELQGDQT 328
Cebus_imitator QPTGFQEMVDIGTWGTSEQSTTLQVDQT 328
Saimiri_boliviensis QPTGFQEMVDIGSWGTSEQSATLQVDQT 328
Nomascus_leucogenys QPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Hylobates_moloch QPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Chlorocebus_sabaeus QPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Papio_anubis QPTGFQEMVDIGTWGTSEQSATLQDDQT 328
Macaca_fascicularis QPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Piliocolobus_tephrosceles QPTGFQEMVDIGTWGNSEQSATLQGDQT 328
Pongo_abelii QPTGFQEMVDTGTWGTSEQSATLQGDQT 328
Homo_sapiens QPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Pan_troglodytes QPTGFQEMVDIGTWGTSEHSATLQGDQT 328
Gorilla_gorilla QPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Dasypus_novemcinctus QLTGIQEMVDIGTWGTREQSPRLEDDHI 328
Choloepus_didactylus QPTGIQEMIDIGIWGTPEQSPRLEGDQL 328
Loxodonta_africana QPTGIQEMIDIGIWGTSEERPVLEGDQT 328
Elephas_maximus_indicus QPTGIQEMIDIGIWGTSEERPVLEGDQT 328
Trichechus_manatus QPTGIQEMIDIGIWGTPEQRPVLEAYQT 328
Orycteropus_afer QPTGIQEVVDIGIWGTTEQSPVLEGDQT 328
Meles_meles LPTGFQEVVDIGTWGPPEQSPASQEGQI 328
Gulo_gulo_luscus LPTGFQEVVDIGTWGPPEQSPASQGDQI 328
Lynx_canadensis LPTGFQEMVDIGTWGTPEQSPTSQRDQM 328
Puma_concolor LPTGFQEMVDIGTWGTPEQSPTSQGDQM 328
Panthera_uncia LPTGFQEMVDIGTWGTPEQSPTSQGDQM 328
Panthera_tigris LPTGFQEMVDIGTWGTPEQSPTSQGDQM 328
Leopardus_geoffroyi LPTGFQEMVDIGTWGIPEQSPTSQGDQM 328
Felis_catus LPTGFQEMVDIGTWGIPEQSPTSQGDQM 328
Puma_yagouaroundi LPTGFQEMVDIGTWGTPEQSPTSQGDQM 328
Acinonyx_jubatus LPTGFQEMVDIGTWGTPEQSPTSQGDQM 328
Lynx_rufus LPTGFQEMVDIGTWGTPEQSPTSQGDQM 328
Ailuropoda_melanoleuca LPTGFQEMVDIGTWGPPEQSPASQGDQI 328
Ursus_arctos LPTGFQEMVDIGTWGPPEQSPASQGDQI 328
Canis_lupus_dingo LPTGFQEMVDIGTWGPPEQSPASQGDQI 328
Vulpes_lagopus LPTGFQEMVDIGTWGPPEQSPASQGDQI 328
Vulpes_vulpes LPTGFQEMVDIGTWGPPEQSPASQGDQI 328
Callorhinus_ursinus LPTGFQEMVDLGTWAPPEQSPASQGDQI 328
Mirounga_leonina LPTGFQEMVDVGTWAPPEQSPASQGDQI 328
Neomonachus_schauinslandi LPTGFQEMVDIGTWAPPEQSPASHGDQI 328
Phoca_vitulina LPTGFQEMVDIGTWAPPEQSPASQGDQI 328
Halichoerus_grypus LPTGFQEMVDIGTWAPPEQSPASQGDQI 328
Vicugna_pacos LPTGFQEMVDLGTWGTSEQSPQLQGDHT 328
Camelus_bactrianus LPTGFQEMVDLGTWGTSEQSPQLQGDHT 328
Camelus_ferus LPTGFQEMVDLGTWGTSEQSPQLQGDHT 328
Balaenoptera_musculus LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Balaenoptera_acutorostrata LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Delphinapterus_leucas LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Monodon_monoceros LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Neophocaena_asiaeorientalis LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Lagenorhynchus_obliquidens LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Tursiops_truncatus LPTGFQEMVDIGTWGTPVQSPTLQGDHM 328
Orcinus_orca LATGFQEMVDIGTWGTPVQSPTLQGDHM 328
Globicephala_melas LPTGFQEMVDIGTWATPVQSPTLQGDHM 328
Equus_caballus LPTGFQETVDIGTWGTLKQSPALQGHPT 328
Equus_quagga LPTGFQETVDIGTWGTLKQSPALQGHPT 328
Equus_asinus LPTGFQETVDIGTWGTLKQSPALQGHPT 328
Rousettus_aegyptiacus LPTGFQETVDIGTWGTPEQSPVLQGDQI 328
Pteropus_alecto LPTGFQETVDIGTWGTPEQSPVLQGDQM 328
Pteropus_giganteus LPTGFQETVDIGTWGTPEQSPVLQGDQM 328
Pipistrellus_kuhlii LPTGSQEPVDIGTWGALEHGPLSPGEQA 328
Molossus_molossus LPTGFQETVDIETWGTLEQSPALQGDKI 328
Artibeus_jamaicensis LPTGFQETVDIGGWGTLEQSPAQQGQQM 328
Desmodus_rotundus LPTGFQEPVDIGTWGTLQQSPALQGEQL 328
Myotis_lucifugus LPTGSQETVDIGTWGALEQSPLLQGEQA 328
Myotis_myotis LPTGSQETVDIGTWGALEQSPLLQGEQT 328
* :* ::

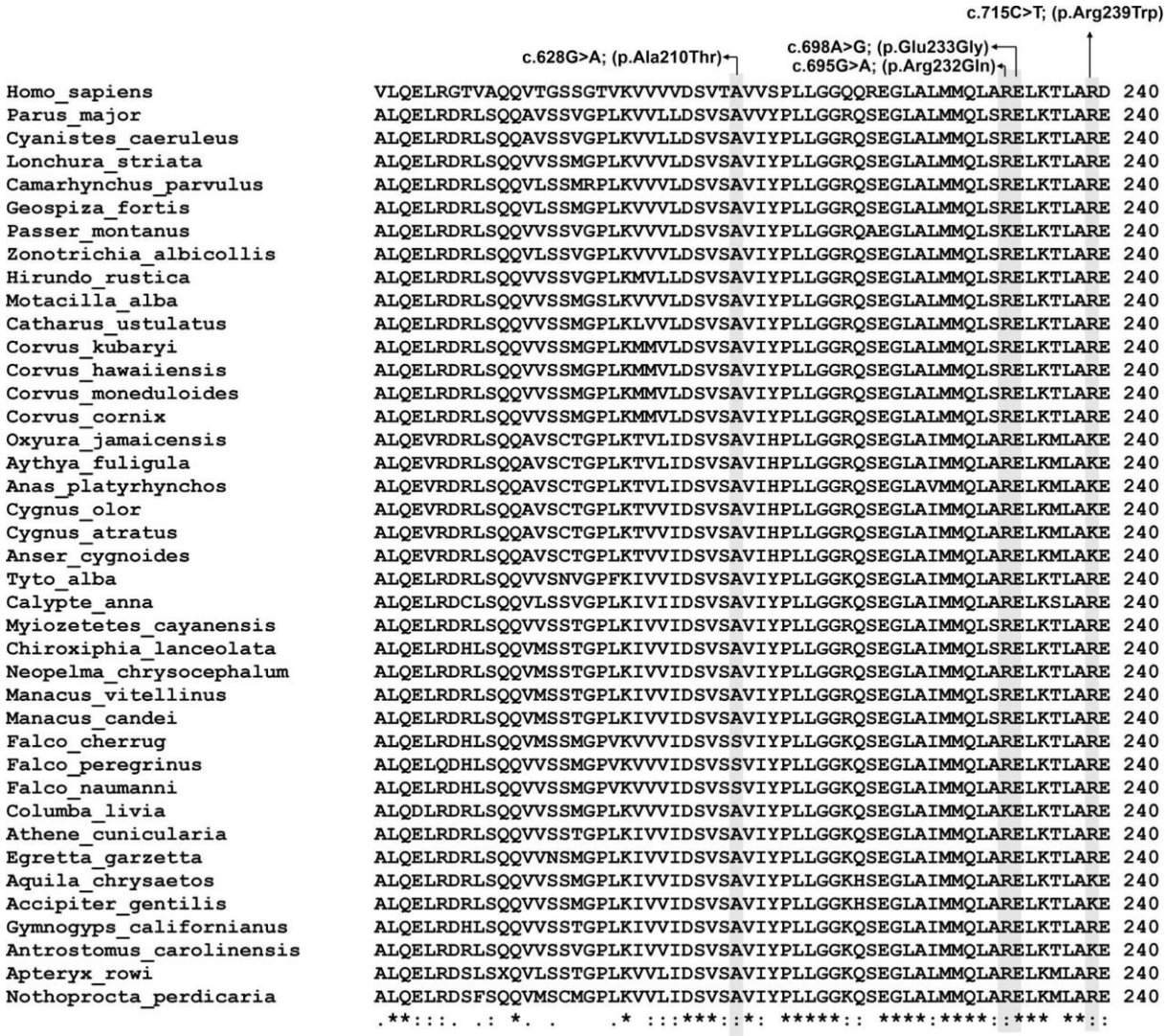
Alinhamento 2: aves

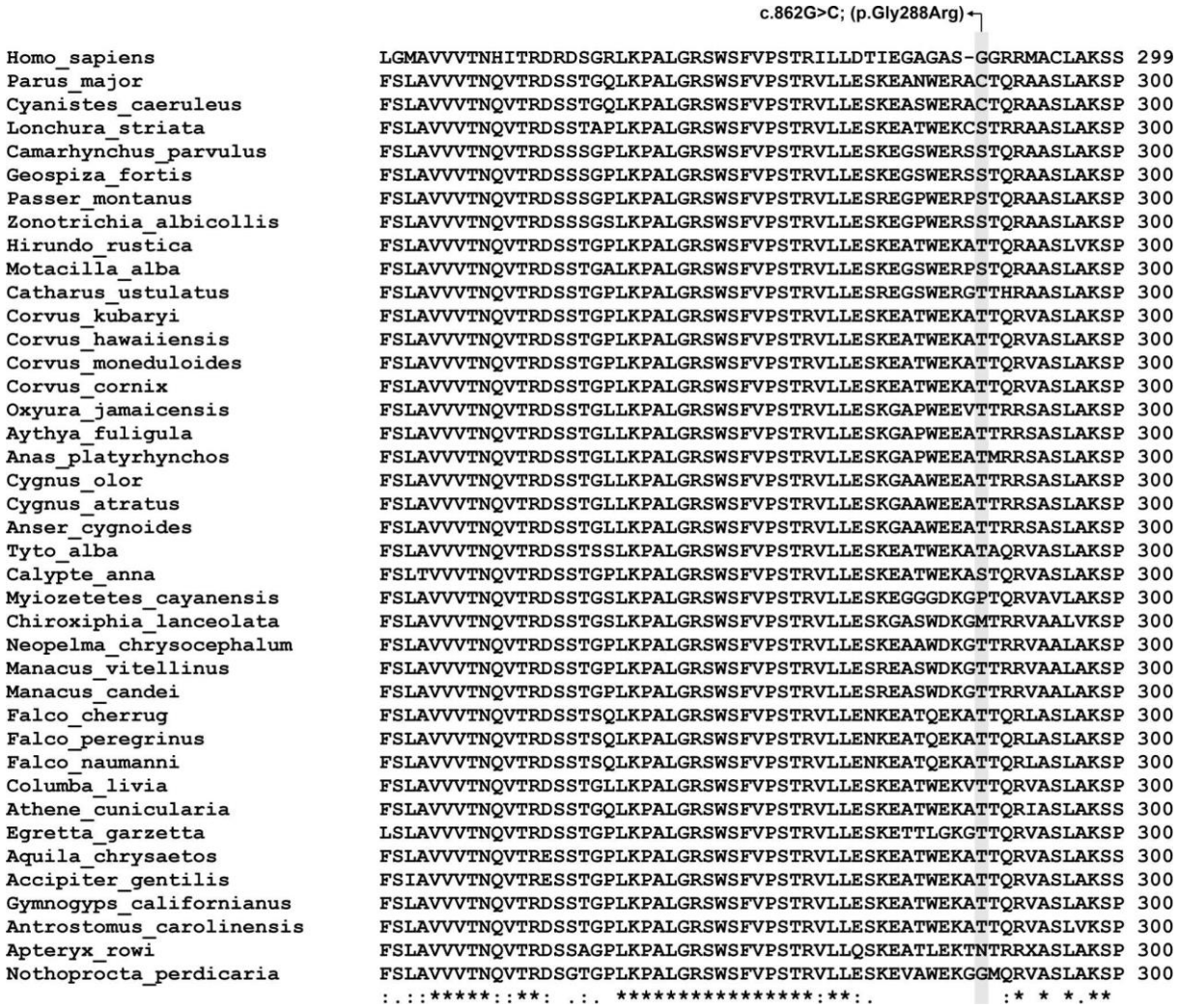
c.26G>C; (p.Cys9Ser) c.137C>G; (p.Ser46Cys) c.140A>G; (p.Tyr47Cys)

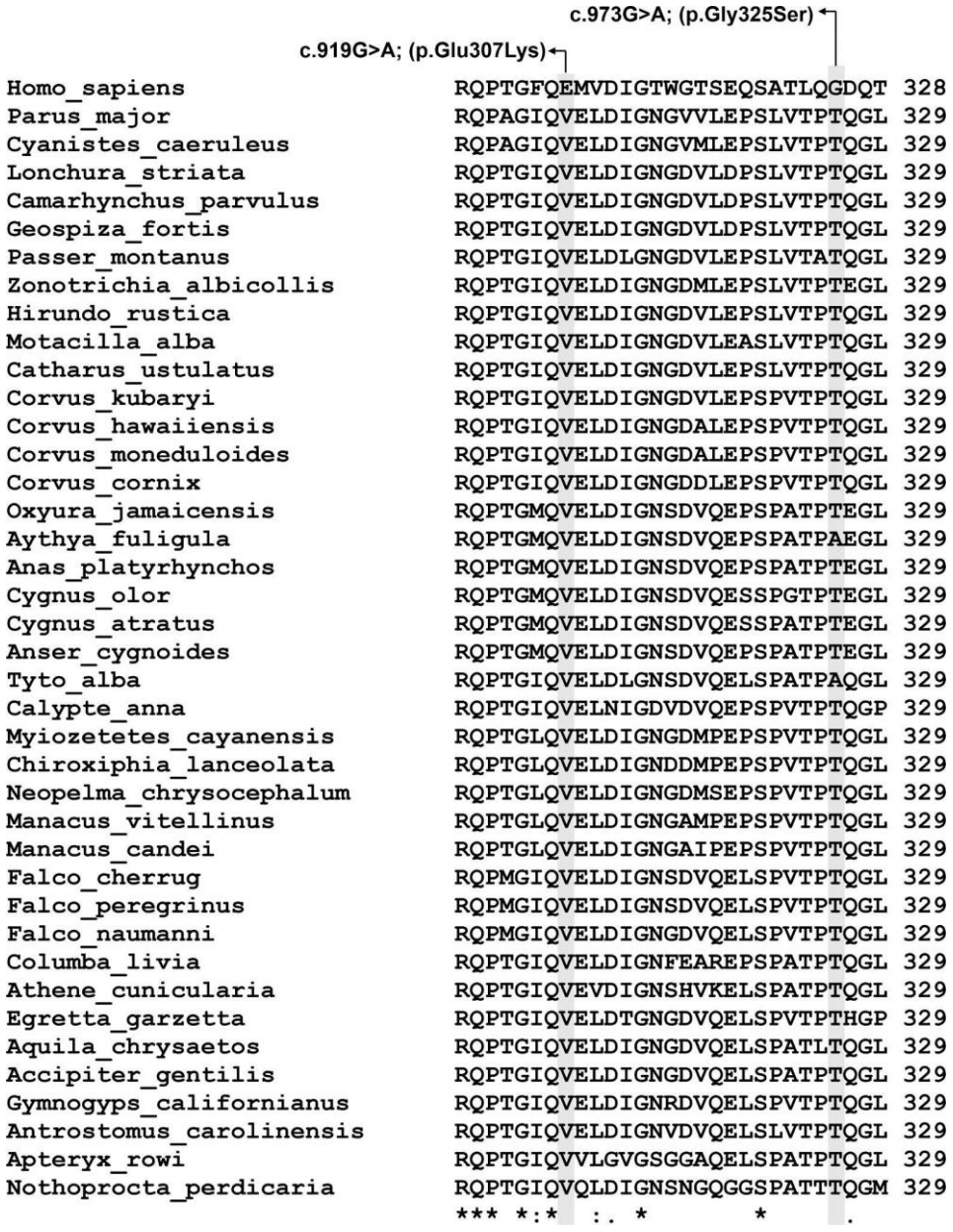
```
Homo_sapiens                MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKALVALRRVLLAQ 60
Parus_major                 MVVLQAGLCPGLTEEMLQLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Cyanistes_caeruleus         MVVLQAGLCPGLTEEMLQLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Lonchura_striata            MVVLRAGLCPGLSEDMIQLLSANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Camarhynchus_parvulus       MVVLRAGLCPGLTEDMIQLLRDNGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Geospiza_fortis             MVVLRAGLCPGLTEDMIQLLRDNGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Passer_montanus             MVLLRAGLCPGLTEDMIQLLRANAIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Zonotrichia_albicollis      MVVLRAGLCPGLTEDMIQLLRDNGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Hirundo_rustica             MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Motacilla_alba              MVVLRAGLCPGLTEDVVQLLRANGIRTVVDFVSSDMEDVAQKCSLSYKALVAVRRVLLAQ 60
Catharus_ustulatus          MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Corvus_kubaryi              MVVLRAGLCPGLTEDMIQLLRANGIKTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Corvus_hawaiiensis          MVVLRAGLCPGLTEDMIQLLRANGIKTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Corvus_moneduloides         MVVLRAGLCPGLTEDMIQLLRANGIKTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Corvus_cornix               MVVLRAGLCPGLTEDMIQLLRANGIKTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Oxyura_jamaicensis          MVVLRAGLCPGLTEEMIQVLRANNVRTVVDFVSSDLEDVSQKCSLSYKALVAVRRVLLAQ 60
Aythya_fuligula             MVVLRAGLCPGLTEEMIQVLRANGVRTVVDFVSSDLEDVSQKCSLSYKALVAVRRVLLAQ 60
Anas_platyrhynchos          MVVLRAGLCPGLTEEMIQVLRANGVRTVVDFVSSDLEDVSQKCSLSYKALVAVRRVLLAQ 60
Cygnus_olor                 MVVLRAGLCPGLTEEMIQVLRANSVRTVVDFVSSDLEDVSQKCSLSYKALVAVRRVLLAQ 60
Cygnus_atratus              MVVLRAGLCPGLTEEMIQVLRANSVRTVVDFVSSDLEDVSQKCSLSYKALVAVRRVLLAQ 60
Anser_cygnoides             MVVLRAGLCPGLTEEMIQVLRANSVRTVVDFVSSDLEDVSQKCSLSYKALVAVRRVLLAQ 60
Tyto_alba                   MVLLRAGLCPGLSRATVQLLGANGVRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Calypte_anna                MVVLRAGLCPGLGEEMLHLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Myiozetetes_cayanensis      MVVLRAGLCPGLTEEMIQLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Chiroxiphia_lanceolata      MVVLRAGLCPGLTEEMIQLLRANGVRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Neopelma_chrysocephalum     MVVLRAGLCPGLTEEMIQLLKANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Manacus_vitellinus          MVVLRAGLCPGLTEEMIQLLRANGIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Manacus_candei              MVVLRAGLCPGLTEEMIQLLRANGIRTVVDFVSLDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Falco_cherrug               MVVLRAGLCPGLTEEMIGLLRASGIRTVVDFVSSDLEDVAQKCSLSYKVLVAVRRVLLAQ 60
Falco_peregrinus            MVVLRAGLCPGLTEEMIGLLRASGIRTVVDFVSSDLEDVAQKCSLSYKVLVAVRRVLLAQ 60
Falco_naumanni              MVVLRAGLCPGLTEEMIGLLRASGIRTVVDFVSSDLEDIAQKCSLSYKVLVAVRRVLLAQ 60
Columba_livia               MVVLRAGLCPGLTEDMIQLLRANGIRTVVDFVSSDMEDVAQKCSLSYKALVAVRRVLLAQ 60
Athene_cunicularia          MVVLRAGLCPGLTEEMIQLLRAGGTRTVVDFVSSDLEDIAQKCSLSYKALVAVRRVLLAQ 60
Egretta_garzetta            MVVLRAGLCPGLTEEMIQLLRARSIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Aquila_chrysaetos           MVVLRAGLCPGLTEEMIQLLRANGIRTVVDFVSSDLEEVAQKCSLSYKALVAVRRVLLAQ 60
Accipiter_gentilis          MVVLRAGLCPGLTEEMIQLLRANGIRTVVDFVSSDLEDIAQKCSLSYKALVAVRRVLLAQ 60
Gymnogyps_californianus     MVVLRAGLCPGLTEEMIQLLRANGIRTVVDFVSSDLEDVAQKCTLSYKALVAVRRVLLAQ 60
Antrostomus_carolinensis    MVVLQAGLCPGLTEEMIQLLRANSIRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Apteryx_rowi                MALLRAALCPGLTAQMVPLLRASGVRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
Nothoprocta_perdicaria      MALLRPGLCAGLTAEMAPLLRASGLRTVVDFVSSDLEDVAQKCSLSYKALVAVRRVLLAQ 60
                            * :*: .** **      :*     :****:** *:*:::*** ****.***:******* 
```


c.196G>A; (p.Val66Met)
c.355T>C; (p.Cys119Arg)
Homo_sapiens FSAFPVNGADLYEELKTSTAILSTGIGSLDKLLDAGLYTGEVTEIVGGPGSGKTQVCLCM 120
Parus_major FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGELTELMGAPGSGKTQVCLGI 120
Cyanistes_caeruleus FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGELTELVGAPGSGKTQVCLGI 120
Lonchura_striata FSAFPANGADLYEELKSSTAILPTGSSSLDQLLDSGLYTGEVMELMGAPGTGKTQVCLGI 120
Camarhynchus_parvulus FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVMELMGAPGTGKTQVCLGI 120
Geospiza_fortis FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVMELMGAPGTGKTQVCLGI 120
Passer_montanus FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGTGKTQVCLGI 120
Zonotrichia_albicollis FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVMELMGAPGTGKTQVCLGI 120
Hirundo_rustica FSAFPANGADLYEELKSSTAILPTGSPSLDRLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Motacilla_alba FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGTGKTQVCLGI 120
Catharus_ustulatus FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Corvus_kubaryi FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Corvus_hawaiiensis FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Corvus_moneduloides FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Corvus_cornix FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGASGSGKTQVCLGI 120
Oxyura_jamaicensis FSAFPANGADLYEELKSSTAILPTGSPSLDQMLDSGLYTGEVTELMGAPGSGKTQVCLSI 120
Aythya_fuligula FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLSI 120
Anas_platyrhynchos FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLSI 120
Cygnus_olor FSAFPANGADLYEELKSSTAILSTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLSI 120
Cygnus_atratus FSAFPANGADLYEELKSSTAILSTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLSI 120
Anser_cygnoides FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGTPGSGKTQVCLSI 120
Tyto_alba FSAFPTNGADLYEELKSSTAILPTGNPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGV 120
Calypte_anna FSAFPANGADLYEELKSTTAILPTGIPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Myiozetetes_cayanensis FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Chiroxiphia_lanceolata FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Neopelma_chrysocephalum FSAFPANGADLYEELKSSTAILPTGNPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Manacus_vitellinus FSAFPANGADLYEELKSSTAILPTGNASLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Manacus_candei FSAFPANGADLYEELKSSTAILPTGNPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Falco_cherrug FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Falco_peregrinus FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Falco_naumanni FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Columba_livia FSAFPANGADLYEELKSSMAILPTGSPSLDKLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Athene_cunicularia FSAFPTNGAVLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Egretta_garzetta FSAFPANGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Aquila_chrysaetos FSAFPTNGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Accipiter_gentilis FSAFPINGADLYEELKSSTAILPTGSPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Gymnogyps_californianus FSAFPANGADLYEELKSSTAILPTGSPRLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Antrostomus_carolinensis FSAFPANGADLYEELKSSTAILPTGNPSLDQLLDSGLYTGEVTELMGAPGSGKTQVCLGI 120
Apteryx_rowi FSAFPANGADLYEELKSSTAILSTGIPSLDKLLDSGLYTGEVTELMGAPGSGKTQVCLSI 120
Nothoprocta_perdicaria FSAFPANGADLYEELKSSTAILSTGISSLDKLLDSGLYTGEVTEIMGVPGSGKTQVCLSI 120

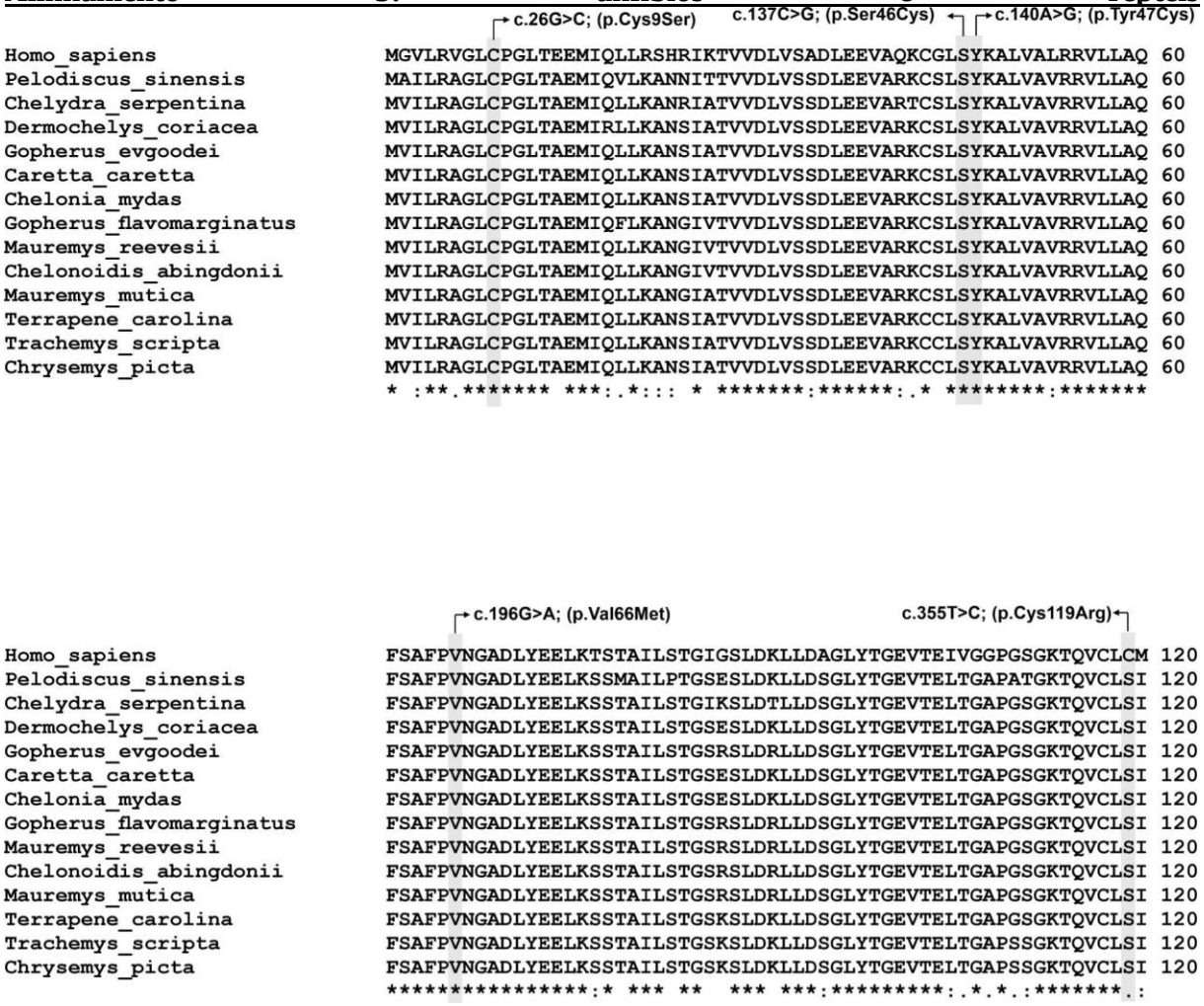

c.413A>G; (p.Asn138Ser)
c.493C>T; (p.Arg165Trp)
Homo_sapiens AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD 180
Parus_major AASVSLGLKQHVLFLDSTGGFTASRLYQILQAQTEDEEEQLEALQRVQVVRVFNIYELLG 180
Cyanistes_caeruleus AASVSLGLKQHVLFLDSTGGFTASRLYQILQAQTEDEEEQLEALQRVQVVRVFNIYELLG 180
Lonchura_striata AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRVRVERVFNIYEVLR 180
Camarhynchus_parvulus AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRVQVQRVFNIYEVLR 180
Geospiza_fortis AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRVQVQRVFNIYEVLR 180
Passer_montanus AASVSLGLKQHVLFLDSTGGFTASRLCQMLQAQTEDEEEQLEALQRVQVERVFNIYEVLG 180
Zonotrichia_albicollis AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRVQVQRVFNIYEVLR 180
Hirundo_rustica AASVSLGLRQHVLFLDSTAGFTASRLYQILRAQTEDEEEQLEALQRIQLVRVFNIYELLR 180
Motacilla_alba AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRVQVQRVFNVYEVLS 180
Catharus_ustulatus AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRIQVVRVFNIYELLS 180
Corvus_kubaryi AASVSLGLKQHVLFLDSTGGFTASRLYQMLRAQTEDEEEQLEALQRIQVVRVFSIYELLS 180
Corvus_hawaiiensis AASVSLSLKQHVLFLDSTGGFTASRLYQMLRAQTEDEEEQLEALQRIQVVRVFSIYELLS 180
Corvus_moneduloides AASVSLGLKQHVLFLDSTGGFTASRLYQMLRAQTEDEEEQLEALQRIQVVCVFSIYELLS 180
Corvus_cornix AASVSLGLKQHVLFLDSTGGFTASRLYQMLRAQTEDEEEQLEALQRIQVVRVFSIYELLS 180
Oxyura_jamaicensis AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQAEDEEEQLEALQRIQVARVFDVYEMLS 180
Aythya_fuligula AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQAEDEEEQLEALQRIQVARVFDVYEMLS 180
Anas_platyrhynchos AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQAEDEEEQLEALQRIQVARVFDVYEMLS 180
Cygnus_olor AASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQAEDEEEQLEALQRIQVARVFDVYEMLS 180
Cygnus_atratus TASVSLGLKQHVLFLDSTGGFTASRLYQMLQAQAEDEEEQLEALQRIQVARVFDVYEMLS 180
Anser_cygnoides AASVSLGLKQHVLFLDSTGGFTASRLYQMLRAQAEDEEEQLEALQRIQVARVFDVYEMLS 180
Tyto_alba AASVSLALKQHVLFLDSTAGFTASRLYQMLQAQAGDKEEQLEALQRIQVVRVFNIYEMLS 180
Calypte_anna AATMSLSLKHRVLFLDSTGGLTASRLYQMLQARVEDEEEQLEALQRIQVVRVFDIYEMLS 180
Myiozetetes_cayanensis AASVSLSLKQHVVFLDSTGGFTASRLYQMLQAQTEDEEEQLEALQRIQVVRVFDVYELLS 180
Chiroxiphia_lanceolata AASVSLGLKQDVLFLDSTGGFIASRLYQMLQAQTEDEEEQLEALQRIQVVRVFDVYEMLS 180
Neopelma_chrysocephalum AASVSLGLKQHVLFLDSTGGFIASRLYQMLQAQTEDEEEQLEALQRIQVFRVFDVYEMLS 180
Manacus_vitellinus AASVSLSLKQQVLFLDSTGGFIASRLYQMLQAQTEDEEEQLEALQRIQVARVFDVYEMLS 180
Manacus_candei AASVSLSLKQQVLFLDSTGGFIASRLYQMLQAQTEDEEEQLEALQRIQVARVFDVYEMLS 180
Falco_cherrug AVSVSLGLKQHVLFLDSTGGFTASRVYQMLRAQTEDEEEQLEALQRIQVVRVFDIYEVLS 180
Falco_peregrinus AVSVSLGLKQHVLFLDSTGGFTASRVYQMLRAQTEDEEEQLEALQRIQVVRVFDIYEVLS 180
Falco_naumanni SVSVSLGLKQHVLFLDSTGGFTASRVYQMLRAQTEDEEEQLEALQRIQVVRVFDIYEVLS 180
Columba_livia AASVSLGLKQHVLFLDSTGGFTASRLYQMLQARTEDEEEQLEALQRIQVVRAFDIYEMLS 180
Athene_cunicularia AASVSLGLKQPVLFLDSTGGFTASRLYQILQAQAEDKEEQLEALQRIQVVRVFNIYEMLS 180
Egretta_garzetta AASVSLGLKQHVLFLDSTGGFTASRVYQMLRAQTEDEEEQLEALQRIQVVRVFDIYEMLS 180
Aquila_chrysaetos AASVSLGLKQHVLFLDSTGGFTASRLYQILRARAEDGEEQLEALQRIQVVRVFNIYEMLS 180
Accipiter_gentilis AASVSLGLKQHVLFLDSTGGFTASRLYQILRARAEDGEEQLEALQRIQVVRVFNIYEMLS 180
Gymnogyps_californianus AASVSLGLKQHVLFLDSTGGFTASRLYQMLRARTEDEEEQMEALQRIQVVRVFDIYEMLS 180
Antrostomus_carolinensis AASVSLSLKQHVLFLDSTAGFTASRLYQMLQARAEDEEEQLEALQRIQVIRVFDIYEMLS 180
Apteryx_rowi AASMSLGLKQHVLFLDSTGGFTASRLYQMLQAQAEDEEEQLEALQRVQVTRVFDVYEMLR 180
Nothoprocta_perdicaria AASVSLGLKQHVLFLDSTGGFTASRLHQMLQAEAEDEEEQLEALQRVQVARVFNIYEMLG 180

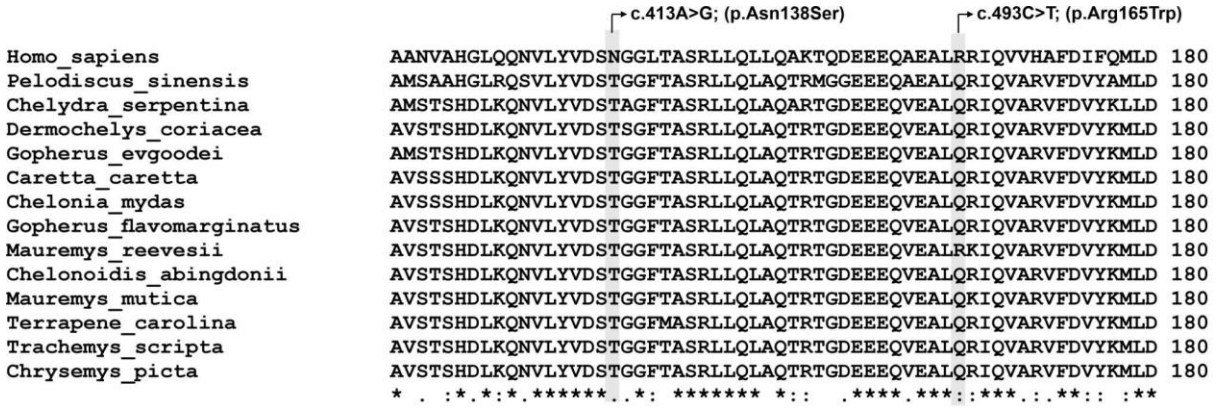

c.715C>T; (p.Arg239Trp)
c.628G>A; (p.Ala210Thr)
c.698A>G; (p.Glu233Gly)
c.695G>A; (p.Arg232Gln)
Homo_sapiens VLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Parus_major ALQELRDRLSQQAVSSVGPLKVVLLDSVSAVVYPLLGGRQSEGLALMMQLSRELKTLARE 240
Cyanistes_caeruleus ALQELRDRLSQQAVSSVGPLKVVLLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Lonchura_striata ALQELRDRLSQQVVSSMGPLKVVVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Camarhynchus_parvulus ALQELRDRLSQQVLSSMRPLKVVVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Geospiza_fortis ALQELRDRLSQQVLSSMGPLKVVVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Passer_montanus ALQELRDRLSQQVVSSVGPLKVVVLDSVSAVIYPLLGGRQAEGLALMMQLSKELKTLARE 240
Zonotrichia_albicollis ALQELRDRLSQQVLSSVGPLKVVVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Hirundo_rustica ALQELRDRLSQQVVSSVGPLKMVLLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Motacilla_alba ALQELRDRLSQQVVSSMGSLKVVVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Catharus_ustulatus ALQELRDRLSQQVVSSMGPLKLVVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Corvus_kubaryi ALQELRDRLSQQVVSSMGPLKMMVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Corvus_hawaiiensis ALQELRDRLSQQVVSSMGPLKMMVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Corvus_moneduloides ALQELRDRLSQQVVSSMGPLKMMVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Corvus_cornix ALQELRDRLSQQVVSSMGPLKMMVLDSVSAVIYPLLGGRQSEGLALMMQLSRELKTLARE 240
Oxyura_jamaicensis ALQEVRDRLSQQAVSCTGPLKTVLIDSVSAVIHPLLGGRQSEGLAIMMQLARELKMLAKE 240
Aythya_fuligula ALQEVRDRLSQQAVSCTGPLKTVLIDSVSAVIHPLLGGRQSEGLAIMMQLARELKMLAKE 240
Anas_platyrhynchos ALQEVRDRLSQQAVSCTGPLKTVLIDSVSAVIHPLLGGRQSEGLAVMMQLARELKMLAKE 240
Cygnus_olor ALQEVRDRLSQQAVSCTGPLKTVVIDSVSAVIHPLLGGRQSEGLAIMMQLARELKMLAKE 240
Cygnus_atratus ALQEVRDRLSQQAVSCTGPLKTVVIDSVSAVIHPLLGGRQSEGLAIMMQLARELKMLAKE 240
Anser_cygnoides ALQEVRDRLSQQAVSCTGPLKTVVIDSVSAVIHPLLGGRQSEGLAIMMQLARELKMLAKE 240
Tyto_alba ALQELRDRLSQQVVSNVGPFKIVVIDSVSAVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Calypte_anna ALQELRDCLSQQVLSSVGPLKIVIIDSVSAVIYPLLGGKQSEGLAIMMQLARELKSLARE 240
Myiozetetes_cayanensis ALQELRDRLSQQVVSSTGPLKIVVIDSVSAVIYPLLGGRQSEGLAIMMQLSRELKTLARE 240
Chiroxiphia_lanceolata ALQELRDHLSQQVMSSTGPLKIVVIDSVSAVIYPLLGGRQSEGLAIMMQLSRELKTLARE 240
Neopelma_chrysocephalum ALQELRDRLSQQVMSSTGPLKIVVIDSVSAVIYPLLGGRQSEGLAIMMQLARELKTLARE 240
Manacus_vitellinus ALQELRDRLSQQVMSSTGPLKIVVIDSVSAVIYPLLGGRQSEGLAIMMQLSRELKTLARE 240
Manacus_candei ALQELRDRLSQQVMSSTGPLKIVVIDSVSAVIYPLLGGRQSEGLAIMMQLSRELKTLARE 240
Falco_cherrug ALQELRDHLSQQVMSSMGPVKVVVIDSVSSVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Falco_peregrinus ALQELQDHLSQQVVSSMGPVKVVVIDSVSSVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Falco_naumanni ALQELRDHLSQQVVSSMGPVKVVVIDSVSSVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Columba_livia ALQDLRDRLSQQVVSSMGPLKVVVIDSVSAVIYPLLGGKQSEGLAIMMQLAKELKTLARE 240
Athene_cunicularia ALQELRDRLSQQVVSSTGPLKIVVIDSVSAVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Egretta_garzetta ALQELRDRLSQQVVNSMGPLKIVVIDSVSAVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Aquila_chrysaetos ALQELRDRLSQQVVSSMGPLKIVVIDSVSAVIYPLLGGKHSEGLAIMMQLARELKTLAKE 240
Accipiter_gentilis ALQELRDRLSQQVVSSMGPLKIVVIDSVSAVIYPLLGGKHSEGLAIMMQLARELKTLAKE 240
Gymnogyps_californianus ALQELRDHLSQQVVSSTGPLKIVVIDSVSAVIYPLLGGKQSEGLAIMMQLARELKTLARE 240
Antrostomus_carolinensis ALQELRDRLSQQVVSSVGPLKIVVIDSVSAVIYPLLGGKQSEGLAIMMQLARELKTLAKE 240
Apteryx_rowi ALQELRDSLSXQVLSSTGPLKVVLIDSVSAVIYPLLGGKQSEGLALMMQLARELKMLARE 240
Nothoprocta_perdicaria ALQELRDSFSQQVMSCMGPLKVVLIDSVSAVIYPLLGGKQSEGLALMMQLARELKMLARE 240

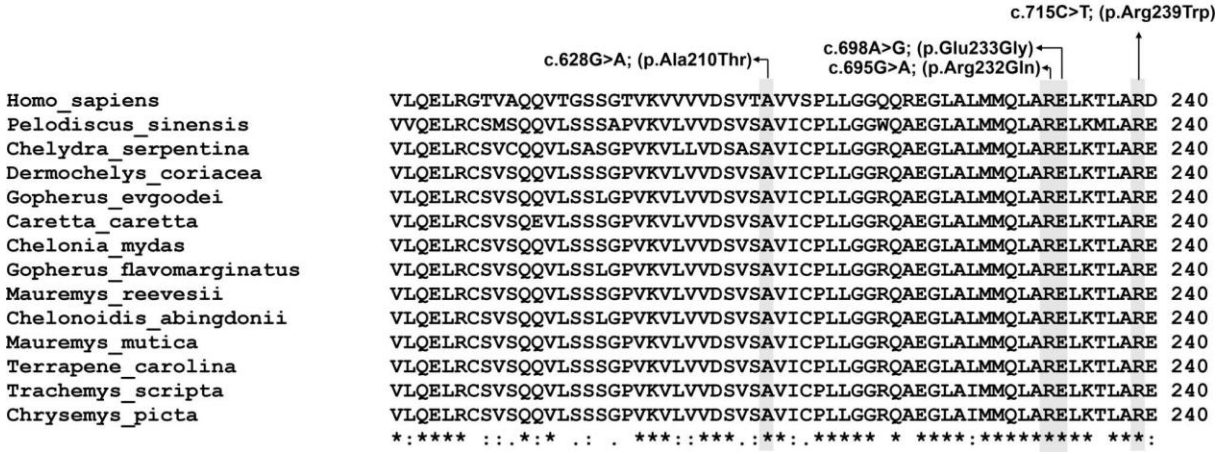

c.862G>C; (p.Gly288Arg)
Homo_sapiens LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGAS-GGRRMACLAKSS 299
Parus_major FSLAVVVTNQVTRDSSTGQLKPALGRSWSFVPSTRVLLESKEANWERACTQRAASLAKSP 300
Cyanistes_caeruleus FSLAVVVTNQVTRDSSTGQLKPALGRSWSFVPSTRVLLESKEASWERACTQRAASLAKSP 300
Lonchura_striata FSLAVVVTNQVTRDSSTAPLKPALGRSWSFVPSTRVLLESKEATWEKCSTRRAASLAKSP 300
Camarhynchus_parvulus FSLAVVVTNQVTRDSSSGPLKPALGRSWSFVPSTRVLLESKEGSWERSSTQRAASLAKSP 300
Geospiza_fortis FSLAVVVTNQVTRDSSSGPLKPALGRSWSFVPSTRVLLESKEGSWERSSTQRAASLAKSP 300
Passer_montanus FSLAVVVTNQVTRDSSSGPLKPALGRSWSFVPSTRVLLESREGPWERPSTQRAASLAKSP 300
Zonotrichia_albicollis FSLAVVVTNQVTRDSSSGSLKPALGRSWSFVPSTRVLLESKEGPWERSSTQRAASLAKSP 300
Hirundo_rustica FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRAASLVKSP 300
Motacilla_alba FSLAVVVTNQVTRDSSTGALKPALGRSWSFVPSTRVLLESKEGSWERPSTQRAASLAKSP 300
Catharus_ustulatus FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESREGSWERGTTHRAASLAKSP 300
Corvus_kubaryi FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSP 300
Corvus_hawaiiensis FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSP 300
Corvus_moneduloides FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSP 300
Corvus_cornix FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSP 300
Oxyura_jamaicensis FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKGAPWEEVTTRRSASLAKSP 300
Aythya_fuligula FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKGAPWEEATTRRSASLAKSP 300
Anas_platyrhynchos FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKGAPWEEATMRRSASLAKSP 300
Cygnus_olor FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKGAAWEEATTRRSASLAKSP 300
Cygnus_atratus FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKGAAWEEATTRRSASLAKSP 300
Anser_cygnoides FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKGAAWEEATTRRSASLAKSP 300
Tyto_alba FSLAVVVTNQVTRDSSTSSLKPALGRSWSFVPSTRVLLESKEATWEKATAQRVASLAKSP 300
Calypte_anna FSLTVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKASTQRVASLAKSP 300
Myiozetetes_cayanensis FSLAVVVTNQVTRDSSTGSLKPALGRSWSFVPSTRVLLESKEGGGDKGPTQRVAVLAKSP 300
Chiroxiphia_lanceolata FSLAVVVTNQVTRDSSTGSLKPALGRSWSFVPSTRVLLESKGASWDKGMTRRVAALVKSP 300
Neopelma_chrysocephalum FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEAAWDKGTTRRVAALAKSP 300
Manacus_vitellinus FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESREASWDKGTTRRVAALAKSP 300
Manacus_candei FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESREASWDKGTTRRVAALAKSP 300
Falco_cherrug FSLAVVVTNQVTRDSSTSQLKPALGRSWSFVPSTRVLLENKEATQEKATTQRLASLAKSP 300
Falco_peregrinus FSLAVVVTNQVTRDSSTSQLKPALGRSWSFVPSTRVLLENKEATQEKATTQRLASLAKSP 300
Falco_naumanni FSLAVVVTNQVTRDSSTSQLKPALGRSWSFVPSTRVLLENKEATQEKATTQRLASLAKSP 300
Columba_livia FSLAVVVTNQVTRDSSTGLLKPALGRSWSFVPSTRVLLESKEATWEKVTTQRVASLAKSP 300
Athene_cunicularia FSLAVVVTNQVTRDSSTGQLKPALGRSWSFVPSTRVLLESKEATWEKATTQRIASLAKSS 300
Egretta_garzetta LSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKETTLGKGTTQRVASLAKSP 300
Aquila_chrysaetos FSLAVVVTNQVTRESSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSS 300
Accipiter_gentilis FSIAVVVTNQVTRESSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSS 300
Gymnogyps_californianus FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLAKSP 300
Antrostomus_carolinensis FSLAVVVTNQVTRDSSTGPLKPALGRSWSFVPSTRVLLESKEATWEKATTQRVASLVKSP 300
Apteryx_rowi FSLAVVVTNQVTRDSSAGPLKPALGRSWSFVPSTRVLLQSKEATLEKTNTRRXASLAKSP 300
Nothoprocta_perdicaria FSLAVVVTNQVTRDSGTGPLKPALGRSWSFVPSTRVLLESKEVAWEKGGMQRVASLAKSP 300


c.973G>A; (p.Gly325Ser)
c.919G>A; (p.Glu307Lys)
Homo_sapiens RQPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Parus_major RQPAGIQVELDIGNGVVLEPSLVTPTQGL 329
Cyanistes_caeruleus RQPAGIQVELDIGNGVMLEPSLVTPTQGL 329
Lonchura_striata RQPTGIQVELDIGNGDVLDPSLVTPTQGL 329
Camarhynchus_parvulus RQPTGIQVELDIGNGDVLDPSLVTPTQGL 329
Geospiza_fortis RQPTGIQVELDIGNGDVLDPSLVTPTQGL 329
Passer_montanus RQPTGIQVELDLGNGDVLEPSLVTATQGL 329
Zonotrichia_albicollis RQPTGIQVELDIGNGDMLEPSLVTPTEGL 329
Hirundo_rustica RQPTGIQVELDIGNGDVLEPSLVTPTQGL 329
Motacilla_alba RQPTGIQVELDIGNGDVLEASLVTPTQGL 329
Catharus_ustulatus RQPTGIQVELDIGNGDVLEPSLVTPTQGL 329
Corvus_kubaryi RQPTGIQVELDIGNGDVLEPSPVTPTQGL 329
Corvus_hawaiiensis RQPTGIQVELDIGNGDALEPSPVTPTQGL 329
Corvus_moneduloides RQPTGIQVELDIGNGDALEPSPVTPTQGL 329
Corvus_cornix RQPTGIQVELDIGNGDDLEPSPVTPTQGL 329
Oxyura_jamaicensis RQPTGMQVELDIGNSDVQEPSPATPTEGL 329
Aythya_fuligula RQPTGMQVELDIGNSDVQEPSPATPAEGL 329
Anas_platyrhynchos RQPTGMQVELDIGNSDVQEPSPATPTEGL 329
Cygnus_olor RQPTGMQVELDIGNSDVQESSPGTPTEGL 329
Cygnus_atratus RQPTGMQVELDIGNSDVQESSPATPTEGL 329
Anser_cygnoides RQPTGMQVELDIGNSDVQEPSPATPTEGL 329
Tyto_alba RQPTGIQVELDLGNSDVQELSPATPAQGL 329
Calypte_anna RQPTGIQVELNIGDVDVQEPSPVTPTQGP 329
Myiozetetes_cayanensis RQPTGLQVELDIGNGDMPEPSPVTPTQGL 329
Chiroxiphia_lanceolata RQPTGLQVELDIGNDDMPEPSPVTPTQGL 329
Neopelma_chrysocephalum RQPTGLQVELDIGNGDMSEPSPVTPTQGL 329
Manacus_vitellinus RQPTGLQVELDIGNGAMPEPSPVTPTQGL 329
Manacus_candei RQPTGLQVELDIGNGAIPEPSPVTPTQGL 329
Falco_cherrug RQPMGIQVELDIGNSDVQELSPVTPTQGL 329
Falco_peregrinus RQPMGIQVELDIGNSDVQELSPVTPTQGL 329
Falco_naumanni RQPMGIQVELDIGNGDVQELSPVTPTQGL 329
Columba_livia RQPTGIQVELDIGNFEAREPSPATPTQGL 329
Athene_cunicularia RQPTGIQVEVDIGNSHVKELSPATPTQGL 329
Egretta_garzetta RQPTGIQVELDTGNGDVQELSPVTPTHGP 329
Aquila_chrysaetos RQPTGIQVELDIGNGDVQELSPATLTQGL 329
Accipiter_gentilis RQPTGIQVELDIGNGDVQELSPATPTQGL 329
Gymnogyps_californianus RQPTGIQVELDIGNRDVQELSPVTPTQGL 329
Antrostomus_carolinensis RQPTGIQVELDIGNVDVQELSLVTPTQGL 329
Apteryx_rowi RQPTGIQVVLGVGSGGAQELSPATPTQGL 329
Nothoprocta_perdicaria RQPTGIQVQLDIGNSNGQGGSPATTTQGM 329
*** *:* :. * * .

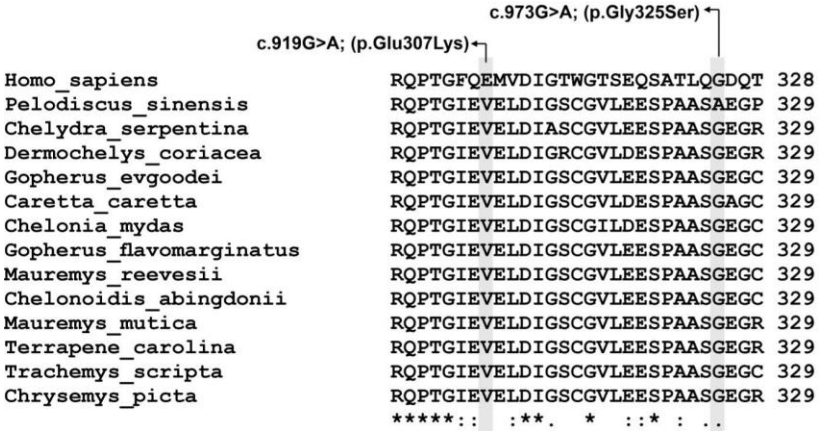
Alinhamento 3: anfíbios e répteis



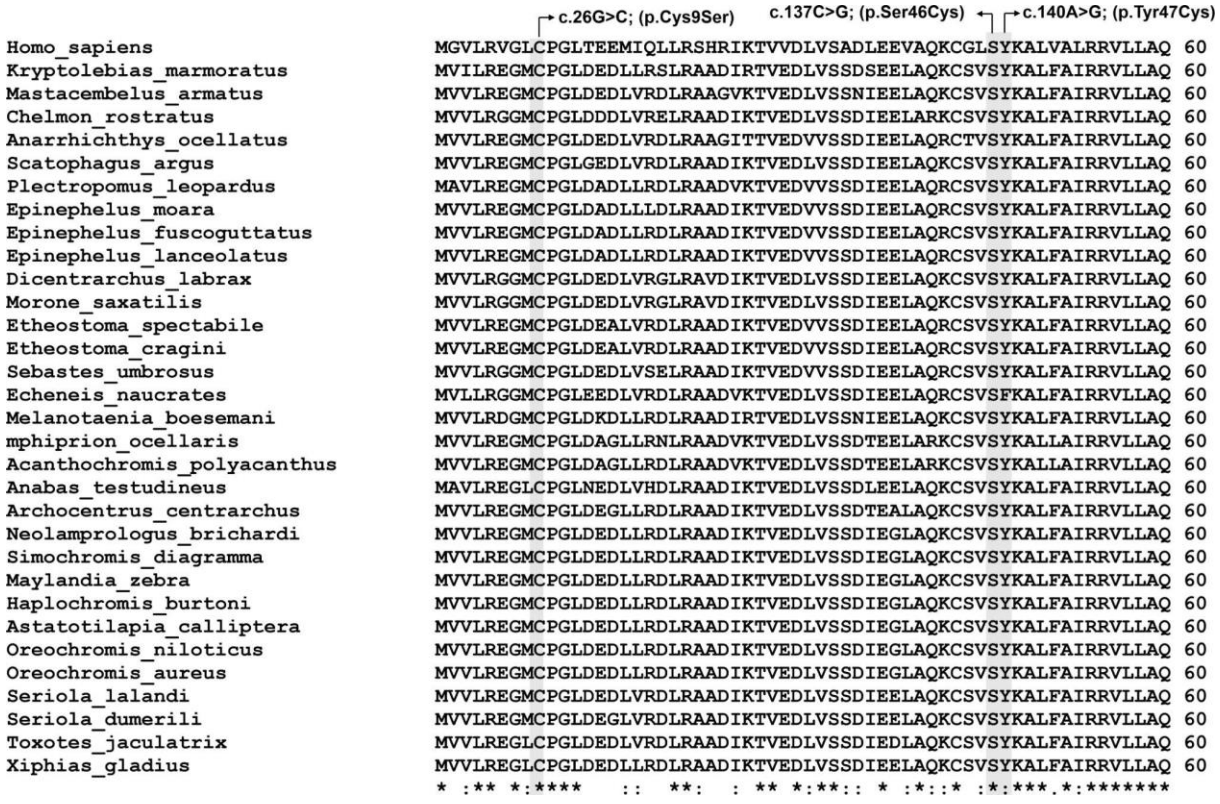

c.413A>G; (p.Asn138Ser)
c.493C>T; (p.Arg165Trp)
Homo_sapiens AANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQAEALRRIQVVHAFDIFQMLD 180
Pelodiscus_sinensis AMSAAHGLRQSVLYVDSTGGFTASRLLQLAQTRMGGEEEQAEALQRIQVARVFDVYAMLD 180
Chelydra_serpentina AMSTSHDLKQNVLYVDSTAGFTASRLLQLAQARTGDEEEQVEALQRIQVARVFDVYKLLD 180
Dermochelys_coriacea AVSTSHDLKQNVLYVDSTSGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Gopherus_evgoodei AMSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Caretta_caretta AVSSSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Chelonia_mydas AVSSSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Gopherus_flavomarginatus AVSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Mauremys_reevesii AVSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALRKIQVARVFDVYKMLD 180
Chelonoidis_abingdonii AVSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Mauremys_mutica AVSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQKIQVARVFDVYKMLD 180
Terrapene_carolina AVSTSHDLKQNVLYVDSTGGFMASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Trachemys_scripta AVSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180
Chrysemys_picta AVSTSHDLKQNVLYVDSTGGFTASRLLQLAQTRTGDEEEQVEALQRIQVARVFDVYKMLD 180

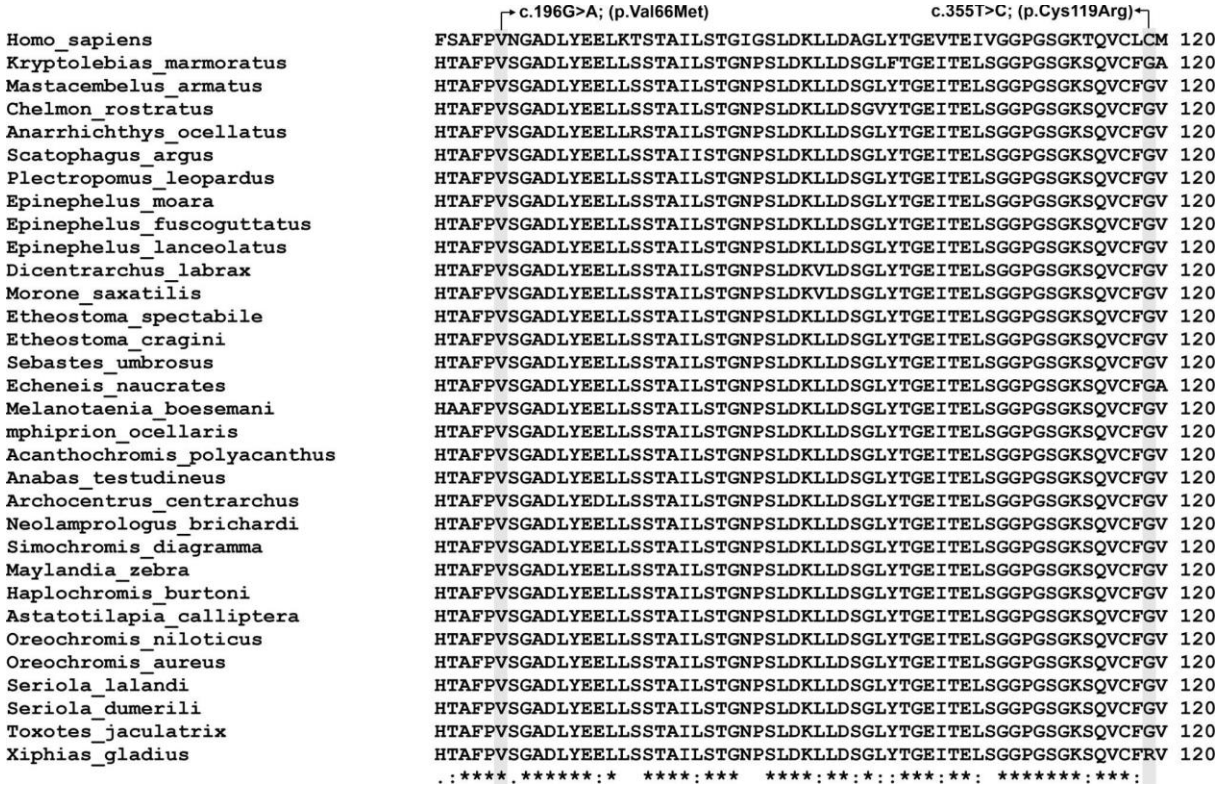

c.715C>T; (p.Arg239Trp)
c.628G>A; (p.Ala210Thr)
c.698A>G; (p.Glu233Gly)
c.695G>A; (p.Arg232Gln)
Homo_sapiens VLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQREGLALMMQLARELKTLARD 240
Pelodiscus_sinensis VVQELRCSMSQQVLSSSAPVKVLVVDSVSAVICPLLGGWQAEGLALMMQLARELKMLARE 240
Chelydra_serpentina VLQELRCSVCQQVLSASGPVKVLLVDSASAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Dermochelys_coriacea VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Gopherus_evgoodei VLQELRCSVSQQVLSSLGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Caretta_caretta VLQELRCSVSQEVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Chelonia_mydas VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Gopherus_flavomarginatus VLQELRCSVSQQVLSSLGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Mauremys_reevesii VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Chelonoidis_abingdonii VLQELRCSVSQQVLSSLGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Mauremys_mutica VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Terrapene_carolina VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLALMMQLARELKTLARE 240
Trachemys_scripta VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLAIMMQLARELKTLARE 240
Chrysemys_picta VLQELRCSVSQQVLSSSGPVKVLVVDSVSAVICPLLGGRQAEGLAIMMQLARELKTLARE 240

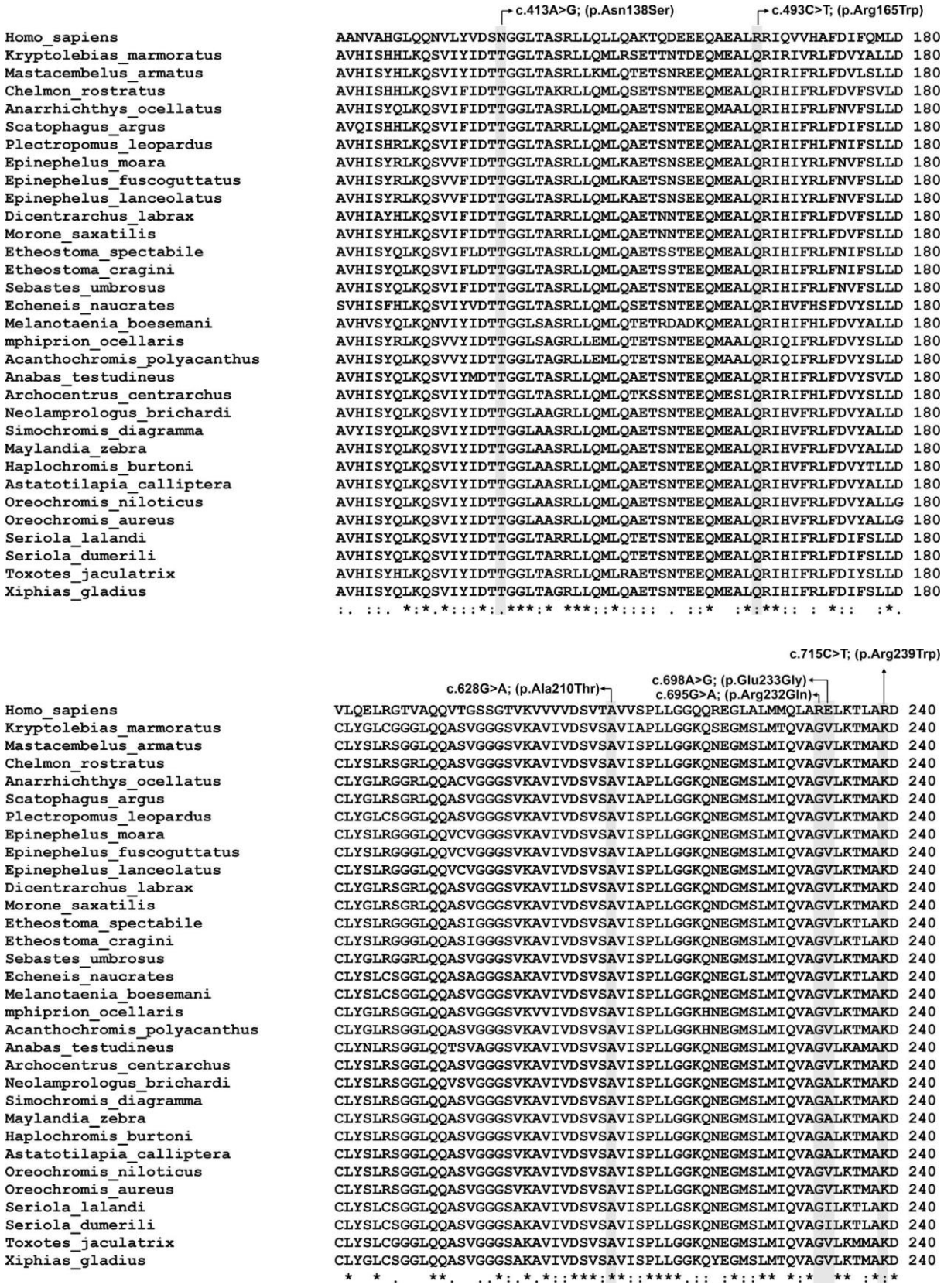

c.862G>C; (p.Gly288Arg)
Homo_sapiens LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGA-SGGRRMACLAKSS 299
Pelodiscus_sinensis LSMAVVVTNHVTRESSSGHLKPALGRSWSFVPSTRVLLESRAGPGGNASTHRVATLTKSS 300
Chelydra_serpentina LSMAVVVTNHVTKDSSSGHLKPALGRSWSFVPSTRVLLESREGTWGKPSTRRLASLTKSS 300
Dermochelys_coriacea LNMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESRGSTWGKASTCRIASLTKSP 300
Gopherus_evgoodei LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTRRIASLTKSS 300
Caretta_caretta LNVAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTRRIASLTKSP 300
Chelonia_mydas LNVAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTRRIASLTKSP 300
Gopherus_flavomarginatus LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTRRIASLTKSS 300
Mauremys_reevesii LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTRRIASLTKSS 300
Chelonoidis_abingdonii LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTCRIASLTKSS 300
Mauremys_mutica LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGKASTRRIASLTKSS 300
Terrapene_carolina LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESRESTWGQASTRRIASLTKSP 300
Trachemys_scripta LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGQASTRRIASLTKSP 300
Chrysemys_picta LSMAVVVTNHVTRDGSSGHLKPALGRSWSFVPSTRVLLESREGTWGQASTHRIASLTKSP 300

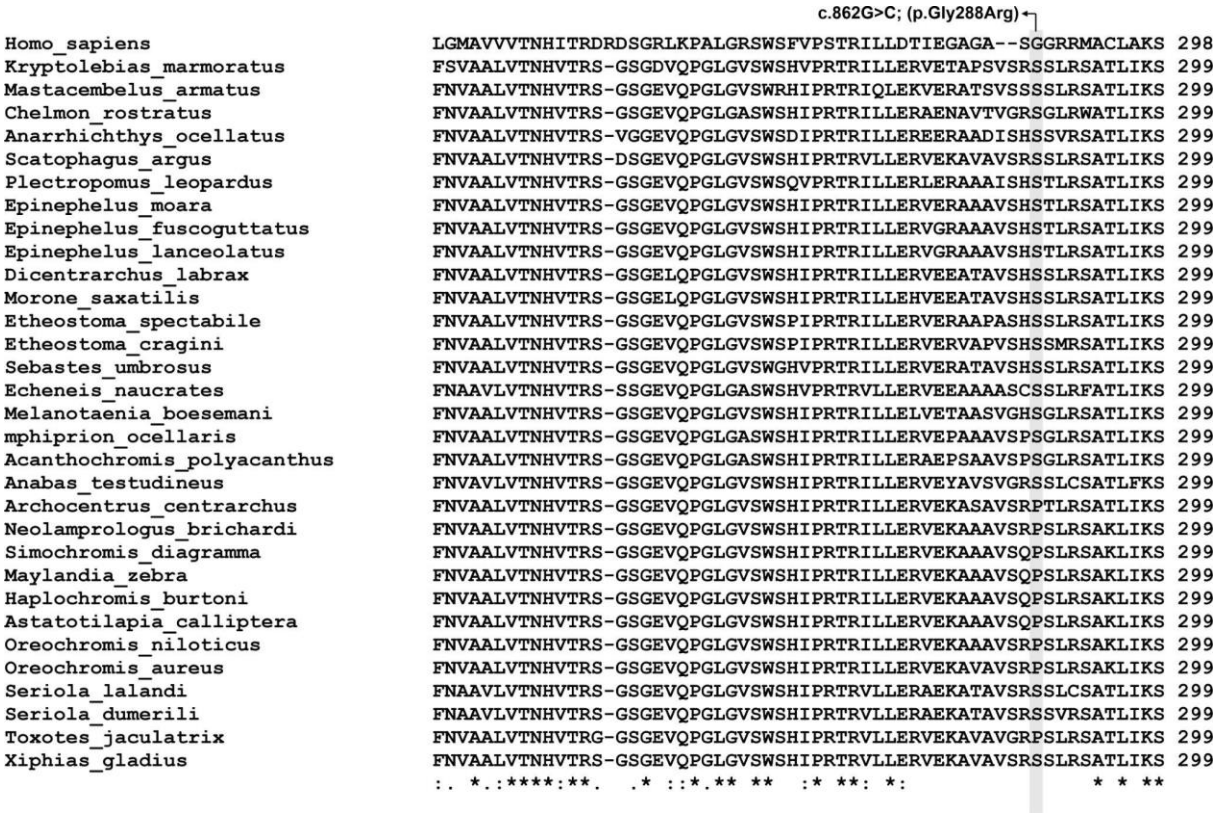

c.973G>A; (p.Gly325Ser)
c.919G>A; (p.Glu307Lys)
Homo_sapiens RQPTGFQEMVDIGTWGTSEQSATLQGDQT 328
Pelodiscus_sinensis RQPTGIEVELDIGSCGVLEESPAASAEGP 329
Chelydra_serpentina RQPTGIEVELDIASCGVLEESPAASGEGR 329
Dermochelys_coriacea RQPTGIEVELDIGRCGVLDESPAASGEGR 329
Gopherus_evgoodei RQPTGIEVELDIGSCGVLEESPAASGEGC 329
Caretta_caretta RQPTGIEVELDIGSCGVLDESPAASGAGC 329
Chelonia_mydas RQPTGIEVELDIGSCGILDESPAASGEGC 329
Gopherus_flavomarginatus RQPTGIEVELDIGSCGVLEESPAASGEGC 329
Mauremys_reevesii RQPTGIEVELDIGSCGVLEESPAASGEGC 329
Chelonoidis_abingdonii RQPTGIEVELDIGSCGVLEESPAASGEGC 329
Mauremys_mutica RQPTGIEVELDIGSCGVLEESPAASGEGR 329
Terrapene_carolina RQPTGIEVELDIGSCGVLEESPAASGEGR 329
Trachemys_scripta RQPTGIEVELDIGSCGVLEESPAASGEGC 329
Chrysemys_picta RQPTGIEVELDIGSCGVLEESPAASGEGR 329

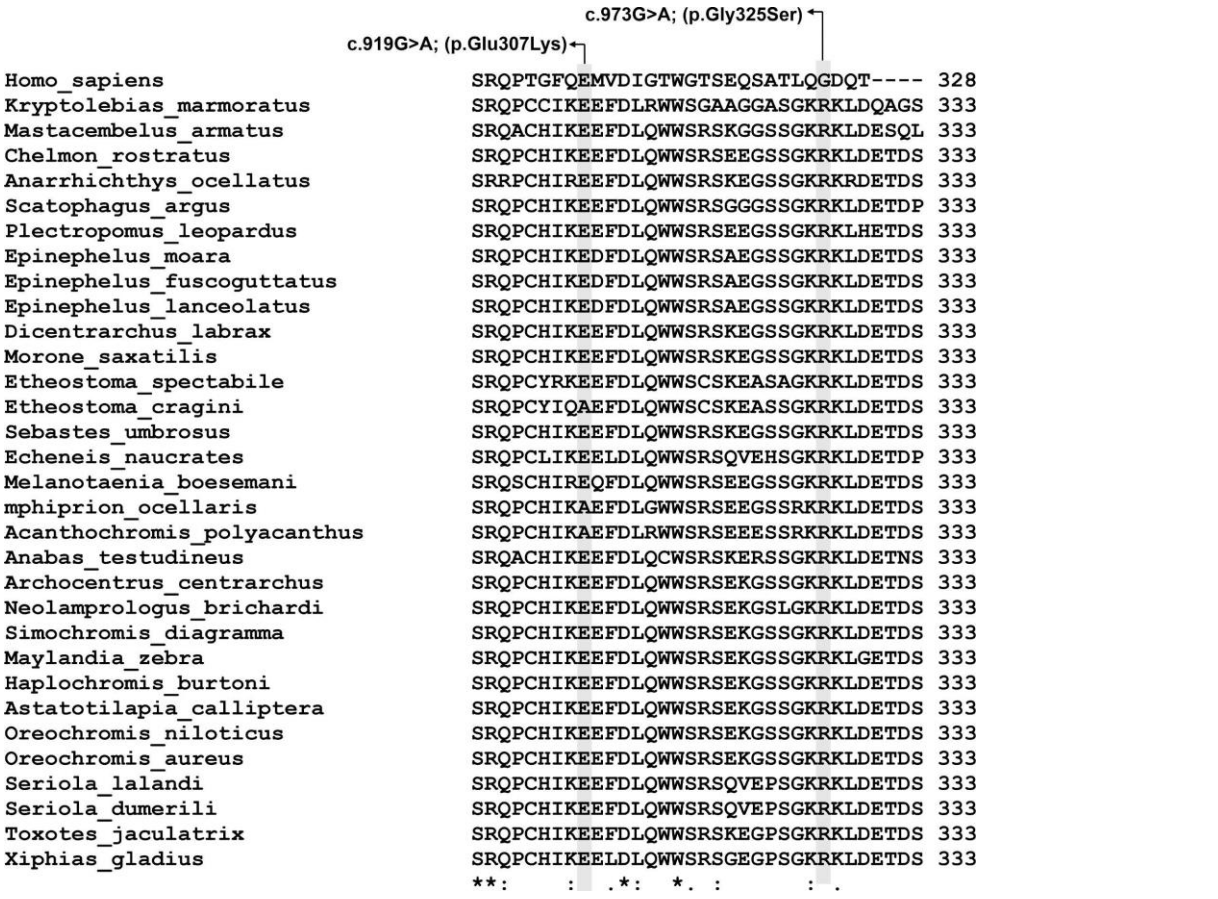
Alinhamento 4: peixes





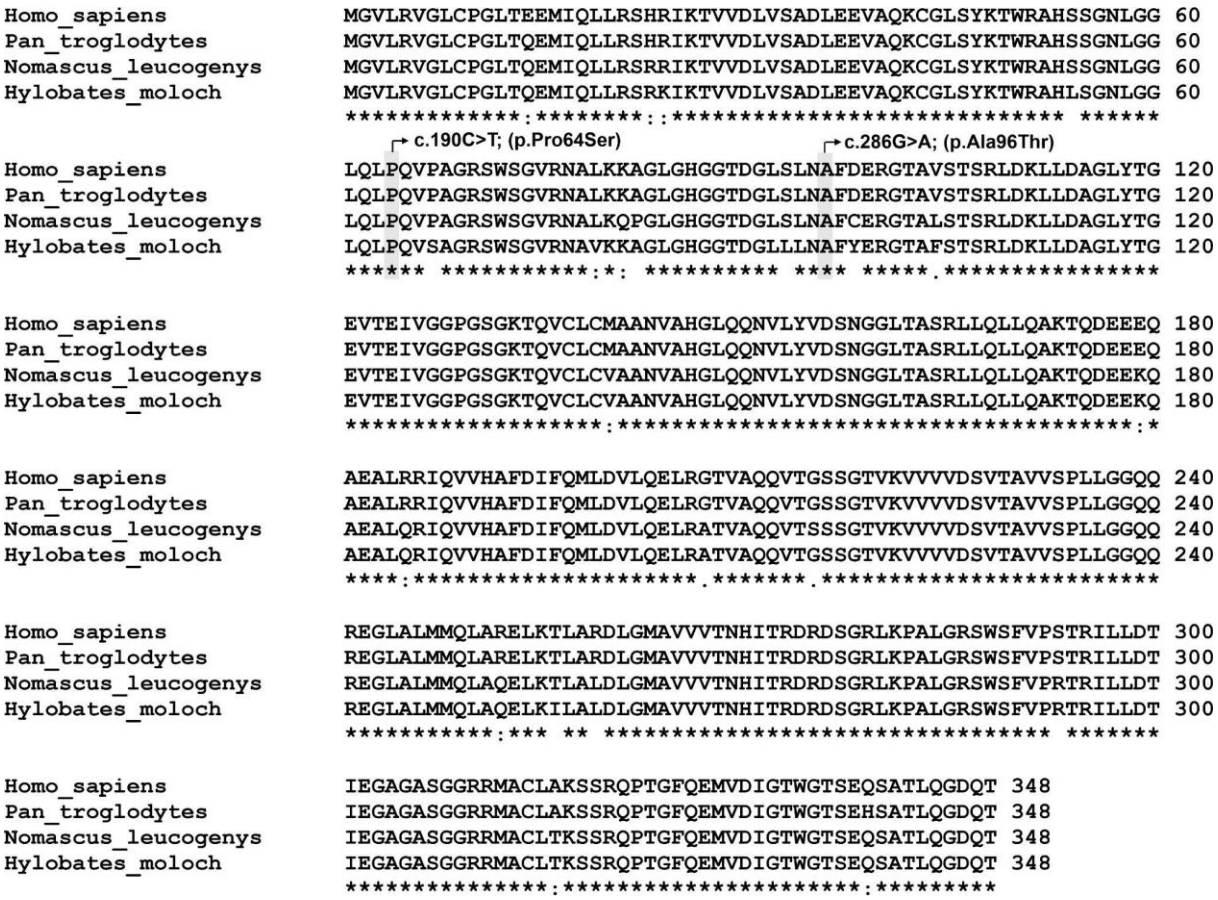

c.413A>G; (p.Asn138Ser)
c.493C>T; (p.Arg165Trp)
c.715C>T; (p.Arg239Trp)
c.628G>A; (p.Ala210Thr)
c.698A>G; (p.Glu233Gly)
c.695G>A; (p.Arg232Gln)


c.862G>C; (p.Gly288Arg)
Homo_sapiens LGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDTIEGAGA--SGGRRMACLAKS 298
Kryptolebias_marmoratus FSVAALVTNHVTRS-GSGDVQPGLGVSWSHVPRTRILLERVETAPSVSRSSLRSATLIKS 299
Mastacembelus_armatus FNVAALVTNHVTRS-GSGEVQPGLGVSWRHIPRTRIQLEKVERATSVSSSSLRSATLIKS 299
Chelmon_rostratus FNVAALVTNHVTRS-GSGEVQPGLGASWSHIPRTRILLERAENAVTVGRSGLRWATLIKS 299
Anarrhichthys_ocellatus FNVAALVTNHVTRS-VGGEVQPGLGVSWSDIPRTRILLEREERAADISHSSVRSATLIKS 299
Scatophagus_argus FNVAALVTNHVTRS-DSGEVQPGLGVSWSHIPRTRVLLERVEKAVAVSRSSLRSATLIKS 299
Plectropomus_leopardus FNVAALVTNHVTRS-GSGEVQPGLGVSWSQVPRTRILLERLERAAAISHSTLRSATLIKS 299
Epinephelus_moara FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVERAAAVSHSTLRSATLIKS 299
Epinephelus_fuscoguttatus FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVGRAAAVSHSTLRSATLIKS 299
Epinephelus_lanceolatus FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVGRAAAVSHSTLRSATLIKS 299
Dicentrarchus_labrax FNVAALVTNHVTRS-GSGELQPGLGVSWSHIPRTRILLERVEEATAVSHSSLRSATLIKS 299
Morone_saxatilis FNVAALVTNHVTRS-GSGELQPGLGVSWSHIPRTRILLEHVEEATAVSHSSLRSATLIKS 299
Etheostoma_spectabile FNVAALVTNHVTRS-GSGEVQPGLGVSWSPIPRTRILLERVERAAPASHSSLRSATLIKS 299
Etheostoma_cragini FNVAALVTNHVTRS-GSGEVQPGLGVSWSPIPRTRILLERVERVAPVSHSSMRSATLIKS 299
Sebastes_umbrosus FNVAALVTNHVTRS-GSGEVQPGLGVSWGHVPRTRILLERVERATAVSHSSLRSATLIKS 299
Echeneis_naucrates FNAAVLVTNHVTRS-SSGEVQPGLGASWSHVPRTRVLLERVEEAAAASCSSLRFATLIKS 299
Melanotaenia_boesemani FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLELVETAASVGHSGLRSATLIKS 299
mphiprion_ocellaris FNVAALVTNHVTRS-GSGEVQPGLGASWSHIPRTRILLERVEPAAAVSPSGLRSATLIKS 299
Acanthochromis_polyacanthus FNVAALVTNHVTRS-GSGEVQPGLGASWSHIPRTRILLERAEPSAAVSPSGLRSATLIKS 299
Anabas_testudineus FNVAVLVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEYAVSVGRSSLCSATLFKS 299
Archocentrus_centrarchus FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKASAVSRPTLRSATLIKS 299
Neolamprologus_brichardi FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAAAVSRPSLRSAKLIKS 299
Simochromis_diagramma FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAAAVSQPSLRSAKLIKS 299
Maylandia_zebra FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAAAVSQPSLRSAKLIKS 299
Haplochromis_burtoni FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAAAVSQPSLRSAKLIKS 299
Astatotilapia_calliptera FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAAAVSQPSLRSAKLIKS 299
Oreochromis_niloticus FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAAAVSRPSLRSAKLIKS 299
Oreochromis_aureus FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRILLERVEKAVAVSRPSLRSAKLIKS 299
Seriola_lalandi FNAAVLVTNHVTRS-GSGEVQPGLGVSWSHIPRTRVLLERAEKATAVSRSSLCSATLIKS 299
Seriola_dumerili FNAAVLVTNHVTRS-GSGEVQPGLGVSWSHIPRTRVLLERAEKATAVSRSSVRSATLIKS 299
Toxotes_jaculatrix FNVAALVTNHVTRG-GSGEVQPGLGVSWSHIPRTRVLLERVEKAVAVGRPSLRSATLIKS 299
Xiphias_gladius FNVAALVTNHVTRS-GSGEVQPGLGVSWSHIPRTRVLLERVEKAVAVSRSSLRSATLIKS 299
:. *.:****:**. .* ::*.** ** :* **: *: * * **


c.973G>A; (p.Gly325Ser)
c.919G>A; (p.Glu307Lys)
Homo_sapiens SRQPTGFQEMVDIGTWGTSEQSATLQGDQT---- 328
Kryptolebias_marmoratus SRQPCCIKEEFDLRWWSGAAGGASGKRKLDQAGS 333
Mastacembelus_armatus SRQACHIKEEFDLQWWSRSKGGSSGKRKLDESQL 333
Chelmon_rostratus SRQPCHIKEEFDLQWWSRSEEGSSGKRKLDETDS 333
Anarrhichthys_ocellatus SRRPCHIREEFDLQWWSRSKEGSSGKRKRDETDS 333
Scatophagus_argus SRQPCHIKEEFDLQWWSRSGGGSSGKRKLDETDP 333
Plectropomus_leopardus SRQPCHIKEEFDLQWWSRSEEGSSGKRKLHETDS 333
Epinephelus_moara SRQPCHIKEDFDLQWWSRSAEGSSGKRKLDETDS 333
Epinephelus_fuscoguttatus SRQPCHIKEDFDLQWWSRSAEGSSGKRKLDETDS 333
Epinephelus_lanceolatus SRQPCHIKEDFDLQWWSRSAEGSSGKRKLDETDS 333
Dicentrarchus_labrax SRQPCHIKEEFDLQWWSRSKEGSSGKRKLDETDS 333
Morone_saxatilis SRQPCHIKEEFDLQWWSRSKEGSSGKRKLDETDS 333
Etheostoma_spectabile SRQPCYRKEEFDLQWWSCSKEASAGKRKLDETDS 333
Etheostoma_cragini SRQPCYIQAEFDLQWWSCSKEASSGKRKLDETDS 333
Sebastes_umbrosus SRQPCHIKEEFDLQWWSRSKEGSSGKRKLDETDS 333
Echeneis_naucrates SRQPCLIKEELDLQWWSRSQVEHSGKRKLDETDP 333
Melanotaenia_boesemani SRQSCHIREQFDLQWWSRSEEGSSGKRKLDETDS 333
mphiprion_ocellaris SRQPCHIKAEFDLGWWSRSEEGSSRKRKLDETDS 333
Acanthochromis_polyacanthus SRQPCHIKAEFDLRWWSRSEEESSRKRKLDETDS 333
Anabas_testudineus SRQACHIKEEFDLQCWSRSKERSSGKRKLDETNS 333
Archocentrus_centrarchus SRQPCHIKEEFDLQWWSRSEKGSSGKRKLDETDS 333
Neolamprologus_brichardi SRQPCHIKEEFDLQWWSRSEKGSLGKRKLDETDS 333
Simochromis_diagramma SRQPCHIKEEFDLQWWSRSEKGSSGKRKLDETDS 333
Maylandia_zebra SRQPCHIKEEFDLQWWSRSEKGSSGKRKLGETDS 333
Haplochromis_burtoni SRQPCHIKEEFDLQWWSRSEKGSSGKRKLDETDS 333
Astatotilapia_calliptera SRQPCHIKEEFDLQWWSRSEKGSSGKRKLDETDS 333
Oreochromis_niloticus SRQPCHIKEEFDLQWWSRSEKGSSGKRKLDETDS 333
Oreochromis_aureus SRQPCHIKEEFDLQWWSRSEKGSSGKRKLDETDS 333
Seriola_lalandi SRQPCHIKEEFDLQWWSRSQVEPSGKRKLDETDS 333
Seriola_dumerili SRQPCHIKEEFDLQWWSRSQVEPSGKRKLDETDS 333
Toxotes_jaculatrix SRQPCHIKEEFDLQWWSRSKEGPSGKRKLDETDS 333
Xiphias_gladius SRQPCHIKEELDLQWWSRSGEGPSGKRKLDETDS 333
**: : .*: *. : : .

ANEXO 5 - Alinhamento múltiplo da isoforma 8 (O75771-8)

Alinhamento 5: mamíferos




```
Homo_sapiens            MGVLRVGLCPGLTEEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKTWRAHSSGNLGG 60
Pan_troglodytes         MGVLRVGLCPGLTQEMIQLLRSHRIKTVVDLVSADLEEVAQKCGLSYKTWRAHSSGNLGG 60
Nomascus_leucogenys     MGVLRVGLCPGLTQEMIQLLRSRRIKTVVDLVSADLEEVAQKCGLSYKTWRAHSSGNLGG 60
Hylobates_moloch        MGVLRVGLCPGLTQEMIQLLRSRKIKTVVDLVSADLEEVAQKCGLSYKTWRAHLSGNLGG 60
                        ************:********::***************************** ******

Homo_sapiens            LQLPQVPAGRSWSGVRNALKKAGLGHGGTDGLSLNAFDERGTAVSTSRLDKLLDAGLYTG 120
Pan_troglodytes         LQLPQVPAGRSWSGVRNALKKAGLGHGGTDGLSLNAFDERGTAVSTSRLDKLLDAGLYTG 120
Nomascus_leucogenys     LQLPQVPAGRSWSGVRNALKQPGLGHGGTDGLSLNAFCERGTALSTSRLDKLLDAGLYTG 120
Hylobates_moloch        LQLPQVSAGRSWSGVRNAVKKAGLGHGGTDGLLLNAFYERGTAFSTSRLDKLLDAGLYTG 120
                        ****** ***********:*: ********** **** *****.**************

Homo_sapiens            EVTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQ 180
Pan_troglodytes         EVTEIVGGPGSGKTQVCLCMAANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEEQ 180
Nomascus_leucogenys     EVTEIVGGPGSGKTQVCLCVAANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEKQ 180
Hylobates_moloch        EVTEIVGGPGSGKTQVCLCVAANVAHGLQQNVLYVDSNGGLTASRLLQLLQAKTQDEEKQ 180
                        ******************:*************************************:*

Homo_sapiens            AEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQ 240
Pan_troglodytes         AEALRRIQVVHAFDIFQMLDVLQELRGTVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQ 240
Nomascus_leucogenys     AEALQRIQVVHAFDIFQMLDVLQELRATVAQQVTSSSGTVKVVVVDSVTAVVSPLLGGQQ 240
Hylobates_moloch        AEALQRIQVVHAFDIFQMLDVLQELRATVAQQVTGSSGTVKVVVVDSVTAVVSPLLGGQQ 240
                        ****:********************.*******.************************

Homo_sapiens            REGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDT 300
Pan_troglodytes         REGLALMMQLARELKTLARDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPSTRILLDT 300
Nomascus_leucogenys     REGLALMMQLAQELKTLALDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPRTRILLDT 300
Hylobates_moloch        REGLALMMQLAQELKILALDLGMAVVVTNHITRDRDSGRLKPALGRSWSFVPRTRILLDT 300
                        ***********:*** ** ********************************* *******

Homo_sapiens            IEGAGASGGRRMACLAKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT 348
Pan_troglodytes         IEGAGASGGRRMACLAKSSRQPTGFQEMVDIGTWGTSEHSATLQGDQT 348
Nomascus_leucogenys     IEGAGASGGRRMACLTKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT 348
Hylobates_moloch        IEGAGASGGRRMACLTKSSRQPTGFQEMVDIGTWGTSEQSATLQGDQT 348
                        ***************:**********************:*********
```