



Ministério da Saúde

Instituto Nacional de Câncer

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação Em Oncologia

**AVALIAÇÃO FUNCIONAL DO IMPACTO DE VARIANTES DE
SENTIDO TROCADO E GERADAS POR INSERÇÕES E/OU
DELEÇÕES QUE RESPEITAM O QUADRO DE LEITURA NA
PROTEÍNA BRCA1**

ANA PAULA FERNANDES PAZ DOS SANTOS

Orientador: Marcelo Alex de Carvalho

Coorientador: Thales da Costa Nepomuceno

Rio de Janeiro
2021



Ministério da Saúde

Instituto Nacional de Câncer

INSTITUTO NACIONAL DE CÂNCER

Pós-Graduação Em Oncologia

ANA PAULA FERNANDES PAZ DOS SANTOS

AVALIAÇÃO FUNCIONAL DO IMPACTO DE VARIANTES DE SENTIDO TROCADO E GERADAS POR INSERÇÕES E/OU DELEÇÕES QUE RESPEITAM O QUADRO DE LEITURA NA PROTEÍNA BRCA1

Dissertação apresentada ao Programa de Pós-Graduação em Oncologia do Instituto Nacional de Câncer, como requisito parcial para obtenção do título de Mestre em Oncologia.

Orientador: Marcelo Alex de Carvalho

Coorientador: Thales da Costa Nepomuceno

Rio de janeiro

2021

S237a Santos, Ana Paula Fernandes Paz dos
Avaliação funcional do impacto de variantes de sentido trocado e geradas por inserções e/ou deleções que respeitam o quadro de leitura na proteína BRCA1. / Ana Paula Fernandes Paz dos Santos. – Rio de Janeiro, 2021.
87 f.: il.

Dissertação (Mestrado em Oncologia) – Programa de Pós-Graduação em Oncologia,
Instituto Nacional de Câncer José Alencar Gomes da Silva, 2021.

Orientador: Dr. Marcelo Alex de Carvalho

Coorientador: Dr. Thales da Costa Nepomuceno

1. Neoplasias. 2. Neoplasias da Mama. 3. Proteína BRCA1. 4. Variantes Farmacogenômicos.
I. Carvalho, Marcelo Alex de (Orient.). II. Nepomuceno, Thales da Costa (CoOrient.).
III. Instituto Nacional de Câncer José Alencar Gomes da Silva. IV. Título.

CDD 616.99404249

Catálogo na fonte

Kátia Simões CRB7/5952



Ministério da Saúde

Instituto Nacional de Câncer

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação Em Oncologia

ANA PAULA FERNANDES PAZ DOS SANTOS

AVALIAÇÃO FUNCIONAL DO IMPACTO DE VARIANTES DE SENTIDO TROCADO E GERADAS POR INSERÇÕES E/OU DELEÇÕES QUE RESPEITAM O QUADRO DE LEITURA NA PROTEÍNA BRCA1

Orientador: Marcelo Alex de Carvalho

Coorientador: Thales da Costa Nepomuceno

Aprovada em: ____/____/____

EXAMINADORES:

Dr. Miguel Angelo Martins Moreira – INCA - Titular

Dra. Rosane Vianna-Jorge – UFRJ – Titular

Dr. Rafael Dias Mesquita – UFRJ – Titular

Dra. Tuane Cristine Ramos Goncalves Vieira – UFRJ – Suplente

Dr. Julio Cesar Madureira de Freitas Junior – INCA - Suplente

Rio de Janeiro
2021

“Do not go gentle into that good night.”

Dylan Thomas

Agradecimentos

Sou grata a mim, à parte boa de mim, que não me deixou desistir.

Aos meus pais, por passarem por tudo comigo, e só eles realmente sabem o que é “tudo”. E minha terapeuta. E minha psiquiatra. E minhas gatas.

Na verdade, eu tenho muita gente para agradecer. Em 2019 eu senti como se tivesse vivido dois anos em um. No primeiro ano eu tenho que agradecer à Babi, à Rafaela, e à Bárbara. Vocês três foram as únicas que ficaram ao meu lado real na hora do perrengue. Se eu estivesse completamente sozinha lá, tenho certeza de que eu não estaria aqui agora, escrevendo esse texto.

No segundo ano eu sou grata a mais gente do que achei que seria... A todos da pesquisa clínica, na verdade! A recepção por vocês naquela salinha de estudos foi surpreendentemente linda e calorosa.

Eu agradeço e devo muito ao meu grupo de pesquisa, à Anna, Thiago, Vanessa, Thales, Fernanda e Vitória. Esse grupo é uma família. Não aquele tipo de família que se reúne no Natal e faz treta todo ano, mas aquela família que te zoa, mas te ama e te apoia em tudo, e te ensina. Ao Thales principalmente, meu coorientador, que me ensinou tudo o que eu tinha que saber na bancada, que me guiou e ajudou e brigou comigo e se preocupou até o final. Eu não tenho dúvidas de que você vai ser um ótimo orientador quando formar seu próprio grupo.

Sou MUITO grata, em caixa alta, ao meu orientador Marcelo Alex. Você se deparou com essa aluna de mestrado com 9 meses para defender batendo na sua porta, e você me acolheu, e me formou para ser um membro do seu grupo em todos os aspectos. Horas e horas de conversa na sua sala, de muito apoio, de muitos conselhos e muito incentivo. Mas muito incentivo mesmo. O que você vê em mim como aluna de pós-graduação parece óbvio pra você, mas é e sempre foi o mais difícil de enxergar para mim. Obrigada por despejar toda a sua sinceridade, ela foi mais motivadora do que você pensa.

Agradeço aos amigos que fiz na UFF, os que ainda estão ao meu lado. Ana Paula (não, não sou eu), Jessica, Maíra, Pedro e Tate. Amigos não são só companhias pra festas e barzinhos, amigos são principalmente ouvintes, e até apoios em alguns momentos. O apoio de vocês é fundamental na minha vida, e sem essa base

emocional que vocês me provêm todos os dias eu não teria chegado até aqui. Eu sei que posso contar com vocês sempre, da mesma forma que contei durante meu TCC, como eu pude contar agora, e como poderei nos próximos passos. E não tenham dúvida nenhuma de que tudo é recíproco. Eu amo muito todos vocês.

Obrigada, Jéssica Cruz, ou “Jéssica do INCA”. Amiga, você fez muito mais por mim do que possivelmente um familiar faria. E a gente foi família por alguns meses né? Foi bom demais dividir apartamento com você por aqueles meses, nos encher de pizza e depois passar mal, vendo filmes e séries que a gente já viu. Da mesma forma, todos os momentos em que você me pegou tristonha e chorando, em que você só dizia “ô bicha”, e passava a mão na minha cabeça. Você está marcada no meu coração.

Insiro em meus agradecimentos algo que eu não imaginei que colocaria no início dessa jornada: meus amigos que fiz na quarentena, que em conjunto chamarei sim de Arabollo, porque se eu citar todos vocês eu fico sem espaço. Fugir da minha zona de conforto foi muito importante e muito bom pra mim em 2020, e eu devo muito disso a vocês.

Agradeço à banca examinadora por aceitarem meu convite, e por contribuírem para esse momento da minha formação.

Por fim, agradeço ao Programa de Pós-Graduação em Oncologia do INCA por me aceitar como aluna e contribuir de forma tão grandiosa com a minha formação.



Ministério da Saúde

Instituto Nacional de Câncer

AVALIAÇÃO FUNCIONAL DO IMPACTO DE VARIANTES DE SENTIDO TROCADO E GERADAS POR INSERÇÕES E/OU DELEÇÕES QUE RESPEITAM O QUADRO DE LEITURA NA PROTEÍNA BRCA1

RESUMO

BRCA1 (Breast Cancer 1, *early onset*) foi o primeiro gene relacionado à susceptibilidade ao câncer de mama e ovário. Seu produto proteico homônimo desempenha um papel central na resposta ao dano de DNA (RDD). *BRCA1* encerra um domínio *coiled-coil* e dois domínios BRCT (BRCT1 e BRCT2) em tandem (tBRCT) em sua região C-terminal, separados por uma região desestruturada. Os domínios BRCT, quando em tandem, se comportam como uma única unidade estrutural. Esse domínio é crucial para a supressão tumoral mediada por *BRCA1*, tendo em vista que a perda de 11 resíduos de aminoácido na borda C-terminal da proteína resulta na predisposição ao câncer. Logo, variantes sem sentido (*nonsense*) ou que alteram o quadro de leitura codificante são geralmente classificadas como patogênicas. A interpretação do impacto de variantes de sentido trocado (*missense*) e de pequenas deleções e/ou inserções (indels) que respeitam o quadro de leitura da proteína não é simples. Essas variantes, em especial aquelas com baixa frequência na população, tendem a ser classificadas como variantes de significado incerto (*Variants of Uncertain Significance*; VUS). Estudos funcionais representam uma poderosa ferramenta na determinação do impacto de VUS na funcionalidade da proteína. Nosso grupo demonstrou que a integridade do domínio tBRCT de *BRCA1* está diretamente relacionada à sua capacidade de ativar a transcrição quando fusionada a um domínio heterólogo de ligação ao DNA (ensaio de TA). Nos últimos anos, foram avaliadas mais de 300 variantes *missense* localizados na porção C-terminal de *BRCA1* (região compreendida entre os resíduos de aminoácido 1396 e 1863, codificados pelos éxons 13 ao 24) através dessa ferramenta. Contudo, pouco se sabe sobre o impacto funcional de indels que respeitam o quadro de leitura localizadas na região C-terminal de *BRCA1*. Nosso estudo tem foco na caracterização do impacto funcional dessas variantes, bem como na avaliação da importância da borda C-terminal do domínio BRCT2 para manutenção da integridade do domínio tBRCT. Indels previamente reportadas na população foram identificadas através do banco de dados disponível na plataforma *BRCA Exchange*. Nós selecionamos e geramos 16 variantes localizadas entre os éxons 13 e 24 de *BRCA1*. Também foram geradas 6 variantes *nonsense* na porção C-terminal do domínio tBRCT. Essas variantes foram analisadas funcionalmente através do ensaio de TA, conduzido em células humanas da linhagem HEK293FT, e os níveis proteicos avaliados através de rotinas de *immunoblotting*. As variantes localizadas no domínio tBRCT atividade transcricional inferior à 50% do controle selvagem. No entanto, as variantes localizadas na região desestruturada apresentam atividades que variam entre 50% e 160% do comportamento do controle. Interessantemente, a perda de um único resíduo de aminoácido na borda C-terminal do BRCT2 (p.Y1863) resulta na redução de 50% da atividade. Esse comportamento é mantido até a perda do quinto resíduo estudado (p.P1859). A deleção de resíduos adicionais é acompanhada por uma redução significativa da atividade transcricional. Coletivamente, nossos dados mostram o valor de interrogar o impacto funcional dessas variantes. Uma melhor compreensão do seu comportamento é a chave para a compreensão da organização e funções estruturais do tBRCT de *BRCA1*.

Palavras-chave: Câncer, Câncer de mama, *BRCA1*, Variantes de significado incerto



Ministério da Saúde

Instituto Nacional de Câncer

AVALIAÇÃO FUNCIONAL DO IMPACTO DE VARIANTES DE SENTIDO TROCADO E GERADAS POR INSERÇÕES E/OU DELEÇÕES QUE RESPEITAM O QUADRO DE LEITURA NA PROTEÍNA BRCA1

ABSTRACT

BRCA1 (Breast Cancer 1, early onset) was the first gene linked to the hereditary breast and ovarian cancer syndrome. Its homonymous protein product plays a key role in the DNA damage response (RDD). *BRCA1* encloses a coiled-coil and two tandem BRCTs (BRCT1, BRCT2; tBRCT) on its C-terminal region, both separated by a disordered region. The tBRCT domains function as a single structural unit, which is crucial for *BRCA1*-mediated tumor suppression as the loss of its eleven amino acid residues on the protein C-terminal border results in cancer predisposition. Thus, nonsense or frameshift variants are usually classified as pathogenic. The functional implications of missense and small in-frame deletions and/or insertions (indels) are more difficult to predict. These variants, in special the ones that are less frequent in the population, are classified as variants of uncertain significance (VUS). Functional assays are powerful tools for assessing the impact of VUS on specific protein function. Our group showed that *BRCA1* tBRCT domain integrity correlates to its transcriptional activity when fused to a heterologous DNA binding domain (TA assay). We have previously evaluated more than 300 missense variants enclosed within the C-terminal region (1396-1863 amino acid residues, coded by exons 13 to 24) of *BRCA1* using the TA assay. However, little is known about the functional impact of small in-frame indels and nonsense mutations on the C-terminal border of the tBRCT domain. Our study is focused on the functional characterization of these variants and the relevance of the C-terminal border of the BRCT2 domains in the integrity maintenance of the tBRCT. In-frame indels previously reported in the population were identified through the BRCA Exchange platform. We selected and generated all 16 in-frame indels reported in the 13/24 region. We also generated 6 nonsense truncation variants in the C-terminal border of the tBRCT domain. The functional impact of variants was interrogated by the TA assay in human HEK293FT cells. Protein expression was evaluated by immunoblotting. Indels located within the tBRCT domain presented less than 50% of the wild-type activity. However, variants in the disordered region presented activity ranging from 50% to 160% of the control. Interestingly, the loss of a single amino acid residue in the C-terminal border of the BRCT2 (p.Y1863) resulted in a 50% decrease in TA activity. This activity is sustained by the loss of the other five residues (p.P1859). The deletion of additional amino acid residues results in a marked reduction in TA. Collectively, our data show the value of interrogating the functional impact of these variants. A better understanding of their behavior is key to the understanding of *BRCA1* tBRCT structural organization and functions.

keywords: Cancer, Breast Cancer, *BRCA1*, Variants of uncertain significance

ÍNDICE

1. INTRODUÇÃO	1
1.1. Câncer.....	1
1.2. Câncer de Mama.....	1
1.3. BRCA1	4
1.3.1. Funções da proteína BRCA1	9
1.4. Variantes genéticas de BRCA1	12
1.5. Ensaio Funcionais para a classificação de VUS de BRCA1	16
2. JUSTIFICATIVA	23
3. OBJETIVOS	26
3.1. Objetivo geral	26
3.2. Objetivos específicos	26
4. METODOLOGIA.....	27
4.1. Identificação e seleção das variantes descritas	27
4.2. Avaliação de impacto funcional de variantes através de modelos de predição <i>in silico</i>	27
4.3. Geração das variantes	27
4.4. Indução de eletrocompetência bacteriana	32
4.5. Transformação bacteriana	32
4.6. Extração plasmidial	32
4.7. Sequenciamento	33
4.8. Cultura de células humanas	33
4.9. Transfecção de células humanas	34
4.10. Ensaio de ativação transcricional (TA)	34
4.11. Avaliação de níveis proteicos	36
4.11.1. Extratos proteicos	36
4.11.2. <i>Immunoblottings</i>	37
5. RESULTADOS	38
5.1. Seleção das variantes	38
5.2. Avaliação <i>in silico</i> das variantes selecionadas	38
5.3. Avaliação da capacidade de ativação transcricional	42
5.3.1. Capacidade de ativação transcricional das variantes <i>missense</i>	42
5.3.2. Capacidade de ativação transcricional das variantes <i>de deleções de códon único</i>	44

5.3.3. Capacidade de ativação transcricional das variantes <i>nonsense</i> na borda C-terminal de BRCA1	44
5.4. Avaliação dos níveis proteicos das variantes	46
6. DISCUSSÃO	50
7. CONCLUSÃO.....	59
8. REFERÊNCIAS.....	60

LISTA DE TABELAS

Tabela 4.1. Oligonucleotídeos iniciadores (primers) utilizados na geração de variantes para o estudo	30
Tabela 5.1. Classificação das VUS de troca de aminoácido de acordo com <i>Align-GVGD</i>	41

LISTA DE FIGURAS

Figura 1.1. Classificação dos tipos de câncer mais incidentes no Brasil.	2
Figura 1.2. Estrutura da proteína BRCA1.	5
Figura 1.3. Estrutura do tBRCT de BRCA1.	8
Figura 1.4. Representação esquemática dos complexos compostos por BRCA1 através de seu domínio tBRCT.	11
Figura 1.5. Frequência alélica de BRCA1.	14
Figura 1.6. Esquema do ensaio de ativação transcricional.	20
Figura 2.1: Inserções e deleções que respeitam o quadro de leitura.	24
Figura 4.1. Representação esquemática da construção pcDNA3-GAL4 _{DBD} :BRCA1 _{13/24}	28
Figura 4.2. Esquema representativo da rotina de mutagênese sítio-dirigida.	31
Figura 4.3. Representação esquemática dos plasmídeos pG5 _{luc} e pGR-TK.	35
Figura 5.1. Visão geral das variantes de classificação incerta (VUS) de BRCA1.	39
Figura 5.2. Representação esquemática da localização das variantes selecionadas ao longo da proteína de fusão GAL4 _{DBD} :BRCA1 _{13/24}	40
Figura 5.3. Análise da capacidade de ativação transcricional das variantes <i>missense</i>	43
Figura 5.4. Análise da capacidade transcricional das variantes com deleção de códon único	45
Figura 5.5. Análise da capacidade transcricional das variantes nonsense da porção C-terminal do tBRCT de BRCA1	47
Figura 5.6. Perfil das proteínas quiméricas GAL4 _{DBD} :BRCA1 _{13/24}	48
Figura 6.1. Conservação de aminoácidos entre as espécies	52

LISTA DE ANEXOS

Anexo I: Lista de variantes indels que respeitam o quadro de leitura na região codificante de BRCA1	69
---	----

LISTA DE ABREVIATURAS

°C – graus Celsius

ATM- *Ataxia Telangiectasia Mutated*

BACH1 – *BTB and CNC homology 1*

BARD1 – *BRCA1- associated ring domain 1*

BRCA1 – *Breast cancer 1, early onset*

BRCA2 – *Breast cancer 2, early onset*

BRCC36/45 – *BRCA1/BRCA2-Containing Complex Subunit*

BRCT – *BRCA1 C terminus*

CDKN1A – *Cyclin Dependent Kinase Inhibitor 1A*

CHEK2 – *Checkpoint Kinase 2*

CM – Câncer de mama

CO – Câncer de ovário

CtIP – *C-terminal Binding Protein (CtBP) Interacting Protein*

DBD – *DNA binding domain*

DMEM – *Dulbecco's Modified Eagle's Medium*

DNA – *Ácido desoxirribonucleico*

DO – densidade óptica

DSB – *Double Strand Break* (quebra de dupla fita)

EDTA - *ácido etilenodiamina tetraacético*

g – gravidade ($g = 9,8 \text{ cm/s}^2$)

GADD45 – *Growth Arrest And DNA Damage Inducible Alpha*

HBOC – *Hereditary breast and ovarian cancer* (Síndrome hereditária de câncer de mama e ovário)

HCl – *ácido clorídrico*

HR – *Recombinação homóloga*

IARC – *International Agency for Research on Cancer*

IgG – *imunoglobulina G*

INCA – *Instituto Nacional de Câncer*

Kb – *quilobase*

KCl – *Cloreto de potássio*

kDa – quilo-Dalton

kV – quilovolts

LB – Luria-Bertani

M - molar

MERIT40 – *Mediator of Rap80 Interactions and Targeting 40 KDa*

mL – mililitro

mM – milimolar

MRE11 – *Meiotic Recombination 11*

NaCl – Cloreto de sódio

NaOH – Hidróxido de Sódio

NBS1 – *Nijmegen Breakage Syndrome 1*

NCBI - *National Center for Biotechnology Information*

NHEJ – *Non-Homologous End Joining* (Junção não homóloga de pontas)

NLS – *Nuclear Localization Signal*

nm - nanômetros

PALB2 – *Partner and Localizer of BRCA2*

PARP – *Poli (ADP-ribose) polymerase*

PARPi – inibidores de PARP

PBS – tampão fosfato-salino

PCR – *Polymerase Chain Reaction* (reação em cadeia da polimerase)

PEI - polietilenimina

pH – potencial hidrogeniônico

PTEN – Phosphatase and Tensin Homolog

PVDF – difluoreto de polivinilideno

p/v – peso por volume

RAP80 – *Receptor-associated Protein 80*

Rb – *Retinoblastoma protein*

RDD – Reparo ao Dano de DNA

RING – *Real Interesting New Gene*

RNA – ácido ribonucleico

RNA pol. II – RNA polimerase II

RIPA - *Radioimmunoprecipitation assay buffer*
SDS – dodecil sulfato de sódio
SDS-PAGE- *SDS-PolyAcrilamide Gel Electrophoresis*
TA – *Transcription activation* (ativação transcricional)
TBE – Tris-Borato-EDTA
tBRCT – Domínio BRCT em tandem
TBS – *Tris-Buffered Saline*
TopBP1 – *DNA Topoisomerase II Binding Protein 1*
TP53 – *Tumor Protein p53*
v/v – volume por volume
VUS – *Variants of uncertain significance*
WHO – *World Health Organization*
WT – *wild-type* (selvagem)
 μg – micrograma
 μL – microlitro
 μM - micromolar

1. INTRODUÇÃO

1.1. Câncer

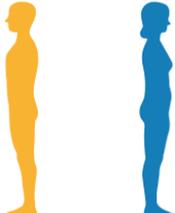
A definição de câncer engloba um conjunto de mais de 100 doenças caracterizadas pelo acúmulo de mutações que culminam na regulação alterada da proliferação celular. Eventualmente, as células tumorais se tornam capazes de invadir e colonizar tecidos e órgãos, comprometendo seu funcionamento (HANAHAN & WEINBERG, 2011).

O câncer é um problema de saúde pública mundial, tendo sido responsável por mais de 9 milhões de mortes no ano de 2018, segundo dados da organização mundial da saúde. Estima-se um aumento de 70% de novos casos até 2030 (WHO, 2018). Esse fenômeno pode ser explicado pelo aumento da expectativa de vida da população e da exposição a fatores ambientais, como álcool e cigarro (FERLAY *et al.*, 2013). No Brasil, para o ano de 2020, são esperados 600 mil novos casos, com cerca de 200 mil óbitos segundo dados do Instituto Nacional de Câncer (INCA) (INCA, 2020).

Esse cenário torna o estudo e a compreensão dos mecanismos de progressão da doença essenciais para o desenvolvimento de novas abordagens preventivas e terapêuticas, além da criação de políticas públicas de controle e prevenção.

1.2. Câncer de Mama

O câncer de mama (CM) é a neoplasia mais frequente em mulheres no Brasil e no mundo, sendo responsável por aproximadamente 30% de todos os casos (desconsiderando o câncer de pele não melanoma). Dados do INCA estimam 66.280 novos casos para 2020 no Brasil, e uma taxa de mortalidade de 16,4% (INCA, 2020) (**Figura 1.1**).

Localização Primária	Casos	%	Homens	Mulheres	Localização Primária	Casos	%
Próstata	65.840	29,2%			Mama feminina	66.280	29,7%
Cólon e reto	20.520	9,1%			Cólon e reto	20.470	9,2%
Traqueia, brônquio e pulmão	17.760	7,9%			Colo do útero	16.590	7,4%
Estômago	13.360	5,9%			Traqueia, brônquio e pulmão	12.440	5,6%
Cavidade oral	11.180	5,0%			Glândula tireoide	11.950	5,4%
Esôfago	8.690	3,9%			Estômago	7.870	3,5%
Bexiga	7.590	3,4%			Ovário	6.650	3,0%
Linfoma não Hodgkin	6.580	2,9%			Corpo do útero	6.540	2,9%
Laringe	6.470	2,9%			Linfoma não Hodgkin	5.450	2,4%
Leucemias	5.920	2,6%			Sistema nervoso central	5.220	2,3%

*Números arredondados para múltiplos de 10.

Figura 1.1. Classificação dos tipos de câncer mais incidentes no Brasil. Ranqueamento dos dez tipos de câncer mais incidentes para homens e para mulheres em 2020, exceto câncer de pele não melanoma (Adaptado de INCA, 2020).

Os fatores de risco para o desenvolvimento dessa neoplasia estão associados à exposição hormonal, com contribuição direta do número de ciclos ovarianos ao longo da vida da mulher (SUN *et al.*, 2017). Dessa forma, mulheres que possuem menarca antes dos 12 anos apresentam risco aumentado em 5%, e mulheres que entram em menopausa depois dos 55 anos, possuem risco 2 vezes maior de desenvolver CM. A nuliparidade e o uso de medidas contraceptivas também aumentam o risco do surgimento do câncer (ROJAS & STUCKEY, 2016). Dos fatores de risco individuais para o CM esporádico, o mais importante é a idade. Mulheres com mais de 50 anos apresentam maior risco de desenvolvimento do CM se comparado com outras faixas etárias (INCA, 2020).

Estima-se que 10% dos casos de CM apresentam origem hereditária. Mulheres com histórico familiar apresentam risco aumentado de desenvolvimento da doença (até 85%) se comparado com o risco da população em geral (13%) (PRUTHI, 2010; ENGEL & FISCHER, 2015).

Nos anos 1990, estudos baseados em famílias com histórico de câncer foram essenciais para o estabelecimento da relação entre o CM e os fatores genéticos responsáveis pela hereditariedade. Hall e colaboradores (1990) estabeleceram a ligação entre o desenvolvimento do CM e um locus gênico presente no cromossomo 17 (HALL *et al.*, 1990). Em 1993, outro estudo confirmou a relação entre a forma hereditária do CM e o cromossomo 17, consolidando a ideia de um gene associado a essa susceptibilidade (EASTON *et al.*, 1993). No ano seguinte, Miki e colaboradores (1994) mapearam e clonaram o primeiro gene determinante para a predisposição hereditária ao CM, *BRCA1* (**B**reast **C**ancer 1, *early onset*) (MIKI *et al.*, 1994). Ainda em 1994, foi identificado um segundo gene associado à susceptibilidade ao CM, *BRCA2* (**B**reast **C**ancer 2, *early onset*) (WOOSTER *et al.*, 1994).

Mutações em *BRCA1* e *BRCA2* observadas nas linhagens germinativas são as principais causas da síndrome hereditária de CM e ovário (**H**ereditary **b**reast and **o**varian **c**ancer, HBOC) (ECONOMOPOULOU *et al.*, 2015; SCALIA-WILBUR *et al.*, 2016). Essa é uma síndrome autossômica dominante que é caracterizada pelo surgimento precoce do câncer em mulheres com histórico familiar. Mulheres carreadoras de mutações em *BRCA1* apresentam risco aumentado de 56% a 80% de desenvolvimento da doença entre 50 e 70 anos de idade, enquanto mutações em *BRCA2* elevam o risco em até 60% (CLARK *et al.*, 2012; KOBAYASHI *et al.*, 2013; SHIOVITZ & KORDE, 2015).

Mutações em *BRCA1*, bem como em *BRCA2*, também aumentam o risco de desenvolvimento de outros tipos de câncer, como de próstata, estômago, pâncreas, cólon, leucemias e linfomas (MERSCH *et al.*, 2014; CAVANAGH & ROGERS, 2015; PRAKASH *et al.*, 2015). A importância de *BRCA1* e *BRCA2* é corroborada pela noção de que ambos os genes são essenciais para o desenvolvimento embrionário em camundongos (GOWEN *et al.*, 1996; SUZUKI *et al.*, 1997; LIU & LU, 2020).

BRCA1 e *BRCA2* não apresentam homologia na sequência proteica, porém ambos são genes supressores de tumor e possuem papéis fundamentais na manutenção da integridade genômica, através dos mecanismos de resposta a danos no DNA (RDD) (PAUL & PAUL, 2014).

Mutações em *BRCA1* e *BRCA2* são responsáveis por cerca de 50% dos casos de HBOC. O restante dos casos é decorrente de mutações em outros genes de risco variável. Portadores de mutações em *PALB2* (***Partner and Localizer of BRCA2***), por exemplo, também apresentam alto risco de desenvolvimento de CM. Curiosamente, *PALB2* também atua no RDD, sendo responsável por mediar a interação entre *BRCA1* e *BRCA2*. Alterações em outros genes, como *TP53* (***Tumor Protein 53***), *ATM* (***Ataxia Telangiectasia Mutated***) *CHEK2* (***Checkpoint Kinase 2***) atribuem risco moderado para o desenvolvimento do CM, no entanto também contribuem para o desenvolvimento de outros tipos de câncer (KOBAYASHI *et al.*, 2013; MERSCH *et al.*, 2014; ECONOMOPOULOU *et al.*, 2015; SCALIA-WILBUR *et al.*, 2016; COUCH *et al.*, 2017).

1.3. *BRCA1*

O gene *BRCA1* é composto por mais de 80 kb, se localiza no cromossomo 17 (17q21) e é formado por 24 éxons, sendo 22 codificantes, que dão origem a um transcrito de 7.1 kb que, quando traduzido, corresponde a uma proteína de 220 kDa composta por 1863 resíduos de aminoácidos (MIKI *et al.*, 1994; SMITH *et al.*, 1996; ROSEN & PISHIVAIAN, 2014) (**Figura 1.2**).

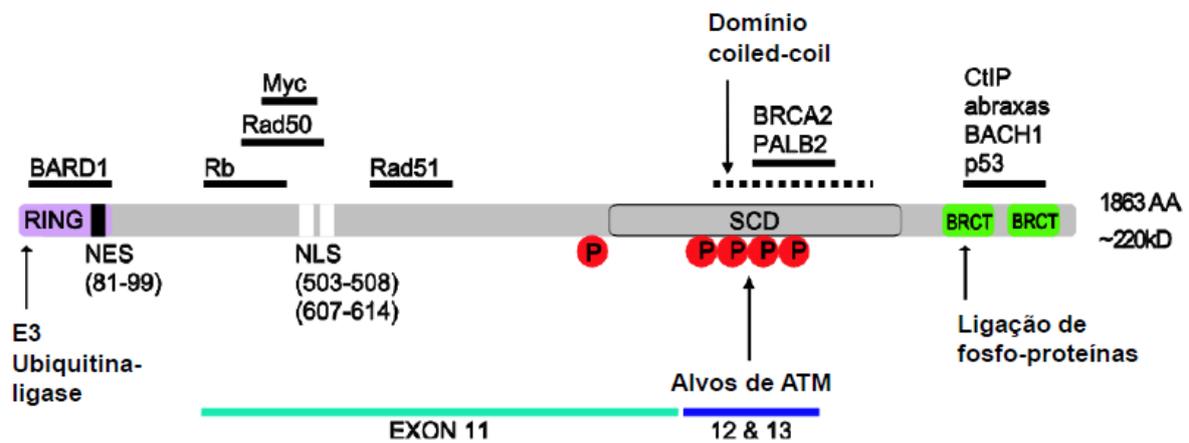


Figura 1.2. Estrutura da proteína BRCA1. A região lilás mostra o domínio *RING finger*, os retângulos em verde representam os domínios BRCT em tandem. Em branco, os sinais de localização nuclear (*nuclear localization signal*; NLS) e em preto o sinal de exportação nuclear (*nuclear exporting signal*; NES). Em vermelho, os sítios de fosforilação (P) e a região SCD (retângulo cinza) determina a região rica em serinas fosforiláveis. As principais proteínas que interagem com BRCA1 estão discriminadas por linhas pretas que indicam suas respectivas regiões de interação (Adaptado de Clark *et al.*, 2012).

Dados epidemiológicos sugerem que a ocorrência de mutações em diferentes regiões de *BRCA1* determinam prevalências distintas no desenvolvimento do CM ou de câncer de ovário (CO). Observa-se que mutações na região central, que encerra o éxon 11, são mais frequentes em casos de CO, enquanto mutações nas regiões que codificam os domínios RING (*Really Interesting New Gene*)-finger, o motivo NLS (*Nuclear Localization Signal*), e os domínios BRCT (*BRCA1 C-Terminal*) são mais frequentes em casos de CM (**Figura 1.2**) (REBBECK *et al.*, 2015).

A região N-terminal de *BRCA1* encerra o domínio RING-finger, compreendido entre os resíduos de aminoácidos 24 a 64 da proteína (PAUL & PAUL, 2015). Esse domínio é bastante conservado entre os diferentes ortólogos de *BRCA1*, e confere a atividade E3 ubiquitina ligase à proteína. O domínio RING também é responsável por promover a interação de *BRCA1* com seu principal parceiro, *BARD1* (*BRCA1-Associated RING Domain protein 1*). A formação do heterodímero *BRCA1/BARD1* potencializa a atividade E3 de *BRCA1*, que é responsável pela ubiquitinação da forma fosforilada da holoenzima RNA polimerase II (RNA pol II), bem como da histona H2A (STARITA *et al.*, 2005; KALB *et al.*, 2014). Porém, ainda não está claro o papel dessa função enzimática no controle de supressão tumoral desempenhado por *BRCA1* e *BARD1*. Dados da literatura demonstram que o heterodímero *BRCA1/BARD1* catalisa a formação de cadeias de poliubiquitina ligadas através do resíduo de lisina na posição 6 (WU-BAER *et al.*, 2003). Esse tipo de sinalização é observado em células tratadas com radiação ionizante, o que sugere que *BRCA1* possa atuar na RDD através de sua atividade E3 ubiquitina ligase (ELIA *et al.*, 2015). *BARD1* também é responsável pela manutenção de *BRCA1* no núcleo através da oclusão do sinal de exportação nuclear (resíduos 81 a 99) adjacente ao domínio RING (FABBRO *et al.*, 2002, SAVAGE & HARKIN, 2015) (**Figura 1.2**).

A região central de *BRCA1*, codificada majoritariamente pelo éxon 11, equivale a aproximadamente 65% do seguimento proteico (CLARK *et al.*, 2012). Essa região não apresenta estrutura tridimensional bem definida. Esse tipo de organização é tipicamente encontrado em proteínas que servem de arcabouço para formação de complexos, onde as regiões menos ordenadas permitem a flexibilização da ligação entre domínios da mesma molécula, facilitando a formação de interações (SAVAGE & HARKIN, 2015). (**Figura 1.2**)

A região central interage com diversas proteínas, como *Rb* (*Retinoblastoma protein*), *c-Myc*, *RAD50* e *RAD51*, atores importantes para o controle do RDD, do ciclo celular e da transcrição (CLARK *et al.*, 2012; PAUL & PAUL, 2015). Ainda nesta região,

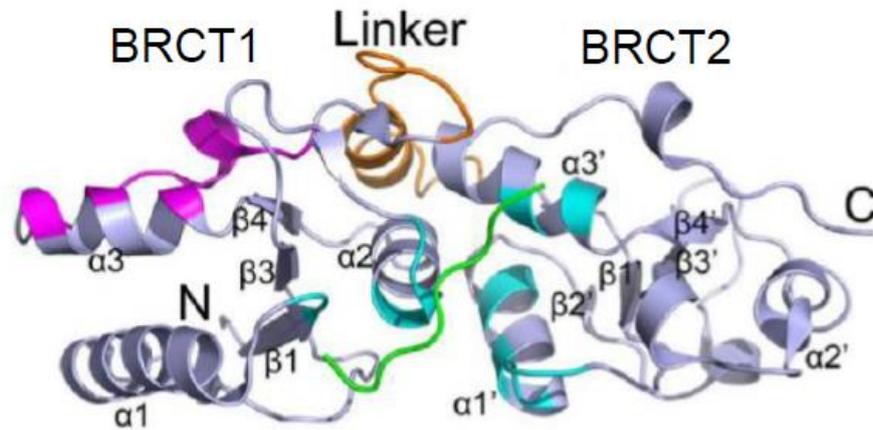
destacam-se a presença de dois NLS (resíduos 503 a 509 e 606 a 615) e sítios de fosforilação mediados pela RDD (resíduos de serina nas posições 988, 1423, 1457 e 1525, por exemplo). A fosforilação desses resíduos é importante para que ocorra a parada do ciclo celular, reparo dos danos no DNA e morte celular por apoptose PARAMESWARAN *et al.*, 2014; YI *et al.*, 2014) (**Figura 1.2**).

Próximo à região C-terminal, se localiza o motivo *coiled-coil*, fundamental para a interação com PALB2 (**Figura 1.2**). A heterodimerização BRCA1/PALB2 é importante para a formação do complexo BRCA1/PALB2/BRCA2 que é o responsável pelo recrutamento da recombinase efetora RAD51 para os sítios de dano no DNA (KAKAROUGKAS & JEGGO, 2013; YI *et al.*, 2014; JIANG & GREENBERG, 2015; PRAKASH *et al.*, 2015).

A porção C-terminal encerra os domínios BRCT, primeiramente identificados em BRCA1 mas presentes em diferentes proteínas codificadas pelo genoma humano. Apesar de não possuírem atividade enzimática intrínseca, os domínios BRCT são fundamentais para a organização de complexo proteicos, em especial na RDD, participando na sinalização de dano e da maquinaria de reparo (SHENG *et al.*, 2011). O BRCT se organiza como um domínio globular singular, composto por quatro folhas- β centrais paralelas flanqueadas por três hélices- α (MESQUITA *et al.*, 2010; SHENG *et al.*, 2011) (**Figura 1.3**). O segmento da proteína que separa o motivo *coiled-coil* dos domínios BRCT de BRCA1 é caracterizada por ser uma região não estruturada, sem função definida até então.

Os domínios BRCT de BRCA1 se apresentam em tandem (tBRCT) (**Figura 1.3**). Quando nessa organização, esses domínios se comportam como uma única unidade estrutural capaz de reconhecer e se ligar a proteínas fosforiladas (MESQUITA *et al.*, 2010).

A)



B)

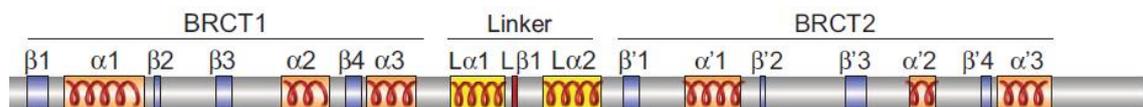


Figura 1.3. Estrutura do tBRCT de BRCA1. (A) Representação esquemática da estrutura tridimensional do tBRCT, mostrando as duas unidades ligadas pela região *linker*, em marrom. Em azul a região de interação entre os domínios BRCT, regiões em rosa e laranja são de interação com outras proteínas e, destacado em verde, um peptídeo fosforilado. (Adaptado de PDB ID: 1T15) (B) Esquema linear do tBRCT, mostrando as estruturas secundárias formadas ao longo do domínio (Adaptados de Sheng *et al.*, 2011 e Woods *et al.*, 2016).

A porção C-terminal de BRCA1 é essencial na manutenção de suas funções, tendo em vista que a perda dos últimos 11 resíduos de aminoácidos da proteína (ex. Variante p.Y1853X) é suficiente para sua disfunção, resultando no aumento de risco de desenvolvimento de CM (MIKI *et al.*, 1994; MONTEIRO *et al.*, 1996; PHELAN *et al.*, 2005; CARVALHO *et al.*, 2007). A arquitetura do domínio tBRCT de BRCA1 permite o reconhecimento de proteínas fosforiladas que respeitem a sequência consenso formada por um resíduo de serina fosforilado na posição 0, seguido de dois resíduos quaisquer nas posições +1 e +2, e um resíduo de fenilalanina na posição +3 (pS₀-X₊₁-X₊₂-F₊₃). Essa interação é mediada pelos resíduos de serina, glicina e lisina nas posições 1655, 1656 e 1702 de BRCA1, respectivamente (MANKE *et al.*, 2003; WILLIAMS *et al.*, 2004; PRAKASH *et al.*, 2015).

1.3.1. Funções da proteína BRCA1

No final da década de 1990, a busca pela função de BRCA1 resultou na identificação de um possível papel associado ao controle de transcrição gênica. Miki e colaboradores (1994) sugeriram essa função devido ao grande número de resíduos ácidos na porção C-terminal da proteína (MIKI *et al.*, 1994).

Monteiro e colaboradores (1996) demonstraram que quando fusionado a um domínio heterólogo de ligação ao DNA, o domínio tBRCT de BRCA1 é capaz de ativar a transcrição de genes repórteres tanto em células humanas quanto em leveduras (MONTEIRO *et al.*, 1996). Posteriormente, foi demonstrado que a interação de BRCA1 com a holoenzima RNA pol. II acontece através de sua associação com a RNA helicase A, corroborando a hipótese da participação de BRCA1 no controle transcricional (SCULLY *et al.*, 1997; ANDERSON *et al.*, 1998; MULLAN *et al.*, 2006). Ao longo dos anos seguintes, diferentes trabalhos demonstraram um papel de BRCA1 no controle transcricional, principalmente de genes envolvidos na RDD, como *CDKN1A* (**C**yclin **D**ependent **K**inase **I**nhibitor **1A**) e *GADD45* (**G**rowth **A**rrest **A**nd **D**N**A** **D**amage **I**nducible **A**lpha) (FAN, *et al.*, 2002; MULLAN *et al.*, 2006; YU *et al.*, 2020). Dados recentes demonstram que BRCA1 encontra-se globalmente associado a regiões transcricionalmente ativas, corroborando uma possível função no controle de expressão (GARDINI, *et al.*, 2014). Apesar desses dados serem consistentes, o impacto dessa função biológica na supressão tumoral mediada por BRCA1 ainda não está clara.

Por outro lado, dados da literatura demonstram um papel central de BRCA1 na manutenção da integridade genômica atuando no reparo de quebras de dupla fita por

recombinação homóloga (HR; *Homologous Recombination*) (revisado em SAVAGE & HARKIN, 2015). Essa função de BRCA1 é fundamental para a supressão tumoral, tendo em vista que nossas células estão constantemente expostas a danos no DNA. Essas lesões podem ter origem endógena e/ou exógena, resultando em diferentes tipos de dano. Como por exemplo, lesões de base única, formação sítios abásicos, quebras de uma única fita, ou da dupla fita do DNA (TUBBS & NUSSENZWEIG, 2017). A quebra de dupla fita (DSB, *Double Strand Break*) é considerada o tipo de lesão mais tóxica por permitir perdas e/ou rearranjos cromossômicos (SAVAGE & HARKIN, 2015; TUBBS & NUSSENZWEIG, 2017).

As DSBs são reparadas essencialmente por dois mecanismos distintos: o reparo por junção de extremidades não homólogas (NHEJ, *Non-Homologous End Joining*), ou a HR. O reparo por NHEJ é uma via propensa a geração de alterações na sequência do DNA. Numa descrição simplista, ela promove apenas a religação das fitas lesionadas. Por outro lado, a HR conduz o reparo usando a cromátide irmã não danificada como molde para a ressíntese da região danificada, possibilitando a manutenção da sequência original (APARACIO *et al.*, 2014; CHATTERJEE & WALKER, 2017).

BRCA1 é uma proteína chave para o reparo por HR, ela participa em diferentes momentos do processo, modulando a formação de diversos complexos proteicos importantes para a transdução da sinalização na RDD. O domínio tBRCT apresenta um papel de destaque na formação desses oligômeros, sendo diretamente responsável pela formação dos complexos BRCA1-A, BRCA1-B e BRCA1-C, através do reconhecimento de regiões fosforiladas nas proteínas Abraxas, BACH1 (*BTB Domain And CNC Homolog 1*) e CtIP (*CtBP Interacting Protein*), respectivamente (SAVAGE & HARKIN, 2015).

Mecanicamente, o comprometimento celular com o reparo por HR, é estimulado pelo processamento inicial da DSB. Esse evento é mediado pela ressecção das extremidades 5' das fitas lesionadas, gerando fitas simples com a extremidade 3'OH livre, que serão utilizadas nas etapas finais do reparo. Esse processamento é promovido pela ação do complexo denominado MRN, composto pelas proteínas MRE11 (*Meiotic Recombination 11*), RAD50 e NBS1 (*Nijmegen Breakage Syndrome 1*), juntamente a CtIP e BRCA1, formando um segundo complexo conhecido como BRCA1-C (**Figura 1.4**) (WANG, 2012; KAKAROUGKAS & JEGGO, 2014; JIANG & GREENBER, 2015).

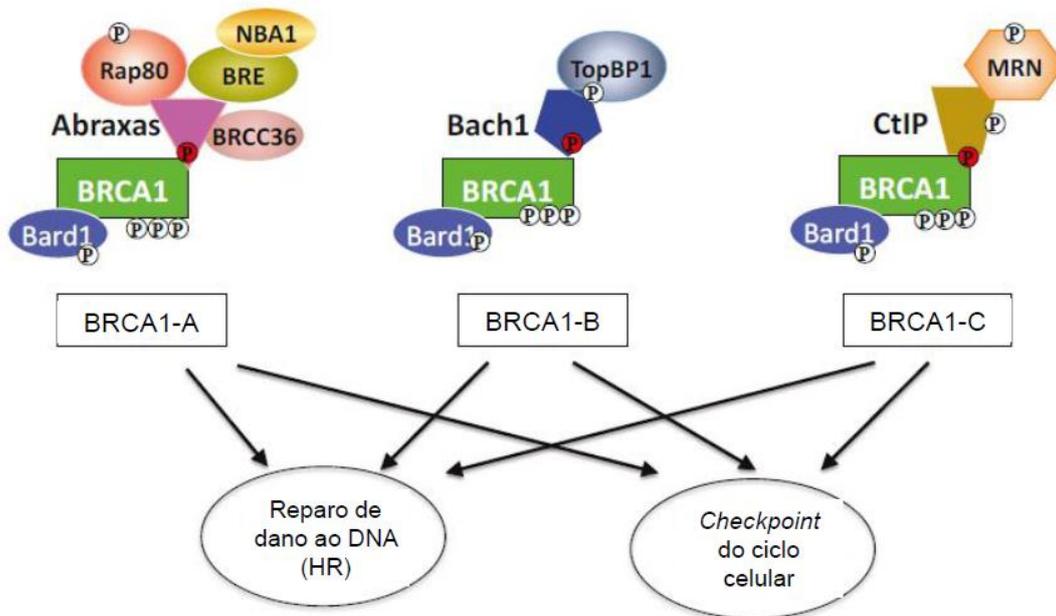


Figura 1.4. Representação esquemática dos complexos de BRCA1 estabelecidos pelo domínio tBRCT. BRCA1-A, modulando a ressecção das fitas; BRCA1-B, que atua no arresto do ciclo celular; e BRCA1-C, complexo central no processo de ressecção das fitas de DNA danificadas durante o processo de HR; as setas indicam em quais processos os complexos atuam. Círculos brancos (P), resíduo de aminoácido fosforilado; em vermelho, o resíduo fosforilado que medeia a interação entre as proteínas dos complexos e BRCA1 (Adaptado de Wang, 2012).

BRCA1 também atua na HR através do complexo BRCA1-A composto por BRCA1, RAP80 (*Receptor-associated Protein 80*), MERIT40 (*Mediator of Rap80 Interactions and Targeting 40 KDa*), BRCC36/45 (*BRCA1/BRCA2-Containing Complex Subunit*) e Abraxas. Esse complexo foi descrito por mediar a parada do ciclo celular na fase G2 mediante danos no DNA, bem como por modular a etapa de ressecção do DNA, prevenindo o processamento excessivo da lesão. Esse controle é importante para prevenir a exposição de regiões de micro-homologia, que podem ser utilizadas em mecanismos de reparo que resultam na perda de segmentos do DNA (HER *et al.*, 2016) (**Figura 1.4**).

BRCA1 também atua através da formação do complexo BRCA1-B, composto por BRCA1, TopBP1 (*DNA Topoisomerase II Binding Protein 1*) e BACH1 que controla a progressão do ciclo celular durante a fase S. Este complexo também atua no reparo por HR, entretanto seu mecanismo de ação ainda não está claro (**Figura 1.4**) (WANG, 2012; APARACIO *et al.*, 2014).

Além do domínio tBRCT, o motivo *coiled-coil* e o domínio RING também são importantes para as funções desempenhadas por BRCA1 na RDD, como mencionado na seção 1.3.

O papel de BRCA1 no reparo por HR é explorado na clínica médica no uso dos inibidores de PARP (*Poly(ADP-Ribose) Polymerase*; PARPi). O racional dessa abordagem terapêutica se dá pelo fato de que a inibição das proteínas da família de PARP (principalmente PARP1) interfere na sinalização de lesões de base única, comuns durante a replicação do DNA. Se não reparadas, essas lesões tendem a evoluir para DSBs. Células onde BRCA1 é funcionalmente deficiente apresentam maior sensibilidade a esse tratamento por apresentarem falha no reparo de DSBs por HR (RANSBURGH *et al.*, 2010; BOUWMAN *et al.*, 2020). Além disso, PARP1 também atua na sinalização inicial do reparo por NHEJ, sendo sua inibição responsável pela disfunção da via, o que corrobora a noção de dependência do reparo por HR na sensibilidade ao tratamento com PARPi (RANSBURGH *et al.*, 2010)

1.4. Variantes genéticas de BRCA1

BRCA1 é um gene que apresenta grande variabilidade alélica, entretanto, as diferentes variantes reportadas apresentam baixa frequência na população (KARCZEWSKI *et al.*, 2020). Desde a identificação do gene, foram reconhecidas mais de 8.500 variantes em linhagens germinativas (*ClinVar*, 2020). No entanto, essas

mutações ocorrem ao longo do gene, sem prevalência de uma região específica (FERNANDES *et al.*, 2019; GOLUBEVA *et al.*, 2019) (**Figura 1.5**).

Segundo o IARC (*International Agency for Research on Cancer*), essas variantes genéticas podem ser classificadas em 5 grupos: (1) benignas, (2) possivelmente benignas, (3) variantes de significado incerto (VUS, *Variants of Uncertain Significance*), (4) possivelmente patogênicas ou (5) patogênicas. A análise de associação genética é considerada o padrão ouro para essa classificação (LINDOR *et al.*, 2011).

A baixa frequência alélica das variantes de *BRCA1* dificulta a realização de associação genética. Uma alternativa de classificação são os modelos multifatoriais que incluem: (a) modelos matemáticos bayesianos de análise de pedigree; (b) análises baseadas em alinhamento de ortólogos; (c) dados de ensaios funcionais; (d) métodos de cocorrência com mutações deletérias; (e) análises computacionais de algoritmos preditores, dentre outros critérios. Todas essas estratégias, mesmo não substituindo os dados clínicos e epidemiológicos, geram informações importantes que, em conjunto, podem servir de base para decisões clínicas (PARSONS *et al.*, 2019; MONTEIRO *et al.*, 2020).

A identificação de indivíduos que carregam mutações patogênicas em *BRCA1* é fundamental para o controle da doença. Uma vez identificados, estes indivíduos podem passar por intervenções clínicas e/ou cirúrgicas que ajudem na prevenção (DAPIC & MONTEIRO, 2006; WINTERS *et al.*, 2016).

Algumas alterações genéticas em regiões codificantes da proteína ou regulatórias do processamento de íntrons permitem inferir perda de função da proteína. É o caso de mutações que resultam na alteração do quadro de leitura codificante da proteína e/ou formação de um códon de parada prematura (LINDOR *et al.*, 2011). Como discutido anteriormente, a perda dos últimos 11 resíduos de aminoácido de *BRCA1* são suficientes para a disfunção da proteína e predisposição à HBOC (MIKI *et al.*, 1994; MONTEIRO *et al.*, 1996; PHELAN *et al.*, 2005; CARVALHO *et al.*, 2007). Alterações que preservam o quadro de leitura da proteína e/ou troca de apenas um único resíduo de aminoácido são mais difíceis de serem interpretadas quanto sua patogenicidade (FERNANDES, *et al.*, 2019).

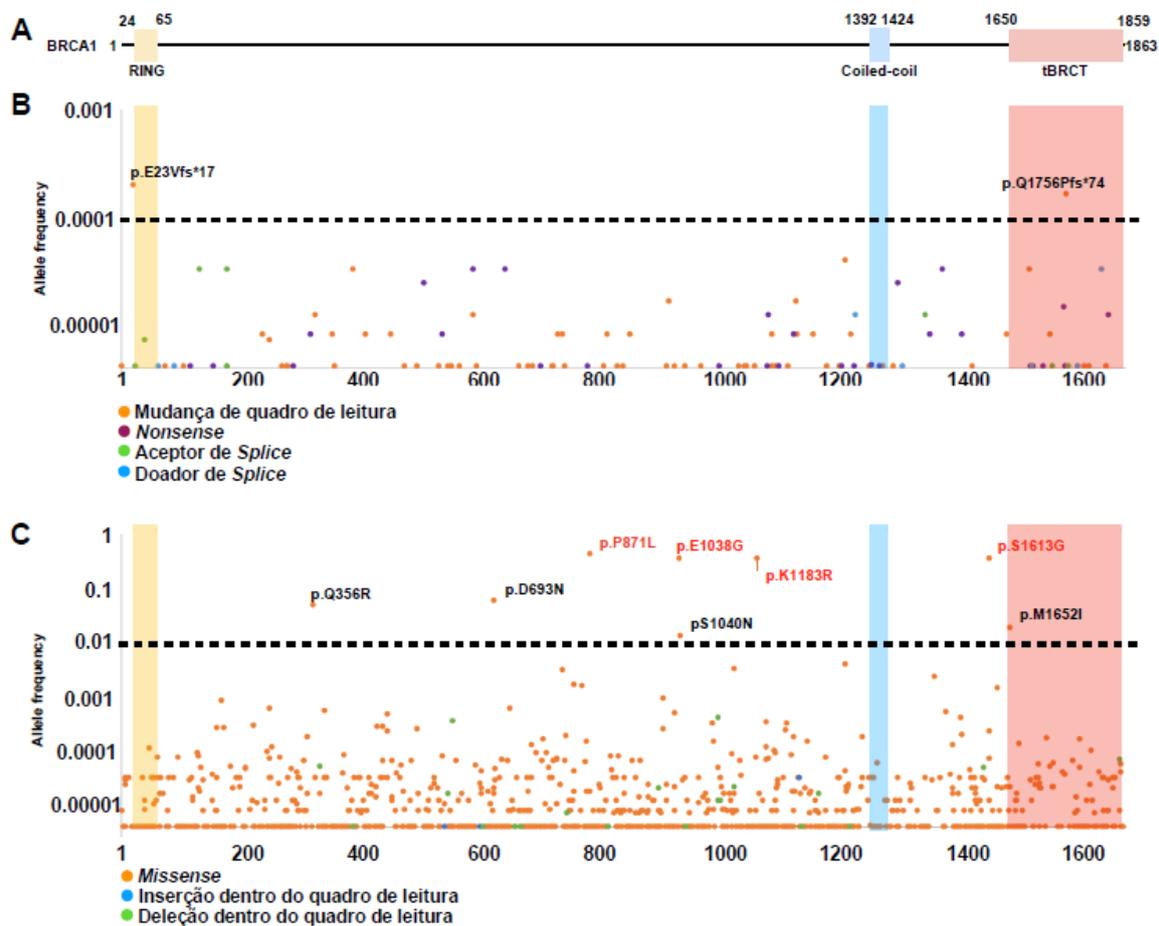


Figura 1.5. Frequência alélica de variantes na região codificante de BRCA1. (A) Representação esquemática da localização dos domínios RING, *coiled-coil*, e tBRCT ao longo da proteína. (B,C) Representação gráfica da frequência das variantes de BRCA1, de acordo com o tipo e mutação, bem como sua localização de acordo com o item (A) (Adaptado de Golubeva *et al.*, 2019).

Importantemente, a identificação de mutações somáticas em *BRCA1* em pacientes que já desenvolveram a doença também é importante pois pode possibilitar condutas terapêuticas específicas como o tratamento com os inibidores de PARP (como discutido na seção 1.3.1) (TURK & WISINSKI, 2018).

Testes para identificação de mutações em *BRCA1* (e outros genes de alto risco, como *BRCA2*, *PALB2*, *PTEN* e *TP53*) na linhagem germinativa são um procedimento recomendado para indivíduos que se enquadram em pelo menos um dos seguintes critérios: (1) ao menos três casos de CM e/ou CO na família; (2) dois casos de CM em parentes próximos, até a terceira geração; (3) dois casos de CM na família diagnosticados antes dos 40 anos; (4) qualquer histórico familiar de CM masculino; (5) ancestralidade judia Ashkenazi (6) histórico de CM, particularmente casos de tumores triplo negativo (negativos para os marcadores: receptor de progesterona, receptor de estrogênio e HER2) diagnosticado antes dos 60 anos de idade; ou (7) CM e/ou CO na mesma paciente (SHIOVITZ & KORDE, 2014, NCCN GUIDELINES, 2020).

A identificação de uma mutação em *BRCA1*, de maneira geral, é realizada através de rotinas de sequenciamento dos éxons e também de suas regiões flangeadoras (bordas intrônicas) (WALLACE, 2016). Essa investigação pode resultar na identificação da sequência selvagem, de variantes previamente classificadas ou de VUS. As variantes genéticas de *BRCA1* identificadas são depositadas em diferentes bancos de dados, como BIC (*Breast Cancer Information Core* - <http://research.nhgri.nih.gov/bic/>), *ClinVar* (<https://www.ncbi.nlm.nih.gov/clinvar>), LOVD (*Leiden Open Variation Database* - <https://www.lovd.nl>) e o *gnomAD* (*Genome Aggregation Database* - <https://gnomad.broadinstitute.org>) (SZABO *et al.*, 2000; LANDRUM *et al.*, 2018; KARCZEWSKI *et al.*, 2020). Mais recentemente, foi desenvolvido a plataforma *BRCA Exchange*, que compila dados genéticos de variantes de *BRCA1* e *BRCA2* disponíveis nos diferentes bancos citados anteriormente em um único sítio (<https://brcaexchange.org>) (CLINE *et al.*, 2018).

A identificação da sequência referência, de variantes benignas ou possivelmente benignas, por si só, não atribui ausência de riscos. Esses resultados não excluem a possibilidade de mutações patogênicas em regiões não codificantes de *BRCA1* bem como em outros genes de predisposição à doença (WALLACE, 2016). A identificação de uma variante patogênica ou possivelmente patogênica permite que o indivíduo portador tome medidas preventivas cabíveis que vão desde o

monitoramento mais frequente, passando por radioterapia e quimioterapia profiláticas, até a mastectomia e ooforectomia preventiva (LEE *et al.*, 2020).

A identificação de VUS é um cenário ímpar, pois a impossibilidade de inferência do impacto funcional dessas mutações através da informação genética dificulta a determinação de sua associação com o câncer (MONTEIRO *et al.*, 2021). A maioria das VUS é composta por mutações de sentido trocado (*missense*), entretanto, pequenas deleções e/ou inserções que preservam o quadro de leitura da proteína também apresentam impacto funcional incerto. Além desse cenário, como discutido anteriormente, essas variantes são identificadas em baixa frequência na população (<1:10.000) dificultando a análise por associação genética (FERNANDES *et al.*, 2019).

Nesses casos, ensaios funcionais se apresentam como ferramentas importantes para identificação do impacto dessas variantes nas funções biológicas de BRCA1.

1.5. Ensaios Funcionais para a classificação de VUS de BRCA1

Segundo dados do *ClinVar*, as VUS correspondem a cerca de 35% das variantes de *BRCA1* reportadas, demonstrando a importância de se obter informações que possam contribuir para a reclassificação dessas variantes. Os estudos funcionais podem contribuir com uma referência primária de comportamento das variantes, oferecendo dados adicionais que, combinados com outras informações disponíveis, podem auxiliar as avaliações de risco (MONTEIRO *et al.*, 2021; LYRA *et al.* 2020).

Os ensaios funcionais visam interrogar o impacto de uma determinada variante em uma função biológica e/ou domínio específico relevante para o estabelecimento do fenótipo em questão (CARVALHO *et al.*, 2007; MONTEIRO, *et al.*, 2021). Do ponto de vista da interpretação de disfunção proteica, é importante ressaltar que os dados funcionais se restringem ao fenótipo e/ou domínio avaliado. No caso de BRCA1, existem diversos ensaios focados nos domínios da proteína, bem como análises compreensivas valendo-se da proteína cadeia-completa em suas funções associadas à manutenção da integridade genômica (LYRA *et al.*, 2020).

BRCA1 apresenta domínios de papel já bem descritos, como *RING-finger*, *coiled-coil*, e o tBRCT. Parte dos ensaios funcionais que buscam avaliar o impacto de variantes em BRCA1, se restringem a essas regiões da proteína. Nesses casos, o desenho experimental leva em consideração as funções estabelecidas de cada domínio. A integridade do domínio RING é fundamental para a manutenção da atividade E3 ubiquitina ligase de BRCA1, bem como para a interação com BARD1.

Dessa forma, o desenho experimental dos ensaios focados nessa região, buscam avaliar o impacto das variantes nessas funções (MORRIS *et al.*, 2006; STARITA *et al.*, 2015). O ensaio mais bem estabelecido na literatura para a avaliação da capacidade de heterodimerização de BRCA1/BARD1 é o de interação proteína-proteína por dois híbridos em levedura, já tendo avaliado mais de 1500 variantes nessa região (MORRIS *et al.*, 2006; STARITA *et al.*, 2015). Já a atividade E3 ligase de BRCA1 é avaliada através de ensaios de ubiquitinação *in vitro*. Os dados disponíveis na literatura corroboram a importância da interação com BARD1 na função E3 de BRCA1, de forma que ainda não foram descritas variantes que impactem a dimerização dessas proteínas, mas mantenham a atividade enzimática de BRCA1 (RANSBURGH *et al.*, 2010; CHRISTOU *et al.*, 2014; ANANTHA *et al.*, 2017; CALECA *et al.*, 2019). Por outro lado, já foi observado que algumas variantes que não sustentam a atividade enzimática de BRCA1, não interferem na interação com BARD1 (MORRIS *et al.*, 2006; STARITA *et al.*, 2015).

O motivo *coiled-coil* de BRCA1 é responsável por promover o recrutamento de PALB2, BRCA2 e RAD51 para os sítios de danos ao DNA através da heterodimerização com PALB2 (NEPOMUCENO *et al.*, 2017; DUCY *et al.*, 2019). Assim como ocorre com BRCA1, carregadores de mutações patogênicas em *PALB2* apresentam maior risco de desenvolvimento de CM, corroborando o papel da interação BRCA1/PALB2 na supressão tumoral (NEPOMUCENO *et al.*, 2017). O único trabalho que avalia o impacto de mutações no *coiled-coil* de BRCA1 o faz através da capacidade de interação com PALB2. Esse trabalho utilizou um sistema quantitativo de dois híbridos em células de mamífero e interrogou o impacto de apenas 8 variantes nessa região (WOODS *et al.*, 2016). A carência de trabalhos focados no motivo *coiled-coil* pode ser explicado pela ausência de variantes classificadas como patogênicas nessa região, que possam ser utilizadas para validação dos ensaios. Entretanto, do ponto de vista de PALB2, já foi demonstrado que uma variante classificada como possivelmente patogênica (ex. p.L35P) é capaz de interferir na interação com BRCA1, sugerindo que esse possa ser um mecanismo de predisposição ao desenvolvimento do CM por meio de mutações no motivo *coiled-coil* de BRCA1 (FOO *et al.*, 2017).

Os ensaios funcionais relacionados ao domínio tBRCT são bastante difundidos e apresentam um conjunto maior de variáveis avaliadas (LYRA *et al.*, 2020). Os ensaios focados no tBRCT buscam interrogar características previamente estabelecidas do domínio que permitam inferir sua disfunção. Durante a sinalização de resposta a danos no DNA, BRCA1 é responsável pela formação de diversos

complexos mediado pelos domínios tBRCT por meio do reconhecimento de proteínas fosforiladas. É através da formação desses complexos (BRCA1-A, -B e -C) que BRCA1 atua na transdução da sinalização que culmina com o reparo fiel por HR. Dessa forma, a avaliação do impacto no reconhecimento de proteínas fosforiladas, bem como na interação com os parceiros de BRCA1, é utilizada como ferramenta para caracterização funcional de variantes localizados no tBRCT (LEE *et al.*, 2010; ANANTHA *et al.*, 2017; PETITALOT *et al.*, 2018;).

A manutenção da integridade do domínio tBRCT é fundamental para o desempenho de suas funções. O domínio, quando íntegro, apresenta resistência ao tratamento com proteases, e variantes que comprometem a organização estrutural do domínio promovem uma maior sensibilidade ao tratamento proteolítico. A avaliação dessa susceptibilidade também é utilizada como ferramenta para determinação do impacto de variantes no domínio tBRCT (LEE *et al.*, 2010). Vale ressaltar que os ensaios de ligação a peptídeos fosforilados, bem como de sensibilidade a proteases, são realizados *in vitro* e que não necessariamente refletem a realidade de um sistema biológico mais complexo.

BRCA1 é evolutivamente conservada e central para a manutenção da homeostase celular, entretanto, eucariotos inferiores (como leveduras) não apresentam ortólogos da proteína. Curiosamente, a expressão da sequência codificante humana de *BRCA1* nessas células resulta no comprometimento da proliferação celular (THOUVENOT *et al.*, 2016). Apesar desse fenótipo de inibição de crescimento de leveduras não ser uma função biológica de BRCA1, ele está diretamente relacionado à integridade do domínio tBRCT, podendo ser utilizada para reportar alterações que resultem na desestruturação desse domínio (THOUVENOT *et al.*, 2016; LYRA *et al.*, 2020)

Durante os anos 1990 foi demonstrado que o domínio tBRCT de BRCA1 interage com a holoenzima RNA pol. II, sendo capaz de promover a transcrição de um gene repórter quando fusionada a um domínio heterólogo de ligação ao DNA (MONTEIRO *et al.*, 1996). Essa capacidade de transativação transcricional também apresenta correlação direta com a integridade do domínio tBRCT (FERNDANDES *et al.*, 2019). Dessa forma, através da construção de um sistema repórter baseado na expressão condicional de genes codificantes de luciferase, essa capacidade de ativar a transcrição pode ser, então, quantificada e utilizada para avaliação indireta do impacto de variantes de *BRCA1* (FERNDANDES *et al.*, 2019). Nosso grupo foi pioneiro na caracterização de variantes através da avaliação da ativação transcricional

(*transcriptional activation*; TA), demonstrando que variantes sabidamente patogênicas apresentam uma redução drástica de TA, ao passo que variantes benignas apresentam comportamento semelhante ao da proteína selvagem (**Figura 1.6**) (CARVALHO *et al.*, 2007). Apesar do ensaio de TA apresentar alta sensibilidade e especificidade, algumas variantes apresentam atividade intermediária, impossibilitando a interpretação da funcionalidade da proteína (FERNANDES *et al.*, 2019).

Para o estabelecimento de classificação com maior acurácia, é utilizado um modelo hierárquico Bayesiano (*VarCall*) que integra os dados de ensaios funcionais de TA e o comportamento das variantes classificadas por análise de ligação (associação genética) para inferir patogenicidade das VUS. O algoritmo *VarCall* acumula as informações de análises anteriores, para que, então, sejam utilizadas como referência para análises futuras, aprimorando a interpretação do comportamento das variantes a cada análise realizada. (IVERSEN *et al.*, 2011; WOODS *et al.*, 2016). Dessa forma, os dados de TA mais recentes do nosso grupo foram capazes de classificar 97% das 347 variantes interrogadas em diferentes momentos (FERNANDES *et al.*, 2019).

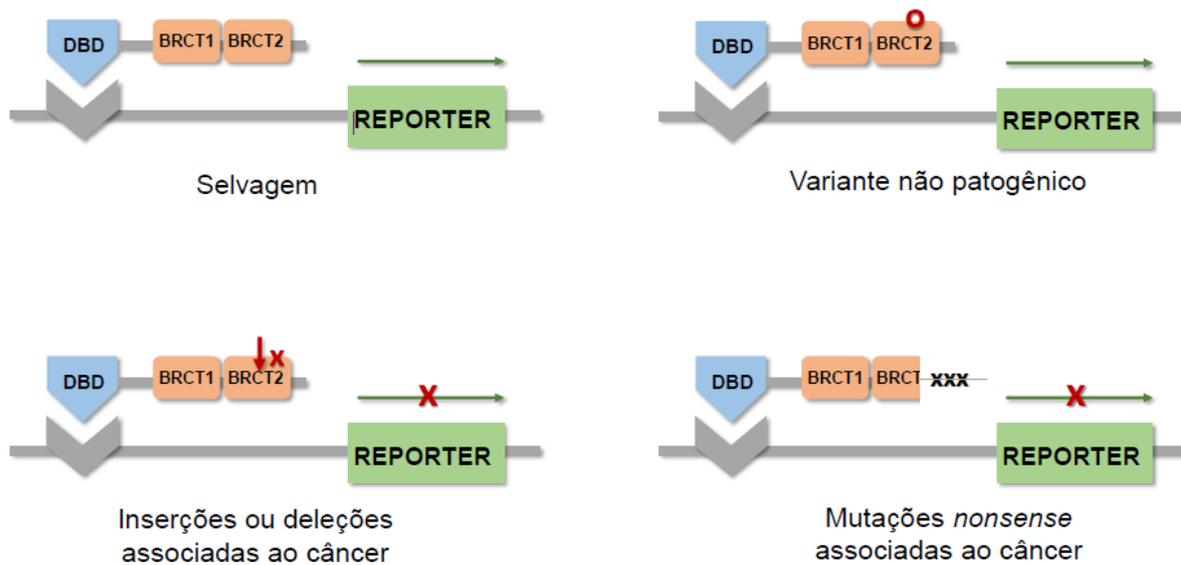


Figura 1.6. Representação esquemática do ensaio de ativação transcricional. Observa-se em (A) o comportamento esperado pela proteína selvagem, em (B) o comportamento para uma variante não patogênico, em (C) para inserções e deleções patogênicas e em (D) para uma variante *nonsense* patogênica (Adaptado de Monteiro *et al.*, 2000).

Ensaio restrito a um domínio específico impossibilita uma avaliação abrangente das funções biológicas da proteína. Nesse sentido, alguns ensaios visam avaliar o papel da proteína completa na biologia de BRCA1. Esses ensaios se baseiam no conceito estabelecido por Scully e colaboradores (1999) que demonstraram que a reconstituição ectópica funcional (usando o alelo selvagem) em células deficientes para *BRCA1* são capazes de restaurar a sensibilidade ao tratamento com radiação ionizante (SCULLY *et al.*, 1999). O mesmo não acontece quando as células são reconstituídas com variantes sabidamente patogênicas, demonstrando que a sensibilidade a danos no DNA está diretamente associada ao papel de BRCA1. Posteriormente, esse conceito foi expandido para avaliar o impacto de VUS na sensibilidade ao tratamento com diferentes agentes genotóxicos, como cisplatina e PARPi (RANSBURGH *et al.*, 2010; BOUWMAN *et al.*, 2020).

Os adutos de DNA gerados pelo tratamento com cisplatina tendem a ser reparados por HR, durante a fase S do ciclo celular. Portanto, o reparo dessas lesões são, em grande extensão, dependentes de BRCA1. Já o tratamento com PARPi apresenta outro racional. Como abordado anteriormente, a inibição das proteínas da família de PARP (principalmente PARP1) resulta no acúmulo de DSBs, que em células deficientes para BRCA1, não são reparados por HR, resultado em maior sensibilidade à esse tratamento (RANSBURGH *et al.*, 2010; BOUWMAN *et al.*, 2020).

De forma geral, os ensaios de sensibilidade a danos no DNA se valem da avaliação indireta da eficiência do reparo por HR. Com o intuito de superar essa questão, foram desenvolvidos sistemas repórteres capazes de quantificar o reparo por HR, possibilitando a avaliação direta do impacto das variantes de BRCA1 na resolução das quebras de dupla fita. Os diferentes sistemas se baseiam na reconstituição de um cassete codificante da proteína verde fluorescente (*Green Fluorescent Protein*; GFP) mediante reparo da quebra da dupla fita de DNA, utilizando de uma região homóloga. Dessa forma, o reparo fiel por HR pode ser quantificado por citometria de fluxo. (YOSHINO *et al.*, 2019).

Com o avanço das tecnologias de sequenciamento, o número de variantes reportadas cresceu consideravelmente nos últimos anos, dificultando a geração de dados funcionais na mesma velocidade (GOLUBEVA *et al.*, 2019). Para isso, foram desenvolvidas abordagens de mutagênese de saturação, que visam gerar todo o conjunto de variações possíveis e analisá-las, agregando informação funcional a variantes ainda não descritas na população. Findlay e colaboradores (2018) desenvolveram uma metodologia baseada em edição genômica de mutagênese por

saturação em células haploides para avaliar todas as variantes resultantes da troca de um único nucleotídeo localizadas nos domínios *RING* e *tBRCT* de BRCA1 (FINDLAY *et al.*, 2018). Esse ensaio se baseia na dependência que células proficientes em p53 apresentam em relação a BRCA1. Dessa forma, a introdução de variantes de BRCA1 disfuncionais, resultam em morte celular. Esse trabalho avaliou aproximadamente 4 mil variantes, representando 96,5% de todas as variantes de substituição de um nucleotídeo possíveis (FINDLAY *et al.*, 2018). É importante ressaltar que esse trabalho não avaliou algumas variantes que apresentam patogenicidade conhecida nessas regiões, interferindo na capacidade de correlacionar a perda de função observada com a inferência de patogenicidade que o método apresenta. Isso se justifica devido ao método de edição genômica utilizado, que não se mostrou eficiente para geração de todas as variantes possíveis nas regiões estudadas.

2. JUSTIFICATIVA

Segundo dados do *BRCA Exchange*, até abril de 2019 haviam sido reportados 7703 variantes localizadas na região codificante de *BRCA1*. Destas, 52,7% resultam em variantes *missense* (derivadas de troca de um único nucleotídeo), 15,7% em alterações sinônimas, e 30,7% apresentam códon de parada prematuro. Podendo o último ser dividido em mutações *nonsense* e variantes que alteram o quadro de leitura codificante, representando 28,7% e 71,3%, respectivamente.

Como apresentado na seção 1.4, o comportamento funcional de algumas variantes é de fácil inferência a partir da análise da sequência genética. Entretanto, variações de um único resíduo de aminoácido apresentam um cenário mais complexo. Apesar de serem o tipo de alteração mais frequente em *BRCA1*, as variantes *missense* são observadas em baixa frequência na população (**Figura 1.5**), impossibilitando a avaliação de risco associado ao desenvolvimento de CM e/ou CO. Sendo assim, existe na literatura um esforço para caracterização funcional desse tipo de variantes através de diferentes ensaios funcionais (discutido na **Seção 1.5**) (LYRA *et al.* 2020).

Uma parcela menor das alterações em *BRCA1* (aproximadamente 1% das variantes reportadas ao *BRCA Exchange*) é composta por indels (**inserções e/ou deleções**) que mantêm o quadro de leitura da proteína (**Figura 2.1**). Esse tipo de alteração pode ser dividido em deleções e/ou inserções de um ou mais códons, ou em deleções e inserções que resultam na troca de um único resíduo de aminoácido (**Figura 2.1**). Esse cenário dificulta a inferência da perda (ou não) de função dessas variantes apenas através da análise da sequência genética. Historicamente, essas variantes foram negligenciadas no que diz respeito a avaliação funcional, existindo apenas dois relatos pontuais na literatura de seu impacto (VALLON-CHRISTERSSON *et al.*, 2001; CARVALHO *et al.*, 2009). A caracterização desse tipo de alteração é importante para a melhor compreensão de seu impacto nas funções de *BRCA1*, e consequentemente na interpretação de patogenicidade. Quando localizadas na região C-terminal da proteína, o estudo dessas variantes também é relevante para a compreensão da organização estrutural dos domínios tBRCT de *BRCA1*.

Selvagem

CAG CTA GCA CGT AGC ATC GTA CGT AGT AGC TAC
Gln Leu Ala Arg Ser Ile Val Arg Ser Ser Tyr

Indels

Deleção de códon único

CAG CTA GCA CGT ATC GTA CGT AGT AGC TAC
Gln Leu Ala Arg Ile Val Arg Ser Ser Tyr

Missense

CAG CTA GCA CGT A^{TG} ATC GTA CGT AGT AGC TAC
Gln Leu Ala Arg **Met Ile Val Arg Ser Ser Tyr**

Deleções

CAG CTA GTA CGT AGT AGC TAC
Gln Leu Val Arg Ser Ser Tyr

Inserções

CAG CTA GCA CGT AGC **AGC** ATC GTA CGT AGT AGC TAC
Gln Leu Ala Arg Ser **Ser Ile Val Arg Ser Ser Tyr**

Figura 2.1: Inserções e deleções que respeitam o quadro de leitura. Exemplo de uma sequência selvagem de nucleotídeos organizada em códons, determinando um polipeptídeo. Abaixo, exemplos de deleção de códon único, deleção e inserção que resulta em uma mutação *missense* a nível proteico, deleção de mais de um códon, e inserção de um códon.

A interpretação de patogenicidade de variantes de BRCA1 que resultam em proteínas truncadas se dá, originalmente, pelo comportamento disfuncional do variante p.Y1853X que leva a perda dos últimos 11 resíduos de aminoácido na borda C-terminal do domínio tBRCT (MIKI *et al.*, 1994; CARVALHO *et al.*, 2007). Pode-se concluir, então, que a perda de mais resíduos (através de uma mutação *nonsense* ou pela alteração do quadro de leitura) sejam prejudiciais para a função de BRCA1.

Mais recentemente, Fernandes e colaboradores (2019) avaliaram os limites da borda C-terminal do domínio tBRCT de BRCA1, melhorando compreensão sobre a importância dessa região para a estabilidade e função da proteína. Esse trabalho caracterizou a perda de até 7 resíduos de aminoácidos na porção C-terminal. Embora tenha sido observado que a perda desses últimos resíduos comprometa a estrutura e função de BRCA1, ainda não está claro qual é o limite funcional da região C-terminal de BRCA1.

Esse trabalho visa caracterizar o impacto funcional de variantes indels e *nonsense* na borda C-terminal de BRCA1, para que possamos compreender melhor o impacto dessas alterações na predisposição ao CM e CO.

3. OBJETIVOS

3.1. Objetivo Geral

Avaliar o impacto funcional de variantes genéticas resultantes de inserções e/ou deleções (indels) que respeitam o quadro aberto de leitura na região codificadora da porção C-terminal de BRCA1.

3.2. Objetivos Específicos

- Identificar e selecionar as indels que respeitam o quadro aberto de leitura já descritas na população através da plataforma *BRCA Exchange*.
- Avaliar as indels que resultem na troca de apenas um resíduo de aminoácido através de ferramenta de predição de impacto *in silico*.
- Avaliar a capacidade de ativação transcricional das variantes selecionadas em células humanas.
- Interrogar o papel dos últimos 6 resíduos de aminoácidos na capacidade de ativação transcricional de BRCA1.

4. METODOLOGIA

4.1. Identificação e seleção das variantes descritas

As variantes de *BRCA1* previamente reportadas na população foram identificadas no repositório global de variantes genéticas de *BRCA1* e *BRCA2* (*BRCA Exchange*; <https://brcaexchange.org>) em 11/04/2019 (CLINE *et al.*, 2018). Foram selecionadas as variantes que apresentam deleções e/ou inserções que respeitam o quadro aberto de leitura codificante de *BRCA1*. A partir desse conjunto, foram identificadas as variantes codificadas dentro da região de estudo (compreendidas nos éxons 13 a 24, resíduos de aminoácido 1396 ao 1863).

4.2. Avaliação de impacto funcional de variantes através de modelos de predição *in silico*

As variantes que apresentam deleção e inserção de nucleotídeos que resultam na troca de um único resíduo de aminoácido foram submetidas ao algoritmo de predição *Align-GVGD*, disponibilizado pela Universidade de Utah (acessível através do sítio <http://agvgd.hci.utah.edu>) para suporte na classificação de patogenicidade (TAVTIGIAN *et al.*, 2005). Todas as análises foram realizadas com a configuração padrão.

4.3. Geração das variantes

A região de estudo compreende os nucleotídeos 4186 a 5592 (NM_007294.4, sequência de referência NCBI). O segmento codificado por essa sequência de nucleotídeos corresponde à região do éxon 13 até o final do éxon 24, porção C-terminal de *BRCA1* (*BRCA1*_{13/24}), compreendendo 468 resíduos de aminoácidos, posições 1396 a 1863.

As rotinas de PCR (*Polymerase Chain Reaction*, reação em cadeia da polimerase) conduzidas para a mutagênese sítio-dirigida foram realizadas tendo como molde a construção plasmidial pcDNA3-GAL4_{DBD}:*BRCA1*_{13/24}, previamente gerada pelo nosso grupo (MULLIS *et al.*, 1986, CARVALHO *et al.*, 2007, FERNANDES *et al.*, 2019) (**Figura 4.1**).

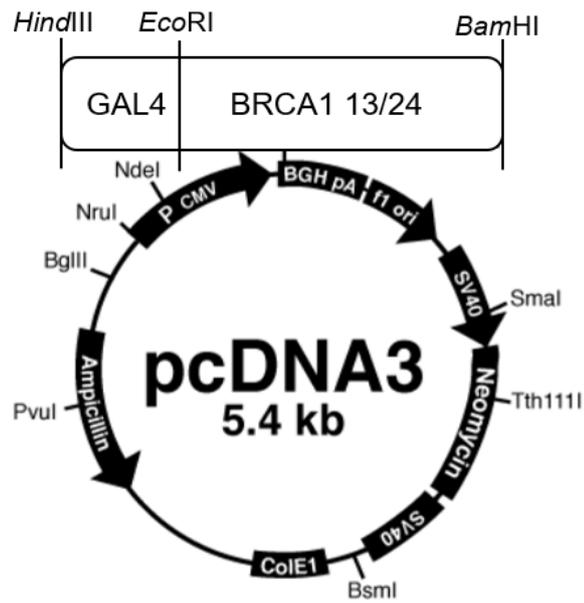


Figura 4.1. Representação esquemática da construção pcDNA3-GAL4^{DBD}-BRCA1_{13/24}. O cassette BRCA1_{13/24}, em fusão com GAL4^{DBD}, clonado direcionalmente nos sítios de *HindIII* e *BamHI*, no vetor de expressão em células de mamíferos pcDNA3. (adaptado de *Invitrogen User Manual*, 2010.)

As reações de amplificação foram conduzidas utilizando a DNA polimerase de alta fidelidade, *PrimeSTAR Max DNA Polymerase* (Takara, Shiga Japão) e oligonucleotídeos iniciadores (*primers*) especificamente desenhados (**Quadro 4.1**). O protocolo de ciclagem é composto por uma etapa inicial de desnaturação (95°C por 2 minutos) para garantir a dissociação completa do molde plasmidial utilizado. A seguir, foram realizados 30 ciclos de amplificação, formados por sucessivas incubações em três temperaturas diferentes: 95°C, para desnaturação da dupla fita; 55°C, para a hibridação dos *primers* no sítio alvo; e 72°C, para a extensão da fita. Ao final dos 30 ciclos, foi realizada mais uma etapa de extensão (72°C por 2 minutos) para garantir a síntese completa da região amplificada.

Os produtos de amplificação (*amplicons*) foram analisados por eletroforese em gel de agarose 0,8% (p/v), em tampão TBE (45mM Tris base, 45mM ácido bórico, 1mM EDTA, todos Merck, Darmstadt, Alemanha). Os *amplicons* foram, posteriormente, tratados com a enzima de restrição *DpnI* (Thermo Fisher Scientific, Massachusetts, EUA), segundo recomendações dos fabricantes. A enzima reconhece e cliva o sítio GA^{m6}TC, quando metilado, a fim de eliminar a construção molde, provenientes de bactérias *dam/dcm* proficiente (**Figura 4.2**). O produto da reação de restrição foi utilizado em rotinas de transformação de bactérias (**Seção 4.5**) para obtenção dos clones mutados.

Quadro 4.1: Oligonucleotídeos iniciadores (*primers*) utilizados na geração de variantes para o estudo (todos Thermo Fisher Scientific).

Variante nucleotídica ^a	Variante proteica	Sentido ^b	Sequência nucleotídica (5'→3')
c.4510_4512delTTA	p.L1504del	S	CCCCTTCTAAATGCCCATCAGATGATAGGT
		AS	TGATGGGCATTTAGAAGGGGATGACCTTTC
c.4579_4580delGAinsAT	p.E1527M	S	GAAACTACCCATCTCAAGAGATGCTCATTAAAG
		AS	CTCTTGAGATGGGTAGTTTCTATTCTGAAG
c.4614_4615delGCinsTT	p.Q1538H	S	GTTGATGTGGAGGAGCAACATTTGGAAGAGTC
		AS	TGTTGCTCCTCCACATCAACAACCTTAATG
c.4676_4678delAGG	p.E1559del	S	CTTGCCAAGGCAAGATCTAGGAACCCCTTA
		AS	CTAGATCTTGCCCTTGGCAAGTAAGATGTTT
c.5017_5019delCAC	p.H1673del	S	ACAAGTTTGCCAGAAAACACATCACTTTAA
		AS	GTGTTTTCTGGCAAACCTTGACACGAGCAT
c.5062_5064delGTT	p.V1688del	S	CTGAAGAGACTACTCATGTTATGAAAACAG
		AS	AACATGAGTAGTCTCTTCAGTAATTAGATT
c.5078_5080delCTG	p.A1693del	S	TGTTGTTATGAAAACAGATGAGTTTGTGTG
		AS	CATCTGTTTTTATAACAACATGAGTAGTCT
c.5092_5093delGAinsAG	p.E1698R	S	CAATGGAAGAAACTACCAAGGTCCAAA
		AS	TTTGGACCTTGGTAGTTTCTTCCATTG
c.5146_5148delTATinsAAA	p.Y1716K	S	GAGGAAAATGGGTAGTTAGCAAATTCTGGGTGA
		AS	GCTAACTACCCATTTTCTCCCGCAATTCC
c.5161_5163delCAG	p.Q1721del	S	TTAGCTATTTCTGGGTGACCTCTATTAAG
		AS	GGTCACCCAGAAATAGCTAACTACCCATTT
c.5181_5183delAAT	p.M1728del	S	CAGTCTATTAAGAAAGAAAGCTGAATGAG
		AS	TTTCTTTCTTTAATAGACTGGGTCACCAG
c.5218_5220delGTGinsAAC	p.V1740N	S	ATTTTGAAGTCAGAGGAGATAACGTCAATGGAA
		AS	ATCTCCTCTGACTTCAAAATCATGCTCATT
c.5234_5235delACinsGG	p.N1745R	S	AGATGTGGTCAATGGAAGAAGGCACCAAGGTC
		AS	TTCTTCCATTGACCACATCTCCTCTGACTT
c.5275_5276delAAinsTG	p.K1759W	S	CAAGAGAATCCCAGGACAGATGGATCTTCAGG
		AS	TCTGTCCTGGGATTCTCTTGCTCGCTTTGG
c.5321_5322delACinsGG	p.N1774R	S	TTGCTATGGGCCCTTACCAGGATGCCACAG
		AS	TGGTGAAGGGCCCATAGCAACAGATTTCTA
c.5529_5530delACinsCA	p.L1844I	S	TGGGTGTTGGACAGTGTAGCCATCTACCAGTG
		AS	GCTACACTGTCCAACACCCACTCTCGGGTC
c.5566_5568delCCCinsTGA	p.P1856X	S	AGCTGGACACCTACCTGATATGACAGATCCCC
		AS	TATCAGGTAGGTGTCCAGCTCCTGGCACTG
c.5569_5571delGACinsTGA	p.Q1857X	S	TGGACACCTACCTGATACCCTGAATCCCCACA
		AS	GGTATCAGGTAGGTGTCCAGCTCCTGGCA
c.5572_5574delATCinsTGA	p.I1858X	S	ACACCTACCTGATACCCAGTGACCCACAGCC
		AS	CTGGGGTATCAGGTAGGTGTCCAGCTCCTG
c.5575_5577delCCCinsTGA	p.P1859X	S	CCTACCTGATACCCAGATCTGACACAGCCACT
		AS	GATCTGGGGTATCAGGTAGGTGTCCAGCTC
c.5578_5580delCACinsTGA	p.H1860X	S	ACCTGATACCCAGATCCCCTGAAGCCACTACT
		AS	GGGGATCTGGGGTATCAGGTAGGTGTCCAG
c.5581_5583delAGCinsTGA	p.S1861X	S	TGATACCCAGATCCCCACTGACACTACTGAG
		AS	GTGGGGGATCTGGGGTATCAGGTAGGTGTCT
c.5584_5586delCACinsTGA	p.H1862X	S	TACCCAGATCCCCACAGCTGATACTGAGGAT
		AS	GCTGTGGGGGATCTGGGGTATCAGGTAGGT
c.5588_5589delACinsGA	p.Y1863X	S	CCAGATCCCCACAGCCACTGATGAGGATCC
		AS	AGTGGCTGTGGGGGATCTGGGGTATCAGGT

a - Nomenclatura de acordo com as regras estabelecidas pela AGVS (*American Genome Variation Society*)

b - S: *primer* senso; AS: *primer* anti-senso

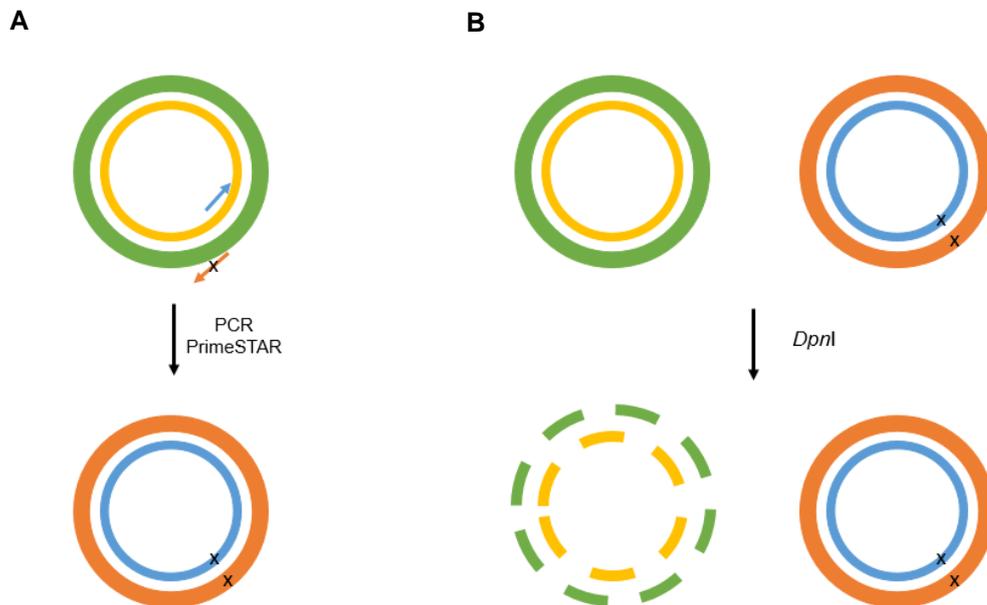


Figura 4.2. Esquema representativo da rotina de mutagênese sítio-dirigida. **A.** As rotinas de PCR foram conduzidas utilizando *primers* (Quadro 3.1) específicos (setas laranja e azul), tendo como molde a construção pcDNA3-GAL4_{DBD}:BRCA1_{13/24} selvagem (verde e amarelo). O produto da reação (laranja e azul) consiste na cópia de todo o vetor contendo a mutação sítio-específica desejada (“X”). **B.** O produto da rotina de PCR foi submetido ao tratamento com *DpnI* para restrição de sítios metilados presentes apenas no molde (adaptado de *QuikChange Lightning Site-Directed Mutagenesis Kit User Manual*, Agilent Technologies, 2015).

4.4. Indução de eletrocompetência bacteriana

As rotinas de manipulação de bactérias foram realizadas utilizando a cepa DH5 α de *Escherichia coli*, de genótipo: F⁻ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* ϕ 80d*lacZ* Δ M15 Δ (*lacZYA-argF*)U169, *hsdR17(rK⁻mK⁺)*, λ ⁻ (MESELSON & YUAN, 1968). A indução de eletrocompetência foi desenvolvida conforme protocolo descrito previamente por Nováková e colaboradores (2014). Resumidamente, as bactérias foram cultivadas em meio SOC (peptona 2% p/v, extrato de levedura 0,5% p/v, 10 mM NaCl, 2,5 mM KCl, 20 mM glicose, todos Sigma-Aldrich, Missouri, EUA), sob constante agitação, a 37°C. O crescimento da cultura foi acompanhado por espectrofotometria através da densidade óptica a 600nm (DO₆₀₀) até a fase exponencial, compreendida entre 0,4 e 0,5 unidades de DO₆₀₀. Em seguida, as bactérias foram mantidas no gelo por 20 minutos e centrifugadas (2.000 xg) por 15 minutos a 4°C. As células foram recuperadas em glicerol 10% (v/v) gelado, no mesmo volume inicial, e centrifugadas (2.000 xg). O *pellet* formado foi ressuspensão em glicerol 10% (v/v) gelado, em 50% do volume inicial e, novamente, a suspensão foi centrifugada (2.000 xg). O *pellet* foi ressuspensão em glicerol 10% (v/v) gelado, em 10% do volume inicial. Essa suspensão de bactérias foi aliquoteada e mantida a -80°C para utilização nas rotinas de transformação (**Seção 4.5**)

4.5. Transformação bacteriana

A transformação de bactérias foi conduzida pelo processo de eletroporação (NEUMANN *et al.*, 1982). Resumidamente, ao DNA de interesse foi adicionada uma alíquota de 50 μ L das bactérias eletrocompetentes (**Seção 4.4**) e imediatamente submetidas ao pulso elétrico de 1.8 kV em eletroporador *MicroPulser* (BIO-RAD Laboratories, Califórnia, Estados Unidos). Em seguida, as bactérias foram recuperadas em 1 mL de meio LB (Luria-Bertani; extrato de levedura 5% p/v, triptona 10% p/v, cloreto de sódio 5% p/v) por 1 hora a 37°C e inoculadas por espalhamento em LB sólido (ágar 4% p/v) suplementado com ampicilina (50 μ g/mL) (todos os reagentes Sigma-Aldrich).

4.6. Extração plasmidial

A extração do DNA plasmidial foi conduzida pelo método de lise alcalina, adaptado de Ausubel (2002). Resumidamente, colônias de bactérias carregando o DNA plasmidial de interesse foram inoculadas individualmente em meio LB

suplementado com ampicilina 50 µg/mL, e incubadas a 37°C sob agitação por 18 horas. Após esse período, 3 mL de cada cultura foi centrifugada a 16.000 xg por 1 minuto, e o *pellet* formado foi ressuspensão em 150 µL de tampão GTE (50 mM glicose, 25 mM Tris-Cl pH 8,0, 10 mM EDTA, todos Merck). A suspensão foi lisada em 150 µL de tampão de lise (0,2 N NaOH, SDS à 1% p/v, ambos Merck), e neutralizada com 150 µL de solução de neutralização (3 M acetato de potássio, Sigma-Aldrich). O produto foi clarificado através de centrifugação (16.000 xg) por 10 minutos e o sobrenadante recuperado e tratado com 5 µL de RNase (50 µg/mL, Sigma-Aldrich). O material foi desproteíno com adição de 150 µL de clorofórmio (Merck), seguido de agitação intensa e centrifugação (16.000 xg por 2 minutos). A fase aquosa foi recuperada, e o DNA precipitado pela adição de 245 µL de isopropanol (Merck) e centrifugação (16.000 xg por 30 minutos). O DNA recém precipitado foi lavado com etanol 70% (v/v), centrifugado (16.000 xg por 15 minutos) e ressuspensão em água deionizada. O material obtido foi utilizado nas rotinas de sequenciamento (**Seção 4.7**).

Alternativamente, quando para uso em rotinas de transfecção de células humanas, os DNAs plasmidiais foram obtidos utilizando o kit *QIAprep Spin Miniprep* (QIAGEN, Hilden, Alemanha), segundo orientações do fabricante.

4.7. Sequenciamento

As construções geradas foram confirmadas através de rotina de sequenciamento automático (método de Sanger) utilizando o kit BigDye Terminator v. 3.1 (Thermo Fisher Scientific), segundo instruções do fabricante. A análise dos produtos da rotina de sequenciamento foi conduzida na plataforma ABI 3130xl (Thermo Fisher Scientific) em colaboração com o Programa de Genética do Centro de Pesquisa do Instituto Nacional de Câncer.

4.8. Cultura de células humanas

Células humanas da linhagem HEK293FT (Invitrogen, Massachusetts, EUA) foram cultivadas em meio DMEM (*Dulbecco's Modified Eagle's Medium*) pH 7,4, suplementado com 10% v/v de soro fetal bovino (SFB; Invitrogen), penicilina (1000 U/L) e estreptomicina (100 mg/mL) (ambos Life Technologies, Califórnia, EUA), a 37°C em atmosfera de 5% v/v de CO₂. A manutenção da cultura (subcultivos) foi realizada por dissociação enzimática com solução de tripsina 0,125% p/v (Sigma-Aldrich) e 1mM EDTA (Merck).

4.9. Transfecção de células humanas

As rotinas de transfecção foram conduzidas utilizando solução 10 μ M de polietilenimina (PEI) (Polysciences Inc, Pennsylvania, EUA) como agente de transfecção (LONGO *et al.*, 2013). O DNA de interesse foi incubado à temperatura ambiente por 30 min com PEI, respeitado a proporção de 3 μ L de PEI para cada 1 μ g de DNA plasmidial, em meio DMEM sem SFB. A mistura de DNA/PEI foi então adicionada às células em cultura para ensaios de TA e *immunoblotting*.

4.10. Ensaio de ativação transcricional (TA)

Células da linhagem HEK293FT foram inoculadas em placas de cultivo de 96 poços ($3,5 \times 10^4$ células / $0,32 \text{ cm}^2$). Após incubação por 24 horas em condições padrão de cultivo (**Seção 4.8**), as células foram submetidas à rotina de transfecção (**Seção 4.9**).

Resumidamente, as células foram cotransfectadas com as construções pcDNA3GAL4_{DBD}:BRCA1_{13/24} (selvagem ou variante), o plasmídeo pG5Luc (contendo o gene repórter: codificante da luciferase de *Photinus pyralis*, sob o controle de operadores de GAL4), e pGR-TK (encerrando o gene repórter codificante da luciferase de *Renilla reniformis*, sob o controle de um promotor constitutivamente ativo HSV-TK) como controle da transfecção (**Figura 4.3**).

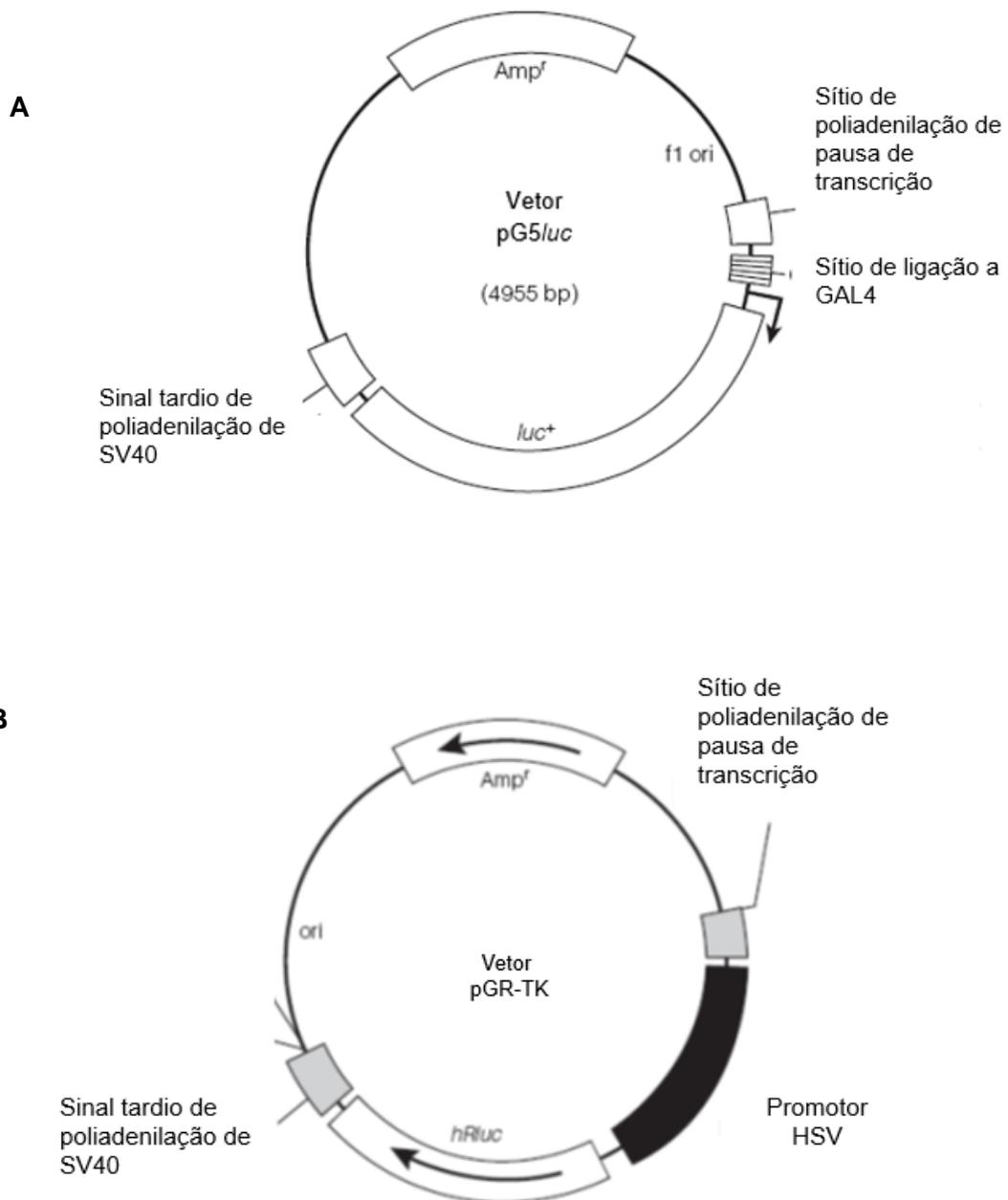


Figura 4.3. Representação esquemática dos plasmídeos pG5luc e pGR-TK. (A) O plasmídeo pG5luc encerra a sequência codificante da enzima luciferase de *P. pyralis* (*luc+*) regulado por 5 operadores de GAL4 (sítios de ligação de GAL4_{DBD}). (B) O plasmídeo pGR-TK encerra a sequência codificante da enzima luciferase de *R. reniformis* (*Rluc*) controlado por um promotor reconhecido de forma constitutiva por células humanas, HSV TK. (adaptado de Promega *Two-Hybrid System*, 2009)

Após 24 horas da transfecção, a atividade das enzimas repórteres foi quantificada no luminômetro GloMax[®] (Promega, Wisconsin, EUA) utilizando o kit *Dual-Luciferase[®] Reporter Assay System* (Promega), seguindo as recomendações do fabricante. Resumidamente, depois da remoção do meio do poço de cultivo, as células foram lisadas utilizando 25 µL do tampão *Passive Lysis* presente no kit, e mantidas por 15 minutos sob agitação leve. Em seguida, 40 µL do substrato da luciferase de *P. pyralis* (reagente LARII[®]) foi adicionado a uma alíquota de 20 µL do extrato lisado, a luz emitida foi quantificada. Em seguida, pela adição de 40 µL do reagente Stop & Glo[®], a reação foi interrompida e o substrato da reação catalisada pela luciferase de *R. reniformis* foi disponibilizado. A luz emitida foi novamente quantificada, e atividade relativa foi calculada como a razão das atividades de cada luciferase (*P. pyralis* / *R. reniformis*). Como controles do ensaio foram utilizadas quatro construções codificantes de BRCA1: selvagem (*wild-type; wt*), a variante não patogênica p.S1613G, e as variantes patogênicas p.M1775R e p.Y1853X (FERNANDES *et al.*, 2019; PARSONS *et al.*, 2019).

4.11. Avaliação de níveis proteicos

4.11.1. Extratos proteicos

Foram plaqueadas 5×10^5 células da linhagem HEK293FT em placas de cultivo de 6 poços (9,4 cm² de área). Após incubação por 24 horas em condições padrão de cultura, como descrito na **Seção 4.8**, as células foram submetidas à transfecção (**Seção 4.9**) utilizando a construção pcDNA3GAL4_{DBD}:BRCA1_{13/24} (selvagem ou variante).

As células foram mantidas em cultura por 24 horas e, então, lisadas em tampão RIPA (NP-40 1% v/v, 150 mM NaCl, 10 mM Tris-HCl pH 7,4, 5 mM EDTA, todos Merck) suplementadas com coquetel de inibidores de protease (Sigma-Aldrich). O extrato foi mantido em banho de gelo por 30 minutos e centrifugados a 16.000 xg por 5 minutos a 4°C. O sobrenadante contendo as proteínas totais foi recuperado e armazenado a -80°C.

A concentração total de proteínas nos extratos foi determinada através do ensaio colorimétrico de *Bradford* (1976).

4.11.2. Immunoblottings

Os extratos proteicos obtidos como descrito na **Seção 4.11.1** foram utilizados nas rotinas de *immunoblotting*. Resumidamente, os extratos foram resolvidos em gel desnaturante de poliacrilamida (*SDS-PolyAcrylamide Gel Electrophoresis*, SDS-PAGE) (LAEMMLI, 1970) 8% (p/v) e eletrotransferidos para membranas de difluoreto de polivinilideno (PVDF; GE Healthcare, Illinois, EUA), usando o sistema *Trans-Blot Semidry* (BIO-RAD Laboratories), segundo recomendações do fabricante.

Após a transferência, as proteínas imobilizadas nas membranas foram bloqueadas utilizando leite em pó desnatado 5% (p/v) em TBS (TBS: 50 mM Tris base, 150 mM NaCl) suplementado com Tween-20 0,1% (v/v) (T-TBS), por 1 hora à temperatura ambiente (todos Sigma-Aldrich). Após o bloqueio, as membranas foram incubadas com os anticorpos monoclonais anti-GAL4_{DBD} (IgG de camundongo, sc-46680, Santa Cruz, Califórnia, Estados Unidos) e anti- β -actina (IgG de camundongo, sc-47778, Santa Cruz) diluídos em T-TBS contendo leite desnatado 0,5% (p/v), por 16 horas a 4°C. A membrana foi, em seguida, incubada com o anticorpo policlonal anti-IgG de camundongo conjugado à peroxidase (produzidos em cabra, sc-2005; Santa Cruz), por 1 hora à temperatura ambiente. O imunocomplexo foi revelado usando o sistema *Amersham ECL™ Western Blotting Detection Reagents* (GE Healthcare), seguindo as recomendações do fabricante.

A quimioluminescência foi registrada no sistema de aquisição de imagem *ChemiDoc* (*ChemiDoc Imaging Systems*, BIO-RAD Laboratories).

5. RESULTADOS

5.1. Seleção das variantes

Foram identificadas, através da plataforma *BRCA Exchange* (dados obtidos em 11/04/2019), todas as variantes naturais codificantes de *BRCA1* (7703 variações). Destas, 1758 são consequência de indels, das quais, 72 respeitam o quadro aberto de leitura de *BRCA1* (**Anexo I**). Deste grupo, 26 variantes encontram-se na região de estudo, compreendida entre os éxons 13 e 24. Foram selecionadas 9 variantes cuja deleção/inserção na sequência genômica codifica a troca de um resíduo de aminoácido. Devido ao caráter funcional do estudo, essas variantes serão reportadas a seguir como alterações de sentido trocado (*missense*). Também foram selecionadas para o estudo 7 variantes que resultam na deleção de um único códon (deleção de três nucleotídeos). Dentre elas, a variante p.V1688del já foi previamente avaliada por Vallon-Christersson e colaboradores. Como os dados da literatura sugerem que deleções de grandes porções da proteína resultem em perda de função, essas variantes não foram consideradas para análise (VALLON-CHRISTERSSON *et al.*, 2001).

A fim de caracterizar o limite funcional da borda C-terminal do domínio tBRCT, também foram selecionadas outras 6 mutações sem sentido (*nonsense*) que resultam na perda de até 6 resíduos de aminoácido da proteína (**Figura 5.1**).

A apresentação dos resultados foi subdividida em três conjuntos: (1) variantes *missense* (p.E1527M, p.Q1538H, p.E1698R, p.Y1716K, p.V1740N, p.N1745R, p.K1759W, p.N1774R e p.L1844I); (2) deleções de códon único (p. L1504del, p.E1559del, p.H1673del, p.V1688del, p.A1693del, p.Q1721del e p.M1728del); (3) mutações *nonsense* no C-terminal da proteína (p.I1858X, p.P1859X, p.H1860X, p.S1861X, p.H1862X e p.Y1863X) (**Figura 5.2**).

5.2. Avaliação *in silico* das variantes selecionadas

Variantes de sentido trocado (*missense*) podem ser analisadas quanto sua patogenicidade através de preditores funcionais *in silico* (TAVTIGIAN *et al.*, 2005). Assim, as variantes *missense*, foram avaliadas para classificação de patogenicidade através da ferramenta de predição *in silico*, *Align-GVGD* (disponível através do sítio http://agvgd.hci.utah.edu/agvgd_input).

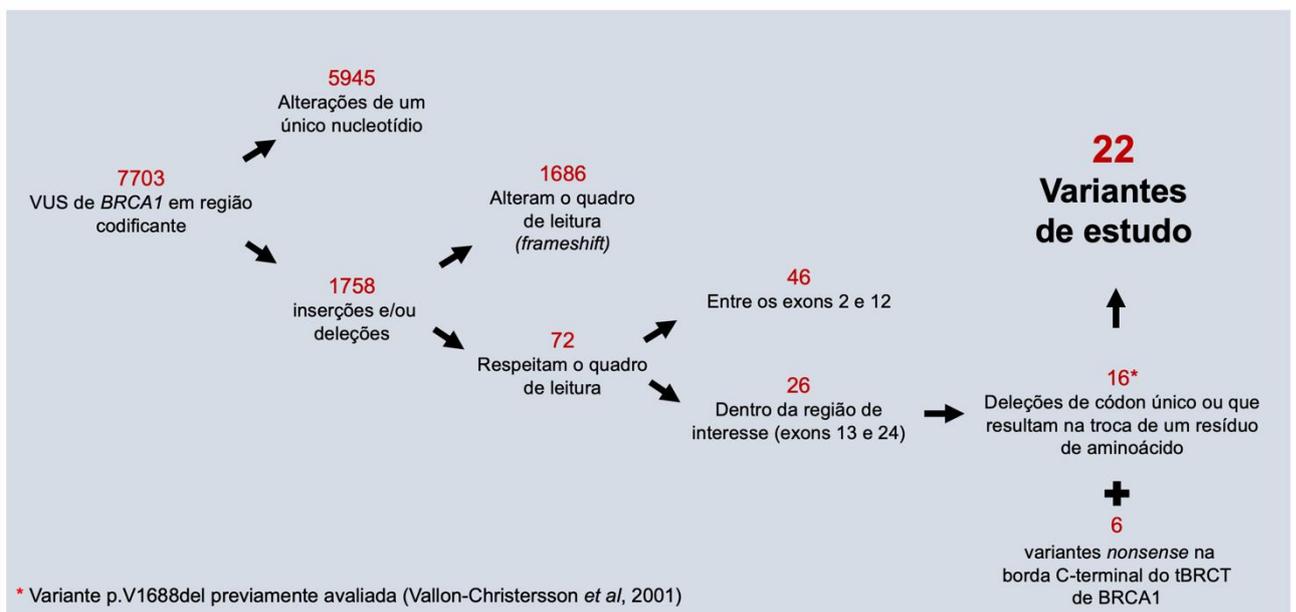


Figura 5.1. Visão geral das variantes de classificação incerta (VUS) de BRCA1. Fluxograma de seleção das variantes para esse estudo (BRCA Exchange, <https://brcaexchange.org>, acesso em 11/04/2019).

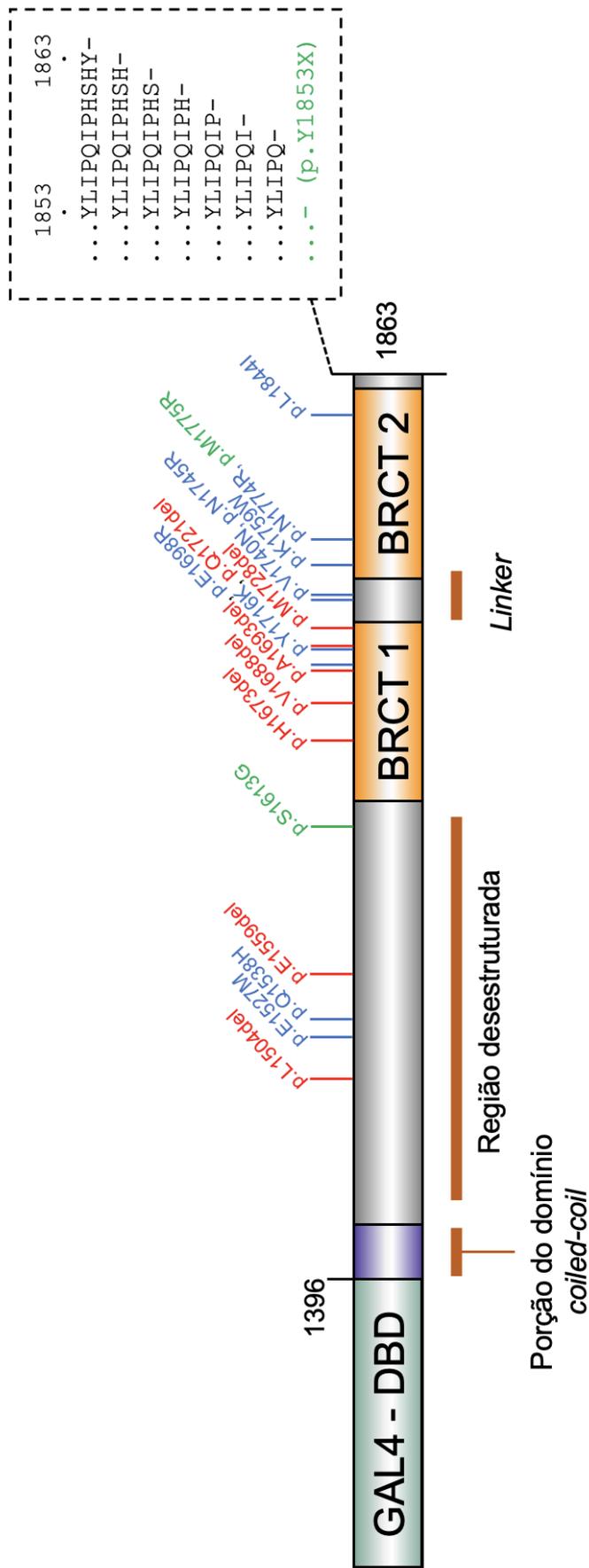


Figura 5.2. Representação esquemática da localização das variantes selecionadas ao longo da proteína de fusão GAL4_{DBD}:BRCA1_{13/24}. Em azul, variantes *missense*. Em vermelho, variantes que resultam na deleção de códon único. Destacado no quadro tracejado, as mutações *nonsense* realizadas na borda C-terminal da proteína BRCT1. Em verde, os controles usados nesse estudo.

A ferramenta *Align-GVGD* realiza alinhamentos múltiplos das sequências proteicas de diferentes ortólogos, que variam de *Pan troglodytes* (chimpanzé) a *Strongylocentrotus purpuratus* (ouriço-do-mar). O algoritmo considera propriedades estruturais, analisa aspectos como conservação dos resíduos de aminoácido pela proximidade taxonômica das espécies e a relevância da região de troca para a função da proteína. Levando em consideração esses parâmetros, o *Align-GVGD* usa um espectro de classificação que varia de enriquecimento de neutralidade até enriquecimento de patogenicidade em intervalos categóricos de risco: *Class C0*, *Class C15*, *Class C25*, *Class C35*, *Class C45*, *Class C55* e *Class C65* (TAVTIGIAN *et al.*, 2005).

Das nove variantes analisadas, 5 foram preditas como não patogênicas (p.E1527M, p.Q1538H, p.K1759W, p.N1774R e p.L1844I; *Class C0*). Apenas as variantes p.E1698R como *Class C45*; p.Y1716K e p.N1745R, como *Class C35*; e p.V1740N, com *Class C65*, foram preditas como potencialmente patogênicas. Dentre elas, a p.V1740N possuiu predição de maior patogenicidade (**Tabela 5.1**).

Tabela 5.1. Classificação das VUS de troca de aminoácido de acordo com *Align-GVGD*

Variação nucleotídica	Variantes proteicas	Classificação de risco (Align GVGD)
c.4579_4580delGAinsAT	p.E1527M	<i>Class C0</i>
c.4614_4615delGCinsTT	p.Q1538H	<i>Class C0</i>
c.5092_5093delGAinsAG	p.E1698R	<i>Class C45</i>
c.5146_5148delTATinsAAA	p.Y1716K	<i>Class C35</i>
c.5218_5220delGTGinsAAC	p.V1740N	<i>Class C65</i>
c.5234_5235delACinsGG	p.N1745R	<i>Class C35</i>
c.5275_5276delAAinsTG	p.K1759W	<i>Class C0</i>
c.5321_5322delACinsGG	p.N1774R	<i>Class C0</i>
c.5529_5530delACinsCA	p.L1844I	<i>Class C0</i>

5.3. Avaliação da capacidade de ativação transcricional

Todas as variantes selecionadas para o estudo foram geradas com sucesso, através da rotina de mutagênese sítio-dirigida, utilizando os *primers* mostrados no quadro 3.1, e o vetor pcDNA3-GAL4_{DBD}:BRCA1_{13/24} como molde.

As variantes selecionadas foram analisadas através do ensaio de ativação transcricional, avaliando sua capacidade de recrutamento da RNA polimerase II e indução da transcrição do sistema repórter que codifica a enzima luciferase, em modelo de células humanas. Os ensaios foram realizados em pelo menos três experimentos independentes, em quadruplicata; e os gráficos gerados são representativos dos experimentos conduzidos.

As atividades foram avaliadas comparativamente, usando como controles construções codificantes de variantes situados na porção C-terminal da proteína previamente caracterizados, além da sequência selvagem (MONTEIRO *et al.*, 1996; PHELAN *et al.*, 2005; CARVALHO *et al.*, 2007). Como controle de não patogenicidade foi utilizada a variante *missense* p.S1613G (DUNNING *et al.*, 1997). Como controles patogênicos, foram utilizadas as variantes p.M1775R e p.Y1853X (MIKI *et al.*, 1994).

5.3.1. Capacidade de ativação transcricional das variantes *missense*

Dentre as 9 variantes *missense* selecionadas, apenas as localizadas na região desestruturada apresentaram comportamento (atividade) semelhante ou superior ao selvagem. A variante p.Q1538H apresentou comportamento semelhante ao controle p.S1613G (variante não patogênico), com 115,8% de atividade relativa ao controle selvagem. Enquanto a variante p.E1527M apresentou 161,6% de atividade (**Figura 5.3**).

Todas as variantes situadas no domínio tBRCT apresentaram atividade inferior a 55% do controle selvagem. A variante p.Y1716K (4,8%), localizada no BRCT1, apresentou atividade semelhante às variantes p.M1775R e p.Y1853X (controles negativos, patogênicos). As variantes p.E1698R (40,5%), p.V1740N (20%), p.N1745R (37,5%), p.N1774R (33,2%), e p.L1844I (41%), presentes ao longo do tBRCT, apresentaram valores superiores ao dos controles patogênicos, no entanto não alcançaram 50% da atividade da proteína selvagem. Já a variante p.K1759W, situada na região *linker* do tBRCT apresentou 53,8% de atividade (**Figura 5.3**).

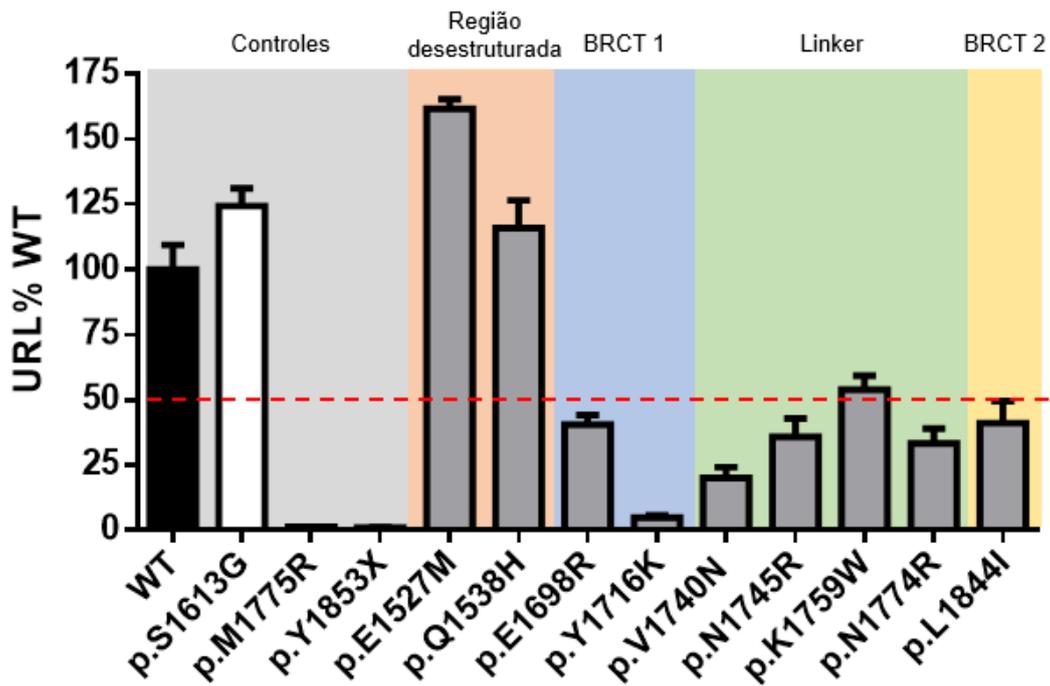


Figura 5.3. Análise da capacidade de ativação transcricional das variantes *missense*. Gráfico representativo dos experimentos de ativação transcricional realizados com as variantes do tipo *missense*. O eixo Y representa a porcentagem de atividade de cada variante em relação ao controle selvagem (*Wild Type*; WT). URL%WT = Unidade Relativa de Luz a 100% do selvagem. O controle selvagem (WT) está indicado pela barra preta, o variante não patogênico (p.S1613G), pela barra branca. As diferentes cores de fundo delimitam as regiões onde as variantes estão localizadas na estrutura proteica de BRCA1, indicadas na parte superior do gráfico. A linha vermelha indica 50% da atividade do controle selvagem.

Dois terços dos resultados apresentados estão de acordo com a predição *in silico* (**Tabela 5.1**). Apenas as variantes p.K1759W, p.N1774R, p.L1844I apresentaram resultados conflitantes, tendo sido preditas como *Class C0* (baixo risco, neutralidade) e apresentado atividade transcricional inferior à 55% do controle selvagem.

5.3.2. Capacidade de ativação transcricional das variantes de deleções de códon único

Dentre as 7 variantes que resultam na deleção de um único códon, as variantes localizadas no BRCT1, apresentaram atividade semelhantes aos valores dos controles patogênicos (p.M1775R e p.Y1853X). A saber: p.H1673del (2%), p.V1688del (0,9%), p.A1693del (0,7%), p.Q1721del (1,1%) e p.M1728del (2,3%). Por outro lado, as variantes p.L1504del (55,1%) e p.E1559del (113,2%), localizadas na região desestruturada, apresentaram atividade superior a 50% do controle selvagem (**Figura 5.4**).

Coletivamente, esses dados corroboram a importância da organização estrutural do domínio tBRCT na atividade transcricional da porção C-terminal de BRCA1, sugerindo que a região desestruturada tolere a perda de alguns resíduos de aminoácido.

5.3.3. Capacidade de ativação transcricional das variantes *nonsense* na borda C-terminal de BRCA1

A literatura mostra de forma consistente que a perda dos 11 últimos resíduos de aminoácido na extremidade C-terminal de BRCA1 resulta no aumento de risco de desenvolvimento de CM (MIKI *et al.*, 1994; MONTEIRO *et al.*, 1996; PHELAN *et al.*, 2005; CARVALHO *et al.*, 2007). Dados funcionais recentemente publicados sugerem que os sete resíduos de aminoácidos localizados na borda C-terminal do domínio tBRCT são essenciais para sua integridade (FERNANDES *et al.*, 2019). Para caracterizar o limite funcional de impacto na borda C-terminal do domínio tBRCT foram geradas 6 mutações *nonsense* localizadas entre as posições 1856 e 1863 de BRCA1 (p.P1856X, p.Q1857X, p.I1858X, p.P1859X, p.H1860X, p.S1861X, p.H1862X, p.Y1863X; **Figura 5.2**).

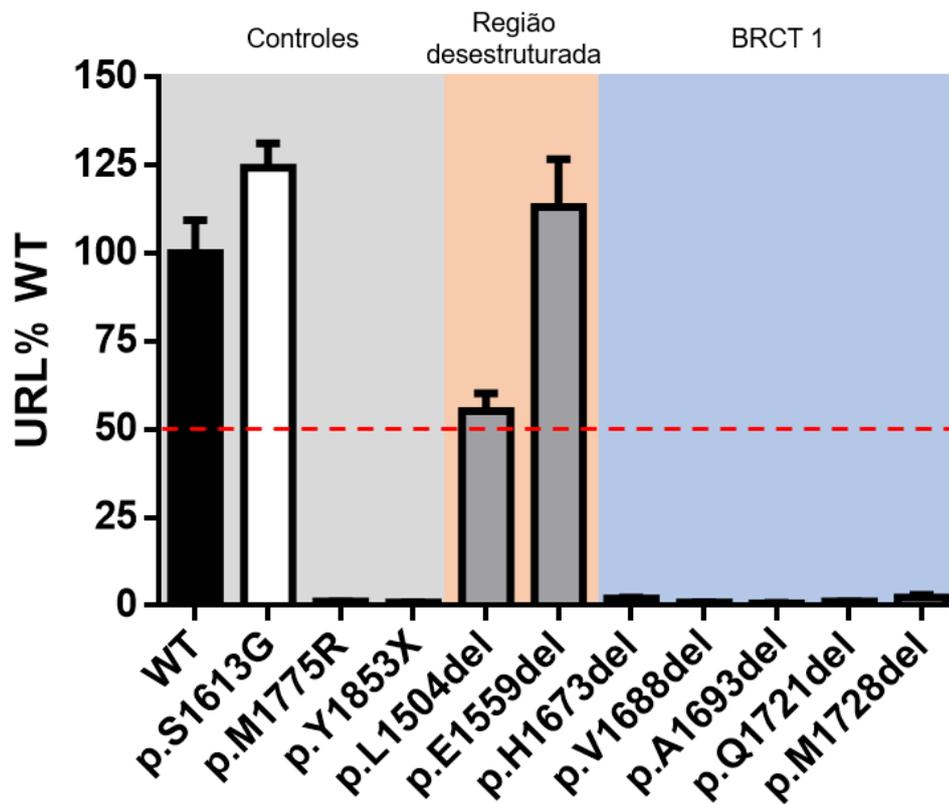


Figura 5.4. Análise da capacidade de ativação transcricional das variantes com deleção de códon único. Gráfico representativo dos experimentos de ativação transcricional realizados com as variantes com deleção de códon único. O eixo Y representa a porcentagem de atividade de cada variante em relação ao controle selvagem (*Wild Type*; WT). URL%WT = Unidade Relativa de Luz a 100% do selvagem. O controle selvagem (WT) está indicado pela barra preta, o não patogênico (p.S1613G), pela barra branca. As diferentes cores de fundo delimitam as regiões onde as variantes estão localizadas na estrutura proteica de BRCA1, representadas na porção superior do gráfico. A linha vermelha indica 50% da atividade do controle selvagem.

As variantes p.P1856X e p.Q1857X apresentaram atividade semelhante a relatada em um trabalho prévio do grupo, tendo atingido 3,8% e 27,4%, respectivamente (**Figura 5.5**) (FERNANDES *et al.*, 2019). A presença dos resíduos p.I1858 e p.P1859 na porção C-terminal resultou em um aumento da atividade em relação a variante p.Q1857X, 37,6% e 48,8%, respectivamente. Entretanto, esse patamar de atividade se manteve nas variantes p.H1860X (46,9%), p.S1861X (48,8%), p.H1862X (45,5%) e p.Y1863X (46,1%), não ultrapassando os 50% de atividade relativa ao controle selvagem (**Figura 5.5**). Esses dados sugerem que os 8 últimos resíduos de aminoácido de BRCA1 são importantes para manutenção da integridade do domínio tBRCT, onde a perda de um único resíduo na porção C-terminal resulta na perda de 50% da atividade transcricional.

5.4. Avaliação dos níveis proteicos das variantes

Os níveis das proteínas quiméricas GAL4_{DBD}:BRCA1_{13/24} (selvagem ou variantes) foram analisados por *immunoblotting*, utilizando anticorpos específicos para o domínio de ligação ao DNA de GAL4 (GAL4_{DBD}) e β -actina (controle de carregamento). A proteína quimérica GAL4_{DBD}:BRCA1_{13/24} foi identificada com peso molecular aproximado de 70kDa, como pode ser observado na **Figura 5.6**.

Como já descrito anteriormente, os níveis proteicos das variantes patogênicas p.M1775R e p.Y1853X foram observados inferiores ao controle não patogênico (p.S1613G) e ao selvagem (**Figura 5.6A**). (CARVALHO *et al.*, 2007).

As variantes *missense* (p.E1527M, p.Q1538H, p.E1698R, p.V1740N, p.N1745R, p.K1759W, p.N1774R) apresentaram níveis proteicos semelhantes ao controle selvagem. A variante p.L1844I apresentou níveis reduzidos quando comparada ao controle selvagem, e apenas a variante p.Y1716K, que apresentou a menor atividade transcricional, não apresentou níveis detectáveis da proteína (**Figura 5.6B**).

Das deleções de códon único, apenas a variante p.E1559del apresentou níveis próximos à construção selvagem. As variantes p.L1504del, p.H1673del, p.V1688del, p.Q1721del e p.M1728del apresentaram níveis proteicos detectáveis, porém reduzidos quando comparados ao controle. Apenas a variante p.A1693del não apresentou níveis proteicos detectáveis (**Figura 5.6C**).

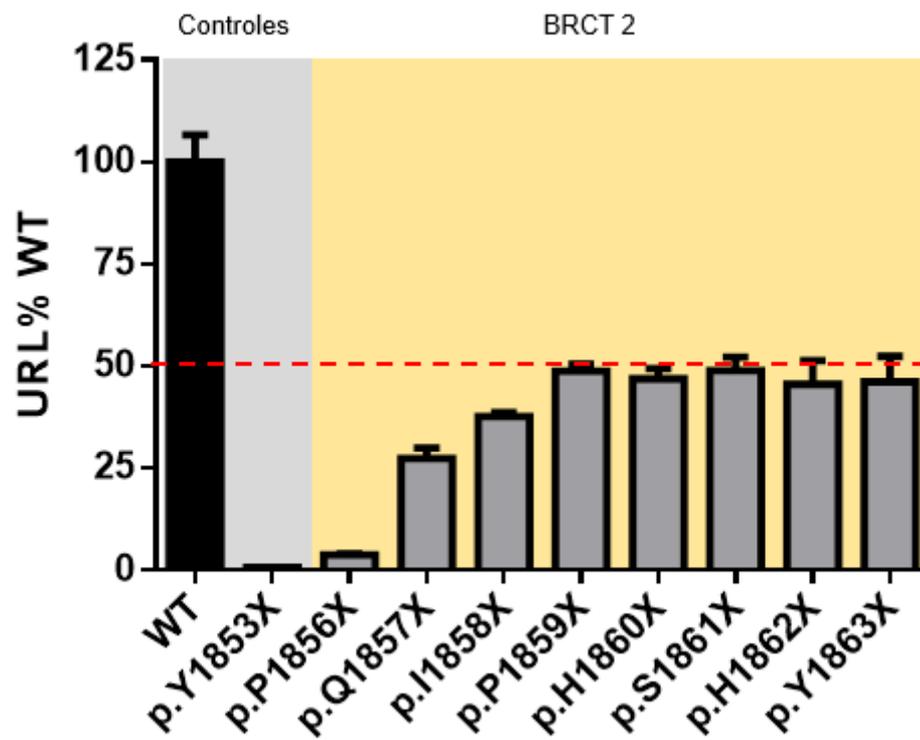


Figura 5.5. Análise da capacidade de ativação transcricional das variantes *nonsense* da porção C-terminal do tBRCT de BRCA1. Gráfico representativo dos experimentos de ativação transcricional realizados com as variantes nonsense. O eixo Y representa a porcentagem de atividade de cada variante em relação ao controle selvagem (*Wild Type*; WT). URL%WT = Unidade Relativa de Luz a 100% do selvagem. O controle selvagem (WT) está indicado pela barra preta. As diferentes cores de fundo delimitam as regiões onde as variantes estão localizadas na estrutura de BRCA1, representadas na porção superior do gráfico. A linha vermelha indica 50% da atividade do controle selvagem.

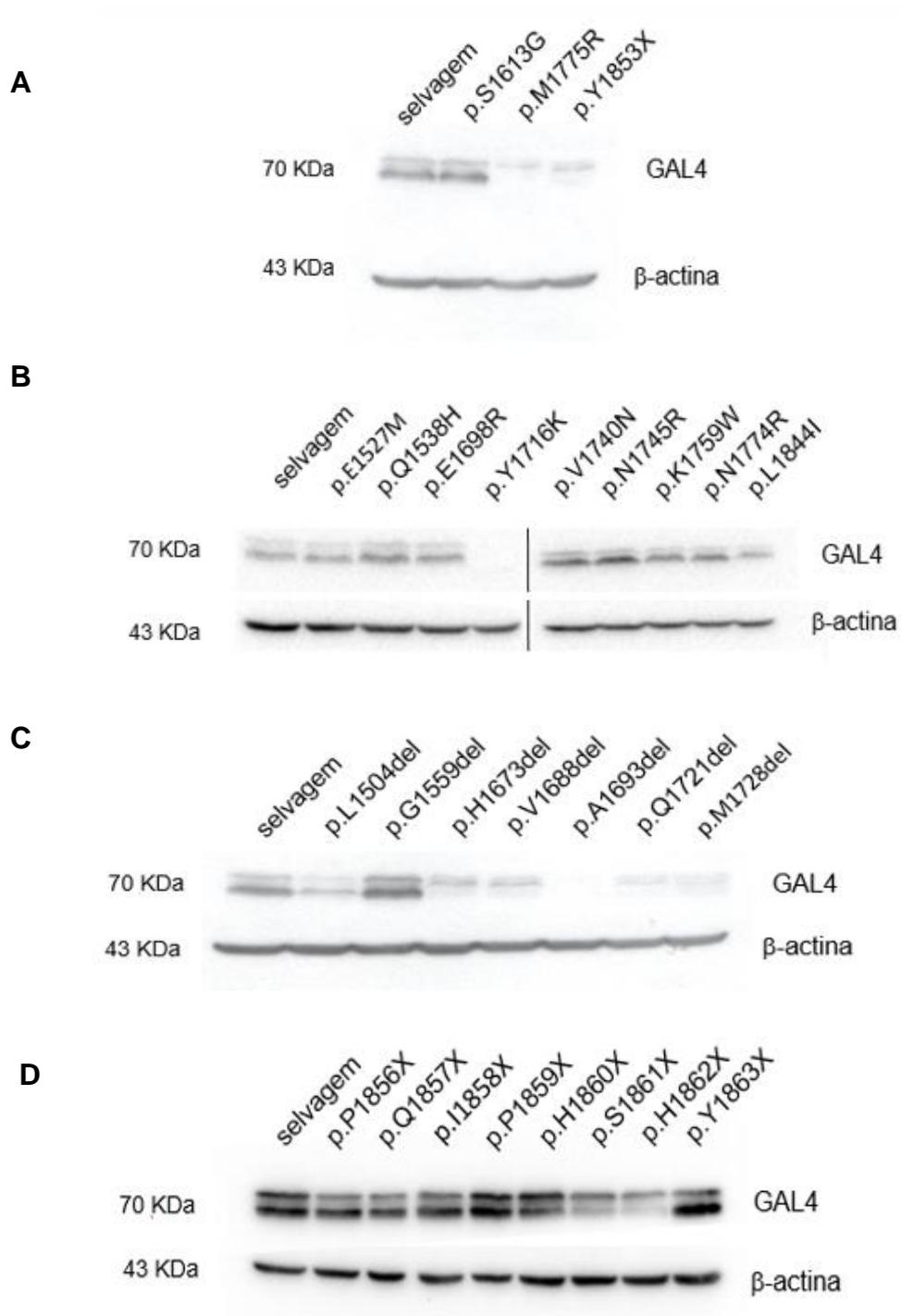


Figura 5.6. Perfil das proteínas quiméricas GAL4^{DBD}:BRCA1_{13/24}. A determinação dos níveis proteicos das variantes estudadas foi realizada através de rotinas de *immunoblotting* utilizando anticorpos anti-GAL4^{DBD} e anti- β -actina (controle de carregamento). **(A)** Níveis proteicos dos controles selvagem, não patogênico (p.S1613G), e patogênicos (p.M1775R e p.Y1853X). **(B)** variantes *missense*, **(C)** variantes com deleção de códon único, e **(D)** das mutações *nonsense* da borda C-terminal do tBRCT de BRCA1.

Dentre as oito variantes *nonsense* geradas na porção C-terminal do tBRCT de BRCA1, seis apresentaram níveis proteicos semelhantes aos do controle selvagem (p.1856X, p.Q1857X, p.I1858X, p.P1859X, p.H1860X e p.Y1863X). Apenas as variantes p.S1861X e p.H1862X resultaram na redução dos níveis proteicos relativos ao controle selvagem (**Figura 5.6D**).

6. DISCUSSÃO

A observação de que a existência de histórico familiar de CM aumenta o risco de desenvolvimento da neoplasia data do final dos anos 1970 (KELSEY, 1979; HURST, 2014). Duas décadas depois, em 1994, foi estabelecida a primeira relação direta entre um *locus* gênico e casos de CM com perfil hereditário, levando à identificação de *BRCA1* como o primeiro gene determinante de predisposição à doença e se tornando um tópico de relevância para a compreensão da biologia do CM (MIKI *et al.*, 1994).

As primeiras variações em *BRCA1* relacionadas à predisposição hereditária no CM foram mutações do tipo *nonsense* (p.Q1313X), *missense* (p.M1775R) e indels (p.C24fs6X – deleção de 11 aminoácidos e p.Q1756fs1829X – inserção de um nucleotídeo) (MIKI *et al.*, 1994). As variantes do tipo *missense* compõem cerca de 50% das mutações reportadas (*ClinVar*, 2020). A maior parte destas variações ocorrem em baixa frequência na população, dificultando estudos de associação com poder estatístico e a consequente inferência de relação da variante com o câncer. O desafio em classificá-las também se deve ao fato de resultarem na troca de um único resíduo de aminoácido, dificultando a interpretação de seu impacto na estrutura e função da proteína (GOLUBEVA *et al.*, 2019). Com isso, a maioria dos depósitos de variantes *missense* em bancos de dados são classificadas como VUS (*ClinVar*, 2020).

Nesse cenário, as indels que respeitam o quadro aberto de leitura de tradução também se apresentam como um desafio de classificação quanto a sua patogenicidade. Nosso trabalho foca no estudo funcional dessas variantes de *BRCA1*, a fim de contribuir para o entendimento de seu impacto funcional.

As variantes identificadas no banco de dados do BRCA Exchange corroboram a observação de que não há uma prevalência de mutações em uma região específica de *BRCA1*. Através de uma varredura inicial, identificamos indels ao longo de toda a região codificante (**Anexo I**), o que se reflete na porção estudada compreendida entre os éxons 13 ao 24 de *BRCA1* (**Figura 5.2**) (GOLUBEVA *et al.*, 2019). Essa região codifica os últimos resíduos de aminoácido do motivo *coiled-coil*, e do domínio tBRCT (subdividido em BRCT1, região *linker* e BRCT2), separados por uma região desestruturada (FERNANDES *et al.*, 2019).

Existem apenas dois estudos que avaliam o impacto funcional de indels na região C-terminal de *BRCA1*, ilustrando a escassez de dados disponíveis na literatura sobre alterações que mantêm o quadro de leitura da proteína. Parte do nosso trabalho visa contribuir para a elucidação do comportamento funcional das variantes indels de

BRCA1 identificadas na população. Nesse contexto, a discussão desse conjunto de dados será guiada por dois pontos principais: (1) o impacto e relevância do produto proteico das variantes indels, e (2) a localização dessas variantes nas diferentes regiões encontradas ao longo da porção C-terminal da proteína.

A rigor, a classificação funcional de variantes de *BRCA1* passa por análises integradas, incluindo o modelo hierárquico Bayesiano do *VarCall*. Do ponto de vista da interpretação de manutenção ou perda de função no nosso conjunto de dados, nos baseamos em um trabalho prévio do nosso grupo, onde Carvalho e colaboradores (2007) demonstraram que um conjunto de variantes sabidamente patogênicas apresentam um comportamento semelhante quando avaliados através dos ensaios de TA, não ultrapassando a marca de 45% da atividade relativa da condição selvagem (controle), enquanto variantes sabidamente benignas apresentam atividade superior a 50%. Dessa forma, consideramos as variantes como disfuncionais quando apresentam atividade inferior aos 45%, assim como interpretamos como manutenção de função as variantes que apresentam atividade superior aos 50% relativos ao controle (CARVALHO *et al.*, 2007).

Alterações causadas por inserções e deleções de um mesmo número de nucleotídeos podem resultar em variantes do tipo *missense*. Isso ocorre quando há a alteração de até três nucleotídeos dentro de um mesmo códon, permitindo que ocorra a troca de um único resíduo de aminoácido. Das 22 variantes selecionadas para esse estudo, 9 se comportam dessa forma (**Anexo I**).

Dentro desse conjunto, apenas a variante p.L1844I resulta na troca de um resíduo de aminoácido com as mesmas características físico-químicas do resíduo original (aminoácidos não polares, não carregados). Entretanto, essa conservação não é suficiente para manutenção da atividade transcricional de *BRCA1*, tendo a variante apresentado perda de função, resultando em 41% de atividade relativa. Por outro lado, o resíduo de leucina na posição 1844 encontra-se conservado em diferentes ortólogos de *BRCA1* (**Figura 6.1**). O que sugere que esse resíduo seja importante para a organização e função do domínio tBRCT.

A variante p.E1698R resulta na troca de um resíduo com cadeia lateral ácida por um resíduo básico. Seria razoável antever que essa alteração poderia resultar no comprometimento da organização estrutural do domínio (mesmo não sendo um comportamento seja mandatório). Nossos dados indicam que a troca de um ácido glutâmico por uma arginina, na posição 1698, resulta na perda de função da proteína, com redução de aproximadamente 60% da atividade transcricional de *BRCA1*.

		1698		1716		1740	
<i>Homo sapiens</i>	TDAEFVCERTLK	YF	VSYFWVTQ	SIKE	EV
<i>Pan troglodytes</i>	TDAEFVCERTLK	YF	VSYFWVTQ	SIKE	EV
<i>Gorilla gorila</i>	TDAEFVCERTLK	YF	VSYFWVTQ	SIKE	EV
<i>Pongo pygmaeus</i>	TDAEFVCERTLK	YF	VSYFWVTQ	SIKE	EV
<i>Macaca mulatta</i>	TDAEFVCERTLK	YF	VSYFWVTQ	SIKE	EV
		1759		1844			
<i>Homo sapiens</i>	KIFRGL	EICCY	VLDSVALY	QCQELD	
<i>Pan troglodytes</i>	KIFRGL	EICCY	VLDSVALY	QCQELD	
<i>Gorilla gorila</i>	KIFRGL	DICCY	VLDSVALY	QCQELD	
<i>Pongo pygmaeus</i>	KIFRGL	EICCY	VLDSVALY	QCQELD	
<i>Macaca mulatta</i>	KIFRGL	EICCY	VLDSVALY	QCQELD	

Figura 6.1. Conservação de aminoácidos entre as espécies. Imagem ilustrando a conservação dos resíduos E1698, Y1716, V1740, K1759 e L1844 entre as espécies *Homo sapiens* (humano), *Pan troglodytes* (chimpanzé), *Gorilla gorila* (gorila), *Pongo pygmaeus* (orangotango) e *Macaca mulatta* (macaco) (Alinhamento de sequências retirados do *Align-GVGD*).

O resíduo E1698 é conservado entre as diferentes espécies (**Figura 6.1**). A importância desse resíduo para as funções de BRCA1 também pode ser evidenciada pelo fato de que todas as variantes *missense* decorrentes da troca de um único nucleotídeo resultaram na letalidade de células haploides proficientes para p53 (FINDLAY *et al.*, 2019).

A troca do resíduo de valina na posição 1740 por um resíduo de asparagina resulta na perda de função com redução de aproximadamente 80% da atividade em relação ao controle selvagem. A V1740 também se encontra em uma região conservada, sugerindo que esse resíduo seja importante para organização e função dos domínios tBRCT (**Figura 6.1**). Do ponto de vista físico-químico, a troca de um resíduo com cadeia lateral apolar por outro polar, não carregado, pode impactar a organização estrutural do domínio por mediar a exposição de uma região hidrofóbica. O mesmo ocorre com a variante p.Y1716K, que substitui o resíduo conservado de tirosina (apolar) por uma lisina (básico). Funcionalmente, essa variante apresenta perda de função, resultando em cerca de 5% de atividade relativa.

A variante p.K1759W apresenta a troca de um resíduo com cadeia lateral básica (lisina) por um resíduo apolar (triptofano) também em uma região conservada (**Figura 6.1**). Entretanto, essa variante apresentou cerca de 54% de atividade, quando comparada à sequência selvagem, sugerindo um comportamento limítrofe de manutenção da função. Levando em consideração o desvio padrão observado em nossas análises ($\pm 5.3\%$), devemos considerar essa variante como um VUS devido à limitação do método utilizado (CARVALHO *et al.*, 2007). Curiosamente, análises prévias do nosso grupo sugerem que essa posição não seja tolerante a alterações, as variantes p.K1759N, p.K1759R e p.K1759M foram classificadas como disfuncionais pelo algoritmo do *VarCall* (FERNANDES *et al.*, 2019). É de nosso interesse expandir nossas análises atuais através do uso *VarCall*, que possivelmente será capaz reclassificar a variante p.K1759W, possivelmente como disfuncional, tendo em vista o comportamento das outras variações nesse resíduo.

Já as variantes p.N1745R, p.N1774R, p.Q1538H e p.E1527M resultam na manutenção da polaridade da cadeia lateral do resíduo selvagem. Essas variantes apresentam um largo espectro de atividade transcricional, variando de 33% à 162% do controle selvagem. Entretanto, essa variação de atividade parece estar mais associada à localização da variante na proteína. Esse tópico será discutido a seguir.

Todas as indels que resultam em uma variante *missense* também foram avaliadas através do algoritmo de predição de patogenicidade *in silico* *Align-GVGD*.

Essa ferramenta avalia as alterações dos resíduos de aminoácidos através de suas características físico-químicas, da comparação entre sequências de ortólogos, além de considerar a estrutura da proteína. Com esse conjunto de informações, o algoritmo visa definir o impacto de tal variante na estrutura da proteína (TAVTIGIAN *et al.*, 2005). Dois terços dos resultados da análise *in silico* estão de acordo com nossos dados funcionais, corroborando a ideia de que as ferramentas de predição, mesmo que valiosas para a interpretação de risco, precisam ser validadas empiricamente. As divergências encontradas estão em sintonia com observações anteriores (FERNANDES *et al.*, 2016; RODRIGUE *et al.*, 2019)

Das 22 variantes selecionadas para esse estudo, 6 resultam na deleção de um único códon, mantendo o quadro aberto de tradução de BRCA1 (**Anexo I**). Segundo a literatura, são dois os estudos que já caracterizaram o impacto funcional de variantes que implicam em deleções no tBRCT, respeitando o quadro de leitura de BRCA1 (CARVALHO *et al.*, 2009; VALLON-CHRISTERSSON *et al.*, 2001). Carvalho e colaboradores (2009) analisaram o comportamento de uma variante natural (Δ éxons 16/17) que resulta na deleção dos éxons 16 e 17. A região codificada por esses éxons encerra a porção C-terminal da região desestruturada, e parte da porção N-terminal do BRCT1. Sua perda resulta na redução drástica da atividade transcricional mediada pelos domínios tBRCT de BRCA1 (CARVALHO *et al.*, 2009). Esse dado sugere que a desestruturação do domínio tBRCT, através da perda de parte do BRCT1, tem um impacto direto na ativação transcricional mediada por BRCA1. Essa observação está de acordo com o entendimento de que o ensaio de TA avalia, de forma indireta, a manutenção da estrutura dos domínios tBRCT (FERNANDES *et al.*, 2019). Entretanto, a variante avaliada por Carvalho e colaboradores (2009) resulta na deleção de uma grande porção da proteína, o que não permite inferir perda de funcionalidade de variantes que resultam apenas na deleção de um único resíduo de aminoácido.

Vallon-Christersson e colaboradores (2001) analisaram um grupo de VUS encontradas em famílias escandinavas. Dentre as variantes estudadas, encontra-se a deleção de um único códon (p.V1688del) dentro do BRCT1 de BRCA1. No trabalho conduzido em 2001, tal variante mostrou atividade transcricional baixa, próximo a 1% da atividade relativa do controle selvagem (VALLON-CHRISTERSSON *et al.*, 2001). Esse dado sugere que, diferente do observado anteriormente, a perda de um único resíduo de aminoácido no domínio tBRCT resulta na sua desestruturação. Essa é a primeira observação empírica do impacto funcional desse tipo de variação em BRCA1. Vale ressaltar que os dados de Monteiro e colaboradores (1996) demonstraram que o

domínio BRCT2, de forma isolada, também apresenta atividade transcricional (MONTEIRO *et al.*, 1996). Entretanto, essa atividade é exacerbada pela organização do domínio tBRCT. Coletivamente, esses dados corroboram a noção de que o domínio tBRCT atue como uma única unidade estrutural (MESQUITA *et al.*, 2010), de forma que a perda de um único resíduo no domínio BRCT1 é capaz de impactar na atividade transcricional de BRCA1, mesmo mantendo a integridade do domínio BRCT2.

Nossos dados demonstram que das 6 deleções analisadas, apenas duas apresentam atividade superior à 55% ao controle selvagem. Entretanto, essa manutenção da atividade transcricional também parece estar associada à posição da variante na proteína. De forma genérica, todas as deleções localizadas no tBRCT de BRCA1, resultam em um impacto severo de função, corroborando as observações da literatura (VALLON-CHRISTERSSON *et al.*, 2001).

As diferentes variantes estudadas apresentam grande variação da atividade transcricional, o que parece estar relacionado à localização dessas alterações ao longo da proteína estudada. Nossos dados demonstram uma clara distinção entre o comportamento das variantes presentes na região desestruturada quando comparadas às localizadas no domínio tBRCT.

A região desestruturada é constituída por sequências peptídicas não suficientemente hidrofóbicas para formação de estruturas bem definidas (BABU, 2016). Todas as variantes situadas nessa região, sendo duas *missense* (p.E1527M e p.Q1538H) e duas deleções de códon único (p.L1504del e p.E1559del), apresentam manutenção de função, com atividades superiores aos 50% do controle selvagem. Isso sugere que essa região seja mais tolerante a mutações. Esses dados corroboram a noção de que a atividade transcricional de BRCA1 é definida pelo domínio tBRCT da proteína (FERNANDES *et al.*, 2019).

Por outro lado, nossos dados indicam que a deleção do códon 1504 (p.L1504del) resulta na perda de aproximadamente 50% da atividade da proteína, sugerindo que o contexto adjacente ao domínio tBRCT seja de alguma forma importante para a regulação dessa atividade. O comportamento da variante p.E1527M (atividade relativa de 161.6%) Corrobora essa observação. Ainda não está claro qual o real impacto biológico de variantes que resultam em aumento de atividade transcricional.

É importante salientar que o ensaio de TA desconsidera outras funções desempenhadas por BRCA1, como a formação de complexos proteicos durante o reparo por HR e, por conseguinte, impossibilitando inferências do impacto dessas

variantes nessas atividades. Uma alternativa de avaliação do impacto dessas variantes no papel de BRCA1 durante a RDD, é a avaliação da capacidade de interação com a proteína CtIP, fundamental para o desempenho das funções de BRCA1 no reparo de DSBs (CARVALHO *et al.*, 2002).

Todas as variantes localizadas no domínio tBRCT, com exceção da p.K1759W, apresentam perda de função, com atividade relativa inferior a 50%. No entanto, as variantes *missense* apresentam atividade superior às de deleção de códon único. Tendo todas as variantes de deleção de códon único resultado em atividade relativa próxima à zero. Isso sugere que, intuitivamente, a deleção de um resíduo de aminoácido, ao invés de uma troca, pode ser mais prejudicial para as funções da proteína. Além disso, é importante observar que pequenas deleções de códon único tenham o mesmo impacto na atividade do tBRCT que deleções de éxons inteiros, como visto no estudo de Carvalho e colaboradores em 2009 (CARVALHO *et al.*, 2009).

Análises estruturais sugerem que o último resíduo de aminoácido que constitui a formação da terceira α -hélice do BRCT2 (α '3; **Figura 1.3**) seja a Leucina na posição 1854 (CLEPPERON *et al.*, 2004). Ainda não está claro o papel dessa porção final da proteína, seja do ponto de vista das interações intramoleculares ou do seu papel na manutenção das funções de BRCA1. Em 2019, Fernandes e colaboradores questionaram a importância das extremidades N- e C-terminal do tBRCT de BRCA1 para manutenção de sua integridade. Nesse estudo, foram geradas variantes *nonsense* na borda C-terminal do BRCT2, resultando na perda de sete (p.Q1857X), oito (p.P1856X) ou nove (p.I1855X) resíduos de aminoácido. A variante p.I1855X apresentou atividade transcricional próxima de zero, sugerindo que essa porção final de BRCA1 seja importante para a manutenção da estrutura do domínio tBRCT. Já as variantes p.P1856X e p.Q1857X apresentaram ganho de atividade, quando comparadas à variante p.I1855X, porém não superaram a marca de 30% relativo ao controle selvagem (Figura 5.4) (FERNANDES *et al.*, 2019).

A análise estrutural do domínio tBRCT demonstra que o resíduo de glutamina na posição 1857, interage com os resíduos de fenilalanina (p.F1821), isoleucina (p.I1824) e glicina (p.G1825), presentes na região da segunda α -hélice do domínio BRCT2 (α '2, **Figura 1.3**), isso também ocorre com o resíduo de prolina na posição 1859, que interage com o resíduo de glutamina na posição 1826, também localizado na α '2 (CLEPPERON *et al.*, 2004). Essas observações sugerem que resíduos localizados a C-terminal da p.Q1857 possam apresentar interações intramoleculares

importantes para a manutenção da estrutura do domínio tBRCT. Entretanto, os dados de cristalografia do domínio tBRCT disponíveis na literatura se limitam aos resíduos 1649 ao 1859 de BRCA1 (Sheng *et al.*, 2011).

Nossos dados corroboram essa hipótese, por demonstrar que a perda de um único resíduo de aminoácido na porção C-terminal de BRCA1 (p.Y1863X) é o suficiente para redução de aproximadamente 50% da atividade transcricional da proteína. Curiosamente, o impacto da ausência da tirosina na posição 1863, é semelhante à perda de até cinco resíduos nessa região da proteína (p.P1859X, p.H1860X, p.S1861X, p.H1862X), sugerindo que o papel desempenhado pela p.Y1863 seja central na manutenção dessa organização estrutural.

Um dos possíveis mecanismos responsáveis pela perda de função de variantes, passa pela instabilidade da proteína produzida, que resulta na redução (eventualmente total) dos níveis proteicos. Nossos dados demonstram que as deleções de um único resíduo de aminoácido resultam, majoritariamente, na redução dos níveis proteicos. O mesmo não ocorre com nossas variantes *missense*. As variantes p.A1693del e p.Y1716K resultam na redução drástica dos níveis proteicos, mostrando níveis não detectáveis no *immunoblot* (**Figura 5.5**). Curiosamente, essas variantes também resultam na perda da capacidade de ativação transcricional (0,7% e 4,8%, respectivamente). Entretanto, algumas variantes que resultam na perda de função são capazes de sustentar o perfil proteico a níveis semelhantes ao controle selvagem (ex. p.V1740N e p.1856X).

Dados prévios do nosso grupo sugerem que inferência de perda de função transcricional mediada pela redução dos níveis proteicos não é absolutamente verdadeira. Fernandes e colaboradores (2019) mostraram que algumas (por exemplo, p.V1740E e p.H1746D) quando avaliadas em modelos de células animais (HEK293FT) apresentam redução dos níveis proteicos e perda de atividade transcricional; entretanto, em leveduras (*Saccharomyces cerevisiae*) cultivadas à 30°C, ocorre o resgate dos níveis proteicos, que não é acompanhado pelo restauro da atividade transcricional, sugerindo que a perda de função das variantes seja independente da manutenção da estabilidade da proteína (FERNANDES *et al.*, 2019). Dessa forma, deve-se ter cautela na interpretação da disfuncionalidade de uma variante apenas considerando a perda/manutenção dos níveis proteicos.

Um ponto importante (e que não deve ser ignorado): o nosso modelo de avaliação funcional se vale da produção ectópica de uma porção de BRCA1 (o C-terminal), formando uma proteína quimérica com uma região heteróloga de ligação ao

DNA (GAL4 DBD) – esse contexto não reflete a realidade do impacto de uma variante nos controles de estabilidade de BRCA1. Dessa forma, a avaliação dos níveis proteicos de uma variante deve ser idealmente conduzida no contexto constitutivo do produto cadeia-completa.

7. CONCLUSÃO

Como mencionado anteriormente, a análise do impacto funcional de indels em BRCA1 foi historicamente negligenciada. Nosso trabalho avalia, pela primeira vez, um conjunto de variantes naturais, resultantes de deleções e/ou inserções, na porção C-terminal de BRCA1.

Através das nossas análises, podemos concluir que o domínio tBRCT é extremamente sensível a deleções, mesmo que apenas de um resíduo de aminoácido. Tendo em vista que todas as deleções se encontram no domínio BRCT1, podemos inferir então que a conservação de sua estrutura é fundamental para a manutenção da integridade do domínio tBRCT de BRCA1.

De forma geral, nossos dados corroboram a noção de que variantes *missense* localizados em regiões conservadas do domínio tBRCT tendem a impactar a sua atividade transcricional. Além de confirmar que a análise *in silico* de predição de patogenicidade apresenta significativas discordâncias com dados empíricos.

Por outro lado, os resultados apresentados apontam que a região desestruturada parece tolerar melhor esse tipo de alteração. Porém, ainda não está claro se o comportamento da variante p.L1504del (redução de atividade) é exclusivo da deleção desse resíduo.

Demonstramos, ainda, o importante papel de todos os resíduos localizados na borda C-terminal do domínio BRCT2 para a manutenção da estabilidade do domínio tBRCT de BRCA1.

Coletivamente, nossos dados reforçam a ideia de que o estudo do comportamento dessas variantes é extremamente valioso para a melhor compreensão da organização estrutural e função desempenhada pelo domínio tBRCT de BRCA1.

8. REFERÊNCIAS

Align-GVGD. Disponível em <<http://agvgd.hci.utah.edu/>>, Acesso realizado em 11/2019.

ALKAILANI, M. *et al.* **A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer.** *NAR Cancer*. 3(1):zcaa040, 2021.

ANANTHA, R. W. *et al.* **Functional and mutational landscapes of BRCA1 for homology-directed repair and therapy resistance.** *eLife*.. 6: e21350, 2017.

ANDERSON, S. F. *et al.* **BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A.** *Nat Genet*. 19(3):254-6, 1998.

APARACIO, T.; BAER, R. & GAUTIER, J. **DNA double-strand break repair pathway choice and cancer.** *DNA Repair (Amst)*. 19:169-75, 2014.

AUSUBEL, F. M. *et al.* **Short Protocols in Molecular Biology**, 5th ed., John Wiley & Sons, 1504 pp, 2002.

BABU, M. M. **The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease.** *Biochem Soc Trans*. 44(5): 1185–1200, 2016.

BRADFORD, M. A. **rapid and sensitive method for the quantification microgram quantities of protein utilizing the principle of proteins-dye binding.** *Anal Biochem*. 72: 248-254, 1976.

BRCA Exchange. Disponível em: <<https://brcaexchange.org>>. Acesso realizado em 04/2019

BOUWMAN, P. *et al.* **Functional Categorization of BRCA1 Variants of Uncertain Clinical Significance in Homologous Recombination Repair Complementation Assays.** *Clin Cancer Res*. 26(17):4559-4568, 2020.

CALECA, L. *et al.* **GFP-Fragment Reassembly Screens for the Functional Characterization of Variants of Uncertain Significance in Protein Interaction Domains of the BRCA1 and BRCA2 Genes.** *Cancers (Basel)*. 11(2):15, 2019.

CARVALHO, M. A.; COUCH, F. J. & MONTEIRO, A. **Functional assays for BRCA1 and BRCA2.** *Int J Biochem Cell Biol*, 39(2):298-310, 2007.

- CARVALHO, M. *et al.* **Determination of Cancer Risk Associated with Germ Line BRCA1 Missense Variants by Functional Analysis.** *Cancer Res.* 67(4):1494–1501, 2007.
- CARVALHO, M. *et al.* **Analysis of a set of missense, frameshift, and in-frame deletion variants of BRCA1.** *Mutat Res.* 660(1-2): 1–11, 2009.
- CAVANAGH, H. & ROGERS, K. **The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers.** *Hered Cancer Clin Pract.* 13(1):16, 2015.
- CHATTERJEE, N. & WALKER, G. C. **Mechanisms of DNA damage, repair, and mutagenesis.** *Environ Mol Mutagen.* 58(5):235-263, 2017.
- CLAPPERTON, J. A. *et al.* **Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer.** *Nat Struct Mol Biol.* 11(6):512-8, 2004.
- CLARK, S. L. *et al.* **Structure-Function Of The Tumor Suppressor BRCA1.** *Comput Struct Biotechnol J.* 1(1):e201204005, 2012.
- CLINE, M. S. *et al.* **BRCA Challenge: BRCA Exchange as a global resource for variants in BRCA1 and BRCA2.** *PLoS Genet.* 14(12):e1007752, 2018.
- CHRISTOU, C. M. *et al.* **The BRCA1 variant p.Ser36Tyr abrogates BRCA1 protein function and potentially confers a moderate risk of breast cancer.** *PLoS One.* 9(4):e93400, 2014.
- COUCH, F. *et al.* **Assessment of Functional Effects of Unclassified Genetic Variants.** *Hum Mutat.* 29(11):1314–1326, 2008.
- DUCY, M. *et al.* **The Tumor Suppressor PALB2: Inside Out.** *Trends Biochem Sci.* 44(3):226-240, 2019.
- DUNNING, A. M. *et al.* **Common BRCA1 variants and susceptibility to breast and ovarian cancer in the general population.** *Hum Mol Genet.* 6(2):285-9, 1997.
- EASTON, D. F. *et al.* **Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium.** *Am J Hum Genet.* 52(4):678-701, 1993.

ECONOMOPOULOU, P.; DIMITRIADIS, G. & PSYRRI, A. **Beyond BRCA: new hereditary breast cancer susceptibility genes.** *Cancer Treat Rev.* (1):1-8, 2015.

ELIA, A. E. H. et al. **Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response.** *Mol Cell.* 59(5):867-81, 2015.

ENGEL, C. & FISHER, C. **Breast cancer risks and risk prediction models.** *Breast Care (Basel).* 10(1):7-12, 2015.

FABBRO, M. et al. **BARD1 induces BRCA1 intranuclear foci formation by increasing RING-dependent BRCA1 nuclear import and inhibiting BRCA1 nuclear export.** *J Biol Chem.* 277(24):21315-24, 2002.

FAN, W. et al. **BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs.** *J Biol Chem.* 277(10):8061-7, 2002.

FERLAY, J. et al. **Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012.** *Eur J Cancer.* 49(6):1374-403, 2013.

FERNANDES, V. C. et al. **Impact of amino acid substitutions at secondary structures in the BRCT domains of the tumor suppressor BRCA1: Implications for clinical annotation.** *J Biol Chem.* 294(15):5980-5992, 2019.

FINDLAY, G. M. et al. **Accurate classification of BRCA1 variants with saturation genome editing.** *Nature.* 562(7726):217-222, 2018.

FOO, T. K. et al. **Compromised BRCA1-PALB2 interaction is associated with breast cancer risk.** *Oncogene.* 36(29):4161-4170, 2017.

GARDINI, A. et al. **Genome-wide analysis reveals a role for BRCA1 and PALB2 in transcriptional co-activation.** *EMBO J.* Apr 16;33(8):890-905, 2014

GLOBOCAN 2018: **Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2018.** Disponível em: <<http://www.globocan.iarc.fr/>>.

GOLUBEVA, V.; NEPOMUCENO, T. C. & MONTEIRO, A. N. A. **Germline Missense Variants in BRCA1: New Trends and Challenges for Clinical Annotation.** *Cancers (Basel).* 11(4):522, 2019.

GOWEN, L. C. et al. **Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities.** *Nat Genet.* 12(2):191-4, 1996.

HALL, J. M. *et al.* **Linkage of early-onset familial breast cancer to chromosome 17q21.** *Science.* 250(4988):1684-9, 1990.

HANAHAH, D. & WEINBERG, R. A. **Hallmarks of cancer: the next generation.** *Cell,* 144(5): 646-74, 2011.

HER, J. *et al.* **Factors forming the BRCA1-A complex orchestrate BRCA1 recruitment to the sites of DNA damage.** *Acta Biochim Biophys Sin (Shanghai).* 48(7):658-64, 2016.

HURST, J. **Pioneering geneticist Mary-Claire King receives the 2014 Lasker~Koshland Special Achievement Award in Medical Science.** *J Clin Invest.* 124(10): 4148–4151, 2014.

Estimativa 2020 de Incidência de Câncer no Brasil. INCA, 2020

IVERSEN, E. S. *et al.* **A computational method to classify variants of uncertain significance using functional assay data with application to BRCA1.** *Cancer Epidemiol Biomarkers Prev.* 20(6):1078-88, 2011.

JIANG, Q. & GREENBERG, R. **Deciphering the BRCA1 Tumor Suppressor Network.** *J Biol Chem.* 290(29):17724-32, 2015.

KAKAROUGKAS, A. & JEGGO, P. A. **DNA DSB repair pathway choice: an orchestrated handover mechanism.** *Br J Radiol,* 87(1035):20130685, 2014.

KALB, R. *et al.* **BRCA1 Is a Histone-H2A-Specific Ubiquitin Ligase.** *Cell Rep.* 8(4): 999–1005, 2014.

KARCZEWSKI, K. J. *et al.* **The mutational constraint spectrum quantified from variation in 141,456 humans.** *Nature.* 581(7809):434-443, 2020.

KELSEY, J. L. **A review of the epidemiology of human breast cancer.** *Epidemiol Rev.*1:74-109, 1979.

KOBAYASHI, H. *et al.* **Hereditary breast and ovarian cancer susceptibility genes (review).** *Oncol Rep.* (3):1019-29, 2013.

LAEMMLI, U. K. **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature.* 227(5259):680-5, 1970.

LANDRUM, M. J. *et al.* ClinVar: **improving access to variant interpretations and supporting evidence.** *Nucleic Acids Res.* 46(D1):D1062-D1067, 2018.

- LEE, M. S. *et al.* **Comprehensive analysis of missense variations in the BRCT domain of BRCA1 by structural and functional assays.** *Cancer Res.* 70(12):4880-90, 2010.
- LINDOR, N. M. *et al.* **A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS).** *Hum Mutat.* 33(1):8-21, 2012.
- LIU, Y & LU, L. **BRCA1 and homologous recombination: implications from mouse embryonic development.** *Cell Biosci.* 10:49, 2020.
- LONGO, P. A. *et al.* **Transient mammalian cell transfection with polyethylenimine (PEI).** *Methods Enzymol.* 529:227-40, 2013.
- LYRA, P. C. M. *et al.* **Integration of functional assay data results provides strong evidence for classification of hundreds of BRCA1 variants of uncertain significance.** *Genet Med.* 1-10, 2020.
- MANKE, I. A. *et al.* **BRCT Repeats As Phosphopeptide-Binding Modules Involved in Protein Targeting.** *Science.* 302(5645):636-639, 2003.
- MERSCH, J. *et al.* **Cancers Associated With BRCA1 and BRCA2 Mutations Other Than Breast and Ovarian.** *Cancer.* 121(2):269-75, 2014.
- MESELSON, M. & YUAN, R. **DNA restriction enzyme from E. coli.** *Nature.* 217(5134):1110-4, 1968.
- MESQUITA, R. F. *et al.* **Tandem BRCT Domains: DNA's Praetorian Guard.** *Genes & Cancer.* 1(11):1140-6, 2010.
- MIKI, Y. *et al.* **A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1.** *Science.* 266 (5182):66-71, 1994.
- MONTEIRO, A.; AUGUST, A. & HANAFUSA, H. **Evidence for a transcriptional activation function of BRCA1 C-terminal region.** *Proc. Natl. Acad. Sci.* 93:13595-9, 1996.
- MONTEIRO, A. N. *et al.* **Variants of uncertain clinical significance in hereditary breast and ovarian cancer genes: best practices in functional analysis for clinical annotation.** *J Med Genet.* 57(8):509-518, 2021.
- MORRIS, J. R. *et al.* **Genetic analysis of BRCA1 ubiquitin ligase activity and its relationship to breast cancer susceptibility.** *Hum Mol Genet.* 15(4):599-606, 2006.

- MULLAN, P.B.; QUINN, J. E. & HARKIN, D. P. ***The role of BRCA1 in transcriptional regulation and cell cycle control.*** *Oncogene*. 25:5854–5863, 2006.
- MULLIS, K. et al. ***Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.*** *Cold Spring Harb Symp Quant Biol*. 51 Pt 1:263-73, 1986.
- NCCN Clinical practice guidelines in oncology.*** Disponível em: <https://www.nccn.org/professionals/physician_gls/default.aspx>. Acesso realizado em: 09/2020.
- NEPOMUCENO, T. N. et al. ***The Role of PALB2 in the DNA Damage Response and Cancer Predisposition.*** *Int J Mol Sci*. 18(9):1886, 2017.
- NEUMANN, E. et al. ***Gene transfer into mouse lymphoma cells by electroporation in high electric fields.*** *EMBO J*. 1(7): 841–845, 1982.
- PARAMESWARAN, B. et al. ***Damage-induced BRCA1 phosphorylation by Chk2 contributes to the timing of end resection.*** *Cell Cycle*. 14(3): 437–448, 2015.
- PARSONS, M. T. ***Large scale multifactorial likelihood quantitative analysis of BRCA1 and BRCA2 variants: An ENIGMA resource to support clinical variant classification.*** *Hum Mutat*. 40(9):1557-1578, 2019.
- PAUL, A. E PAUL, S. ***The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers.*** *Front Biosci*. 19: 605–618, 2015.
- PETITALOT, A. et al. ***Combining Homologous Recombination and Phosphopeptide-binding Data to Predict the Impact of BRCA1 BRCT Variants on Cancer Risk.*** *Mol Cancer Res*. 17(1):54-69, 2019.
- PHELAN, C. et al. ***Classification of BRCA1 missense variants of unknown clinical significance.*** *J Med Genet*. 42:138–146, 2005.
- PRAKASH, R.; et al. ***Homologous recombination and human health: the roles of BRCA1, BRCA2 and associated proteins.*** *Cold Spring Harb Perspect Biol*. 7(4): a016600, 2015.
- PRUTHI, S.; GOSTAOUT, B. S. & LINDOR, N. M. ***Identification and Management of Women With BRCA Mutations or Hereditary Predisposition for Breast and Ovarian Cancer.*** *Mayo Clin Proc*. 85(12):1111-20, 2010.

RANSBURGH, D. J. R. *et al.* **Identification of Breast Tumor Mutations in BRCA1 That Abolish Its Function in Homologous DNA Recombination.** *Cancer Res.* 70(3):988-95, 2010.

REBBECK, T. R. *et al.* **Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer.** *JAMA.* 313(13):1347-61, 2015.

ROJAS, K. & STUCKEY, A. **Breast Cancer Epidemiology and Risk Factors.** *Clin Obstet Gynecol.* 59(4):651-672, 2016.

ROSEN, E. M. & PISHIVAIAN, M. J. **Targeting the BRCA1/2 tumor suppressors.** *Curr Drug Targets.* 15(1):17-31, 2014.

SAVAGE, K. & HARKIN, P. **BRCA1, a 'complex' protein involved in the maintenance of genomic stability.** *FEBS J.* 282(4):630-46, 2015.

SCALIA-WILBUR, J. *et al.* **Breast Cancer Risk Assessment: Moving Beyond BRCA 1 and 2.** *Semin Radiat Oncol,* 26:3-8, 2016.

SCULLY, R. *et al.* **Genetic Analysis of BRCA1 Function in a Defined Tumor Cell Line.** *Mol Cell,* 4:1093–9, 1999.

SHENG, Z.; ZHAO, Y. & HUANG, J. **Functional Evolution of BRCT Domains from Binding DNA to Protein.** *Evol Bioinform Online.,* 7:87–97, 2011.

SHIOVITZ, S. E. & KORDE, L. **Genetics of breast cancer: a topic in evolution.** *Annals of Oncology.* 26: 1291–1299, 2015.

SMITH, T. M. *et al.* **Complete genomic sequence and analysis of 117 kb of human DNA containing the gene BRCA1.** *Genome Res.* 6(11):1029-49, 1996.

STARITA, L. M. *et al.* **Massively Parallel Functional Analysis of BRCA1 RING Domain Variants.** *Genetics.* 200(2):413-22, 2015.

SUN, Y.S. *et al.* **Risk Factors and Preventions of Breast Cancer.** *Int J Biol Sci.* 13(11):1387-1397, 2017.

SUZUKI, A. *et al.* **Brca2 is required for embryonic cellular proliferation in the mouse.** *Genes Dev.* 11(10):1242-52, 1997.

SZABO, C. *et al.* **The Breast Cancer Information Core: Database Design, Structure, and Scope.** *Hum Mutat.* 16:123-131, 2000.

TAVTIGIAN, S. V. *et al.* **Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral.** *J Med Genet.* 43(4):295-305, 2006.

THOUVENOT, P. *et al.* **Yeast cells reveal the misfolding and the cellular mislocalization of the human BRCA1 protein.** *J Cell Sci.* 129(23):4366-4378, 2016.

TUBBS, A. & NUSSENZWEIG, A. **Endogenous DNA Damage as a Source of Genomic Instability in Cancer.** *Cell.* 168(4):644-656, 2017.

TURK, A. A. & WISINSKI, K. B. **PARP inhibitors in breast cancer: Bringing synthetic lethality to the bedside.** *Cancer.* 124(12):2498-2506, 2018.

VALLON-CHRISTERSSON, J. *et al.* **Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families.** *Hum Mol Genet.* 10(4):353-60, 2001.

YI, Y.; KANG, H. & BAE, I. **BRCA1 and Oxidative Stress.** *Cancers,* 6(2): 771–795, 2014.

YOSHINO, Y. *et al.* **Evaluation of site-specific homologous recombination activity of BRCA1 by direct quantitation of gene editing efficiency.** *Sci Rep.* 1644, 2019.

YU, Y. *et al.* **RBBP8/CtIP suppresses P21 expression by interacting with CtBP and BRCA1 in gastric cancer.** *Oncogene.* 39(6):1273-1289, 2020.

WALLACE, A. J. **New challenges for BRCA testing: a view from the diagnostic laboratory.** *Eur J Hum Genet.* 24 Suppl 1(Suppl 1):S10-8, 2016.

WANG, B. **BRCA1 tumor suppressor network: focusing on its tail.** *Cell & Biosci.* 2:6, 2012.

WILLIAMS, R. S. *et al.* **Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1.** *Nat Struct Mol Biol.* 11(6):519-25, 2004.

WINTERS, Z. E. *et al.* **Systematic review of the impact of breast-conserving surgery on cancer outcomes of multiple ipsilateral breast cancers.** *BJS Open.* 2(4): 162–174, 2018.

WOODS, N.; *et al.* **Functional assays provide a robust tool for the clinical annotation of genetic variants of uncertain significance.** *npj Genomic Med* 1, 16001, 2016.

WOOSTER, R. *et al.* **Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.** *Science*. 265(5181):2088-90, 1994.

WU-BAER, F. *et al.* **The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin.** *J Biol Chem*. 278(37):34743–6. 2003.

Anexo I

Lista de variantes indels que respeitam o quadro de leitura na região codificante de BRCA1

Variante Nucleotídica	Variante Proteica
c.5218_5220delGTGinsAAC	p.(Val1740Asn)
c.3392_3393delATinsTA	p.(Asp1131Val)
c.5581_5582insCCCCCCCCA	p.(His1860_Ser1861insThrProPro)
c.5218_5220delGTGinsAAC	p.(Val1740Asn)
c.4614_4615delGCinsTT	p.(Gln1538His)
c.2611_2612delCCinsGT	p.(Pro871Val)
c.5219_5224delTGGTCA	p.(Val1740_Asn1742delinsAsp)
c.1763_1764delGCinsAA	p.(Ser588Lys)
c.1846_1848delTCT	p.(Ser616del)
c.5218_5220delGTGinsAAC	p.(Val1740Asn)
c.5092_5093delGAinsAG	p.(Glu1698Arg)
c.2227_2229delAAT	p.(Asn743del)
c.1446_1448delTAT	p.(Ile483del)
c.4807_4821delCCCCAATTGAAAGTT	p.(Pro1603_Val1607del)
c.1106_1108delATG	p.(Asp369del)
c.2211_2213delAGT	p.(Val738del)
c.1200_1206delTGGGGAGinsCTCACATGAACTGTTAGGT	p.(Gly401_Glu402delinsSerHisGluLeuLeuGly)
c.5275_5276delAAinsTG	p.(Lys1759Trp)
c.3772_3774delGAG	p.(Glu1258del)
c.5529_5530delACinsCA	p.(Leu1844Ile)
c.5017_5019delCAC	p.(His1673del)
c.3418_3420delAGT	p.(Ser1140del)
c.1763_1764delGCinsAA	p.(Ser588Lys)
c.5279_5332delTCTTCAGGGGGCTAGAAATCTGTTGCTATGGGCCCTTCACC AACATGCCACAG	p.(Ile1760_Asp1778delinsAsn)
c.1931_1933delGTT	p.(Cys644del)
c.4676_4678delAGG	p.(Glu1559del)
c.5161_5163delCAG	p.(Gln1721del)
c.4579_4580delGAinsAT	p.(Glu1527Met)
c.4614_4615delGCinsTT	p.(Gln1538His)
c.2611_2612delCCinsGT	p.(Pro871Val)
c.4183_4185delCAG	p.(Gln1396del)

c.5078_5080delCTG	p.(Ala1693del)
c.5077_5080delGCTGinsTTGATTCTGC	p.(Ala1693_Glu1694delins LeulleLeuGln)
c.2155_2163delAAAGAATTT	p.(Lys719_Phe721del)
c.1637_1685delTGAATATTACTAATAGTGGTCATGAGAATAAAACAAAAGGTG ATTCTATinsGAAA	p.(Met546_Ile562delinsArg Asn)
c.3481_3492delGAAGATACTAGT	p.(Glu1161_Ser1164del)
c.5219_5224delTGGTCA	p.(Val1740_Asn1742delins Asp)
c.5359_5363delTGTGGinsAGTGA	p.(Cys1787_Gly1788delins SerAsp)
c.1824_1826delGAA	p.(Lys608del)
c.4510_4512delTTA	p.(Leu1504del)
c.5077_5080delGCTGinsTTCATTCTGC	p.(Ala1693_Glu1694delins PhelleLeuGln)
c.4187_4189delAGA	p.(Gln1396del)
c.3234_3236delAAAinsTAT	p.(Lys1079Ile)
c.5075_5152delATGCTGAGTTTGTGTGTGAACGGACACTGAAATATTTTCTAG GAATTGCGGGAGGAAAATGGGTAGTTAGCTATTTCT	p.(Asp1692_Trp1718delins Gly)
c.1769_1771delGTA	p.(Ser590del)
c.3327_3329delAAA	p.(Lys1110del)
c.1763_1764delGCinsTT	p.(Ser588Ile)
c.3782_3784delTAT	p.(Leu1261del)
c.2191_2196delAAAGAA	p.(Lys731_Glu732del)
c.3664_3666delGAG	p.(Glu1222del)
c.3841_3843delCAG	p.(Gln1281del)
c.5580_5581insCCCCCCCC	p.(His1860_Ser1861insPro ProPro)
c.2638_2639delGAinsAC	p.(Glu880Thr)
c.4032_4034delTGA	p.(Asp1344del)
c.5571_5579delGATCCCCCA	p.(Gln1857_Pro1859del)
c.2834_2836delGTA	p.(Ser945del)
c.2483_2485delGCT	p.(Gly828_Phe829delinsV al)
c.2716_2730delAAGGAAGAAAATCAA	p.(Lys906_Gln910del)
c.4063_4065delAAT	p.(Asn1355del)
c.1175_1216delTGTTAGTTCTGATGACTCACATGATGGGGAGTCTGAATCA A	p.(Leu392_Asn406delinsHi s)
c.1983_1991delGCACAGCAG	p.(His662_Arg664del)
c.5565_5573delACCCCAGAT	p.(Gln1857_Pro1859del)

c.4036_4038delGAA	p.(Glu1346del)
c.2068_2082delAAAAGACATGACAGC	p.(Lys690_Ser694del)
c.2235_2236delAGinsCT	p.(Glu745_Asp746delinsAspTyr)
c.5321_5322delACinsGG	p.(Asn1774Arg)
c.3877_3879delGCT	p.(Ala1293del)
c.1504_1518delTTAAAGCGTAAAAGG	p.(Leu502_Arg506del)
c.5234_5235delACinsGG	p.(Asn1745Arg)
c.5425_5430delGTTGTG	p.(Val1809_Val1810del)
c.3344_3346delAAG	p.(Glu1115del)
c.5181_5183delAAT	p.(Met1728del)
c.2018_2023delAACCTG	p.(Glu673_Pro674del)