Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia

VANESSA CÂMARA FERNANDES

Estudo funcional da região não estruturada de ligação entre os domínios BRCT em tandem da proteína BRCA1

Orientadores: Prof. Dr. Guilherme Suarez-Kurtz
Prof. Dr. Marcelo Alex de Carvalho

RIO DE JANEIRO
2016

Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia

VANESSA CÂMARA FERNANDES

Estudo funcional da região não estruturada de ligação entre os domínios BRCT em tandem da proteína BRCA1

Dissertação apresentada ao Instituto Nacional de Câncer como parte dos requisitos para obtenção do título de Mestre em Oncologia

Orientadores: Prof. Dr. Guilherme Suarez-Kurtz
Prof. Dr. Marcelo Alex de Carvalho

RIO DE JANEIRO
2016

F363e Fernandes, Vanessa Câmara.

Estudo funcional da região não estruturada de ligação entre os domínios BRCT em tandem da proteína BRCA1 / Vanessa Câmara Fernandes. – Rio de Janeiro, 2016.
86 f.: il. color.

Dissertação (Mestrado em Oncologia) – Instituto Nacional de Câncer José Alencar Gomes da Silva, 2016.

Orientador: Guilherme Suarez-Kurtz; Marcelo Alex de Carvalho.

1. Genes BRCA1. 2. Variação Estrutural do Genoma. 3. Neoplasias da Mama. I. Suarez-Curtz, Guilherme (Orient.). II. Carvalho, Marcelo Alex de (Orient.). III. Instituto Nacional de Câncer José Alencar Gomes da Silva. IV. Título.

CDD 611.0181663

Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia

VANESSA CÂMARA FERNANDES

Estudo funcional da região não estruturada de ligação entre os domínios BRCT em tandem da proteína BRCA1

ORIENTADORES: Prof. Dr. Guilherme Suarez-Kurtz
Prof. Dr. Marcelo Alex de Carvalho

Aprovada em: _____/_____/_____

EXAMINADORES:

Prof. Dr. Miguel Angelo Martins Moreira - Presidente
Prof. Dr. Turán Peter Ürményi
Prof. Dr. Francisco Meirelles Bastos de Oliveira
Profª. Drª. Sheila Coelho Soares Lima
Profª. Drª. Tuane Cristine Ramos Gonçalves Vieira

RIO DE JANEIRO
2016

Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER

Estudo funcional da região não estruturada de ligação entre os domínios BRCT em tandem da proteína BRCA1

DISSERTAÇÃO DE MESTRADO

Vanessa Câmara Fernandes

RESUMO

As vias de reparo ao dano de DNA (RDD) são essenciais para a manutenção da integridade genômica. Mutações em genes que atuam nessas vias podem contribuir para o desenvolvimento de câncer. *BRCA1* é um gene extremamente polimórfico, codificando uma proteína envolvida em importantes eventos de sinalização nuclear em resposta ao dano de DNA. Portadores de mutações em *BRCA1* que levam a uma proteína disfuncional podem ter o risco aumentado de desenvolvimento de câncer de mama e ovário em até 80%. Alterações em *BRCA1,* tais como inserções, deleções e mutações nonsense, podem ser inferidas como patogênicas, uma vez que a perda dos últimos 11 resíduos de aminoácidos na região C-terminal resulta em uma proteína não funcional. Mutações do tipo missense, assim como inserções e deleções que mantém o quadro de leitura, apresentam dificuldades na classificação. Estudos funcionais representam uma importante ferramenta para a classificação da patogenicidade dos variantes com baixa frequência na população (ou ainda não identificados). Em 2007, nosso grupo estabeleceu um ensaio funcional que correlaciona a atividade transcricional de BRCA1 e a integridade da região C-terminal da proteína (ensaio de TA) – essa região (aminoácidos 1396 a 1863) é capaz de recrutar a holoenzima RNA polimerase II. Assim, é possível correlacionar a atividade de um sistema repórter com a patogenicidade de uma mutação específica. Este estudo tem foco na região de ligação entre os domínios tBRCT, denominada *linker* (aminoácidos 1737 a 1759). Foram preditas todas as mutações naturais missense no *linker* (132 variantes) e, utilizando estratégias de bioinformática (Align-GVGD, SFIT e PolyPhen), foram estabelecidas os níveis de associação ao câncer. Variantes previamente estudados/classificados em outros trabalhos (16) não foram incluídos, resultando em um conjunto de 116 variantes. Somente 23 variantes foram preditos como associados ao câncer por todos algoritmos. Os variantes selecionados para o estudo foram gerados e clonados no vetor pCDNA3, codificando uma proteína de fusão (domínio de ligação ao DNA de GAL4 fusionado a BRCA1 C-terminal). As construções foram utilizadas para obtenção de dados funcionais através do ensaio de TA. Os experimentos foram originalmente conduzidos a 37ºC. A maioria dos variantes (20) apresentou um perfil patogênico, enquanto 3 varinate, todos na posição 1740, apresentaram perfil semelhante aos controles não patogênicos. Esses variantes demonstraram uma variação significativa na atividade transcricional (tanto entre replicatas, quanto experimentos independentes), sugerindo um comportamento termossensível. Assim, todos os variantes foram testados a 30ºC. Os variantes na posição V1740 mantiveram um comportamento não patogênico, no entanto 9 variantes apresentaram aumento estatisticamente significativo na atividade transcricional. Não foram observadas alterações significativas na estrutura secundária do linker como consequência de substituições de aminoácidos quando analisados através do diagrama de Ramachandran. As diferenças observadas a 37ºC e 30ºC sugerem que o *linker* é, possivelmente, um segmento termossensível na estrutura de BRCA1. Todos os dados gerados serão avaliados em modelo matemático (VarCall) para determinar a patogenicidade dos variantes de *BRCA1* abordados nesse estudo.

Palavras-chave: 1. BRCA1; 2. Variantes; 3. Avaliação funcional; 4. Câncer de mama.

Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER

Functional study of BRCA1 non-structured linker region between tandem BRCT domains

DISSERTAÇÃO DE MESTRADO

Vanessa Câmara Fernandes

ABSTRACT

DNA damage repair is crucial for the maintenance of genome integrity and cellular processes. Mutations in genes involved in this pathway may severely contribute to cancer development. *BRCA1* is a highly polymorphic gene and a multi-faceted protein important for nuclear signaling events in response to DNA damage. *BRCA1* germline mutations that lead to a dysfunctional protein can increase the risk of breast and ovarian cancer up to 80%. Abnormalities in *BRCA1* such as deletions, insertions and nonsense mutations might be deduced as pathogenic, since the deletion of only 11 amino acids in the C-terminal results in a compromised protein. Whilst missense mutations, as in frame deletions and insertions, present difficulties in classification. Functional assays represent an important tool to contribute to pathogenicity classification of low frequency (and not yet reported) variants. In 2007, our group established a functional assay that correlates BRCA1 transcriptional activity and its C-terminus integrity (TA assay) - this region (1396-1863 aa) displays the ability to recruit the RNA polymerase II holoenzyme. Therefore, we may correlate the activity of a reporter system with the pathogenicity of a specific mutation. In this work, we focused on the region that links the two tandem BRCT domains (the linker region, 1737-1759 aa). We predicted all possible natural missense mutations in the linker region (132 variants); using bioinformatics strategies (Align-GVGD, SFIT and PolyPhen) mutations were scored for cancer association. Variants already studied/classified (16) were not included, leading to an initial set of 116 variants. Only 23 variants were predicted as cancer associated by all algorithms. We generated and cloned the 23 variants predicted as pathogenic into pCDNA3 vector, coding a fusion protein (GAL4 DNA binding domain fused to BRCA1 C-terminus). The constructs were used to assess functional data through the TA assay. Experiments were originally conducted at 37ºC. Most variants (20) behaved similar to pathogenic controls, whilst 3 variants (all in 1740 position) appear to behave as non-pathogenic controls. Variants in 1740 position demonstrated a significant variation in transcription activity (in replicates and in independent experiments), featuring a possible thermosensitive property. Therefore, all variants were tested at 30ºC. The 1740 position variants sustained a non-pathogenic behavior, however other 9 variants displayed a statistically significant increase in transcriptional activity. We did not observed any significant change in the linker secondary structure due to amino acid change using the Ramachandran diagram approach. The differences in transcription activity observed at 37ºC and 30ºC suggest that the linker region is possibly a thermosensitive segment in BRCA1 structure. The generated data will be evaluated in a mathematical model (the VarCall algorithm) to proper assess the pathogenicity classification of this set of *BRCA1* variants.

Key words: 1. BRCA1; 2. Variants; 3. Functional Assay; 4. Breast cancer.

Agradecimentos

Em primeiro lugar sempre tenho que agradecer a Deus, pois graças a Ele consegui passar por tanta coisa e foi maravilhoso. Realmente eu fiquei muito feliz com o meu mestrado, não termino com o sentimento de até que enfim, termino com o sentimento de felicidade e esperança do que ainda vem pela frente.

Tenho que agradecer e muito a minha família. Meus pais pelo amor incondicional, pela preocupação (de todas as vezes eu saia tarde do INCA) e por me darem todo o suporte até aqui, sem vocês eu jamais conseguiria, também agradeço a minha irmã. Meu cunhado pela pessoa maravilhosa que ele é e por sempre estar disposto a me ajudar. E, talvez, o que mais “sofreu” com meu trabalho, meu namorado, Guilherme. Obrigada e desculpa, por todos “já estou chegando” (e “meia” hora depois eu chego), por todos os compromissos que eu desmarquei porque perdi a hora no lab, por todas as vezes que eu simplesmente queria falar até não poder mais e você só escutou (a gente sabe que não foram poucas) e por ter parado de perguntar até quando vai essa loucura (porque eu e você sabemos que essa loucura não vai ter fim). Muito obrigada!

Agradeço ao meu orientador Guilherme Suarez-Kurtz por me dar a chance de fazer o mestrado no INCA e por confiar em mim para desenvolver essa pesquisa.

Agradeço ao meu coorientador Marcelo Alex de Carvalho que confiou em mim e que não duvidou em nenhum momento que eu poderia concluir esse trabalho. Obrigada, por mais uma vez estar do meu lado e por continuar me ensinando tanto! E obrigada também por ainda não ter me matado! Lembra sempre: feche os olhos e conta até 10 (as vezes até 11).

Agradeço aos meus amigos de grupo, Thales, por esses 10 anos que me atura e por sempre estar ao meu lado, Giuliana, por todas as vezes que a gente ficou até tarde cantando no lab, João, por todas as risadas que você pode proporcionar a alguém (principalmente quando os gráficos possuem cores) Renata, pelas vezes (vezes? Ou sempre?) falar sem parar, Thiago, por todo o chocolate que contrabandeou para mim, Dionathan, por me divertir conversando sobre absolutamente tudo! E aqui também agradeço a nossa agregada, Babi. Pode ser agregada ao grupo, mas pra mim minha amiga, que me fez companhia várias vezes quando eu ia almoçar, comer besteira ou tomar café! E que sempre que precisava falar, reclamar, xingar ou rir estava lá! Aos meus amigos de andar Mateus, Ana Ribeiro, Ana Santoro e Priscila pelos conselhos cafés e risadas. Meu grupo é uma família para mim, as longas horas no lab se tornam menos dolorosas quando você acaba tomando 3 cafés por dia e rindo sempre que alguém fala uma bobagem (que 50% do tempo é de mim que vocês riem).

Não posso deixar de agradecer a uma pessoa muito especial que passou pela minha dissertação, minha “aluna” Ricceli. Obrigada por me ensinar de novo o quanto a ciência é linda e o quanto nós podemos (e devemos) nos surpreender e maravilhar com ela a cada dia, porque quando ensinamos assim formamos pessoas que não se cansam tão facilmente, e ao contrário do que tudo indicava essa fascinação e alegria é sim contagiante.

Agradeço aos amigos que a vida cientifica colocou no meu caminho. Renato, meu eterno coorientador, porque quando eu tenho problemas quem será que eu procuro? Até porque só você entende meu humor negro. Felipe, Maira e Juliene, meus eternos pupilos e minha eterna mestre, fico feliz de compartilhar esse momento com vocês também, afinal pra quem eu corro quando eu preciso beber pra esquecer que o experimento deu errado? Meus amigos do laboratório de genética molecular Stephanie, Michele, Cefas, nem todos ainda estão lá, mas a amizade vai longe, não é?

Tenho que agradecer ao pessoal do SPC, que sempre me dá bom dia super animado e me aturam nos dias que estou elétrica ou de mal humor. A Vera que sempre prepara tudo que está faltando com urgência e sempre resolve os problemas do lab rapidinho e entende quando a gente tem pressa. Ao Fabiano que quebrou MUITOS galhos nesses dois anos, autoclavando material correndo porque um contaminou e precisava fazer um experimento!

Eu tenho que agradecer pela vida que eu tive por trás da dissertação porque ela que me deu sanidade para continuar com a pesquisa. Eu agradeço em especial aos meus amigos Padre Cássio e Padre Robson por sempre estarem prontos a me ajudar! Aos meus amigos Fernanda, Caique, Cintia, Eloa, por aturarem as vezes que eu simplesmente queria falar o que eu fazia e vocês mesmo não entendendo nada ficavam quietos. Aos meus amigos eternos Diogo, Rafael, Raphael, Elcio, Felipe, Thaisa, Gabi, Heric, Renan pela diversão nesses anos, e em especial Haryadylla que todo o tempo sempre me perguntou como eu estava e se eu precisava de ajuda com alguma coisa da dissertação! Aos meus amigos Eric, Fernanda, Eduardo, Carol (meu sobrinho Murilo) e Marcel, por fazerem com que eu desestressasse muitas vezes, curtisse a vida e desligasse da bancada. Muito obrigada a todas essas pessoas, sem a amizade de vocês eu não chegaria até aqui.

Por último queria agradecer ao Carlos e a Maria Carolina, que se tornaram meus amigos nesses dois anos. A Solange Rocha, que passou várias tardes conversando comigo e divertindo a Giuliana e o Thales.

Me desculpem pelo agradecimento gigante, mas eu aprendi, e agora entendi, que por trás de toda dissertação/tese existe uma vida, e como minha pesquisa é muito da minha vida eu tenho que agradecer as pessoas que fizeram esses dois anos possíveis.

“Se esperar pelo momento em que tudo, absolutamente tudo, estiver pronto, talvez nunca comecemos”

Ivan Turgenev

ÍNDICE

1. Introdução 1
1.1. Câncer 1
1.2. Câncer de mama.....1
1.3. BRCA1 4
1.3.1. O papel de BRCA1mediadas pelos domínios BRCT em tandem.....7
1.3.1.1. Interação com BARD1.....7
1.3.1.2. Recombinação homóloga e recombinação não homóloga por junção de extremidades.....7
1.3.1.3. Reparo ao dano de DNA (outras vias).....11
1.3.1.4. Transcrição 12
1.4. Domínios BRCT de BRCA1.....12
1.5. Mutações em BRCA1 e a susceptibilidade ao câncer 15
1.6. Ensaios funcionais para classificação de variantes de BRCA1.....21
1.7. Ensaio de ativação transcricional (*transcriptional assay* – TA).....22
2. Objetivos 24
2.1. Objetivo geral 24
2.2. Objetivos específicos 24
3. Metodologia 25
3.1. Determinação dos variantes naturais missense na região *linker* de BRCA1.....25
3.2. Avaliação dos variantes missense através de modelos de predição *in silico*.....25
3.3. Identificação dos variantes previamente estudados.....25
3.4. Critérios de seleção dos variantes para o estudo.....25
3.5. Geração de construções.....26
3.5.1. Geração dos variantes.....26
3.5.2. Clonagem no vetor pPCR-TOPO.....30
3.5.3. Subclonagem no vetor de expressão pGBT9.....32
3.5.4. Subclonagem no vetor de expressão pCDNA3.....34
3.6. Indução de competência bacteriana.....36
3.7. Transformação bacteriana.....36
3.8. Extração plasmidial.....36
3.9. Sequenciamento 37

3.10. Cultura de células humanas.....37
3.11. Transfecção de células humanas.....37
3.12. Ensaio funcional em células humanas.....40
3.12.1. Ensaio funcional à 37ºC.....40
3.12.2. Ensaio funcional à 30ºC.....41
3.13. Extratos proteicos.....41
3.14. Avaliação de síntese da proteína quimérica.....41
3.15. Diagramas de Ramachandran.....42
3.16. Análises Estatísticas.....42
4. Resultados 43
4.1. Seleção dos variantes.....43
4.2. Geração de variantes e construções plasmidiais.....45
4.3. Avaliação da capacidade de ativação transcricional a 37ºC.....45
4.4. Avaliação da capacidade de ativação transcricional a 30ºC.....50
4.5. Síntese da proteína quimérica.....54
4.6. Análise *in silico* do impacto das substituições na região *linker*.....56
5. Discussão 65
6. Conclusão 75
7. Referências 76
8. Anexos 85

LISTA DE TABELAS

Tabela 3.1 – Oligonucleotídeos iniciadores (*primers*) utilizados na geração de variantes para o estudo.....29

Tabela 4.1 – Análises *in silico* dos variantes naturais localizados na região *linker* de BRCA1 selecionados para o estudo.....44

Tabela 5.1 – Classificação preliminar dos variantes 70

Tabela A1 – Análises *in silico* dos variantes naturais localizados na região *linker* de BRCA1.....86

LISTA DE FIGURAS

Figura 1.1– Distribuição dos tipos de câncer mais incidentes no mundo.....3
Figura 1.2– Distribuição dos tipos de câncer mais incidentes no Brasil.....3
Figura 1.3– Estrutura da proteína BRCA1.....6
Figura 1.4– Complexos formados por BRCA1.....10
Figura 1.5– Esquema simplificado da via de recombinação homóloga.....10
Figura 1.6– Estrutural tridimensional do domínio tBRCT de BRCA1.....14
Figura 1.7– Frequência de variantes de *BRCA1*.....16
Figura 1.8– Representação de um modelo multifatorial para classificação de VNC.....20
Figura 1.9– Esquema do ensaio de ativação transcricional para BRCA1.....23
Figura 3.1– Esquema da rotina de mutagênese sítio-dirigida por sobreposição de fragmentos.....28
Figura 3.2– Representação da construção pPCR-TOPO-BRCA1$_{13/24}$ 31
Figura 3.3– Representação da construção de pGBT9-GAL4DBD:BRCA1$_{13/24}$.33
Figura 3.4– Representação da construção de pCDNA3-BRCA1$_{13/24}$ 35
Figura 3.5– Fluxograma de rotinas realizadas para geração e clonagem das construções utilizadas neste estudo.....35
Figura 3.6– Representação dos plasmídeos pG5*luc* e pGR-TK.....39
Figura 4.1– Esquema da proteína quimérica GAL4DBD:BRCA1$_{13/24}$ e localização dos variantes usados como controles no estudo.....46
Figura 4.2– Análise da atividade de ativação transcricional a 37ºC.....48
Figura 4.3– Análise da atividade de ativação transcricional a 37ºC dos variantes na posição 1740.....49
Figura 4.4– Análise da atividade de ativação transcricional à 30ºC.....51
Figura 4.5 – Comparação da atividade de ativação transcricional a 30ºC e 37ºC.....53
Figura 4.6– Proteínas de fusão GAL4DBD:BRCA1$_{13/24}$ produzidas a 37ºC.....55
Figura 4.7– Proteínas de fusão GAL4DBD:BRCA1$_{13/24}$ produzidas a 30ºC variantes V1740E e H1746D.....55
Figura 4.8– Estrutura tridimensional dos BRCTs.....57
Figura 4.9– Delimitação de áreas do diagrama de Ramachandran.....58

Figura 4.10– Distribuição dos resíduos de aminoácidos do *linker* no gráfico de Ramachandran.....60
Figura 4.11– Perfil de distribuição dos resíduos de aminoácidos no gráfico de Ramachandran.....61
Figura 4.12– Perfil de distribuição dos resíduos de aminoácidos no gráfico de Ramachandran.....62
Figura 4.13– Perfil de distribuição dos resíduos aminoácidos no gráfico de Ramachandran.....63
Figura 4.14– Perfil de distribuição dos resíduos de aminoácidos no gráfico de Ramachandran.....64

LISTA DE ABREVIATURAS

BARD1 – BRCA1- associated ring domain 1

BER – Reparo por excisão de base

BRCA1 – Breast cancer 1, early onset

BRCA2 – Breast cancer 2, early onset

BRCT – BRCA1 C terminus

DNA – Desoxirribonucleic acid

Gly – Glicina

HR – Recombinação homóloga

IARC – Internationl Angency for Research on Cancer

INCA – Instituto Nacional de Câncer

Lys – Lisina

NER – Reparo por excisão de nucleotídeo

NHEJ – Recombinação não homóloga

NIH – National Institute of Health

RDD – Reparo ao Dano de DNA

REα – Receptor de estrogênio α

ROS – Espécies reativas de oxigênio

Ser – Serina

SHCMO – Síndrome hereditária de câncer de mama e ovário

tBRCT – Domínio BRCT em tandem

VNC – Variantes não classificados

WHO – World Health Organization

1. INTRODUÇÃO

1.1. Câncer

Responsável por mais de 8 milhões de mortes, em 2012, e com uma projeção de 17 milhões de mortes em 2030, o câncer se tornou um problema de saúde pública mundial. O crescimento do número de casos de câncer ocorre em razão do aumento global da expectativa de vida, da melhora no diagnóstico e do aumento da exposição à agentes etiológicos, como o cigarro.

A expansão do conhecimento sobre essa doença é essencial para o desenvolvimento de novas terapias, além de auxiliar na construção de políticas para o controle e de prevenção (IARC, 2008; INCA, 2016; WHO, 2016).

No Brasil, para o ano de 2016, foram estimados 600 mil novos casos de câncer, incluindo os casos de câncer pele não melanoma. Desconsiderando esse tipo de câncer, a forma mais incidente em homens é o câncer de próstata (61 mil) e em mulheres o câncer de mama (58 mil) (INCA, 2016).

Atualmente, entende-se o câncer como um processo que compreende múltiplas etapas. Essas etapas refletem alterações genéticas que progridem até a transformação de células normais em células neoplásicas (HANAHAN & WEINBERG, 2011).

1.2. Câncer de mama

O câncer de mama é o mais comum entre as mulheres no Brasil (INCA, 2016) e no mundo (WHO, 2016), apresentando elevada taxa de mortalidade (**Fig. 1.1** e **Fig. 1.2**). O câncer de mama é multifatorial e está intimamente relacionado com o processo de urbanização da sociedade, afetando mais agudamente a parcela da população composta por mulheres de níveis socioeconômicos mais elevados (KEY *et al*, 2001; INCA, 2016).

Os fatores etiológicos envolvidos no desenvolvimento da doença podem ser divididos em dois grupos: os individuais e os ambientais. Entre os fatores ambientais pode-se destacar a exposição à radiação ionizante, mesmo que em baixas doses, que aumenta o risco de desenvolvimento de câncer, não especificamente o câncer de mama. Esse risco é ainda maior caso o indivíduo esteja na puberdade, pois é o período no qual ocorre o desenvolvimento dos seios (KEY *et al*, 2001).

Já a lista de fatores individuais é ampla e variável, porém, possivelmente, o mais importe é a idade. É relativamente raro o desenvolvimento de tumores antes dos 35 anos, mas observa-se um aumento drástico do risco entre os 35 e os 50 anos. Após essa idade risco continua elevado, mas a progressão é menos acentuada (PRUTHI, 2010; INCA, 2016).

Outros fatores individuais incluem o padrão de dieta (alimentos muito calóricos, ricos em proteínas e gorduras de origem animal, carne vermelha e processada, além do consumo de álcool, aumentam o risco) e a vida reprodutiva da mulher - onde o risco é diretamente proporcional ao número de ciclos ovarianos. A menarca precoce, nuliparidade, primeira gestação após os 30 anos, poucos filhos, uso de contraceptivo orais, menopausa tardia e terapia de reposição hormonal também elevam o risco do desenvolvimento de câncer de mama. Já a amamentação e exercícios regulares possuem efeito protetor, diminuindo o risco (KEY *et al*, 2001; PRUTHI, 2010; INCA, 2016; GLOBAL CAN, 2016; IARC, 2016).

Estima-se que a hereditariedade seja responsável por cerca de 10% dos casos de câncer de mama. Mulheres com histórico familiar apresentam risco elevado e em determinados casos pode chegar a até 85%, frente ao risco estimado da população em geral de 13% (PRUTHI, 2010; ENGEL & FISCHER, 2015).

Na década de 90, estudos baseados em famílias com histórico de câncer de mama e, portanto, com risco elevado para essa forma de câncer, foram determinantes para estabelecer uma relação causal entre o câncer de mama e a hereditariedade. Em 1990, Hall e colaboradores estudaram 23 famílias com histórico de câncer familiar e estabeleceram a ligação entre o câncer de mama e um locus gênico presente no cromossomo 17. Em outro estudo envolvendo 214 famílias, Easton e colaboradores (1993), confirmaram a associação do cromossomo 17 e o câncer, determinando a existência de um gene que conferia essa susceptibilidade. Somente em 1994, Miki e colaboradores mapearam e clonaram o primeiro gene de predisposição hereditária ao câncer de mama, *BRCA1* (***Br****east* ***Ca****ncer* ***1***).

Ainda na década de 90, um segundo gene foi relacionado à predisposição hereditária ao câncer de mama, *BRCA2* (***Br****east* ***Ca****ncer* ***2***) (WOOSTER *et al*, 1994).

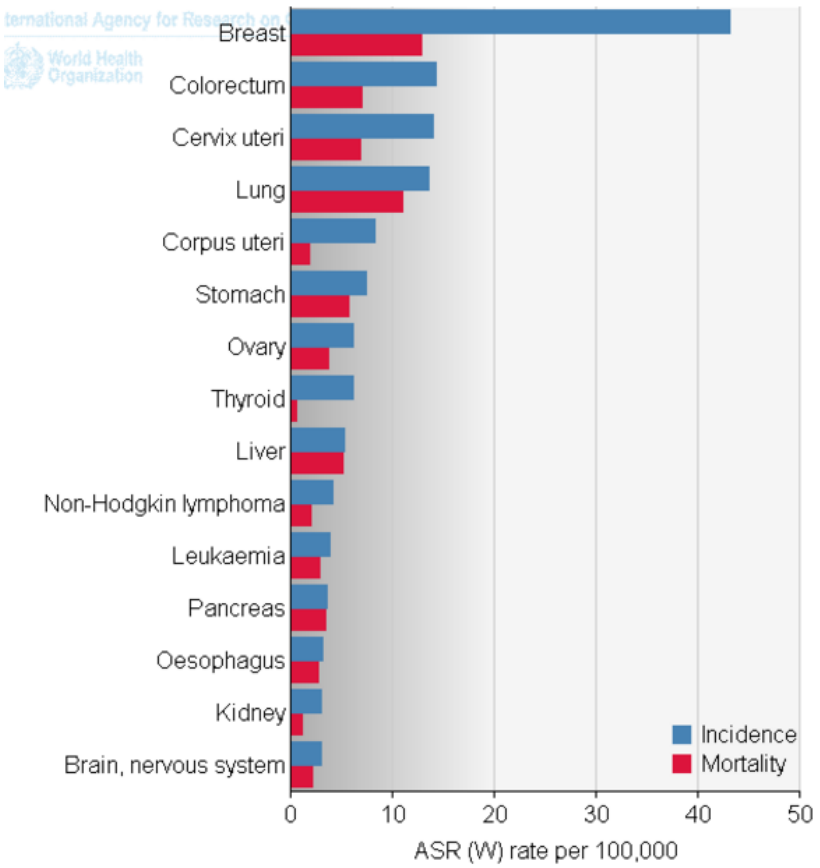


Figura 1.1 – Distribuição dos tipos de câncer mais incidentes no mundo. Estimativa das taxas de incidência e mortalidade globais de câncer para o ano de 2012 para mulheres*. Taxas representadas à cada 100.000 mulheres. Adaptado de WHO, 2016.

(*) exceto câncer de pele não melanoma.

Localização Primária	Casos	%	Homens	Mulheres	Localização Primária	Casos	%
Próstata	61.200	28,6%			Mama feminina	57.960	28,1%
Traqueia, Brônquio e Pulmão	17.330	8,1%			Cólon e Reto	17.620	8,6%
Cólon e Reto	16.660	7,8%			Colo do útero	16.340	7,9%
Estômago	12.920	6,0%			Traqueia, Brônquio e Pulmão	10.890	5,3%
Cavidade Oral	11.140	5,2%			Estômago	7.600	3,7%
Esôfago	7.950	3,7%			Corpo do útero	6.950	3,4%
Bexiga	7.200	3,4%			Ovário	6.150	3,0%
Laringe	6.360	3,0%			Glândula Tireoide	5.870	2,9%
Leucemias	5.540	2,6%			Linfoma não Hodgkin	5.030	2,4%
Sistema Nervoso Central	5.440	2,5%			Sistema Nervoso Central	4.830	2,3%

Figura 1.2 – Distribuição dos tipos de câncer mais incidentes no Brasil. Distribuição dos dez tipos de câncer mais incidentes para homens e para mulheres em 2016*. Adaptado de INCA, 2016.

(*) exceto câncer de pele não melanoma

Mutações em células na linhagem germinativa em *BRCA1* e *BRCA2* são as principais causas da síndrome hereditária de câncer de mama e ovário (SHCMO), responsáveis por até 30% dos casos (ECONOMOPOULOU *et al*, 2015; SCALIA-WILBUR *et al*, 2016). Essa síndrome é uma doença autossômica dominante caracterizada pelo aparecimento de câncer precoce nos indivíduos com histórico familiar - especificamente, parentes de primeiro e segundo graus afetados. Disfunções associadas a *BRCA1* podem aumentar em até 80% o risco de câncer de mama, enquanto que em *BRCA2* elevam o risco em até 60% (CLARK *et al*, 2012; KOBAYASHI *et al*, 2013; SHIOVITZ & KORDE, 2015).

Mutações em outros genes, além de *BRCA1* e *BRCA2*, podem contribuir para aumentar o risco de desenvolvimento de câncer de mama – por exemplo, *ATM*, *CHK2*, *PALB2*, *TP53*, *SKT11* e *PTEN*, no entanto, esses genes estão envolvidos em outras síndromes que não a SHCMO, e não contribuem exclusivamente para o risco em câncer de mama, mas também para outros tipos de câncer. (KOBAYASHI *et al*, 2013; MERSCH *et al*, 2014; ECONOMOPOULOU *et al*, 2015; SCALIA-WILBUR *et al*, 2016).

1.3. BRCA1

O gene *BRCA1* tem mais de 80 kb, se localiza no cromossomo 17 (17q21) e é composto por 24 éxons, sendo 22 codificantes, que dão origem a um transcrito de 7.8 kb, quando traduzido, corresponde a uma proteína com 1863 resíduos de aminoácidos e 220kDa (MIKI *et al*, 1994; SMITH *et al*, 1996; ROSEN & PISHIVAIAN, 2014).

A região N-terminal de BRCA1 (**Fig. 1.3**) encerra o domínio RING-finger (***R****eally* ***I****nsteresting* ***N****ew* ***G****ene*), que corresponde aos resíduos de aminoácidos 24 a 64 (PAUL & PAUL, 2015). Esse domínio é bastante conservado e encontrado em diversas proteínas (SAVAGE & HARKIN, 2015).

A região central da proteína BRCA1, que corresponde aos éxons 11, 12 e parte do éxon 13, equivale a aproximadamente 65% do seguimento proteico (CLARK *et al*, 2012). Essa região é pouco conservada e não possui uma estrutura tridimensional bem definida. Esse tipo de formação é tipicamente encontrada em proteínas que servem de arcabouço para formação de complexos, onde as regiões de desordem permitem a flexibilização da ligação entre domínios da mesma molécula, facilitando diversas interações,

consequência de uma plasticidade conformacional necessária à organização desses complexos (SAVAGE & HARKIN, 2015).

A região central é conhecida por interagir com diversas proteínas, como a proteína retinoblastoma (Rb), c-Myc, RAD50 e RAD51, atores importantes em uma ampla variedade de vias, como reparo ao DNA, controle do ciclo celular e transcrição (CLARK *et al*, 2012; PAUL & PAUL, 2015). Ainda nesta região, pode-se destacar a presença de dois sinais de localização nuclear (resíduos 503 a 509 e 606 a 615) que direcionam a proteína para o núcleo, um sinal de exportação nuclear (resíduos 81 a 99) e sítios de fosforilação (resíduos S1453, S1457, S1525), **Fig. 1.3** (DAPIC & MONTEIRO, 2006; YI *et al*, 2014).

Próximo a região C-terminal (**Fig. 1.3**), se localiza o domínio *coiled-coil*, importante para a interação com PALB2 e a manutenção da função supressora de tumor de BRCA1 (YI *et al*, 2014; PRAKASH *et al*, 2015).

Na região C-terminal (aminoácidos 1650 a 1863, **Fig. 1.3**), verifica-se a presença de dois domínios BRCT **(BR**CA1 **C-t**erminal) em tandem. Essa região é essencial na manutenção das funções de BRCA1, a simples perda dos últimos 11 resíduos de aminoácidos da proteína é suficiente para que ela não seja funcional (MONTEIRO, *et al*, 1996; CARVALHO *et al*, 2007). Os domínios BRCT contribuem para grande parte das interações realizadas por BRCA1 e sua participação em diferentes complexos (JIANG & GREENBERG, 2015).

Disfunções de *BRCA1* em linhagens germinativas estão associadas aos casos hereditários de câncer de mama, no entanto, além de mutações nas linhagens germinativas, outras alterações no gene *BRCA1* podem levar ao desenvolvimento do câncer, como mutações somáticas e alterações epigenéticas. Observa-se que entre 50-70% dos casos de câncer de mama esporádicos há perda de pelo menos um alelo de *BRCA1*, o que sugere que essa perda possa ser uma etapa importante no desenvolvimento do tumor esporádico (CROPP *et al*, 1994; KOBAYASHI *et al*, 2013; DOWNS & WANG, 2015).

Mutações no gene *BRCA1* podem também aumentar o risco de desenvolvimento de outros tipos de câncer, como de próstata, de estômago, de pâncreas, de cólon, leucemias e linfomas - o risco varia entre 2% e 60%, sendo o estômago e o pâncreas os sítios de maior risco (MERSCH *et al*, 2014; CAVANAGH & ROGERS, 2015; PRAKASH *et al*, 2015).

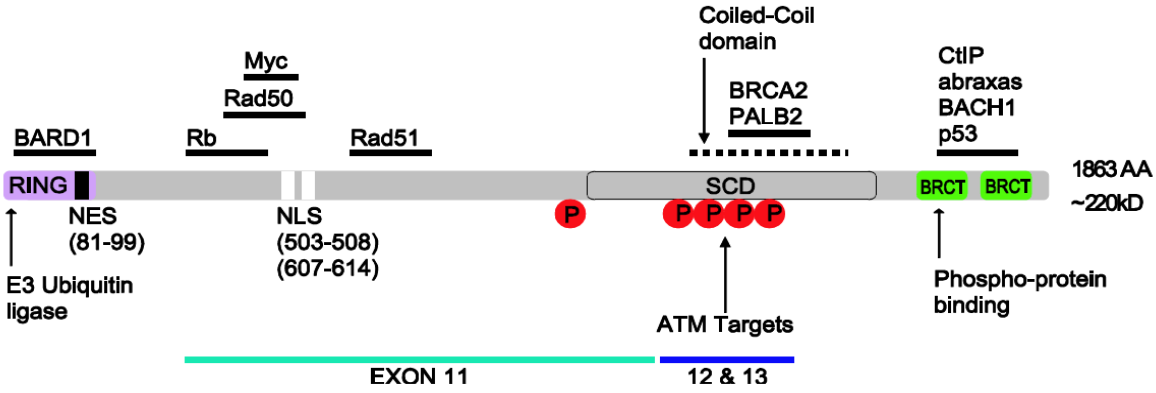


Figura 1.3 Estrutura da proteína BRCA1. O retângulo em roxo representa o domínio RING finger, os retângulos em verde representam os domínios BRCT, em branco os sinais de localização nuclear (NLS), em preto o sinal de exportação nuclear (NES) e em vermelho os sítios de fosforilação (P). SCD – região rica em serinas fosforiláveis. Proteínas interatoras de BRCA1 estão discriminadas por linhas pretas que indicam suas respectivas regiões de interação. Adaptado de Clark *et al*, 2012.

1.3.1. O papel de BRCA1

BRCA1 é uma proteína multifuncional que possui diversos parceiros de interação em diferentes compartimentos celulares, exercendo papéis essenciais em processos envolvidos com a manutenção da integridade genômica, como a sinalização do dano ao DNA, o reparo do DNA, o controle do ciclo celular, a ubiquitinação de proteínas, o remodelamento de cromatina, a apoptose, mecanismos de processamento do RNA e de ativação transcricional (MINAMI *et al*, 2014; SAVAGE & HARKIN, 2015; ALLI & FORD, 2015).

Possivelmente devido a sua habilidade de interagir com diferentes proteínas, sua exata função em muitos desses processos continua indeterminada (PAUL & PAUL, 2015). De fato, mais de 100 interações proteicas de BRCA1 já foram reportadas, e é através dessas interações que BRCA1 atua como uma plataforma de coordenação de múltiplos complexos com diferentes funções (SAVAGE & HARKIN, 2015).

1.3.1.1. Interação com BARD1

A interação mais bem descrita de BRCA1 ocorre com a proteína BARD1 (**B**RCA1 **A**ssociated **R**ING **D**omain protein **1**), formando o heterodímero BRCA1/BARD1, um complexo reconhecido como uma E3 ubiquitina ligase. A interação entre essas duas proteínas, que ocorre através de seus domínios RING-finger (presentes na porção N-terminal de ambas), é necessária tanto para a manutenção da localização nuclear, como para a estabilidade de ambas. O complexo BRCA1/BARD1 tem uma série de substratos de ubiquitinação demonstrados *in vitro*, no entanto poucos substratos foram demonstrados *in vivo* (JIANG & GREENBERG, 2015; SAVAGE & HARKIN, 2015).

1.3.1.2. Recombinação homóloga e recombinação não homóloga por junção de extremidades

A funcionalidade de BRCA1 afeta diretamente as vias de reparo ao dano de DNA, pois ele ocupa posições centrais nesses processos. O reparo ao dano de DNA (RDD), quando ocorre quebras da dupla fita (*double strand breaks* – DSB), é crucial para a manutenção da integridade genômica, isso porque, na região da quebra é possível religar duas extremidades não relacionadas do DNA,

podendo ocasionar perdas ou amplificações genômicas, assim como rearranjos cromossômicos (SAVAGE & HARKIN, 2015).

Os DSBs são reparados por dois mecanismos distintos: recombinação não homóloga por junção de extremidades (**n**on **h**omologous **e**nd **j**oing – NHEJ) e recombinação homóloga (**h**omologous **r**ecombination – HR). A HR utiliza a cromátide irmã não danificada como molde para restaurar a sequência danificada ou perdida, e BRCA1 participa de ao menos 4 diferentes complexos nesse processo (ZHANG & POWELL, 2005; APARACIO *et al*, 2014).

Para o reparo através da via de HR, a extremidade que contém o fosfato livre é reconhecida pelas proteínas MRE11, RAD50, NBS1, BRCA1 e CtIP, que formam o complexo BRCA1-C (**Fig. 1.4**). Esse complexo é responsável pela resseção da fita e geração de uma fita simples com uma extremidade 3' com uma hidroxila livre (**Fig. 1.5**). (WANG, 2012; KAKAROUGKAS & JEGGO, 2014; JIANG & GREENBER, 2015). A fita simples gerada é rapidamente recoberta pela proteína RPA (**Fig. 1.5**), que posteriormente é substituída pela proteína RAD51, num processo dependente do complexo BRCA1-D, composto pelas proteínas: RAD54, PALB2, BRCA2 e BRCA1 (**Fig. 1.4**). Quando RAD51 está associado à fita, ele promove a invasão da dupla fita da cromátide irmã que é usada como molde e, ao final, a invasão é resolvida e as extremidades são religadas (**Fig. 1.5**) (KAKAROUGKAS & JEGGO, 2013; JIANG & GREENBERG, 2015).

Outro complexo importante durante o reparo da DSB é o complexo BRCA1-A composto BRCA1, RAP80, MERIT40, BRCC36/45 e Abraxas (**Fig 1.4**). Esse complexo foi recentemente reportado como necessário para inibição da ressecção do DNA, sugerindo que, em alguns contextos, BRCA1 pode funcionar de modo a limitar ou prevenir uma ressecção excessiva (SAVAGE & HARKIN, 2015).

BRCA1 atua também em um complexo para controle do ciclo celular, o complexo BRCA1-B, composto por BRCA1, TopBP1 e BACH1 (**Fig. 1.4**). Esse complexo atua no arresto da fase S do ciclo celular e na HR, no entanto o mecanismo de ação desse complexo permanece incerto. Como a HR requer a cromátide irmã como molde para o reparo, essa via acontece somente durante as fases S/G2 do ciclo celular (WANG, 2012; APARACIO *et al*, 2014).

Quando a DSB ocorre na fase G1 do ciclo celular e, portanto, não dispõem da cromátide irmã para a HR, ela pode ser resolvida através da NHEJ.

Resumidamente, um complexo proteico reconhece a lesão e liga as duas extremidades lesionadas, independente de uma possível perda. O papel de BRCA1 na NHEJ, se existente, permanece desconhecido. Apesar de haver dados que sugerem defeitos na via de NHEJ em fibroblastos embrionários mesenquimais de camundongos deficientes em BRCA1, outros resultados sugerem que a deficiência de BRCA1 em células tronco embrionárias diminuem os níveis de HR e aumentam a frequência de NHEJ (SAVAGE & HARKIN, 2015).

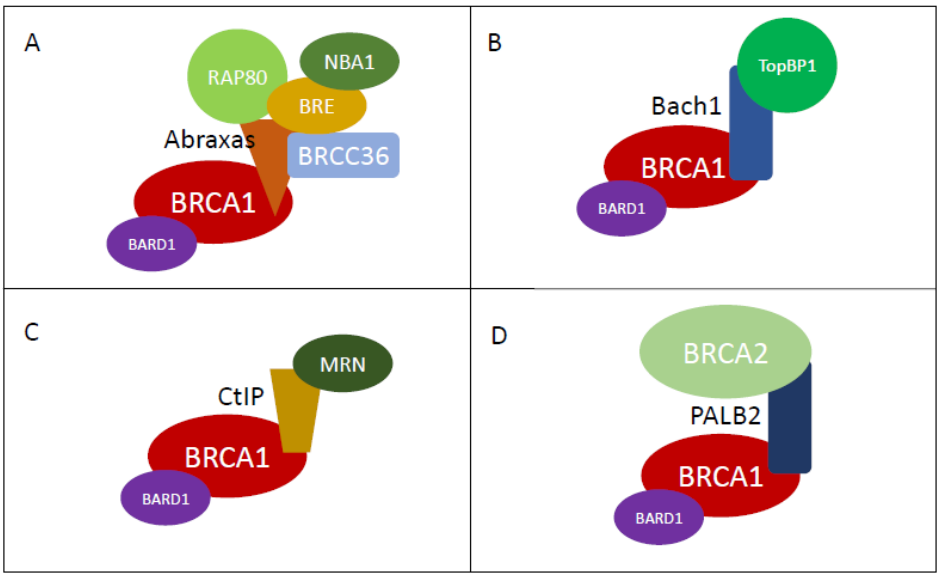


Figura 1.4 – Complexos formados por BRCA1. Os quatro complexos formados por BRCA1 atuam de diferentes formas nas vias de reparo e regulação do ciclo celular. Os complexos BRCA1-A (**A**), BRCA1-C (**C**) e BRCA1-D (**D**) participam da via da HR, com papeis distintos. O complexo (**B**) BRCA1-B tem papel no controle do ciclo celular e tem relação com a via de HR. Adaptado de Wang, 2012.

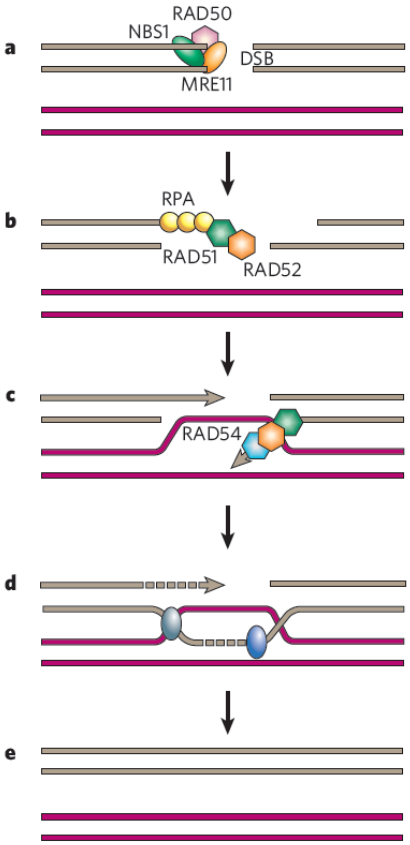


Figura 1.5 – Esquema simplificado da via de recombinação homóloga. (**a**) O complexo MRN reconhece a quebra da dupla fita e realiza a ressecção da fita com a extremidade com fosfato livre. (**b**) A simples fita gerada é recoberta pela proteína RPA, que por ação do complexo BRCA1-PALB2-BRCA2 é substituída pela proteína RAD51. (**c**) Com auxílio das proteínas RAD54 e RAD52 ocorre a invasão da dupla fita da cromátide irmã, (**d**) as fitas são pareadas e a fita complementar é sintetizada. (**e**) O produto final é uma dupla fita idêntica ao original. Downs *et al*, 2007.

1.3.1.3. Reparo ao dano de DNA (outras vias)

Lesões que não induzem distorções significativas na estrutura do DNA, como as distorções observadas nas DSBs, que incluem lesões em uma única base, sítios abásicos (onde a base nitrogenada foi removida) e quebras de uma única fita, são reparadas por mecanismos distintos chamados de reparo por excisão de base (**B**ase **E**xcision **R**epair – BER) e reparo por excisão de nucleotídeo (**N**ucleotide **E**xcision **R**epair – NER) (ALLI & FORD, 2015).

Esses danos são frequentemente gerados por fontes endógenas, mas ocasionalmente podem ter origem na ação de agentes externos. Os agentes de dano incluem espécies reativas de oxigênio (ROS – **r**eactive **o**xygen **s**pecies), deaminações espontâneas de bases, erros de replicação, agentes alquilantes, hidrólises e radiação. Todas essas formas de dano ao DNA podem induzir mutagênese caso não reparadas, levando à quebra de dupla fita, instabilidade genômica e, em último caso, tumorigênese (YI *et al*, 2014).

A via de BER se baseia no reconhecimento e excisão da base danificada e substituição pela base correspondente correta. A função exata de BRCA1 na via de BER ainda é pouco conhecida, entretanto seu papel parece ser semelhante ao observado em outros mecanismos de reparo, nos quais BRCA1 apresenta funções de regulação transcricional e de mediador de interações proteicas (ALBARAKATI *et al*, 2015; ALLI & FORD, 2015).

O processo de NER envolve a excisão de um fragmento de 24-32 pares de base que flaqueia o dano, seguido da ressíntese, usando a fita complementar como molde. Essa via também é usada para o reparo de DNA durante processos de transcrição, conhecido como reparo acoplado à transcrição, processo onde as proteínas MSH2 e MSH6 estão envolvidas. BRCA1 está relacionado com a via de NER devido a sua ação no complexo BASC (*BRCA1-associated genome surveillance complex*), que inclui MLH1, ATM, BLM, o complexo MRN, MSH2 e MSH6 (SAVAGE & HARKIN, 2015).

O complexo BASC regula ainda a expressão de *NRF2*, um fator de transcrição crítico para transcrição de genes antioxidantes, como *ARE*, *NQO1* e *HMOX1*, na presença de ROS (WANG & DI, 2014).

1.3.1.4. Transcrição

O papel de BRCA1 na transcrição foi primeiramente sugerido por Miki e colaboradores em 1994, em função do número elevado de resíduos ácidos em sua região C-terminal. Monteiro e colaboradores (1996) demonstraram que uma construção contendo a região C-terminal de BRCA1 fusionada a domínio de ligação ao DNA é capaz de ativar a transcrição de genes repórteres tanto células humanas quanto em leveduras.

Embora a proteína BRCA1 possua um domínio de ligação ao DNA, ele não se liga a uma sequência específica. Sua atuação em processos transcricionais parece ocorrer através da interação com proteínas envolvidas nesses processos, como por exemplo, a proteína RbAp46 (*Retinoblastoma-Associated protein 46*) e histonas deacetilases 1 e 2 (YI *et al*, 2014).

BRCA1 também parece controlar a transcrição mediada pelo receptor de estrogênio α (REα). O maior efeito dos RE é o estimulo proliferativo. A inibição do REα mediada por BRCA1 ocorre tanto através da inibição direta da proteína REα ou através da inibição de seus efetores subsequentes (YI *et al*, 2014; WANG & DI, 2014).

Na apoptose, BRCA1 parece atuar como um modulador, tanto positivo quando negativo. Mediante dano ao DNA, BRCA1 é fosforilado por ATM e ATR, de modo a modular a atividade de p53 e, consequentemente, a transcrição de *p21*. Além disso, BRCA1 também exerce papel através da formação de um complexo com a subunidade p65 da proteína NF-κB - esse complexo se liga a promotores de genes anti-apoptóticos, incluindo *BCL2* e *XIAP* (SAVAGE & HARKIN, 2015).

1.4. Domínios BRCT de BRCA1

Os domínios BRCT são encontrados em diversos organismos: desde bactérias até humanos. Apesar de não apresentar atividade enzimática intrínseca, esses domínios são essenciais na organização de complexos proteicos nos sítios de dano durante o RDD, sendo importantes tanto para geração dos sinais de dano quanto para a maquinaria de reparo (SHENG *et al*, 2011).

O domínio BRCT pode ocorrer em uma única unidade, como em PARP1 ou em múltiplas unidades, como em TOPBP1 (GLOVER *et al*, 2004). Nas

proteínas onde o domínio BRCT ocorre em múltiplas unidades, o mesmo tende a se organizar em pares de unidades próximas ou justapostas, como acontece em BRCA1, e são chamados de BRCTs em tandem (tBRCT). O domínio tBRCT parece funcionar como uma única unidade estrutural e parece estar relacionado com reparo do dano ao DNA (MESQUITA *et al*, 2010).

O domínio BRCT se apresenta com uma estrutura globular composta por quatro folhas-β centrais paralelas flanqueadas por três α-hélices (**Fig. 1.6**). Quando presente na forma de tandem, os domínios são ligados por uma região denominada *linker* e permanecem lado a lado, em uma estrutura onde a região C-terminal do primeiro fica localizada mais próxima da região N-terminal do segundo (MESQUITA *et al*, 2010; SHENG *et al*, 2011).

A região C-terminal da proteína BRCA1 possui um tBRCT (aminoácidos 1650 a 1863) com um *linker* de 23 resíduos de aminoácidos (**Fig. 1.6**). A arquitetura do domínio tBRCT permite o reconhecimento de serinas fosforiladas através dos resíduos de aminoácidos Ser 1655, Gly 1656 e Lys 1702. O tamanho e a rigidez do núcleo hidrofóbico da interface entre os dois BRCTs determinam a sequência consenso reconhecida, que no caso de BRCA1 é a sequência pS-X-X-F (WILLIAMS *et al*, 2004; PRAKASH *et al* 2015).

A importância das interações entre BRCA1 e proteínas fosforiladas através dos domínios BRCTs é enfatizada pela correlação entre a perturbação da estrutura do domínio tBRCT e a susceptibilidade de camundongos e humanos ao câncer, além de redução das taxas de recombinação homóloga. (CLARK *et al*, 2012; PRAKASH *et al* 2015).

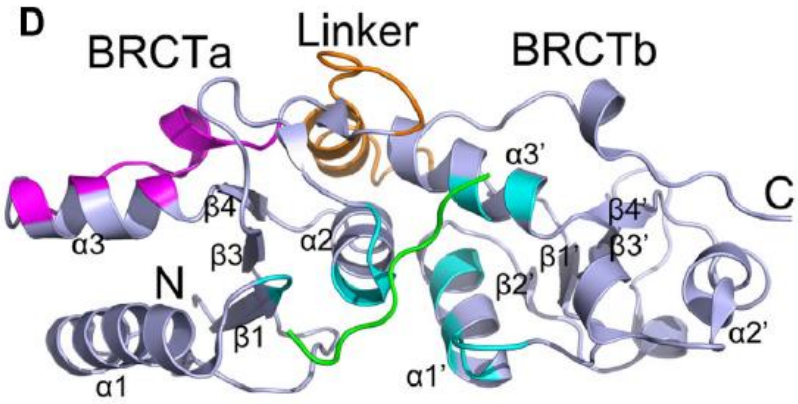


Figura 1.6 – Estrutura tridimensional do domínio tBRCT de BRCA1. Em azul a região de interação entre os domínio BRCT, regiões em rosa e laranja são de interação com outras proteínas e em verde um peptídeo fosforilado. Sheng *et al*, 2011.

1.5. Mutações em BRCA1 e a susceptibilidade ao câncer

Desde a década de 1990, quando *BRCA1* foi identificado, foram reconhecidos cerca de 1800 variantes em linhagens germinativas (IARC, 2016). A identificação de indivíduos carreadores de mutações associadas ao desenvolvimento de câncer de mama e ovário é indispensável para estratégias de prevenção, sejam intervenções clinicas ou mudanças de hábitos de vida (DAPIC & MONTEIRO, 2006).

A frequência de alelos de *BRCA1* mutados é significativamente alta na população (**Fig. 1.7**), no entanto, a frequência de indivíduos homozigotos mutados de *BRCA1* é extremamente baixa - isso porque *BRCA1* atua como um supressor tumoral, e a integridade de *BRCA1* está correlaciona com estabilidade genômica (PRAKASH *et al*, 2015; IARC, 2016).

O câncer de mama hereditário associado à *BRCA1* segue o modelo de inativação bialélica, proposto pela “hipótese dos dois eventos” (*two-hit hypothesis*, KNUDSON, 1971). Esse modelo propõe que o gene de predisposição herdado possui um alelo não funcional e, portanto, somente é necessário um evento mutacional para a perda do alelo funcional e, consequentemente, o desenvolvimento da doença (LUX, *et al*, 2006).

Até recentemente, apenas um único indivíduo bialélico, com duas mutações consideradas patogênicas (consistentes com letalidade embrionária em camundongos – MOYNAHAN, 2002), foi reportado (DOMCHEK *et al*, 2013). Provavelmente esse indivíduo sobreviveu até a fase adulta porque um de seus alelos tinha comportamento hipomórfico, entretanto, a paciente desenvolveu câncer de ovário precocemente. Em 2015, um segundo individuo bialélico para *BRCA1* foi identificado, nesse caso, uma paciente com câncer de mama, portadora de anomalias congênitas, correlacionadas ao quadro de anemia de Fanconi (PRAKASH *et al*, 2015; SAWYER *et al*, 2015).

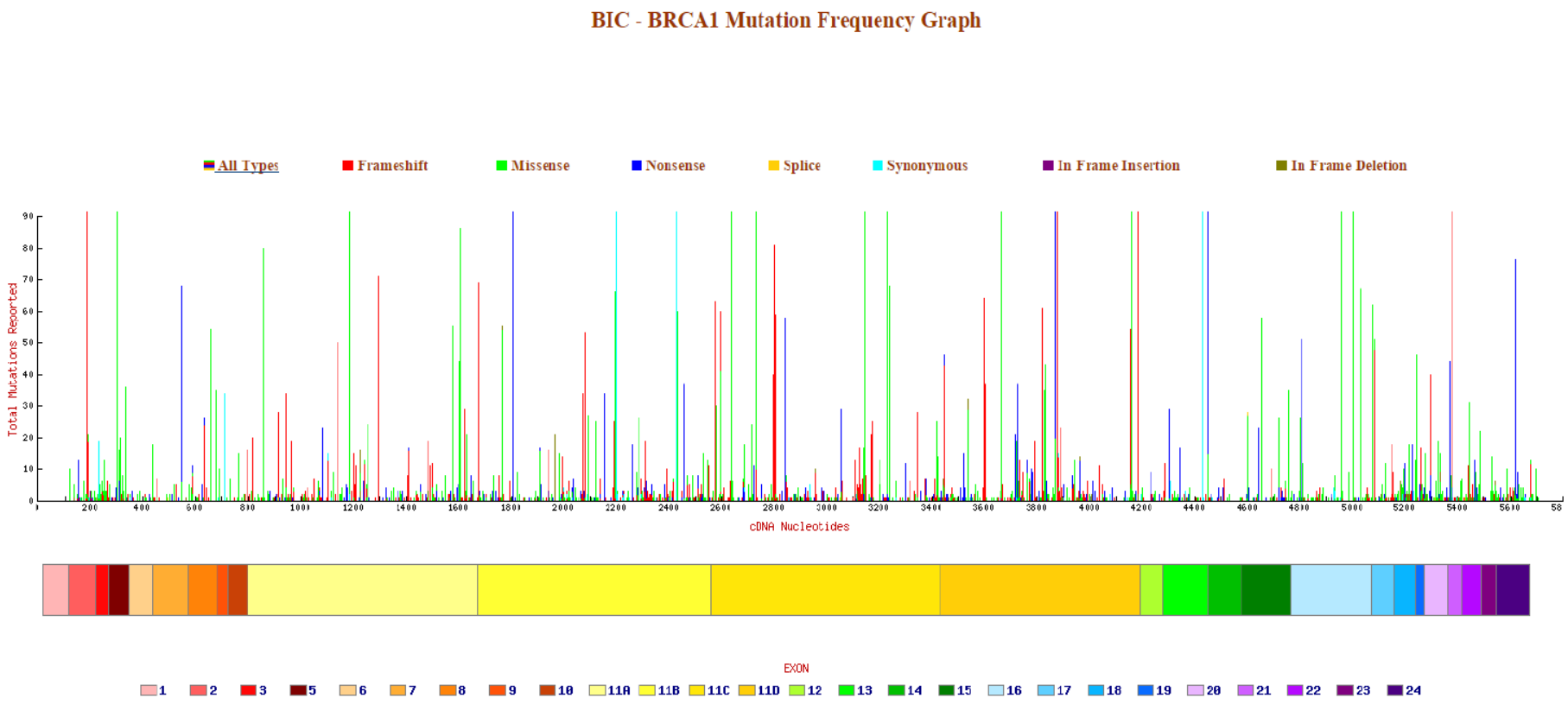


Figura 1.7 – Frequência de variantes de *BRCA1*. O gráfico mostra a frequência de variantes de *BRCA1* reportados na população (discriminados nos respectivos éxons). No eixo x observa-se a posição do nucleotídeo ao longo do cDNA de *BRCA1*; e no eixo y, a frequência das mutações. As cores representam os tipos de mutação (designadas na porção superior do gráfico). Breast Cancer Information Core, IARC, 2016.

Atualmente, testes para identificação de mutações em *BRCA1* são um procedimento padrão para indivíduos quem se enquadram em um dos seguintes critérios: (1) ao menos três casos de câncer de mama e/ou ovário na família; (2) dois casos de câncer de mama em parentes próximos; (3) dois casos de câncer de mama na família diagnosticados antes dos 40 anos; (4) qualquer histórico de câncer de mama em homem, que já possua histórico familiar; (5) ancestralidade judia Ashkenazi com histórico de câncer de mama, particularmente casos de câncer de mama triplo negativo (negativos para os marcadores: receptor de progesterona, receptor de estrogênio e HER2) diagnosticado antes dos 60 anos de idade; ou (6) câncer de mama ou ovário na mesma paciente (SHIOVITZ & KORDE, 2014).

O banco de dados do *Breast Cancer Information Core* (BIC) é um arquivo público de mutações em *BRCA1* e *BRCA2*, administrado pelo *National Institute of Health* (NIH) dos Estados Unidos, onde ficam armazenadas todas as mutações reportadas ao NIH.

A identificação de uma mutação em *BRCA1*, de maneira geral, é realizada através de rotinas de sequenciamento dos éxons e também de suas regiões flanqueadoras (bordas). Qualquer mutação fora dessas regiões não é detectada, e ainda que fossem encontradas, não há uma sequência consenso para identificar possíveis sítios mutados (MILLOT *et al*, 2012). A partir do sequenciamento é possível obter um dos seguintes resultados:

- **Nenhuma mutação identificada**

Caso o indivíduo não possua histórico de câncer de mama na família, significa que o risco de câncer é o mesmo que a população em geral. Porém, caso o indivíduo tenha histórico familiar, existem dois pontos que são importantes: (1) a ausência de mutações em *BRCA1* não exclui a possibilidade de mutações em outros genes de alta ou média penetrância; (2) o sequenciamento não tem cobertura total do gene *BRCA1* (MILLOT *et al*, 2012).

- **Positivo para uma mutação patogênica**

Esses indivíduos possuem predisposição para câncer de mama e ovário hereditários e um elevado risco para vários outros tipos de câncer. Isso permite

que parentes sejam testados e que sejam tomadas as medidas preventivas (CALÒ *et al*, 2010; MILLOT *et al*, 2012).

- **Identificação de um variante não patogênico**

Caso não haja histórico de câncer na família, significa que o risco de câncer é o mesmo que a população em geral. Se houver histórico de câncer deve-se levar em consideração que a ausência de mutações em *BRCA1* não exclui a possibilidade de mutações em outros genes de alta ou média penetrância. (MILLOT *et al*, 2012).

- **Identificação de uma mutação que não possui classificação de patogenicidade**

Não há dados clínicos e epidemiológicos para determinar a patogenicidade do variante, portanto, não há informações sobre sua associação ao câncer (CALÒ *et al*, 2010).

De maneira geral, na população, as mutações do tipo nonsense e missense são as mais comuns. As mutações tipo nonsense, assim como as inserções e deleções, são geralmente associadas ao câncer, isto porque resultam em mudanças do quadro de leitura ou geração de códons de parada prematuros, permitindo uma inferência quanto ao comprometimento funcional da proteína. (CARVALHO *et al*, 2009; CALÒ *et al*, 2010; MILLOT *et al*, 2012).

As mutações do tipo missense e deleções e inserções, que mantêm o quadro de leitura, são de difícil classificação, pois o comprometimento da estrutura da proteína é indeterminado. A associação dessas mutações ao câncer depende de análises populacionais, que são comprometidas pela dificuldade na identificação de famílias acometidas em número adequado para uma análise com apropriado poder estatístico. É comum que variantes desse grupo permaneçam sem definição quanto a sua relação com o câncer, sendo denominados como variantes não classificados (VNC).

Uma alternativa de classificação são os modelos multifatorias 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 coocorrência com mutações deletérias; (e) análises computacionais

de algoritmos preditores, dentre outros critérios (**Fig. 1.8**). Todas essas estratégias, mesmo não substituindo os dados clínicos e epidemiológicos, geram informações importantes que, em conjunto, podem servir de base para decisões clínicas (CARVALHO *et al*, 2007; CALÒ *et al*, 2010; MILLOT *et al,* 2012).

Figura 1.8 – Representação de um modelo multifatorial para classificação de VNC. Para aumentar as informações sobre VNC são somadas características encontradas em coorrência ao variante e todas as informações que possam ser coletadas de estudos populacionais e ensaios *in vitro*. Adaptado de CALÒ *et al*, 2010.

1.6. Ensaios funcionais para classificação de variantes de BRCA1

Os VNCs correspondem a mais de 35% das mutações reportadas ao BIC. A incapacidade de determinar quais mutações são associadas ao câncer e quais não são associadas geram problemas na avaliação de risco, dificultando as decisões clínicas para tratamentos preventivos (CARVALHO *et al*, 2007; COUCH *et al*, 2008; CALÒ *et al*, 2010).

Os estudos funcionais podem contribuir com uma classificação inicial, pois eles se baseiam na avaliação das funções bioquímicas da proteína na presença da mutação. Esses ensaios servem como uma referência primária de comportamento dos variantes, oferecendo dados adicionais que, combinados com outras informações disponíveis, podem auxiliar as avaliações de risco (CARVALHO *et al*, 2007; COUCH *et al*, 2008).

Os estudos funcionais utilizam, em sua grande maioria, uma região da proteína. A região C-terminal de BRCA1 pode ser avaliada de acordo com sua susceptibilidade a proteases (LEE *et al*, 2010; MILLOT *et al*, 2012), sua capacidade de inibição de crescimento de leveduras (CARVALHO *et al*, 2007) ou sua ligação a fosfopeptídeos (LEE *et al*, 2010; MILLOT *et al*, 2012), dentre outros ensaios. A região N-terminal pode ser avaliada por sua capacidade de ligação à proteína BARD1, bem como a capacidade de ubiquitinação de substratos pelo heterodímero BRCA1/BARD1 (MORRIS *et al*, 2006), além outros ensaios possíveis.

Ao avaliar um domínio da proteína BRCA1, muitas vezes são encontradas dificuldades em extrapolar os dados para o contexto clínico. No entanto, dois ensaios foram desenvolvidos de modo a testar funções de BRCA1 no contexto completo da proteína. Esses dois ensaios utilizam células funcionalmente deficientes para BRCA1, e se baseiam em reconstituir o status de BRCA1 a partir da expressão de variantes do gene, avaliando a capacidade de reparo ao dano de DNA. Enquanto o ensaio (1), de recombinação homóloga direta, avalia a capacidade restaurar a estrutura codificante da proteína GFP (sistema repórter), em uma célula que possui integrado em seu genoma dois alelos de *GFP* truncados (inativos), após uma lesão de dupla fita pontualmente gerada em um deles (RANSBURGH *et al*, 2010), o ensaio (2), de sensibilidade à radiação ionizante, expõe as células à radiação e correlaciona sua capacidade de

recuperação com as funções de BRCA1 (SCULLY *et al*, 1999; BOUWMAN *et al*, 2013).

1.7. Ensaio de ativação transcricional (*transcriptional assay* – TA)

O ensaio de TA se baseia na capacidade de recrutamento da RNA polimerase II pela região do tBRCT de BRCA1, que quando fusionado à um domínio de ligação ao DNA tem a capacidade de ativar a transcrição de um gene repórter (MONTEIRO *et al*, 1996). O ensaio correlaciona a integridade do domínio tBRCT com a capacidade de ativação transcricional (LEE *et al*, 2010; MILLOT *et al*, 2012).

O ensaio TA consiste na produção de uma proteína quimérica, produto da fusão da região de ligação ao DNA LexA (de bactéria) ou GAL4 (de levedura) à região C-terminal de BRCA1 (resíduos 1396 a 1863) na presença de uma construção com um sistema repórter, um gene sob o controle de uma região promotora que contenha sequências para a ligação de LexA ou GAL4, levando a expressão do gene, **Fig. 1.9** (CARVALHO *et* al, 2007; CALÒ *et al*, 2010).

Em células humanas, o ensaio é realizado de maneira quantitativa, usando a proteína de fusão GAL4DBD:BRCA1 e como gene repórter uma *luciferase* de *P. pyralis* (vagalume) sobre o controle de um promotor responsivo à GAL4. Para controlar as variações na eficiência de transfecção e viabilidade celular, é usado como controle interno uma *luciferase* de *R. Reniformis (á*gua viva) sob o controle de um promotor constitutivo (CARVALHO *et al*, 2007; LEE *et al*, 2010; MILLOT *et al*, 2012).

O ensaio foi validado utilizando variantes sabidamente associado e não associados ao câncer (classificados através de estudos populacionais) Um conjunto de mais de 50 variantes localizados no tBRCT de BRCA1 já foram analisados através desse ensaio, demonstrando estreita correlação entre a integridade do domínio tBRCT e a capacidade de transcrição do gene repórter (CARVALHO *et al*, 2007; CALÒ *et al*, 2010).

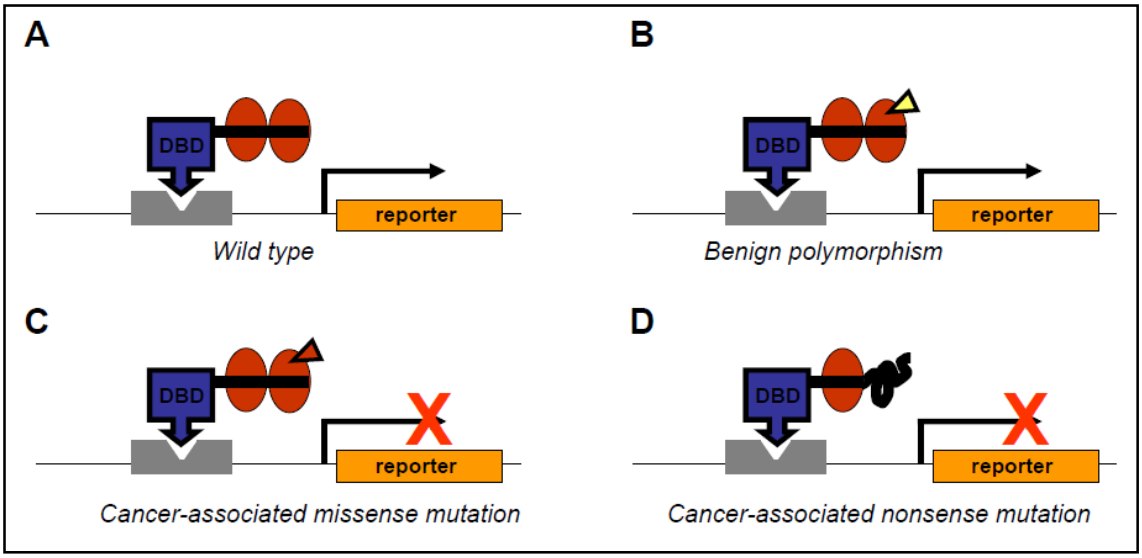


Figura 1.9 – Esquema do ensaio de ativação transcricional para BRCA1. No esquema pode-se observar em (**a**) o comportamento esperado pela proteína selvagem, em (**b**) para um variante missense não patogênico, em (**c**) um variante missense patogênico e em (**d**) um variante nonsense patogênico. Adaptado de Monteiro *et al*, 2000.

2. OBJETIVO

2.1. Objetivo Geral

Avaliar a capacidade de ativação transcricional, como medida de atividade biológica e de integridade estrutural, de variantes missense localizados na região *linker* dos domínios BRCT de BRCA1.

2.2. Objetivos Específicos

2.2.1. Determinar os possíveis variantes naturais missense de BRCA1 na região linker;

2.2.2. Identificar os variantes já descritos na população e/ou previamente estudados;

2.2.3. Avaliar o impacto da variação missense na estabilidade e função da proteína através de algoritmos de predição *in silico* (Align GVGD, SIFT, PolyPhen);

2.2.4. Estabelecer o critério de exclusão e selecionar os variantes para o estudo;

2.2.5. Gerar as construções codificantes dos variantes selecionados para o estudo;

2.2.6. Avaliar a capacidade de ativação transcricional dos variantes de selecionados para o estudo;

2.2.7. Investigar o impacto da temperatura na capacidade de ativação transcricional dos variantes estudados;

2.2.8. Investigar o impacto da variação missense na estrutura secundária da proteína através de análise de predição *in silico*.

3. METODOLOGIA

3.1. Determinação dos variantes naturais missense na região *linker* de BRCA1

Todas as possíveis trocas de um único nucleotídeo na sequência codificante da região *linker* foram verificadas. Trocas que resultaram em variantes missense com desdobramento de substituição de aminoácido foram, inicialmente, incluídas no estudo.

3.2. Avaliação dos variantes missense através de modelos de predição *in silico*

Todos os variantes naturais missense com substituição de aminoácido identificados na região do *linker* foram submetidos aos algoritmos de predição: Align GVGD, disponibilizado pela *International Agency for Research on Cancer - IARC* (*Breast Cancer Genes IARC database – BRCA database*; http://brca.iarc.fr); SIFT, disponibilizado pelo J. Craig Venter Institute (http://sift.jcvi.org/); PolyPhen, disponibilizado pelo *Harvard Medical School, Division of Genetics* (http://genetics.bwh.harvard.edu/pph/data/); e PolyPhen-2, também disponibilizado pelo *Harvard Medical School, Division of Genetics* (http://genetics.bwh.harvard.edu/pph2/), em substituição ao algoritmo PolyPhen.

A sequência proteica modificada de BRCA1 foi submetida aos diferentes algoritmos de predição. Todos os algoritmos se baseiam em alinhamentos múltiplos de sequências de ortólogos, diferindo entre si, essencialmente, nos critérios de escolha das variáveis que são analisadas e nos valores que são atribuídos a cada característica.

3.3. Identificação dos variantes previamente estudados

A identificação de variantes analisados em trabalhos prévios foi investigada no banco de dados do BIC (*Breast Cancer Information Core* - http://research.nhgri.nih.gov/bic/). Esse banco hospeda a plataforma BRCA1 CIRCOS (JHURANEY *et al*, 2014 - https://research.nhgri.nih.gov/projects/bic/circos/index.shtml), que é uma compilação de dados funcionais de todos os variantes de BRCA1 já estudados.

3.4. Critérios de seleção dos variantes para o estudo

Foram selecionados para o estudo todos os possíveis variantes naturais missense identificados na região *linker* de BRCA1 que apresentaram predições

in silico com elevado risco de comprometimento da estrutura da proteína (algoritmos Align GVGD, SIFT, PolyPhen) e que ainda não tivessem sido objeto de estudos anteriores.

3.5. Geração de construções

3.5.1. Geração dos variantes

As construções codificantes de BRCA1 geradas para o estudo compreendem os nucleotídeos 4185 a 5592 (sequência de referência: GenBank U14680.1; http://www.ncbi.nlm.nih.gov/nuccore/U14680). 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.

Os variantes selecionados para o estudo foram gerados por rotinas de mutagênese sítio-dirigida por sobreposição de fragmentos, **Fig. 3.1** (HO *et al;* 1989). A mutagênese por sobreposição de fragmentos emprega a técnica de PCR (*Polymerase Chain Reaction*) (SAIKI *et al;* 1988) para inserir mutações específicas em uma sequência nucleotídica. Resumidamente, em rotinas de PCR distintas são amplificados dois fragmentos da sequência alvo, Fragmento A e Fragmento B. Cada reação utiliza um oligonucleotídeo iniciador (*primer)* que hibridiza em uma das extremidades da seqüência molde (*primers* “UX13” ou “24ENDT”) e um *primer* de mutagênese que hibridiza na região da mutação e contém uma base alterada em relação à seqüência molde, correspondente a mutação de interesse (*primers* “Mut FW” ou “Mut RV”). Como os produtos das reações, Fragmento A e Fragmento B, são parcialmente complementares, é possível então combinar os fragmentos utilizando-os como molde para geração e amplificação da sequência mutante em conjunto com os *primers* “UX13” e “24ENDT” (encerrando sítios de restrição pra *Eco*RI e *Bam*HI, respectivamente; **Fig. 3.1**).

Para as rotinas de mutagênese foram desenhados quarenta e seis *primers*. Para a amplificação da região de estudo (éxon 13 ao éxon 24) foram desenhados dois *primers* (UX13 e 24ENDT, os mesmos utilizados nas reações de mutagênese) que amplificam a sequência codificante de BRCA1 do nucleotídeo 4185 ao 5592, correspondente à região limitada aos resíduos de aminoácidos 1396 a 1863 (**Quadro 3.1**).

A construção plasmidial pCDNA3-BRCA1 13/24, encerrando a sequência codificante correspondente à região C-terminal de BRCA1, previamente gerada pelo nosso grupo (CARVALHO *et al*, 2007), foi usada como molde para as reações de mutagênese. As reações de amplificação foram conduzidas utilizando uma DNA polimerase de alta fidelidade, Pfx50 DNA Polymerase (Invitrogen, Califórnia, Estados Unidos), de acordo com as especificações do fabricante.

Os produtos de amplificação foram analisados por eletroforese em gel de agarose 1% ou 1,5% (p/v), dependendo do tamanho do fragmento gerado, em tampão TBE (45mM Tris base, 45mM ácido bórico, 1mM EDTA, todos Merck, Darmstadt, Alemanha).

Os produtos de amplificação das rotinas de PCR de mutagênese foram purificados a partir de eletroforese preparativa em gel de agarose (como anteriormente descrito) a fim de se evitar contaminação com o molde e com produtos indesejados da reação de amplificação. Os fragmentos foram purificados utilizando *kit* comercial GFX PCR DNA and Gel Purification Kit (GE Healthcare, Buckinghamshire, Reino Unido), de acordo com instruções do fabricante.

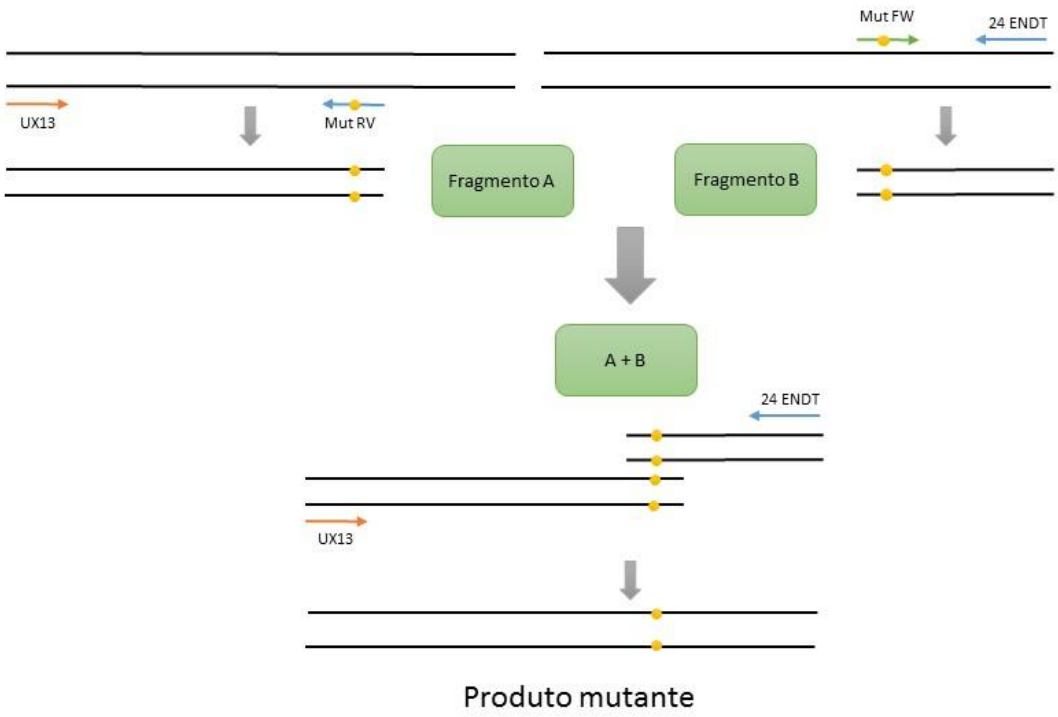


Figura 3.1 – Esquema da rotina de mutagênese sítio-dirigida por sobreposição de fragmentos. A dupla fita de DNA molde está representada por linhas pretas e os *primers* estão representados por linhas com setas na orientação 5'→3'. Os *primers* estão denotados por seus nomes e os produtos de PCR por Fragmento A e B. Na terceira reação, os Fragmentos A e B desnaturam, hibridizam através de sua região complementar e são estendidos pela DNA polimerase, formando o produto recombinante mutante. Na presença dos *primers* UX13 e 24ENDT o produto recombinante mutante é amplificado na rotina de PCR. Os nucleotídeos modificados estão indicados em amarelo. Adaptado de Ho *et al;* 1989.

Tabela 3.1: Oligonucleotídeos iniciadores (*primers*) utilizados na geração de variantes para o estudo (Todos Life Technologies, Califórnia, Estados Unidos).

Mutação	*Primer*	Sentido	Sequência nucleotídica (5'→3')
-	UX13	Senso	CGGAATTCCAGAGGGATACCATGCAA
-	24ENDT	Anti-senso	GCGGATCCTCAGTAGTGGCTGTGGGGGAT
G1738V	G1738V FW	Senso	TTTGAAGTCAGAGTAGATGTGGTCAAT
	G1738V RV	Anti-senso	ATTGACCACATCTACTCTGACTTCAAA
D1739H	D1739H FW	Senso	TGAAGTCAGAGGACATGTGGTCAATGG
	D1739H RV	Anti-senso	CCATTGACCACATGTCCTCTGACTTCA
D1739A	D1739A FW	Senso	TGAAGTCAGAGGACATGTGGTCAATGG
	D1739A RV	Anti-senso	CCATTGACCACATGTCCTCTGACTTCA
V1740A	V1740A FW	Senso	GTCAGAGGAGATGCGGTCAATGGAAGA
	V1740A RV	Anti-senso	TCTTCCATTGACCGCATCTCCTCTGAC
V1740E	V1740E FW	Senso	GTCAGAGGAGATGAGGTCAATGGAAGA
	V1740E RV	Anti-senso	TCTTCCATTGACCTCATCTCCTCTGAC
V1740G	V1740G FW	Senso	GTCAGAGGAGATGGGGTCAATGGAAGA
	V1740G RV	Anti-senso	TCTTCCATTGACCCCATCTCCTCTGAC
G1743R	G1743R FW	Senso	AGATGTGGTCAATCGAAGAAACCACCA
	G1743R RV	Anti-senso	TGGTGGTTTCTTCGATTGACCACATCT
G1743V	G1743V FW	Senso	GATGTGGTCAATGTAAGAAACCACCAA
	G1743V RV	Anti-senso	TTGGTGGTTTCTTACATTGACCACATC
G1743E	G1743E FW	Senso	GATGTGGTCAATGAAAGAAACCACCAA
	G1743E RV	Anti-senso	TTGGTGGTTTCTTTCATTGACCACATC
H1746Y	H1746Y FW	Senso	CAATGGAAGAAACTACCAAGGTCCAAA
	H1746Y RV	Anti-senso	TTTGGACCTTGGTAGTTTCTTCCATTG
H1746D	H1746D FW	Senso	CAATGGAAGAAACGACCAAGGTCCAAA
	H1746D RV	Anti-senso	TTTGGACCTTGGTCGTTTCTTCCATTG
H1746L	H1746L FW	Senso	AATGGAAGAAACCTCCAAGGTCCAAAG
	H1746L RV	Anti-senso	CTTTGGACCTTGGAGGTTTCTTCCATT
H1746P	H1746P FW	Senso	AATGGAAGAAACCCCCAAGGTCCAAAG
	H1746P RV	Anti-senso	CTTTGGACCTTGGGGGTTTCTTCCATT
G1748R	G1748R FW	Senso	AAGAAACCACCAACGTCCAAAGCGAGC
	G1748R RV	Anti-senso	GCTCGCTTTGGACGTTGGTGGTTTCTT
G1748C	G1748C FW	Senso	AAGAAACCACCAATGTCCAAAGCGAGC
	G1748C RV	Anti-senso	GCTCGCTTTGGACATTGGTGGTTTCTT
G1748V	G1748V FW	Senso	AGAAACCACCAAGTTCCAAAGCGAGCA
	G1748V RV	Anti-senso	TGCTCGCTTTGGAACTTGGTGGTTTCT
G1748D	G1748D FW	Senso	AGAAACCACCAAGATCCAAAGCGAGCA
	G1748D RV	Anti-senso	TGCTCGCTTTGGATCTTGGTGGTTTCT
P1749S	P1749S FW	Senso	AAACCACCAAGGTTCAAAGCGAGCAAG
	P1749S RV	Anti-senso	CTTGCTCGCTTTGAACCTTGGTGGTTT
P1749L	P1749L FW	Senso	AACCACCAAGGTCTAAAGCGAGCAAGA
	P1749L RV	Anti-senso	TCTTGCTCGCTTTAGACCTTGGTGGTT
P1749Q	P1749Q FW	Senso	AACCACCAAGGTCAAAAGCGAGCAAGA
	P1749Q RV	Anti-senso	TCTTGCTCGCTTTTGACCTTGGTGGTT
R1753G	R1753G FW	Senso	CAAAGCGAGCAGGAGAATCCCAG
	R1753G RV	Anti-senso	CTGGGATTCTCCTGCTCGCTTTG
R1753I	R1753I FW	Senso	CCAAAGCGAGCAATAGAATCCCAGGAC
	R1753I RV	Anti-senso	GTCCTGGGATTCTATTGCTCGCTTTGG
R1753S	R1753S FW	Senso	AAGCGAGCAAGCGAATCCCAGGA
	R1753S RV	Anti-senso	TCCTGGGATTCGCTTGCTCGCTT

3.5.2. Clonagem no vetor pPCR-TOPO

Os produtos das rotinas de amplificação da região codificante de BRCA1 correspondente aos resíduos de aminoácidos 1396 a 1863 ($BRCA1_{13/24}$) dos variantes foram clonados no vetor pPCR-TOPO (**Fig. 3.2**) utilizando o kit TOPO® PCR Cloning (Life Technologies), de acordo com as instruções do fabricante para produtos gerados com Pfx 50. As construções geradas no vetor pPCR-TOPO foram utilizados em rotinas de transformação e seleção de bactérias *Escherichia coli* DH5'α, de acordo com instruções do fabricante – em meio seletivo contendo ampicilina, X-GAL e IPTG (todos Sigma-Aldrich, Missouri, Estados Unidos).

As construções pPCR-TOPO-$BRCA1_{13/24}$ foram confirmadas por sequenciamento automático, como descrito no item 3.9.

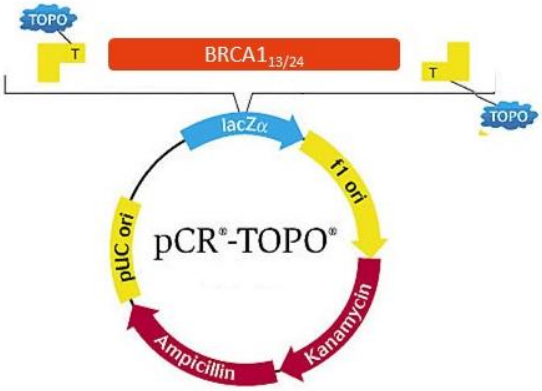


Figura 3.2 – Representação da construção pPCR-TOPO-BRCA1$_{13/24}$. Os produtos obtidos pelas rotinas de mutagênese e amplificação (cassete BRCA1$_{13/24}$, indicado em vermelho) foram clonados no vetor pPCR-TOPO®.

3.5.3. Subclonagem no vetor de expressão pGBT9

As construções pPCR-TOPO $BRCA1_{13/24}$ encerrando os diferentes variantes do estudo e o vetor pGBT9 (Clontech, Califórnia, Estados Unidos) foram submetidos à digestão com as enzimas *Eco*RI e *Bam*HI (ambos Thermo Fisher Scientific, Massachusetts, Estados Unidos). O produto das digestões foi resolvido por eletroforese em gel de agarose 1% p/v, e o cassete $BRCA1_{13/24}$, bem como o vetor pGBT9 linearizado, foram purificados utilizando *kit* comercial GFX PCR DNA and Gel Purification Kit como descrito no item 3.5.1.

O cassete $BRCA1_{13/24}$ foi subclonado no vetor pGBT9 no mesmo quadro de leitura do domínio de ligação ao DNA de GAL4 (GAL4DBD), gerando a construção pGBT9-GAL4DBD:$BRCA1_{13/24}$, **Fig. 3.3**. Para tanto, foi usada a enzima T4 DNA ligase (Promega, Wisconsin, Estados Unidos), seguindo as recomendações do fabricante

Os produtos recombinantes gerados nas reações de ligação foram utilizados em rotina de transformação e seleção de bactérias *E. coli* DH5'α (como descrito no item 3.7).

As construções no vetor pGBT9 foram confirmadas por sequenciamento automático, como descrito no item 3.9.

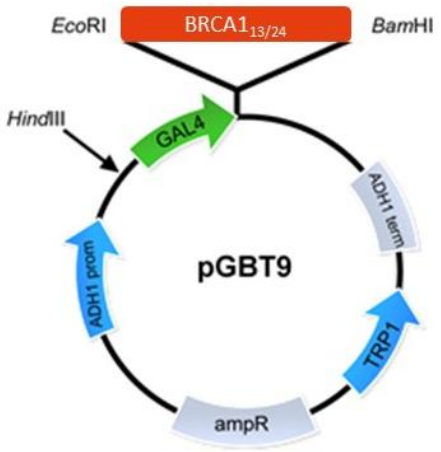


Figura 3.3 – Representação da construção de pGBT9-GAL4DBD:BRCA1$_{13/24}$. O cassete BRCA1$_{13/24}$ (em vermelho) originalmente clonado em pPCR-TOPO® foi subclonado direcionalmente no vetor pGBT9 (sítios de *Eco*RI e *Bam*HI) no mesmo quadro de leitura que a sequência codificante do GAL4DBD (indicado em verde).

3.5.4. Subclonagem no vetor de expressão pCDNA3

As construções pGBT9-GAL4DBD:BRCA1$_{13/24}$ foram digeridas com as enzimas *Hind*III (Thermo Fisher Scientific) e *Bam*HI para obtenção do cassete GAL4DBD:BRCA1$_{13/24}$, de forma análoga à descrita no item 3.5.3. O produto purificado foi subclonado no vetor de expressão em célula de mamífero pCDNA3, codificando a proteína quimérica GAL4DBD:BRCA1$_{13/24}$, também de forma análoga à descrita no item 3.5.3 (**Fig. 3.4**).

Os produtos recombinantes gerados nas reações de ligação foram utilizados em rotina de transformação e seleção de bactérias *E. coli* DH5'α (como descrito no item 3.7).

As construções pCDNA3-BRCA1$_{13/24}$ foram confirmadas por sequenciamento automático, como descrito no item 3.9, e utilizadas no ensaio de ativação transcricional em células humanas (item 3.11).

A sequência de eventos das rotinas de clonagem e subclonagem está representada esquematicamente na **Fig. 3.5**.

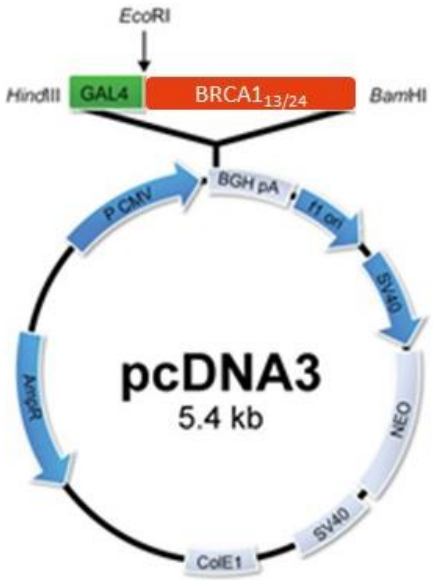


Figura 3.4 – Representação da construção de pCDNA3-GAL4DBD:BRCA1$_{13/24}$. O cassete BRCA1$_{13/24}$ (em vermelho), originalmente clonado em pGBT9, foi subclonado direcionalmente no vetor pcDNA3 (sítios de *Hind*II e *Bam*HI) em fusão com GAL4DBD.

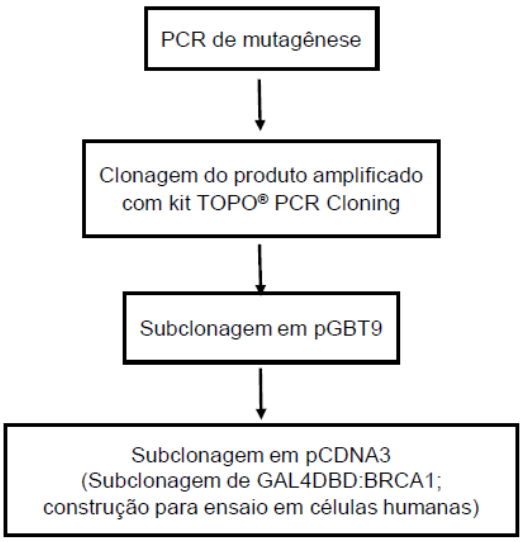


Figura 3.5 – Fluxograma das rotinas. Descrição do processo de clonagem das construções utilizadas neste estudo, desde as rotinas de mutagênese até a clonagem dos fragmentos vetor pCDNA3.

3.6. Indução de competência bacteriana

Bactérias eletrocompetentes foram obtidas conforme descrito previamente por Nováková e colaboradores (2014). Resumidamente, bactérias *E. coli* da cepa DH5'α, foram cultivadas em meio SOC (peptona 2% p/v, extrato de levedura 0,5 % p/v, NaCl 10mM, KCl 2,5mM, glicose 20mM, todos Sigma Aldrich), sob constante agitação a 37ºC. O crescimento da cultura foi acompanhado até a fase exponencial (DO_{600nm} de aproximadamente 0,5 unidades). A cultura foi, então, mantida no gelo por 20 minutos e, posteriormente, centrifugada (2.000 xg / 15 minutos a 4ºC). Após a centrifugação, as células foram recuperadas em glicerol 10% (v/v) gelado, no mesmo volume inicial da cultura, e novamente centrifugadas. O *pellet* obtido foi ressuspenso em glicerol 10% (v/v) gelado, na metade do volume original da cultura. Por fim, realizou-se uma nova centrifugação, recuperando as células em em glicerol 10% (v/v) gelado, em 1/10 do volume original da cultura. A suspensão de bactérias foi aliquotada e estocada a -80°C.

3.7. Transformação bacteriana

A transformação de bactérias usando as construções plasmidiais foi conduzida pelo processo de eletroporação (NEUMANN *et al*, 1982). Resumidamente, para a transformação, as bactérias eletrocompetentes foram incubadas com o DNA plasmidial em banho de gelo, transferidas para uma cubeta própria e submetidas a um pulso elétrico rápido de 1.8 kV em eletroporador MicroPulser (BIO-RAD Califórnia, Estados Unidos). As bactérias foram, então, recuperadas em meio LB a 37°C sob agitação; posteriormente, as células foram plaqueadas em meio LB (extrato de levedura 5% p/v, triptona 10% p/v, cloreto de sódio 5% p/v) sólido contendo ampicilina (10µg/mL) (todos Sigma Aldrich).

3.8. Extração plasmidial

A extração do DNA plasmidial foi conduzida pelo método de lise alcalina, segundo Ausubel (2002). Resumidamente, as bactérias carregando o plasmídeo de interesse foram ressuspensas em tampão GTE (glicose 50mM, Tris-Cl 25mM

pH 8,0, EDTA 10mM, todos Merck) contendo RNAse (Sigma Aldrich), lisadas em solução de hidróxido de sódio/SDS (NaOH 0.2M, SDS 1% p/v, ambos Merck) e neutralizadas em tampão de acetato de potássio (ácido acético glacial 29,5 mL, pérolas de KOH até pH 4,8, H_2O para 100mL, ambos Merck). O material foi desproteinizado com clorofórmio, e precipitado em isopropanol (ambos Merck).

Para as rotinas de sequenciamento e transfecção de células humanas as preparações plasmidiais foram obtidas utilizando o kit *QIAprep Spin Miniprep* (QIAGEN, Hiden, Alemanha), de acordo com as recomendações do fabricante.

3.9. Sequenciamento

As construções geradas foram confirmadas através de rotina de sequenciamento automático 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 3130*xl* (Thermo Fisher Scientific) em colaboração com o Programa de Genética do Centro de Pesquisa do Instituto Nacional de Câncer.

3.10. Cultura de células humanas

A linhagem humana HEK293FT (Invitrogen, Massachusetts, EUA) foi cultivada em meio de cultura DMEM (Life Technologies) suplementado com 10% v/v de soro fetal bovino (SFB) (Invitrogen) e antibióticos penicilina (1000 U/L) e estreptomicina (100mg/mL) (ambos Life Techonologies) à 37ºC em atmosfera contendo 5% v/v de CO_2. Os subcultivos foram realizados a cada 48-72 horas, quando as culturas apresentavam de 80-90% de confluência da área cultivada, através de dissociação enzimática com solução de tripsina 0,125% p/v e EDTA 0,01% p/v (ambos Sigma).

3.11. Transfecção de células humanas

Células humanas (linhagem HEK293FT) foram cotransfectadas com a construção pCDNA3-BRCA1$_{13/24}$ (selvagem ou variante), juntamente com o plasmídeo pG5luc, que encerra o sistema repórter codificante para luciferase de *Photinus pyralis* sob o controle de operadores GAL4 (**Fig. 3.6-A**), e o plasmídeo pGR-TK, que encerra a sequência codificante para a luciferase de *Renilla reniformis* (controle interno de transfecção), sob o controle de um promotor

constitutivo HSV-TK (**Fig. 3.6-B**) utilizando o reagente polietilenimina (PEI) (Polysciences Inc, Pannsylvania, EUA).

Para o ensaio de ativação transcricional, as células HEK293FT foram inoculadas em placa de 96 poços ($5x10^4$ células/poço) e para obtenção de extrato proteico foram inoculadas em placas de 35mm de diâmetro ($5x10^5$ células/poço). Após 24 horas do plaqueamento, as células foram tratadas com a mistura de transfecção: DNA, PEI e PBS (proporção de 3 µL de PEI para cada 1 µg de DNA plasmidial) previamente incubada a temperatura ambiente por 30 minutos.

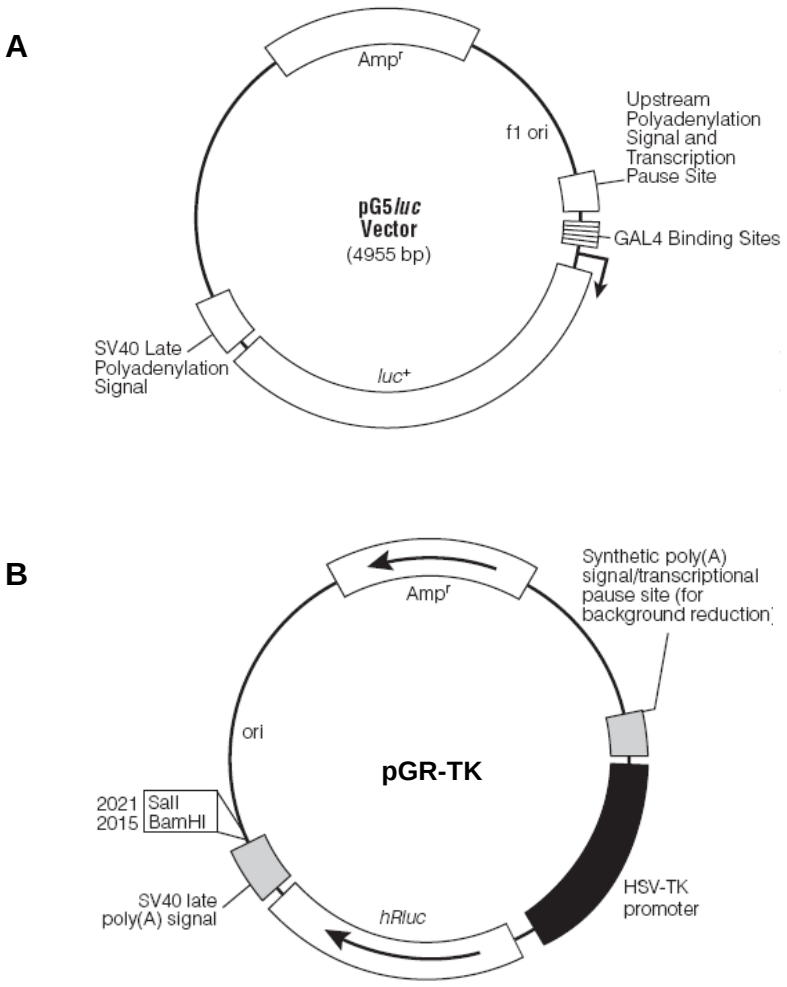


Figura 3.6 – Representação dos plasmídeos pG5*luc* e pGR-TK. (**A**) O plasmídeo pG5*luc* encerra a sequência codificante da enzima luciferase de *P. pyralis* (*luc*+) regulado por 5 operadores de GAL4 (sítios de ligação a GAL4). (**B**) O plasmídeo pGR-TK encerra a sequência codificante da enzima luciferase de *R. reniformis* (R*luc*) controlado por um promotor reconhecido por células humanas (constitutivo), HSV TK.

3.12. Ensaio funcional em células humanas

A capacidade de ativação transcricional da proteína quimérica GAL4DBD:BRCA1$_{13/24}$ foi determinada pela razão entre a atividade da luciferase de *P. pyralis* (produzida pelo sistema repórter) e da luciferase de *R. reniformis* (produzida pelo controle interno de transfecção), como descrito por Vallon-Christersson e colaboradores (2001) e Carvalho e colaboradores (2007).

A determinação quantitativa da capacidade de ativação transcricional da proteína quimérica GAL4DBD:BRCA1$_{13/24}$ foi realizada usando o *Dual-Luciferase® Reporter Assay System* (Promega), seguindo as recomendações do fabricante. Resumidamente, 24 ou 48 horas após a transfecção, as células HEK293FT foram tratadas com o tampão de lise, disponibilizado pelo fabricante, e uma alíquota do lisado foi transferida para um microtubo de 1,5mL para leitura na plataforma GloMax (Promega). A cada amostra foi adicionado o reagente LARII®, também disponibilizado pelo fabricante, que contém luciferina-D. Este é convertido pela luciferase de *P. pyralis*, produzindo oxiluciferina e fótons. A emissão dos fótons é captada pelo luminômetro e medida de forma acumulativa por 1 segundo. Em seguida é adicionado o reagente Stop and Glo®, também fornecido pelo fabricante, que interrompe a primeira reação e disponibiliza a coelenterazina, substrato da segunda reação, catalisada pela luciferase de *R. reniformis.* A reação forma coelenteramida e fótons, também mensurados de forma acumulativa por 1 segundo.

As construções utilizadas como controles no ensaio de ativação transcricional foram previamente geradas por Carvalho e colaboradores (2007), são elas: pCDNA3-BRCA1$_{13/24}$ selvagem, S1613G, Y1853X e M1775R.

3.12.1. Ensaio funcional à 37ºC

Células HEK293FT inoculadas em placa de 96 poços foram transfectadas, mantidas à 37ºC em atmosfera de 5% de CO_2 e meio DMEM com bicarbonato de sódio à 44mM (3,7g/L, como nas instruções do fabricante). Após 24 horas de incubação, as células foram lisadas e a capacidade de ativação transcricional foi avaliada como descrito no **item 3.12**.

3.12.2. Ensaio funcional à 30ºC

Células HEK293FT inoculadas em placa de 96 poços foram transfectadas, mantidas à 30ºC em atmosfera sem injeção de CO_2 e em meio DMEM com bicarbonato de sódio à 2mM (0,2g/L). Após 48 horas de incubação, as células foram lisadas e a capacidade de ativação transcricional foi avaliada como descrito no **item 3.12**.

3.13. Extratos proteicos

O extrato proteico total de células HEK293FT foi obtido 24 e 48 horas após a transfecção com as construções pCDNA3-BRCA1$_{13/24}$, após a incubação à 37ºC ou 30ºC, respectivamente. As células foram recolhidas mecanicamente e ressuspensas em tampão RIPA (NaCl 150 mM, Tris-HCl 10 mM pH 7.4, EDTA 5 mM, dodecil sulfato de sódio 0.1%, Triton X-100 1%, deoxicolato de sódio 0.1%, todos Merck) contendo coquetel de inibidores de protease (Sigma-Aldrich). Os lisados foram, então, incubados em banho de gelo por 30 minutos e centrifugados por 5 minutos à 4ºC e 14.000 RPM (microcentrífuga Hettich Mikro 22R). Os sobrenadantes foram recuperados e armazenados à – 80ºC.

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

3.14. Avaliação de síntese da proteína quimérica

A avaliação do perfil de produção das proteínas quiméricas GAL4DBD:BRCA1$_{13/24}$ foi conduzida por *Western blotting* seguido por imunodetecção (BURNETTE, 1981). Os extratos celulares totais foram resolvidos por rotinas de SDS-PAGE (*SDS-PolyAcrilamide Gel Electrophoresis*) 10% (p/v) e eletrotransferidos para membranas de difluoreto de polivinilideno (PVDF) (GE Healthcare), usando o sistema Trans-Blot Semidry (BIORAD), segundo recomendações do fabricante. Concluída a transferência as membranas foram bloqueadas com leite em pó desnatado 5% em TBS-Tween (TBS: Tris base 50mM, NaCl 150mM; Tween: 0,1%, todos Merck), por 1 hora à temperatura ambiente. Posteriormente, as membranas foram incubadas com os anticorpos monoclonais anti-GAL4DBD (IgG de camundongo, sc-46680, Santa Cruz, Califórnia, Estados Unidos) diluídos em TBS-Tween suplementado com leite desnatado 0,5%, por 16 horas à 4ºC. 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 horas à temperatura ambiente. O *immunoblot* foi revelado utilizando o kit *Amersham ECL™ Western Blotting Detection Reagents* (GE Healthcare) e a membrana foi exposta ao filme radiográfico *Amershan Hyperfilm ECL* (GE Healthcare).

3.15. Diagramas de Ramachandran

A análise da geometria global da estrutura dos tBRCTs foi realizada através do programa Swiss PDB Viewer, SPDBV, (disponível em http://spdbv.vital-it.ch/disclaim.html). Como base, foi empregada a estrutura 1JNX (WILLIAMS *et al*, 2001) – correspondente à análise da sequência selvagem do tBRCT de BRCA1 por difração de raio X – depositada no RCSB PDB (http://www.rcsb.org/pdb/home/home.do).

Para a análise inicial foi desconsiderado o átomo de níquel que se encontrava cristalizado junto com a proteína. Após submeter a sequência ao programa SPDBV, foram determinadas as possíveis pontes de hidrogênio a serem estabelecidas e, para a obtenção da conformação mais favorável possível, foi minimizada a energia do sistema. Esse modelo foi usado para obtenção do diagrama de Ramachandran do tBRCT selvagem.

A partir do modelo tridimensional selvagem foram realizadas as substituições (mutações) pontuais necessárias para as análises. Após cada substituição as pontes de hidrogênio foram redefinidas e a conformação mais favorável possível foi obtida com a minimização de energia. Ao final, o diagrama de Ramachandran dos variantes foi gerado e armazenado.

3.16. Análises Estatísticas

Os dados foram avaliados estatisticamente empregando a análise de variância (two-way ANOVA) e pós teste Bonferroni utilizando o programa GraphPad Prism v. 6.0 para Windows (GraphPad Software, San Diego California USA; www.graphpad.com).

4. Resultados

4.1. Seleção dos variantes

Para a condução deste estudo, primeiramente, foram identificados todos os possíveis variantes naturais missense localizados na região *linker* de BRCA1. Através da simulação da substituição de cada um dos 69 nucleotídeos que compõe a sequência codificante dessa região, foi possível a identificação de 132 variantes missense com desdobramento de substituição de aminoácido. Todos foram submetidos aos algoritmos de predição *in silico*, Align GVGD, SIFT e PolyPhen (**Anexo 1**, **Tabela A1**).

O Align GVGD foi usado em uma rotina de análise de maior estringência, selecionando o conjunto de alinhamentos de referência que incluem desde a sequência humana (*H. sapiens*) até o ortólogo em ouriço do mar (*Strongylocentrotus purpuratus*). As ferramentas SIFT e PolyPhen foram usadas sem alteração dos parâmetros iniciais recomendados.

Os resultados obtidos a partir do uso dos algoritmos foram usados como critérios de seleção para o estudo; assim, apenas variantes com elevado risco de comprometimento da estrutura da proteína foram incluídos (classificação pelo Align GVGD = *C65*; pelo SIFT = *Probably Damaging*; PolyPhen = *Damaging*).

Dentre os 132 variantes originalmente identificados, 16 variantes já haviam sido objeto de estudo descrito na literatura e, independentemente de sua classificação pelos algoritmos de predição, foram descartados de nosso estudo. Dos 116 variantes remanescentes, 23 variantes apresentaram um perfil preditivo patogênico nos três algoritmos de predição utilizados e foram selecionados para nosso estudo (**Tabela 4.1**).

Ao longo do desenvolvimento deste estudo, o algoritmo de predição PolyPhen foi reestruturado e, atualmente, nomeado como PolyPhen-2. Os variantes previamente selecionados para o estudo foram submetidos a uma nova análise com os atuais parâmetros considerados por esse algoritmo. Diferente da classificação anterior, 4 variantes que apresentaram um perfil preditivo patogênico foram classificados como benignos (**Tabela 4.1**).

Tabela 4.1 – Análises *in silico* dos variantes naturais localizados na região *linker* de BRCA1 selecionados para o estudo. Após a submissão aos algoritmos de predição Align GVGD, SIFT, PolyPhen e PolyPhen-2 os variantes com maior risco de comprometimento da estrutura proteica foram selecionados para o estudo. C65 – risco elevado; Damaging (não tolerada, prejudicial); Possibly D –possivelmente não tolerada, prejudicial; Probably D – provavelmente não tolerada, prejudicial; Benign – benigna, não prejudicial.

Variante	Variação nucleotídica	Classificação Align GV/GD	Classificação SIFT	Classificação PolyPhen	Classificação PolyPhen-2
G1738V	G5213T	C65	Damaging	Damaging	Possibily D
D1739H	G5215C	C65	Damaging	Damaging	Possibily D
D1739A	A5216C	C65	Damaging	Damaging	Benign
V1740E	T5219C	C65	Damaging	Damaging	Possibily D
V1740A	T5219A	C65	Damaging	Damaging	Possibily D
V1740G	T5219G	C65	Damaging	Damaging	Probabily D
G1743R	G5227C	C65	Damaging	Damaging	Benign
G1743V	G5228T	C65	Damaging	Damaging	Probabily D
G1743E	G5228A	C65	Damaging	Damaging	Possibily D
H1746Y	C5236T	C65	Damaging	Damaging	Probabily D
H1746D	C5236G	C65	Damaging	Damaging	Possibily D
H1746L	A5237T	C65	Damaging	Damaging	Benign
H1746P	A5237C	C65	Damaging	Damaging	Possibily D
G1748R	G5242C	C65	Damaging	Damaging	Probabily D
G1748C	G5242T	C65	Damaging	Damaging	Probabily D
G1748V	G5243T	C65	Damaging	Damaging	Possibily D
G1748D	G5243A	C65	Damaging	Damaging	Possibily D
P1749S	C5245T	C65	Damaging	Damaging	Possibily D
P1749L	C5246T	C65	Damaging	Damaging	Probabily D
P1749Q	C5246A	C65	Damaging	Damaging	Possibily D
R1753G	A5257G	C65	Damaging	Damaging	Benign
R1753I	A5258T	C65	Damaging	Damaging	Possibily D
R1753S	A5259C	C65	Damaging	Damaging	Possibily D

4.2. Geração de variantes e construções plasmidiais

Todos os variantes selecionados para o estudo foram satisfatoriamente gerados por estratégia de mutagênese sítio-dirigida por sobreposição de fragmentos e, subsequentemente clonados no vetor pPCR-TOPO. A partir das construções em pPCR-TOPO, o cassete $BRCA1_{13/24}$ foi subclonado no vetor de expressão pGBT9 no mesmo quadro de leitura do domínio de ligação ao DNA de GAL4 (GAL4DBD), gerando as construções pGBT9-GAL4DBD:$BRCA1_{13/24}$. Por fim, foram geradas as construções pCDNA3-$BRCA1_{13/24}$, onde o cassete GAL4DBD:$BRCA1_{13/24}$ foi removido das construções pGBT9-GAL4DBD:$BRCA1_{13/24}$ e subclonado no vetor pCDNA3.

As construções geradas foram confirmadas por sequenciamento automático.

4.3. Avaliação da capacidade de ativação transcricional a 37ºC

Todos os variantes selecionados para o estudo foram analisados através do ensaio de ativação transcricional (TA), segundo sua capacidade de indução de transcrição de um sistema repórter (que codifica a enzima luciferase) em células humanas.

As atividades foram avaliadas de modo comparativo, 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).

Os controles positivos correspondem à proteína selvagem (WT) e um variante não patogênico descrito na população, S1613G (DUNNING *et al*, 1997). Enquanto os controles negativos correspondem aos variantes M1775R (variante missense patogênico, descrito por MIKI *et al*, 1994) e Y1853X (variante nonsense patogênico, descrito por FRIEDMAN *et al*, 1994). Os variantes utilizados foram originalmente classificados através de análises populacionais e, posteriormente por ensaios funcionais, inclusive o TA (CARVALHO *et al*, 2007) . A localização desses variantes está representada na **Fig. 4.1A**. O posicionamento dos variantes analisados neste estudo estão representados na **Fig. 4.1B.**

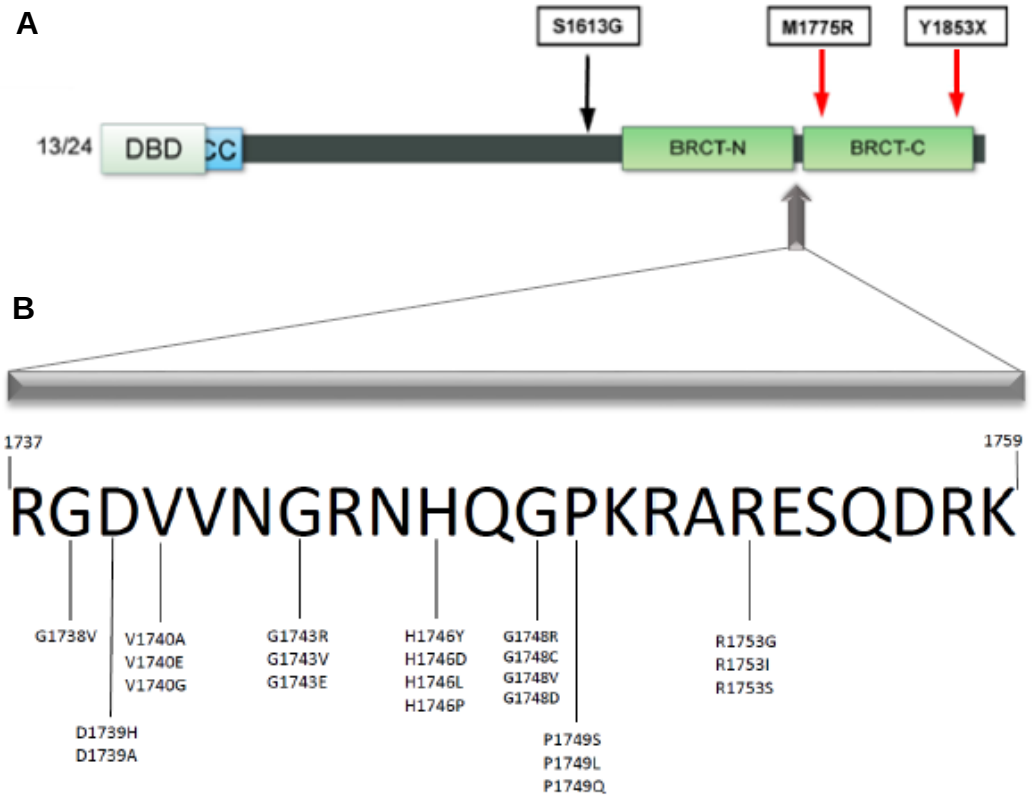


Figura 4.1 – Esquema da proteína quimérica GAL4DBD:BRCA1$_{13/24}$ e localização dos variantes usados como controles no estudo e dos variantes estudados. (A) A figura demonstra a proteína formada pela fusão da região codificada pelos éxons 13 a 24 de *BRCA1* e o domínio de ligação ao DNA de GAL4. O retângulo azul indica localização do domínio *coiled-coil* (CC), parcialmente presente. Os retângulos em verde delimitam a região dos domínios BRCT. A seta na região inferior da figura indica a localização da região de estudo (*linker*). A seta em preta indica a localização do variante não patogênico (controle positivo). As setas em vermelho indicam a localização dos variantes patogênicos (controles negativos). **(B)** Representação da região dos aminoácidos 1737 a 1759. Os 23 aminoácidos que compõem o *linker* estão representados por suas letras correspondentes e as linhas abaixo dos aminoácidos indicam os variantes de cada posição.

A capacidade de ativação transcricional dos variantes foi avaliada em pelo menos 3 experimentos independentes, sempre em quadruplicata. O gráfico representado na **Fig. 4.2** mostra dados representativos de um experimento para a análise dos variantes G1738V, D1739H, D1739A, G1743E, H1746Y, H1746D, H1746L, G1748R, G1748C, G1748V, G1748D, P1749L, P1749Q, R1753G, R1753I e R1753S. De maneira geral, esses variantes apresentaram perfil semelhante aos controles negativos (variantes patogênicos), com taxas de ativação transcricional que variaram de 0,77% a 9,08% em relação ao controle selvagem - os variantes controle Y1853X e M1775R apresentaram taxas de 0,44% e 0,54%, respectivamente.

Cabe ressaltar que os variantes G1743R (24,64%), G1743V (19,32%), H1746P (18,72%) e P1749S (12,65%) apresentaram atividades mais elevadas – ainda assim, seu perfil de atividade se assemelha aos dos variantes patogênicos (controles negativos), segundo critérios definidos por Carvalho e colaboradores (2007). Esses dados estão em consonância com as predições *in silico*.

Diferente dos dados obtidos nas análises *in silico*, os variantes localizados na posição 1740 (V1740A, V1740E e V1740G) apresentaram atividades elevadas, eventualmente comparáveis aos controles positivos, no entanto, com variações amplas, não somente entre ensaios independentes, mas também entre replicatas de um mesmo ensaio.

As variações entre os ensaios podem ser observadas na **Fig. 4.3**. O gráfico mostra todos os ensaios realizados com esses variantes (totalizando 4 ensaios independentes). O variante V1740A apresentou atividade relativa variando de 35,9% a 141,32%, o variante V1710E apresentou uma variação de 8,2% a 87,39% enquanto o variante V1740G apresentou atividade relativa de 58,02% a 76,47%. A amplitude das variações de atividade nas quadruplicatas de um mesmo ensaio podem ser observadas pela amplitude do desvio padrão representado.

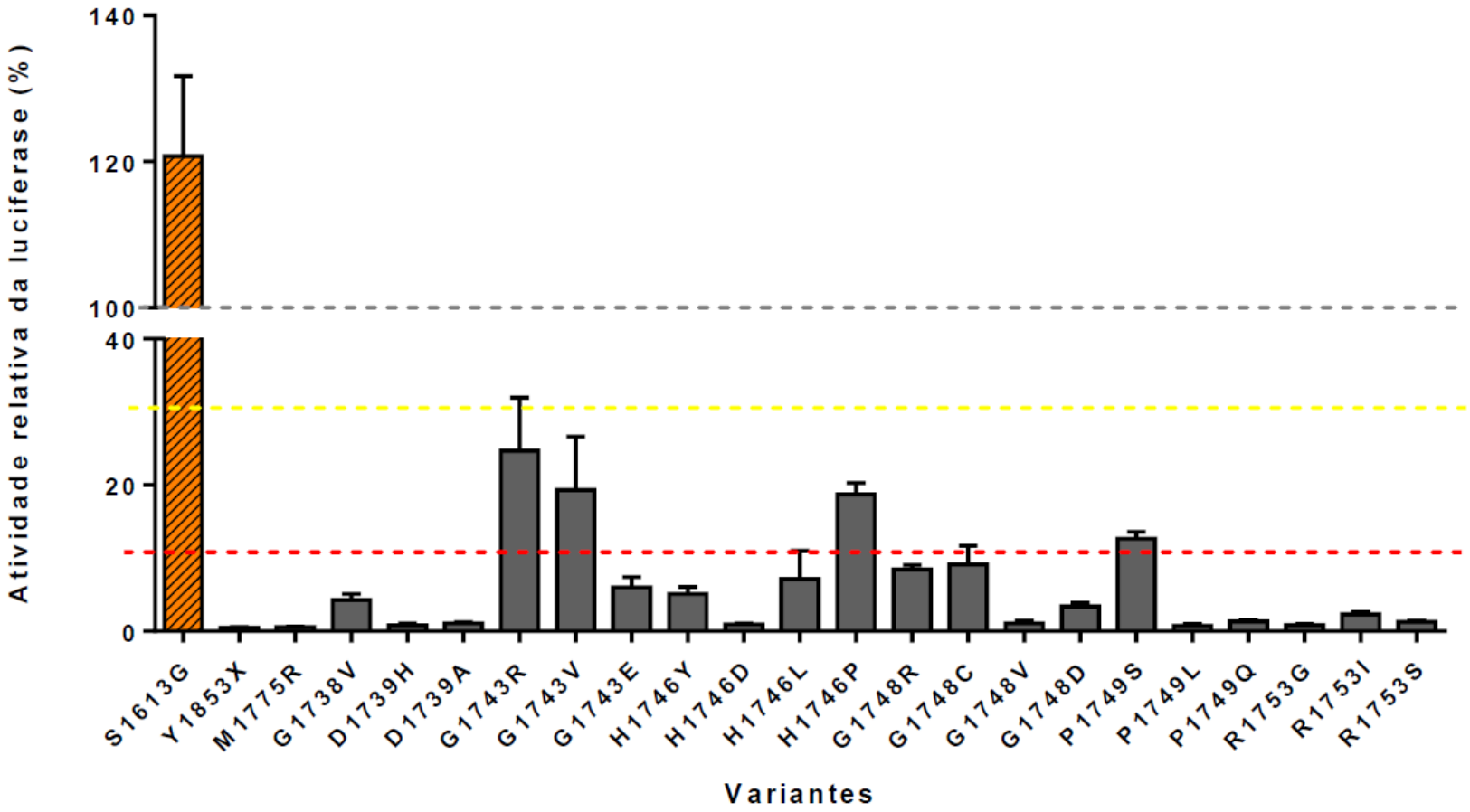


Figura 4.2 – Análise da atividade de ativação transcricional a 37ºC. O gráfico apresentado é representativo de um experimento. As células foram co-transfectadas com construções codificantes para a proteína de fusão GAL4DBD:BRCA1$_{13/24}$ selvagem ou variante, além das construções codificantes do sistema repórter (luciferase). A atividade dos variantes está apresentada em relação à atividade da construção BRCA1 WT, os desvios padrão das quadruplicatas estão indicados nas barras. Os controles estão representados pelas barras hachuradas, sendo o controle positivo o laranja e os controle negativos os vermelhos. As linhas indicam as atividades de: 100% (cinza), 30% (amarelo) e 10% (vermelho).

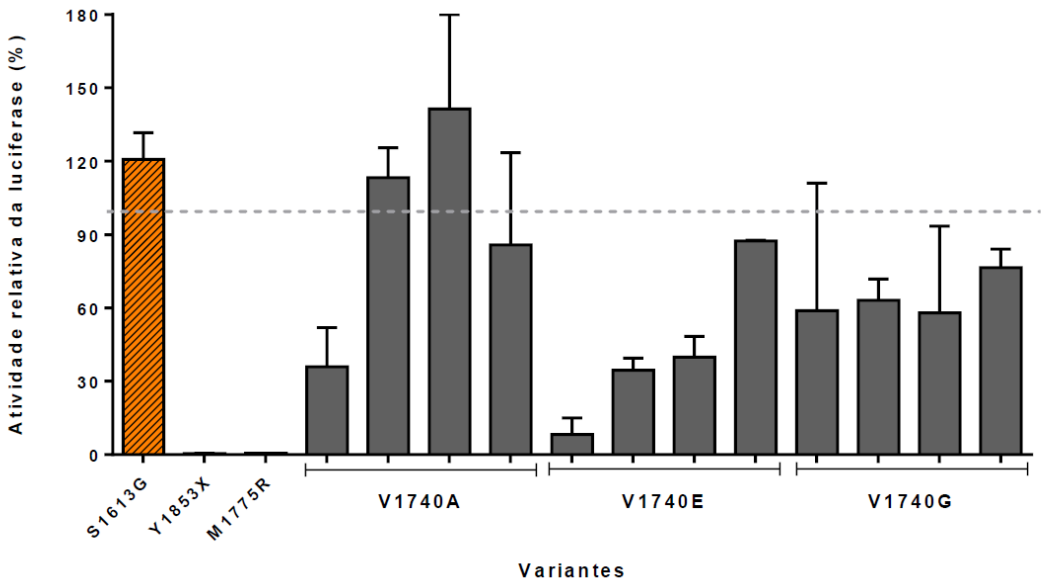


Figura 4.3 – Análise da atividade de ativação transcricional a 37ºC dos variantes na posição 1740. O gráfico apresenta todos os experimentos realizados com os variantes V1740A, V1740E e V1740G. As células foram co-transfectadas com construções codificantes para a proteína de fusão GAL4DBD:BRCA1$_{13/24}$ selvagem ou variante, além das construções codificantes do sistema repórter (luciferase). A atividade dos variantes está apresentada em relação à atividade da construção BRCA1 WT, os desvios padrão das quadruplicatas estão indicados nas barras. Os controles estão representados pelas barras hachuradas, sendo o controle positivo o laranja e os controle negativos os vermelhos. A linha cinza indica a atividade de 100%.

4.4. Avaliação da capacidade de ativação transcricional a 30ºC

A grande variação observada nas análises de ativação transcricional de variantes localizados na posição 1740 (V1740A, V1740E e V1740G) remeteu ao comportamento termossensível previamente descrito para variantes missense localizados na região C-terminal de BRCA1 (CARVALHO *et al*, 2002). Assim, todos os variantes selecionados para o estudo, bem como os controles, foram avaliados quanto à capacidade de ativação transcricional a 30ºC.

A **Fig. 4.4**. mostra os dados de ativação transcricional dos 23 variantes analisados a 30ºC. De maneira geral os variantes G1738V, D1739A, H1746D, G1748R e P1749L reproduziram os baixos níveis de ativação observados a 37ºC, semelhantes aos níveis de atividade dos controles negativos. Cabe ressaltar que esses controles (Y1853X e M1775R), assim como os variantes, apresentaram um discreto aumento de atividade relativa.

Já os variantes D1739H, G1743R, G1743V, G1743E, H1746P, H1746Y, H1746L, G1748C, G1748V, G1748D, P1749S, P1749Q, R1753G, R1753I e R1753S mostraram um aumento na sua capacidade de ativação transcricional quando comparados aos controles, apresentando valores que variaram de 14,4% (R1753S) à 176,27% (H1746L).

Todos os variantes foram avaliados em 2 experimentos independentes (em quadruplicata), sendo apresentado os resultados de um experimento representativo (**Fig. 4.4**. e **4.5.**).

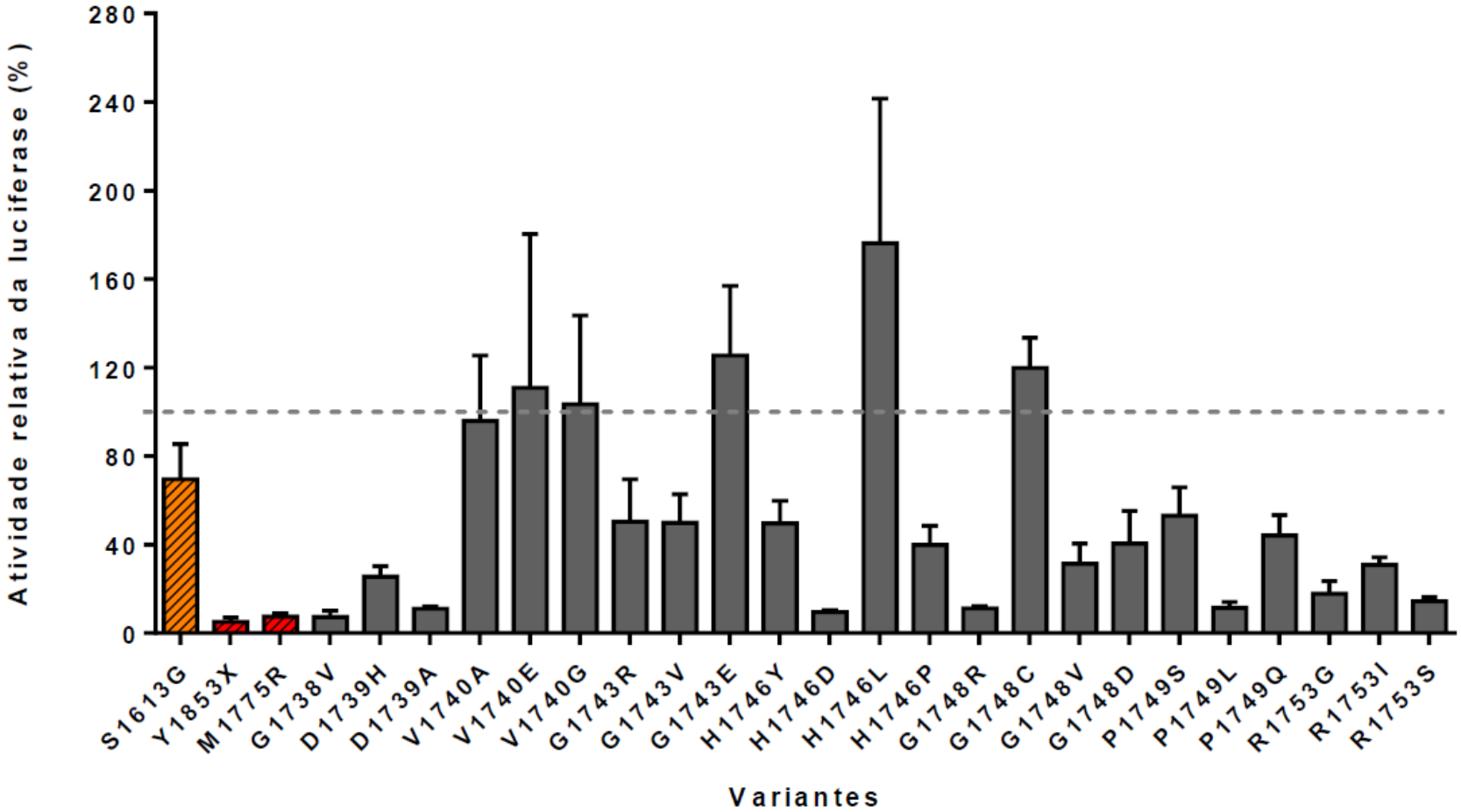
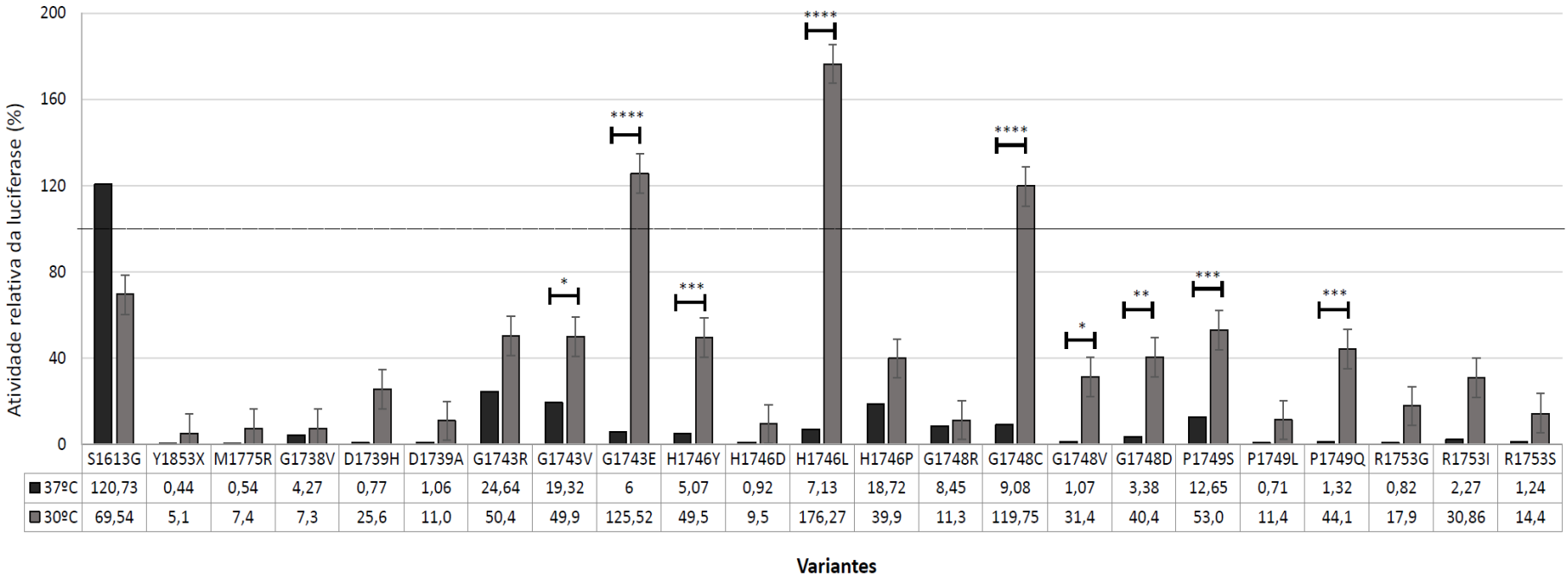


Figura 4.4 – Análise da atividade de ativação transcricional à 30ºC. O gráfico apresentado é representativo de um experimento. As células foram co-transfectadas com construções codificantes para a proteína de fusão GAL4DBD:BRCA1$_{13/24}$ selvagem ou variante, além das construções codificantes do sistema repórter (luciferase). A atividade dos variantes está apresentada em relação à atividade da construção BRCA1 WT, os desvios padrão das quadruplicatas estão indicados nas barras. Os controles estão representados pelas barras hachuradas, sendo o controle positivo o laranja e os controle negativos os vermelhos. A linha em cinza indica a atividade de 100%

Em uma comparação cuidadosa entre os ensaios a 37ºC e 30ºC (**Figura 4.5**), observa-se que apenas 9 variantes apresentam diferenças estatisticamente significativas entre as duas condições de análise: G1743V, G1743E, H1746Y, H1746L, G1748C, G1748V, G1748D, P1749S e P1749Q – foram observadas atividades relativas até 25 vezes maiores que as observadas a 37ºC (H1746L).

Os demais variantes (D1739H, G1743R, H1746P, R1753G, R1753I e R1753S) apresentaram aumentos de atividade relativa a 30ºC quando comparados aos ensaios a 37ºC, no entanto as variações não foram estatisticamente significativas (**Figura 4.5**).

Os variantes na posição 1740 (V1740A, V1740E e V1740G) apresentaram níveis de ativação semelhantes aos controles positivos, no entanto sua ampla variação, previamente observada a 37ºC, também foi observada a 30ºC (**Fig. 4.4**). Os resultados obtidos a 37ºC e 30ºC não foram avaliados comparativamente. As variações observadas nos ensaios independentes e também entre as replicatas de um mesmo ensaio inviabilizaram a determinação da atividade média de forma consistente.



	S1613G	Y1853X	M1775R	G1738V	D1739H	D1739A	G1743R	G1743V	G1743E	H1746Y	H1746D	H1746L	H1746P	G1748R	G1748C	G1748V	G1748D	P1749S	P1749L	P1749Q	R1753G	R1753I	R1753S
37ºC	120,73	0,44	0,54	4,27	0,77	1,06	24,64	19,32	6	5,07	0,92	7,13	18,72	8,45	9,08	1,07	3,38	12,65	0,71	1,32	0,82	2,27	1,24
30ºC	69,54	5,1	7,4	7,3	25,6	11,0	50,4	49,9	125,52	49,5	9,5	176,27	39,9	11,3	119,75	31,4	40,4	53,0	11,4	44,1	17,9	30,86	14,4

Figura 4.5 – Comparação da atividade de ativação transcricional a 30ºC e 37ºC. O gráfico apresentado é representativo de um experimento à cada temperatura. As células foram co-transfectadas com construções codificantes para a proteína de fusão GAL4DBD:BRCA1$_{13/24}$ selvagem ou variante, além das construções codificantes do sistema repórter (luciferase) e mantidas a 30°C ou 37ºC. A atividade dos variantes está apresentada em relação à atividade da construção BRCA1 WT, relativo à temperatura testada, os desvios padrão das quadruplicatas estão indicados nas barras. A linha em cinza indica a atividade de 100%. *0,01<p<0,5; **0,001<p<0,01; ***0,0001<p<0,001; ****p<0,0001.

4.5. Síntese da proteína quimérica

A síntese das proteínas de fusão GAL4DBD:BRCA1$_{13/24}$ foi analisada por Western *blotting* seguido de imunodetecção, com anticorpo específico para o domínio de ligação ao DNA de GAL4 (GAL4 DBD).

Os variantes G1738V, D1739H, D1739A, V1740A, V1740G G1743R, G1743V, H1746Y, H1746P, G1748R, G1748V, G1748D, P1749S, P1749L, P1749Q, R1753G e R1753S apresentaram níveis proteicos detectáveis e como esperado a proteína quimérica GAL4DBD:BRCA1$_{13/24}$ foi identificada com peso molecular aproximado de 70kDa, como pode ser observado na **Fig. 4.6**. As duas formas (banda dupla) observadas podem ser explicadas por eventos de modificação pós-traducional (fosforilação) observados nessa região (CARVALHO *et al*, 2007).

Já os variantes V1740E, G1743E, H1746D, H1746L, G1748R, G1748C e R1753I não apresentaram níveis detectáveis.

Os variantes V1740E e H1746D também tiveram a síntese da proteína de fusão avaliada quando as células foram mantidas à 30ºC por 48 horas após a transfecção, os resultados podem ser observados na **Fig. 4.7**. Diferente do observado anteriormente, níveis detectáveis da proteína quimérica GAL4DBD:BRCA1$_{13/24}$ foram observados com peso molecular esperado para o variante V1740E. Contudo o variante H1746D manteve níveis não detectáveis.

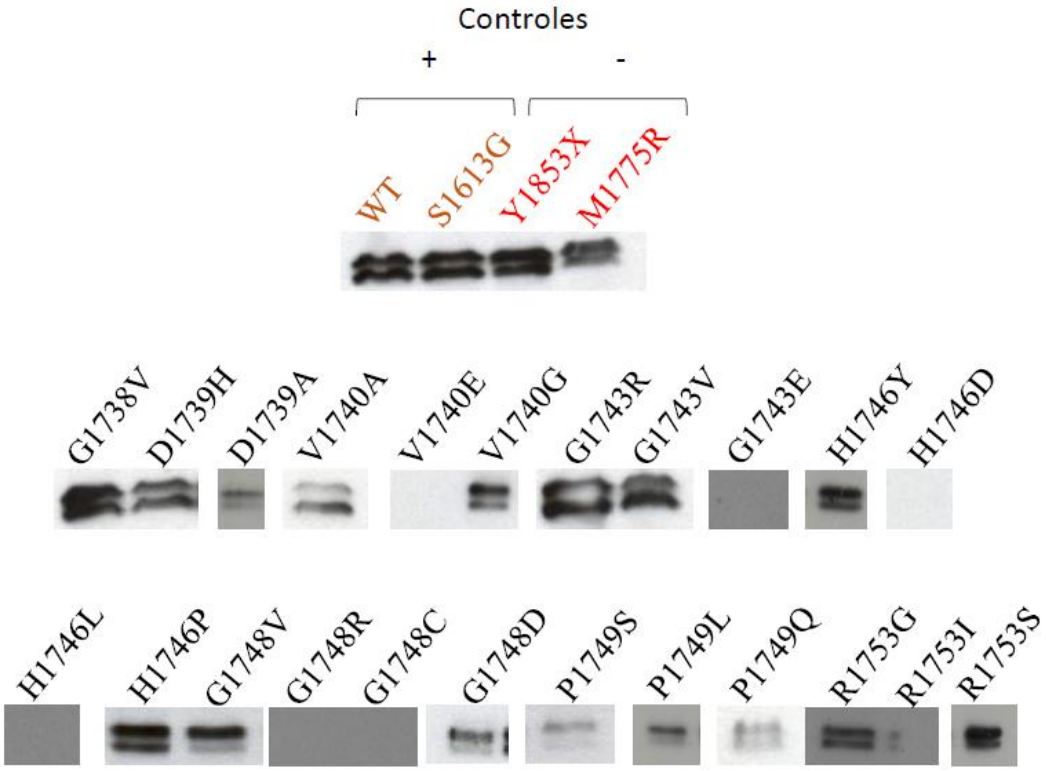


Figura 4.6. Proteínas de fusão GAL4DBD:BRCA1$_{13/24}$ produzidas a 37ºC. As células transfectadas com as construções codificantes das proteínas quiméricas foram mantidas a 37ºC e os extratos foram obtidos 24 horas após a transfecção. O Western blotting seguido por imunodetecção foi realizado utilizando anticorpo anti-GAL4DBD. Os controles positivos estão destacados em laranja e os controles negativos estão destacados em vermelho. Os variantes V1740E, G1743E, H1746D, H1746L, G1748R, G1748C e R1753I não apresentaram níveis detectáveis da proteína GAL4DBD:BRCA1$_{13/24}$.

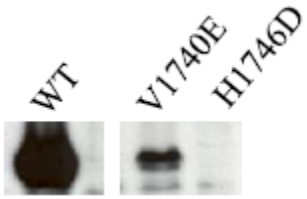


Figura 4.7. Proteínas de fusão GAL4DBD:BRCA1$_{13/24}$ produzidas a 30ºC: variantes V1740E e H1746D. As células transfectadas com as construções codificantes das proteínas quiméricas foram mantidas a 30ºC e os extratos foram obtidos 48 horas após a transfecção. O *Western blotting* seguido por imunodetecção foi realizado utilizando anticorpo anti-GAL4DBD. Os níveis do variante V1740E foram detectados, porém os níveis do variante H1746D continuaram indetectáveis.

4.6. Análise *in silico* do impacto das substituições na região *linker*

A estrutura 3D da região do tBRCT (aminoácidos 1396 a 1863) obtida no programa SPDBV está representada na **Fig 4.8**. A estrutura do *linker* está em consonância com a conformação descrita na literatura, com o segmento organizado em α-hélice mais próximo ao BRCT C-terminal.

O impacto das substituições dos resíduos de aminoácidos na estrutura secundária do domínio tBRCT de BRCA1 (região codificada pelos éxons 16 a 24) foi avaliada através de diagramas de Ramachandran (RAMACHANDRAN *et al*, 1963). O diagrama de Ramachandran permite avaliar o provável posicionamento de cada resíduo de aminoácido com base nos ângulos de suas ligações do carbono α (Phi (φ) e Psi (ψ)). O diagrama gerado pelo programa SPDBV tem como base as regiões de prevalência determinadas por Ramachandran e colaboradores (1963), como mostra a **Fig. 4.9**.

O diagrama de Ramachandran apresenta majoritariamente três regiões. As regiões em amarelo, que indicam zonas permitidas (o conjunto de ângulos favorece a formação de uma determinada estrutura secundária), as regiões em azul indicam regiões aceitáveis (o conjunto de ângulos pode favorecer a formação uma determinada estrutura secundária) e as regiões não delimitadas são zonas não permitidas (o conjunto de ângulos não favorece a formação de estruturas secundárias). A região em amarelo no quadrante superior esquerdo do diagrama demonstra ângulos favoráveis à formação de folha β, a região em amarelo no quadrante inferior esquerdo indica ângulos favoráveis à formação de α-hélice e a região em amarelo no quadrante superior direito indica ângulos que favorecem a formação de uma α-hélice invertida (**Fig 4.9**)

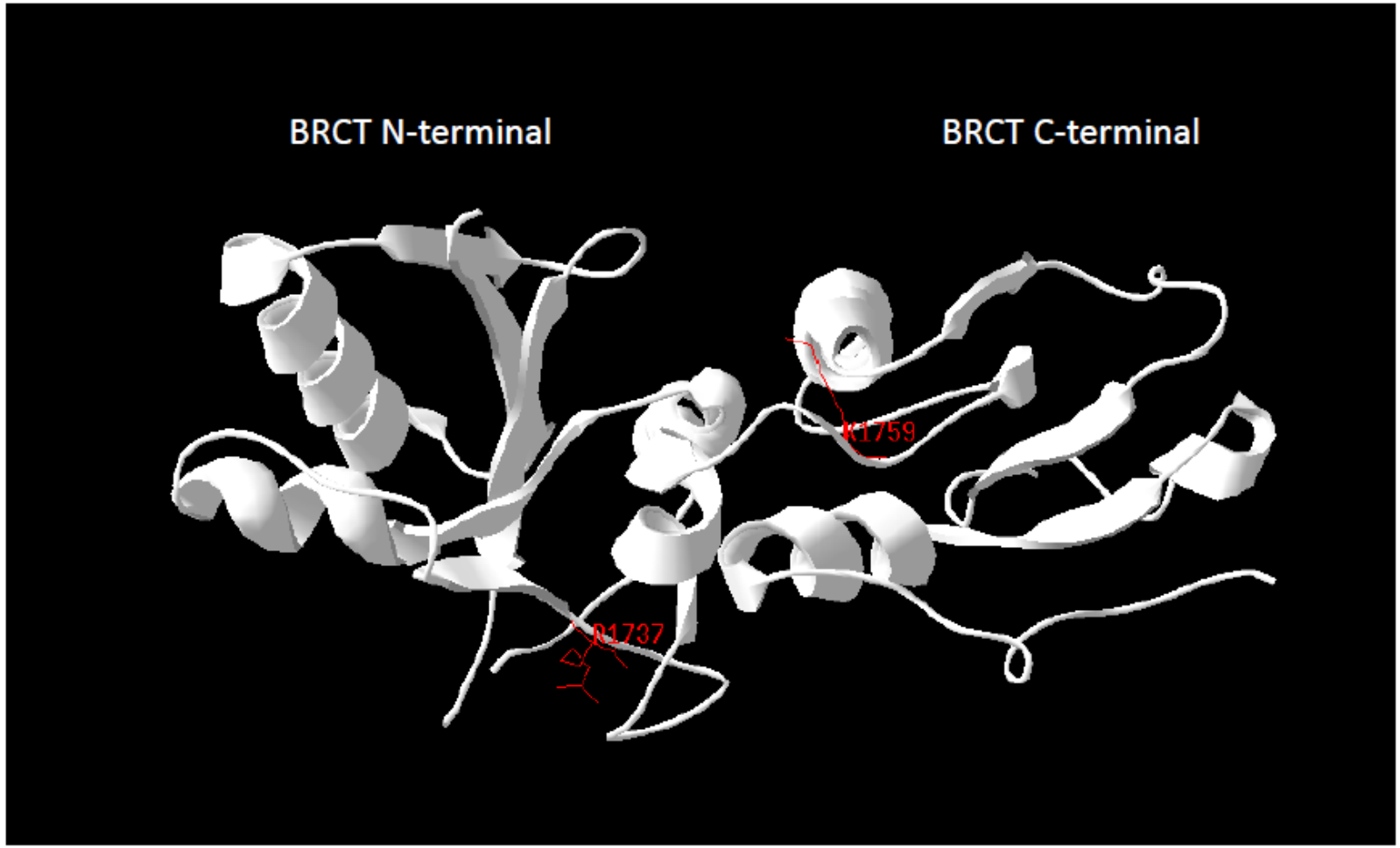


Figura 4.8 – Estrutura tridimensional dos BRCTs. Estrutura tridimensional dos BRCTs, obtida no programa SPDBV com base na estrutura 1JNX, demonstrando as estruturas secundárias formadas. A esquerda o BRCT N-terminal e a direita o BRCT C-terminal. Os resíduos de aminoácidos 1737 e 1759 estão indicados em vermelho delimitando o *linker*.

Figura 4.9 – Delimitação de áreas do diagrama de Ramachandran. O diagrama correlaciona os ângulos Phi, φ, (eixo X) e Psi, ψ, (eixo Y). As áreas delimitadas em amarelo são as regiões permitidas e em azul aceitável. Cada combinação de ângulos favorece a formação de uma estrutura secundária, que estão discriminas dentro das regiões amarelas. Quadrante superior esquerdo favorece a folha β, o quadrante inferior esquerdo favorece a α-hélice e o quadrante superior direito favorece a α-hélice invertida. As regiões externas são denominas não permitidas.

A análise do posicionamento dos 23 aminoácidos que compõem a região do *linker* mostra os resíduos de aminoácidos comprometidos com a formação da α-hélice (G1748 até E1754) (**Fig. 4.10**).

A análise das mutações gerou diagramas individuais dos variantes (**Fig. 4.11 a 4.14**). De maneira geral os variantes que apresentavam resíduos alvo de substituição localizados em regiões não permitidas, depois da substituição, mantiveram seu perfil com ligeiras modificações nos ângulos φ e ψ - sem impacto na estrutura predita para a α-hélice ou nos ângulos dos demais resíduos de aminoácidos (G1738V, D1739H, D1739A, V1740A, V1740E, V1740G, G1743R, G1743V, G1743E) (**Figs. 4.11, 4.12 A-C**).

Uma exceção a esse padrão (localização em região não permitida antes e depois da substituição) são os variantes do resíduo de aminoácido H1746. Os variantes apresentam combinações de ângulos que coincidem com a região aceitável para estruturação de folha β (H1746D, H1746D, H1746L e H1746P), como mostram as **Figs. 4.12 D-F** e **4.13A**.

Os variantes cujos resíduos alvo de substituição estavam originalmente localizados na região determinante de α-hélice, quando modificados, alteraram os ângulos φ e ψ de todos os resíduos de aminoácidos que compõem a estrutura da α-hélice, no entanto, essas alterações não resultaram em modificações preditas da estrutura secundária, uma vez que as combinações de ângulos resultantes ainda favorecem a formação da estrutura em α-hélice (G1748R, G1748C, G1748V, G1748D, P1749S, P1749L, P1749Q, 1753G, 1753I e 1753S) (**Figs. 4.13 B-F** e **4.14A-E**).

Apesar das alterações demonstradas nos diagramas de Ramachandran de alguns variantes, nenhuma das substituições alterou significativamente a combinação de ângulos dos resíduos de aminoácidos, de modo a alterar a estrutura secundária do *linker* no diagrama de Ramachandran.

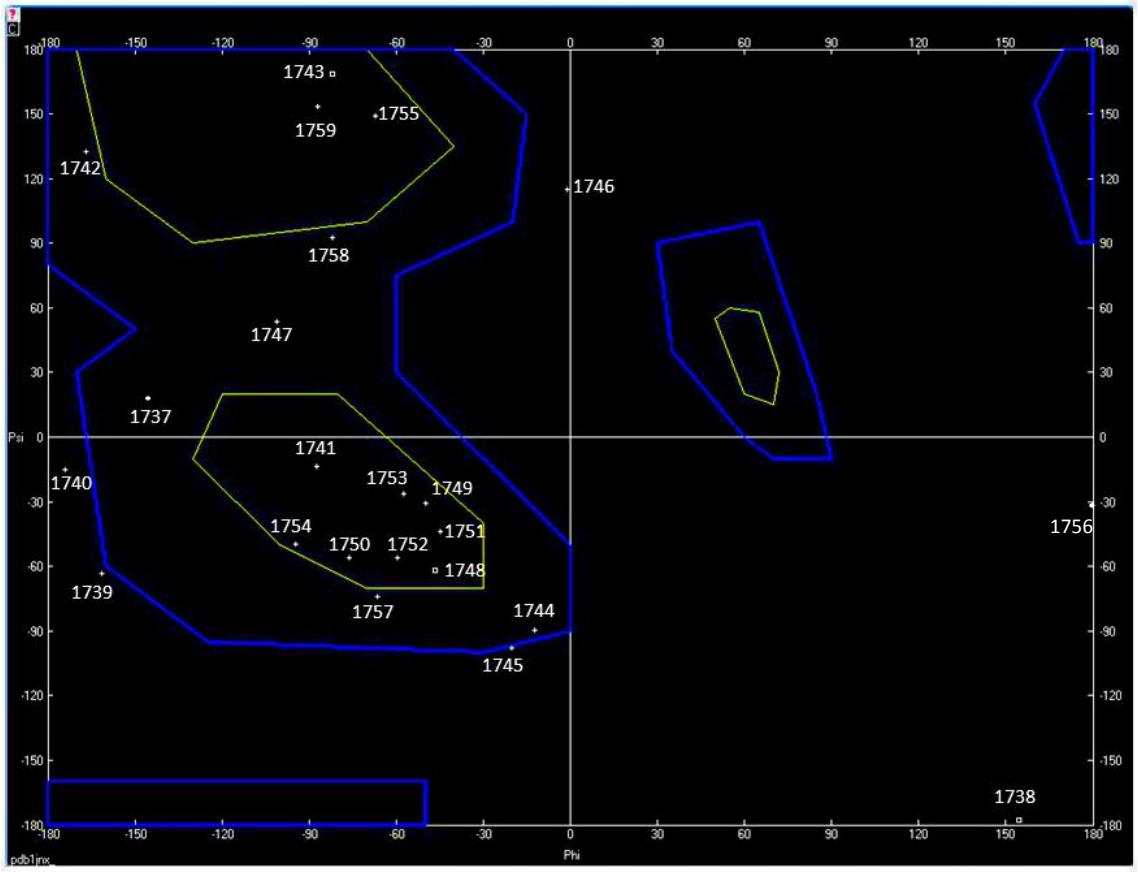


Figura 4.10 – Distribuição dos resíduos de aminoácidos do *linker* no diagrama de Ramachandran. O eixo X corresponde ao ângulo φ e o eixo Y corresponde ao ângulo ψ. No diagrama estão representados os aminoácidos do *linker*. Todos estão representados por pontos, exceto a glicina que é representada por um quadrado. Os números indicam a posição de cada aminoácido na estrutura proteica de BRCA1.

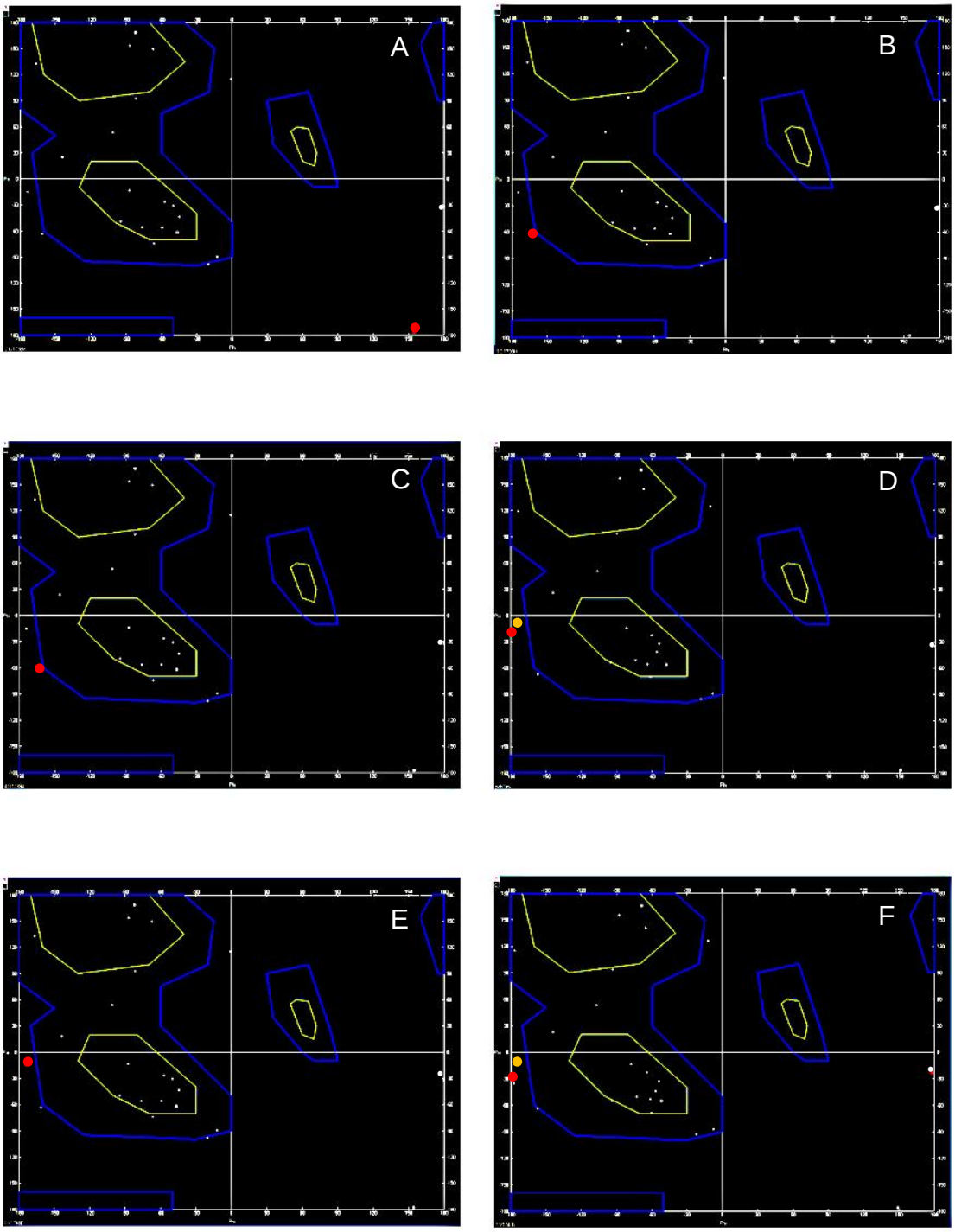


Figura 4.11 – Perfil de distribuição dos resíduos de aminoácidos no diagrama de Ramachandran. O eixo X corresponde ao ângulo φ e o eixo Y corresponde ao ângulo ψ. Os aminoácidos estão representados pelos pontos brancos. O aminoácido substituto está em vermelho e a posição original do aminoácido está em amarelo (a ausência do amarelo indica que a posição ocupada é a mesma). O variante G1738V (**A**) mantém a posição do aminoácido original (quadrante inferior direito). Os variantes D1739H (**B**) e D1739A (**C**) mantêm a posição original para ambas substituições (quadrante inferior esquerdo). Enquanto os variantes V1740A (**D**) e V1740G (**F**) alteram em poucos graus os ângulos de ligação, o variante V1740E (**E**) mantém a posição original (quadrante inferior esquerdo).

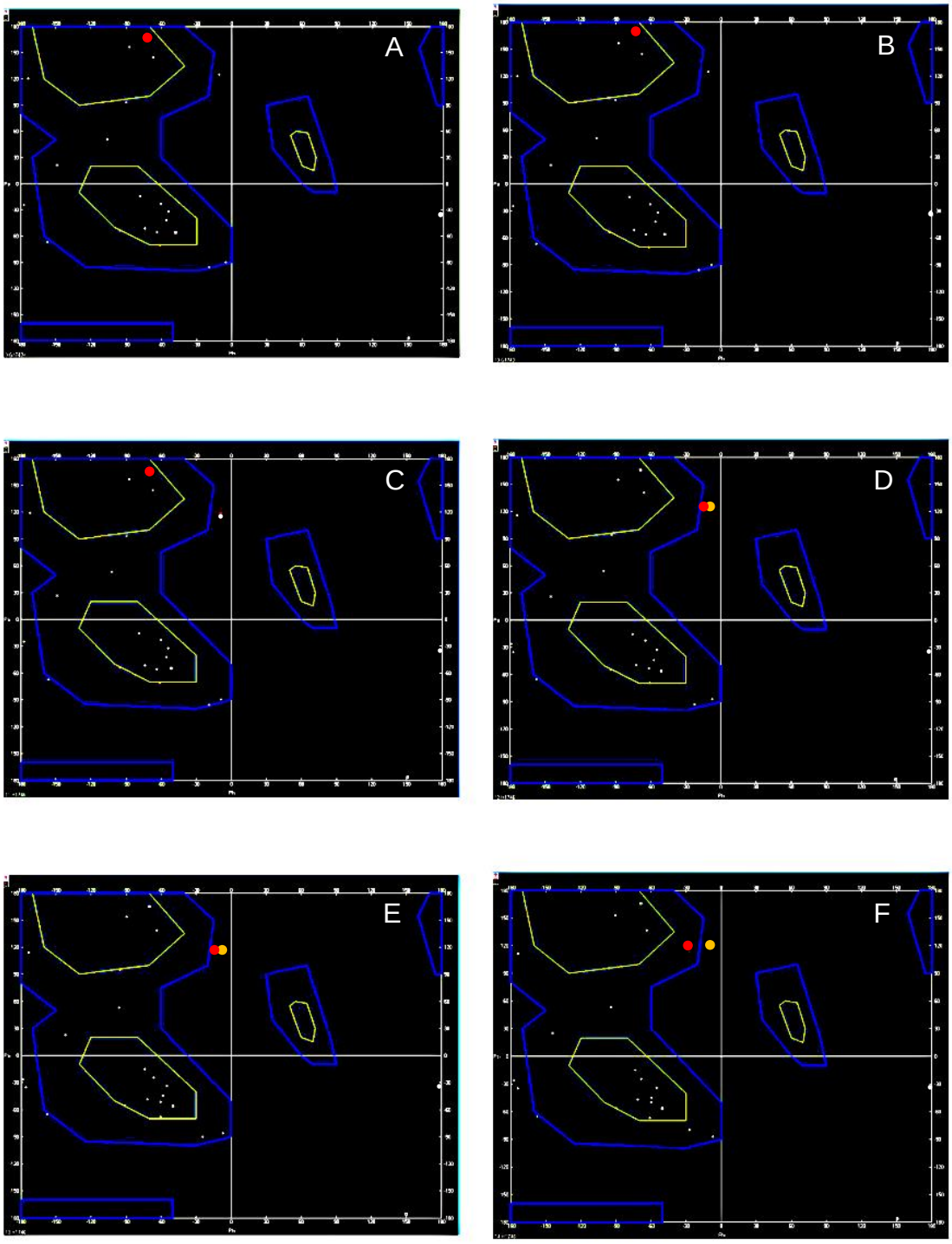


Figura 4.12 – Perfil de distribuição dos resíduos de aminoácidos no diagrama de Ramachandran. O eixo X corresponde ao ângulo φ e o eixo Y corresponde ao ângulo ψ. Os aminoácidos estão representados pelos pontos brancos. O aminoácido substituto está em vermelho e a posição original do aminoácido está em amarelo (a ausência do amarelo indica que a posição ocupada é a mesma). Os variantes G1743R (**A**), G1743V (**B**) e G1743E (**C**) não levam a alterações perceptíveis (quadrante superior esquerdo). Os variantes H1746Y (**D**), H1746D (**E**), H1746L (**F**) conferem uma mudança nos ângulos de ligação, determinando uma combinação de ângulos favorável à estrutura secundária (quadrante superior esquerdo)

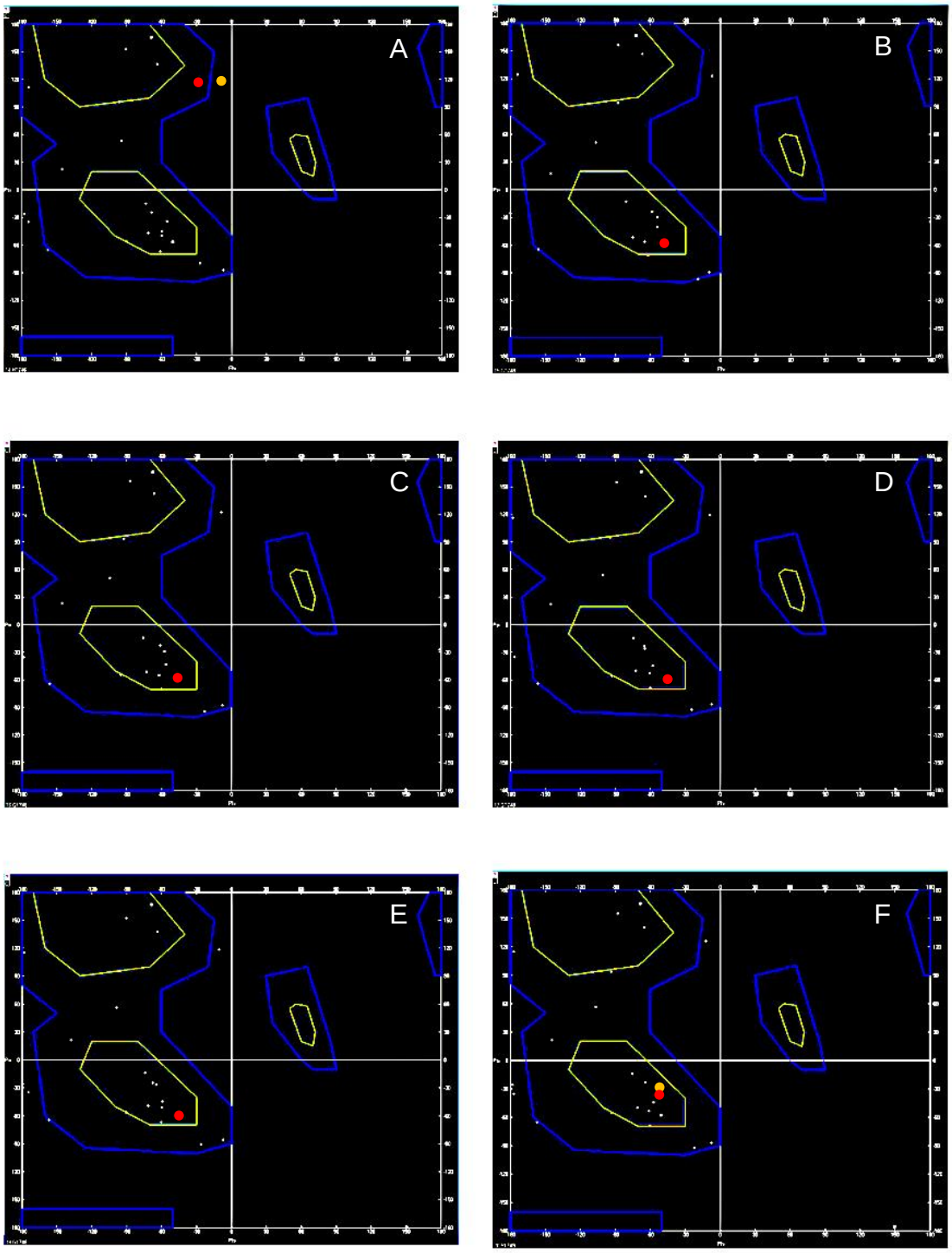


Figura 4.13 – Perfil de distribuição dos resíduos aminoácidos no diagrama de Ramachandran. O eixo X corresponde ao ângulo φ e o eixo Y corresponde ao ângulo ψ. Os aminoácidos estão representados pelos pontos brancos. O aminoácido substituto está em vermelho e a posição original do aminoácido está em amarelo (a ausência do amarelo indica que a posição ocupada é a mesma). O variante H1746P (**A**) determina ângulos aceitáveis, diferente do aminoácido original (quadrante superior esquerdo). Os variantes G1748R (**B**), G1748C (**C**), G1748V (**D**) e G1748D (**E**) ocasionaram alterações nos ângulos ligação dos resíduos da α-hélice, mas sem alterar a combinação ângulos permitida para α-hélice (quadrante inferior esquerdo). O variante P1749S (**F**) modificou os ângulos de ligação dentro da α-hélice, mas manteve a formação da estrutura secundária (quadrante inferior esquerdo).

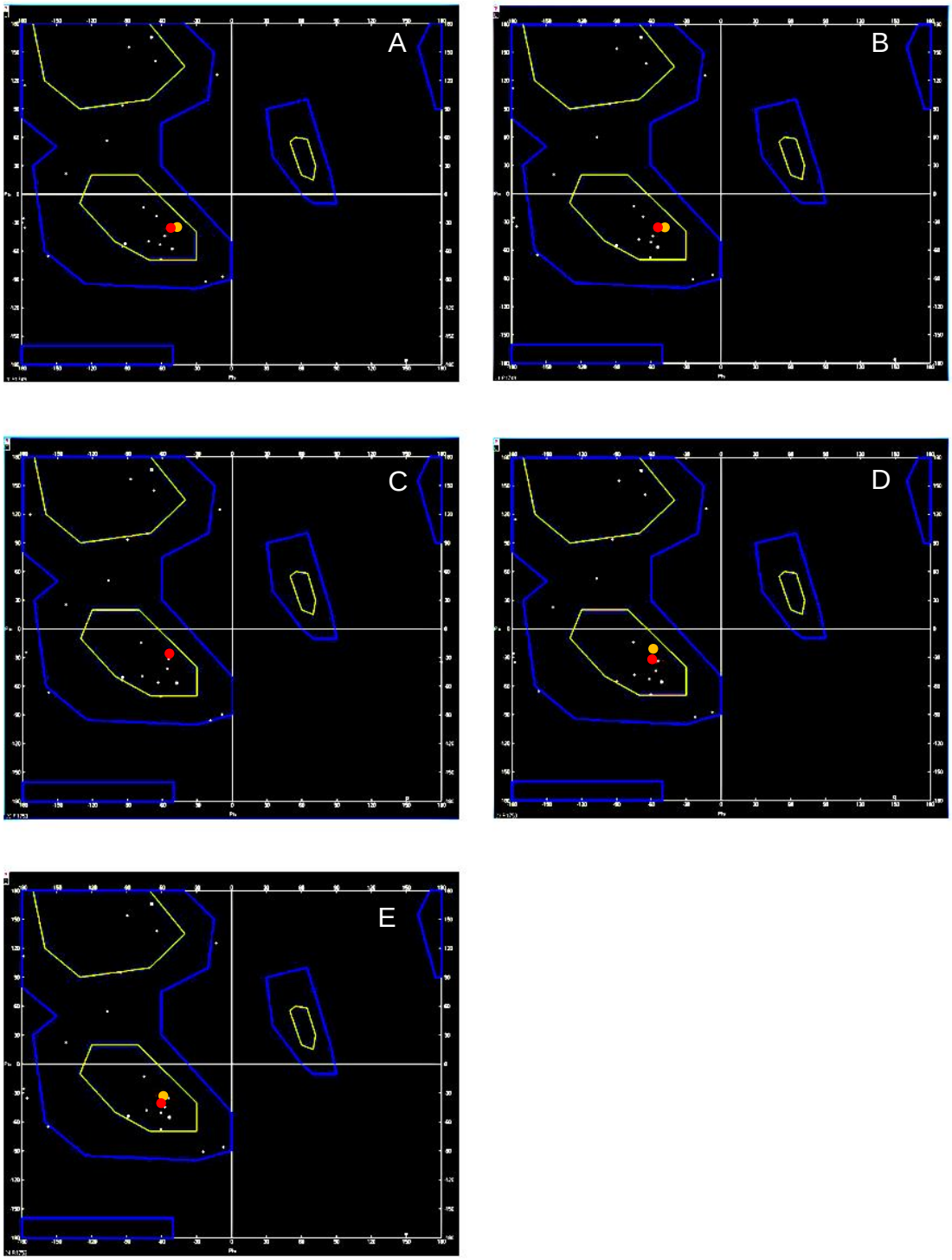


Figura 4.14 – Perfil de distribuição dos resíduos de aminoácidos no diagrama de Ramachandran. O eixo X corresponde ao ângulo φ e o eixo Y corresponde ao ângulo ψ. Os aminoácidos estão representados pelos pontos brancos. O aminoácido substituto está em vermelho e a posição original do aminoácido está em amarelo (a ausência do amarelo indica que a posição é ocupada é a mesma). Os variantes P1749L (**A**) e P1749Q (**B**) modificam os ângulos de ligação dentro da α-hélice, mas mantem a combinação de ângulos permitidos para α-hélice (quadrante inferior esquerdo). Os variantes R1753G (**C**), R1753I (**D**) e R1753S (**E**) levaram a alteração os ângulos dentro da região que favorece a formação da α-hélice, no entanto manteve a combinação de ângulos permitidos para a estrutura.

5. DISCUSSÃO

Desde a década de 70 a relação entre o câncer de mama e a hereditariedade é claramente estabelecida. Nessa época, já era observado que o risco de um indivíduo com histórico familiar de câncer de mama desenvolver essa mesma forma de câncer é maior que o risco encontrado na população (KELSEY, 1979). Na década de 90, o gene *BRCA1* foi identificado, estabelecendo a primeira correlação direta entre um gene e casos de câncer de mama hereditário (HALL *et al*, 1990; EASTON *et al*, 1993, MIKI et al, 1994).

A partir da identificação e clonagem de *BRCA1* (MIKI *et al*, 1994), a determinação da patogenicidade das alterações genéticas encontradas se tornou uma das questões mais importantes a ser esclarecida. Inicialmente, a maioria dos variantes identificados resultava em proteínas menores (truncadas e não funcionais) – consequência de mutações sem sentido (nonsense), inserções ou deleções, todas com estreita relação com o câncer. Até então, poucas alterações do tipo missense haviam sido reportadas (SMITH, 1996). Atualmente, as mutações missense são o segundo tipo mais reportado no BIC, correspondendo a quase 30% do total de entradas (BIC, IARC, 2016).

O conjunto de VNC de *BRCA1* é composto essencialmente por mutações que geram pequenas modificações na sequência primária da proteína (dificultando a inferência acerca do comprometimento da estrutura proteica) e de baixa frequência na população, essas alterações são majoritariamente do tipo missense (EASTON *et al*, 2007).

Determinar o risco de desenvolvimento de câncer para os variantes de *BRCA1* é uma importante estratégia para a prevenção. O carreador de uma mutação patogênica pode optar por rotinas de monitoramento mais frequentes, por cirurgias preventivas ou mesmo terapias profiláticas (EASTON *et al*, 2007).

A identificação de variantes patogênicos de *BRCA1* tem se mostrado importante não somente nas linhagens germinativas, mas também nas células tumorais, auxiliando na escolha de combinações terapêuticas a serem adotadas – o melhor exemplo disso é a susceptibilidade à platina e inibidores de PARP para tumores deficientes de BRCA1 (ROSEN & PISHVAIAN, 2014).

Na tentativa de estimar o risco que os carreadores de VNC tem para o desenvolvimento do câncer, diversos modelos preditores foram desenvolvidos.

Apesar disso, muitas vezes, as informações disponíveis ainda são insuficientes para uma avaliação segura (ENGEL & FISCHER, 2014).

Análise funcional

Os ensaios funcionais foram propostos com o intuito de aumentar o conjunto de dados sobre os VNC de *BRCA1* utilizando o conhecimento prévio derivado das interações e funções descritas para a proteína. Assim, esses ensaios correlacionam as propriedades bioquímicas ou funções de BRCA1 com a integridade estrutural da proteína. (COUCH *et al*, 2008; CARVALHO *et al*, 2009; MILLOT *et al*, 2012; BOWMAN *et al*, 2013).

O ensaio de TA foi incialmente desenvolvido usando o fragmento proteico correspondente à região codificante limitada aos éxons 16 a 24, resíduos de aminoácidos 1560 a 1863 (MONTEIRO *et al*, 1996; MONTEIRO *et al*, 1997; HAYES, *et al*, 2000; Vallon-Christersson *et al*, 2001). Posteriormente, Phelan e colaboradores, em 2005, estenderam a região de cobertura do ensaio aos éxons 13 a 24 (resíduos 1396 a 1863) – sendo validado por Carvalho e colaboradores em 2007 (PHELAN *et al*, 2005; CARVALHO *et al*, 2007; CARVALHO *et al*, 2009; LEE *et al*, 2010; CARVALHO *et al*, 2014).

A análise de TA permite uma correlação entre a atividade transcricional e a integridade dos domínios tBRCT e sua circunvizinhança. Como essa estrutura é fundamental para o papel supressor de tumor de BRCA1, pôde-se consolidar o ensaio de TA como um bom preditor de patogenicidade das mutações em *BRCA1*. No entanto, a classificação de um variante não deve ser definida exclusivamente por um único parâmetro, são necessárias mais informações para que se construa o perfil de patogenicidade (ou não) do variante (CARVALHO *et al*, 2007; CARVALHO *et al*, 2009).

Neste estudo, para a realização do ensaio de TA, foi usada uma construção encerrando um gene repórter (codificante para luciferase de *P. pyralis*, sob o controle de operadores de GAL4- sítios de reconhecimento para o domínio de ligação ao DNA de GAL4). Assim, a produção da luciferase de *P. pyralis* depende não só da capacidade de recrutamento da RNA polimerase II pela proteína GAL4DBD:BRCA1$_{13/24}$, mas também da eficiência de transfecção das construções repórter e pCDNA3-BRCA1$_{13/24}$. Para diminuir a influência da variação dessa eficiência nos resultados foi usado um controle de transfecção,

a construção pGR-TK, que contém o gene codificante da luciferase de *R. Reniformis* controlado pelo promotor da timidina quinase (HSV-TK), apresentando expressão constitutiva e funcionando como um elemento normalizador. Nesse sistema a produção da luciferase *R. Reniformis* depende unicamente da eficiência de transfecção. Logo, o uso da razão entre os dois valores de atividade das luciferases minimiza variações decorrentes da eficiência de transfecção.

Originalmente, ao analisarmos o conjunto de todos os possíveis variantes naturais missense encontrados na região do *linker* de BRCA1, através dos algoritmos de predição *in silico*, observamos 23 variantes com um perfil preditivo patogênico – todos usados neste estudo.

O ensaio de TA a 37ºC (**Fig. 4.2**) revelou que 20 dos variantes estudados (G1738V, D1739H, D1739A, G1743R, G1743V, G1743E, H1746Y, H1746D, H1746L, H1746P, G1748R, G1748C, G1748V, G1748D, P1749S, P1749L, P1749Q, R1753G, R1753I e R1753S) apresentaram atividades semelhantes aos controles negativos e, portanto, em concordância com os resultados preditos originalmente pelos algoritmos.

Curiosamente, revisitando a análise dos 23 variantes no algoritmo Polyphen, mas agora usando sua nova versão (Polyphen-2), 4 variantes que inicialmente foram preditos com um perfil patogênico foram classificados como benignos: D1739A, G1743R, H1746L e R1753G (**Tabela 4.1**) – discordando dos dados funcionais obtidos (**Fig. 4.2**).

Discordâncias entre dados funcionais e algoritmos preditivos não são eventos incomuns, na verdade, também são observadas inconsistências entre predições de diferentes algoritmos para um mesmo variantes – como pode ser observado na plataforma BRCA1 Circos (LEE *et al*, 2010; BIC, 2016).

Os variantes estudados na posição 1740 (V1740A, V1740E e V1740G) apresentaram atividades elevadas no ensaio de TA a 37ºC, comparáveis aos controles não patogênicos, o que poderia ser um indicativo de neutralidade da mutação (no ensaio de TA) (**Fig. 4.3**). Porém, esses variantes exibiram grande variação no ensaio de TA, (dentro do mesmo ensaio, demonstrado pelo amplo desvio padrão, e entre experimentos, com variações de até 10 vezes na atividade), o que poderia indicar um comportamento termossensível.

A experiência prévia do grupo sugere que variantes termossensíveis de BRCA1 costumam apresentar variações amplas do ensaio de TA, não somente entre os ensaios independentes, mas também entre as replicatas de um mesmo ensaio – como o comportamento apresentado pelos variantes na posição 1740. Variantes termossensíveis de *BRCA1* apresentam o fenótipo de ativação transcricional limitada a 37ºC, porém quando testados a 30ºC apresentam transcrição do gene repórter igual ou superior a 30% do controle selvagem (CARVALHO *et al*, 2002; WORLEY *et al*, 2002). Ao contrário do se supunha, os variantes V1740A, V1740E e V1740G não modificaram seu comportamento quando testados a 30ºC, mantendo as variações observadas a 37ºC (amplitude do desvio padrão). É razoável especular que esse comportamento seja inerente a alguma instabilidade estrutural causada pela alteração desse resíduo (**Fig 4.4**).

No ensaio de TA a 30ºC também não foram observadas alterações no comportamento dos variantes G1738V, D1739H, D1739A, G1743R, H1746D, H1746P, G1748R, P1749L, R1753G, R1753I e R1753S, que mantiveram níveis de ativação transcricional semelhantes aos controles negativos (abaixo de 30%).

Por outro lado, os variantes G1743V, G1743E, H1746Y, H1746L, G1748C, G1748V, G1748D, P1749S e P1749Q mostraram um aumento estatisticamente significativo da atividade e níveis de atividade superiores a 30% (**Fig 4.5**). Esses variantes foram considerados termossensíveis, segundo critérios definidos por Carvalho e colaboradores (2002).

A presença de 9 variantes termossensíveis em um grupo de 23 variantes sugere uma possível predisposição à termossensibilidade da região *linker* de BRCA1.

Ainda não existe uma correlação definida entre o comportamento termossensível de variantes de BRCA1 e a fisiologia do câncer. Uma hipótese sugerida é a diferença de susceptibilidade entre o câncer de mama e o de ovário, onde pequenas variações de temperatura observadas entre os órgãos seriam determinantes para a funcionalidade da proteína (WORLEY *et al*, 2002).

Segundo os critérios estabelecidos por Carvalho e colaboradores (2007), variantes de BRCA1 que apresentam atividade transcricional inferior a 45% do controle selvagem seriam considerados patogênicos, enquanto variantes com atividade superior a 50% seriam considerados não patogênicos.

Considerando este critério, e os resultados obtidos neste estudo, pode-se determinar uma classificação preliminar dos variantes analisados, como mostra a **Tabela 5.1**.

Tabela 5.1 – **Classificação preliminar dos variantes**. A classificação preliminar dos variantes com base nos resultados do ensaio de TA a 37ºC, adoptando os critérios de Carvalho e colaboradores (2007) e resumos dos dados obtidos. VNC – Variante não classificado

Variante	Atividade a 37ºC (% selvagem)*	Classificação	Atividade a 30ºC (% selvagem)*	Termosensível?
G1738V	4,3 ±0,82	Patogênico	7,3 ±2,9	Não
D1739H	0,8 ±0,29	Patogênico	25,6 ±4,6	Não
D1739A	1,1 ±0,1	Patogênico	11,0 ±1,0	Não
V1740E	35,9 ±16,1 – 141,3 ±38,6	VNC	95,9 ±29,5	Não
V1740A	8,2 ±6,8 – 87,4 ±0,3	VNC	110,8 ±69,4	Não
V1740G	58,0 ±35,4 – 76,5 ±7,6	Não Patogênico	103,5 ±40,0	Não
G1743R	24,6 ±7,2	Patogênico	50,4 ±19,2	Não
G1743V	19,3 ±7,2	Patogênico	49,9 ±12,9	Sim
G1743E	6,0 ±1,4	Patogênico	125,5 ±31,4	Sim
H1746Y	5,1 ±0,9	Patogênico	49,5 ±10,3	Sim
H1746D	0,9 ±0,1	Patogênico	9,5 ±1,0	Não
H1746L	7,1 ±3,8	Patogênico	176,2 ±65,3	Sim
H1746P	18,7 ±1,5	Patogênico	39,9 ±8,6	Não
G1748R	8,5 ±0,5	Patogênico	11,3 ±0,8	Não
G1748C	9,1 ±2,6	Patogênico	119,7 ±13,7	Sim
G1748V	1,1 ±0,3	Patogênico	31,4 ±9,1	Sim
G1748D	3,4 ±0,4	Patogênico	40,4 ±14,9	Sim
P1749S	12,7 ±0,9	Patogênico	53,0 ±12,9	Sim
P1749L	0,7 ±0,2	Patogênico	11,4 ±2,6	Não
P1749Q	1,3 ±0,1	Patogênico	44,1 ±9,2	Sim
R1753G	0,8 ±0,1	Patogênico	17,9 ±5,6	Não
R1753I	2,3 ±0,3	Patogênico	30,9 ±3,49	Não
R1753S	1,2 ±0,2	Patogênico	14,4 ±1,8	Não

*média de um experimento representativo e desvio padrão.

Uma classificação mais consistente sobre a patogenicidade dos variantes estudados (termossensíveis ou não) será arbitrada após a análise dos dados de atividade transcricional pelo algoritmo VarCall – essa análise se baseia em um modelo matemático Bayesiano hierárquico que consolida (ou não) a significância das diferenças observadas entre controles do estudo e os variantes analisados, considerando todas as séries históricas (dados de ativação transcricional) já realizadas (IVERSEN *et al*, 2011, CARVALHO *et al*, 2014, WOODS *et al*, 2016).

No ensaio de TA, a identificação da adequada produção da proteína quimérica (GAL4DBD:BRCA1$_{13/24}$) é importante para validação da análise, em especial para os variantes que apresentam um comportamento reconhecido como patogênico – com baixa atividade transcricional.

A identificação da proteína foi realizada por rotina de Western *blotting* seguida de imunodetecção. Os variantes G1738V, D1739H, D1739A, V1740A, V1740G G1743R, G1743V, H1746Y, H1746P, G1748V, G1748D, P1749S, P1749L, P1749Q, R1753G e R1753S apresentaram níveis proteicos detectáveis em extratos obtidos de culturas mantidas a 37ºC (**Fig 4.6**). No entanto, nos extratos celulares, transfectados para produção dos variantes V1740E, G1743E, H1746D, H1746L, G1748R, G1748C e R1753I, a proteína GAL4DBD:BRCA1$_{13/24}$ não foi detectada.

A dificuldade de constatar a presença da proteína GAL4DBD:BRCA1$_{13/24}$ não representa, necessariamente, uma ausência de produção da mesma. Algumas mutações podem conferir uma menor estabilidade à proteína ou ao transcrito codificante, diminuindo sua meia vida e consequentemente sua abundância (CARVALHO *et al*, 2007). A quantidade de proteína transitoriamente disponível pode ser suficiente para a ativação da transcrição do gene repórter (cujo produto, a luciferase, é acumulada nas células) permitindo a avaliação de atividade, mas não sua imunodetecção. Em outro cenário, a quantidade de proteína disponível pode não ser suficiente para a ativação da transcrição do gene, apresentando baixos níveis de atividade transcricional. Por fim, há ainda a possibilidade de que quantidade de proteína transitoriamente disponível seja suficiente para uma ativação da transcricional, porém a estrutura da proteína faz dela intrinsecamente disfuncional, não havendo transcrição do gene repórter.

Devido ao perfil termossensível apresentado por alguns variantes, foram realizados ensaios preliminares para detecção dos níveis de produção de dois

variantes mantidos a 30ºC, V1740E e H1746D (que não foram detectados a 37ºC).

O variante V1740E apresentou atividade transcricional variável a 37ºC , o que necessariamente indica a presença da proteína de fusão, no entanto, a não identificação da proteína nas rotinas de imunodetecção sugerem uma meia vida curta (seja do transcrito ou da proteína) o que impede a detecção da quimera proteica (**Fig. 4.3**). No ensaio a 30ºC, foi possível detectar a proteína, como mostra a **Fig. 4.7**. Esse cenário sugere que, mesmo com o comprometimento da estabilidade, a quantidade transitoriamente disponível de proteína foi suficiente para a ativação da transcrição do gene repórter.

Já o variante H1746D exibiu níveis de atividade transcricional baixos, semelhante aos controle negativos, e não apresentou níveis proteicos detectáveis a 37ºC nem a 30ºC. Possivelmente, a mutação compromete severamente a produção da proteína (seja por mecanismos pós-transcricionais ou pós-traducionais), diminui drasticamente sua meia vida, comprometendo sua avaliação no ensaio de TA. Mas não podemos descartar a hipótese da proteína também ser intrinsecamente disfuncional.

Frente a esses resultados iniciais, cabe avaliar a produção dos demais variantes (G1743E, H1746L, G1748R, G1748R e R1753I) a 30ºC.

Análise estrutural

De modo geral a estrutura dos domínos tBRCT apresenta os dois BRCTs associados de maneira que a região N-terminal de um fique voltada para a região C-terminal do outro, como por exemplo a estrutura do tBRCT de BRCA1 (**Fig. 4.8**). A mobilidade entre os dois BRCTs é fundamental para sua ligação a fosfopeptídeos, pois permite uma mudança conformacional dos BRCTs no reconhecimento do peptídeo fosforilado, com uma consequente aproximação dos mesmos (WILLIAMS *et al*, 2004; WU *et al*, 2015).

O núcleo hidrofóbico formado entre os BRCTs e a flexibilidade do *linker* são o que determinam o peptídeo alvo de ligação - no caso de BRCA1, o motivo reconhecido é o pS-X-X-F.

A região do *linker* é fundamental para o posicionamento dos BRCTs e para sua ligação aos fosfopeptídeos, apesar disso, pouco se conhece sobre essa região de BRCA1 (WILLIAMS *et al*, 2004; WU *et al*, 2015).

A **Fig. 4.8** mostra a estrutura do tBRCT de BRCA1, onde na região do *linker* (delimitada pelos os resíduos de aminoácidos R1737 e K1759, indicados em vermelho) destaca-se uma estrutura em α-hélice (composta pelos resíduos de aminoácidos 1748 a 1753), que se apresenta mais próxima ao BRCT C-terminal, como descrito por WILLIAMS e colaboradores (2004).

A disposição dos resíduos de aminoácidos no diagrama de Ramachandran, obtido pela predição dos ângulos φ e ψ, formados pelas ligações do carbono α, são indícios de que essa combinação de ângulos é favorável a uma determinada formação de estrutura secundária, mas não é determinante para a formação da mesma. Para elucidar a estrutura secundária de um segmento proteico é necessário que sejam realizadas análises mais elaboradas do enovelamento da proteína, como, por exemplo, análise da estrutura por ressonância magnética nuclear ou por difração de raio X (BOURNE & WEISSIG, 2009).

O programa SPDBV, usado para a geração dos diagramas de Ramachandran dos variantes, emprega a demarcação de regiões estabelecida por Ramachandran e colaboradores em 1963. Determinando combinações de ângulos permitidos, aceitáveis e não permitidos para formação de estruturas secundárias (**Fig. 4.9**).

A análise de todos os resíduos do *linker* foi realizada para verificar se a alteração de um aminoácido poderia ocasionar a distorção de toda estrutura. Isso porque a substituição de um resíduo com uma cadeia lateral diferente do original pode levar a mudança na circunvizinhança, resultando na desorganização da estrutura.

O diagrama de Ramachandran gerado através do programa SPDBV, define um cenário favorável a estruturação de uma α-hélice na região do *linker* (**Fig. 4.10**), reproduzindo as análises de difração de raio X obtidas por Williams e colaboradores (2004), sustentando o uso desse modelo em uma abordagem *in silico* preditiva sobre o impacto das mutações estudadas na estrutura secundária do *linker*.

A projeção dos aminoácidos presentes no *linker,* no diagrama de Ramachandran, sugere resíduos comprometidos com a formação da alfa hélice (resíduos 1748 ao 1755). Apesar da ausência de outras estruturas secundárias, somente 5 resíduos de aminoácidos possuem ângulos não permitidos para uma

estruturação (G1738, D1739, V1740, H1746 e Q1756), enquanto os outros 18 resíduos de aminoácidos se localizam em regiões permitidas e aceitáveis (V1741, N1742, G1743, R1744, N1745, Q1747, G1748, P1749, K1750, R1751, A1752, R1753, E1754, S1755, D1757, R1758 e K1759) (**Fig. 4.10**).

A avaliação dos variantes através dos diagramas de Ramachandran não demonstrou alterações na estrutura secundária do *linker*, independente do resíduo de aminoácido alterado (resíduos de aminoácidos em regiões permitidas/aceitáveis/não permitidas). As substituições estudadas, mesmo de variantes patogênicos, não alteraram as combinações de ângulos das ligações do carbono α de modo a resultar em uma desorganização da α-hélice ou alteração de outros resíduos de aminoácidos que compõe o *linker*.

O que se pode depreender dos dados gerados é que a manutenção da α-hélice pode não ser suficiente para garantir a integridade do *linker*, e consequentemente a funcionalidade do tBRCT. De acordo com os dados dos diagramas de Ramachandran, a estrutura em α-hélice foi mantida para todos os variantes, no entanto, a maioria das mutações (20 variantes) ocasionaram a disfuncionalidade da proteína.

A avaliação através do diagrama de Ramachandran oferece mais informações sobre o comportamento dos resíduos de aminoácidos que compõem a estrutura do tBRCT. No entanto, para o estudo do *linker*, é possível que variantes nessa região não determinem alterações significativas que possam ser preditas *in silico*, através dessa metodologia. Apesar disso, não se pode descartar o uso do diagrama para prever o comportamento substituições fora do *linker*, fornecendo mais um dado para compor o perfil preditivo dos variantes. É preciso considerar que o diagrama de Ramachandran se limita à avaliação da estrutura secundária. As extrapolações sobre a estrutura terciária são inferidas, mas com poder preditivo limitado. Assim, uma substituição definida como favorável para a manutenção da estrutura secundária não necessariamente garante a manutenção da estrutura terciária ou mesmo a função biológica da proteína.

6. CONCLUSÕES

- Foram identificados 132 possíveis variantes naturais missense na região do *linker*, dos quais 16 já haviam sido previamente estudados/classificados. Entre os 116 variantes restantes, 23 apresentaram perfil patogênico quando analisados por algoritmos de predição (Align GVGD, SIFT, PolyPhen) e foram selecionados para o estudo.
- Os variantes G1738V, D1739H, D1739A, G1743R, G1743V, G1743E, H1746Y, H1746D, H1746L, H1746P, G1748R, G1748C, G1748V, G1748D, P1749S, P1749L, P1749Q, R1753G, R1753I e R1753S apresentam perfil patogênico no ensaio de TA a 37ºC.
- O variante V1740G apresenta perfil não patogênico, no entanto com amplas variações, no ensaio de TA a 37ºC.
- Os variantes V1740A e V1740E apresentaram um comportamento variável, de perfil indeterminado, sugestivo de variantes não classificados.
- Os variantes G1738V, D1739H, D1739A, G1743R, H1746D, H1746P, G1748R, P1749L, R1753G, R1753I e R1753S não apresentaram aumento estatisticamente significativo de atividade quando submetidos ao ensaio de TA a 30ºC.
- Os variantes G1743V, G1743E, H1746Y, H1746L, G1748C, G1748V, G1748D, P1749S e P1749Q apresentaram um aumento estatisticamente significativo de atividade quando submetidos ao ensaio de TA a 30ºC, com atividade superior a 30% do controle selvagem, sendo considerados variantes termossensíveis.
- Os variantes missense estudados na região do *linker* não apresentam distorção da estrutura secundária, quando analisados através dos diagramas de Ramachandran.

7. REFERÊNCIAS

Albarakatia, N.; Abdel-Fatahb, T.; *et al*. Targeting BRCA1-BER deficient breast cancer by ATM or DNA-PKcs blockade either alone or in combination with cisplatin for personalized therapy. **Molecular Oncology**, 9, 204e2 1 7, 2015.

Align GVGD, disponível em <http://agvgd.iarc.fr/agvgd_input.php> acessado em: janeiro, 2016.

Alli, E. & Ford, J. BRCA1: Beyond double-strand break repair. **DNA Repair**, 2015.

Aparicio, T.; Baer, R.; Gautier, J. DNA double-strand break repair pathway choice and câncer. **DNA Repair**, 19: 169–175,2014.

Bourne, Philip & Weissig, Helge. Structural Bioinformatics. Volume 44, Wiley-Liss, New Jersey, 2003.

Bouwman, P.; Gulden, H.; *et al*. A High-Throughput Functional Complementation Assay for Classification of *BRCA1* Missense Variants. **CANCER DISCOVERY**, 2013.

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

BRCA1 CIRCOS, disponível em <https://research.nhgri.nih.gov/projects/bic/circos/index.shtml> acessado em: janeiro, 2016.

Breast Cancer Information Core, disponível em <http://research.nhgri.nih.gov/bic/> acessado em: janeiro, 2016.

Burnette, W. Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. **Anal Biochem**, 112(2): 195-203, 1981.

Calò, V.; Bruno, L.; *et ai*. The Clinical Significance of Unknown Sequence Variants in *BRCA* Genes. **Cancers**, *2*, 1644-1660; 2010.

Carvalho, M.; Billack, B.; *et al*. Mutations in the BRCT Domain Confer Temperature Sensitivity to BRCA1 in Transcription Activation. **Cancer Biology & Therapy**, 1:5, 502-508, 2002.

Carvalho, M.; Couch, F.; Monteiro, A. Functional assays for BRCA1 and BRCA2. **Int J Biochem Cell Biol**, 39(2): 298–310, 2007.

Carvalho, M.; Marsillac, S.; et al. Determination of Cancer Risk Associated with Germ Line BRCA1 Missense Variants by Functional Analysis. **Cancer Res**, 15; 67(4): 1494–1501, 2007.

Carvalho, M.; Pino, M.; et al. Analysis of a set of missense, frameshift, and in-frame deletion variants of BRCA1. **Mutat Res**. 15; 660(1-2): 1–11, 2009.

Carvalho, R.; Abreu, R.; *et al*. Probing Structure-Function Relationships in Missense Variants in the Carboxy-Terminal Region of BRCA1. **PLoS ONE**, 9(5): e97766, 2014.

Cavanagh, H. & Rogers, K. The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers. **Hereditary Cancer in Clinical Practice**, 13:16, 2015.

Clark, S.; Rodriguez, A.; *et al*. Structure-Function of the Tumor Suppressor BRCA1. **Computational and Structural Biotechnology Journal**, volume No: 1, Issue 1, 2012.

Couch, F.; Rasmussen, L.; *et al*. Assessment of Functional Effects of Unclassified Genetic Variants**. Hum Mutat**, 29(11): 1314–1326, 2008.

Cropp, C.; Nevanlinna, H.; *et al*. Evidence for Involvement of BRCA1 in Sporadic Breast Carcinomas. **CANCER RESEARCH**, 54. 2548-2551, 1994.

Đapic', V. & Monteiro, A. Functional Implications of BRCA1 for Early Detection, Prevention, and Treatment of Breast Cancer. **Critical Reviews™ in Eukaryotic Gene Expression**, 16(3):233–252, 2006.

Domchek, S.; Tang, J.; *et al.* Biallelic Deleterious *BRCA1* Mutations in a Woman with Early-Onset Ovarian Cancer. **Cancer Discov**, 3(4): 399–405, 2013.

Downs, J.; Nussenzweig M.; Nussenzweig, A. Chromatin dynamics and the preservation of genetic information. **NATURE**, Vol 447, 21, 2007.

Downs, B. & Wang, S. Epigenetic changes in BRCA1-mutated familial breast cancer. **Cancer Genet**, 208(5):237-40, 2015.

Dunning, A.; Chiano, 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.; Bishop, D.; *et al*. Genetic Linkage Analysis in Familial Breast and Ovarian Cancer: Results from 214 Families. **Am.J. Hum. Genet**., 52:678-701, 1993.

Easton, D.; Deffenbaugh, A.; *et al*. A Systematic Genetic Assessment of 1,433 Sequence Variants of Unknown Clinical Significance in the *BRCA1* and *BRCA2* Breast Cancer–Predisposition Genes. **The American Journal of Human Genetics**, Volume 81, 2007.

Economopoulou, P.; Dimitriadis, G.; Psyrri, A. Beyond BRCA: New hereditary breast cancer susceptibility genes. **Cancer Treatment Reviews**, 41, 1–8, 2015.

Engel, C. & Fischer, C. Breast Cancer Risks and Risk Prediction Models. **Breast Care**, 10:7–12, 2015.

Friedman, L.; Ostermeyer, E.; *et al*. Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. Nat Genet, 8(4):399-404, 1994.

Glover, M. Insights into the molecular basis of human hereditary breast cancer from studies of the BRCA1 BRCT domain. **Familial Cancer**, 5:89–93, 2006.

Glover, M.; Williams R., *et al*. Interactions between BRCT repeats and phophoproteins: tangled up in two. **TRENDS in Biochemical Sciences**, 29(11): 581-585, 2004.

Glover, M.; Williams, S.; Lee, M. Interactions between BRCT repeats and phosphoproteins: tangled up in two. **TRENDS in Biochemical Sciences**, vol.29, No.11, 2004.

Hall, J.; Lee, M.; *et al*. Linkage of Early-Onset Familial Breast Cancer to Chromosome 17q21. **SCIENCE**, VOL. 250, 1990.

Hanahan, D. & Weinberg, R. Hallmarks of cancer: the next generation. **Cell**, 144: 646-674, 2011.

Hanahan, D. & Weinberg, R. The hallmarks of cancer. **Cell**, 100: 57-70, 2000.

Hayes, F.; Cayanan, C.; *et al*. Functional Assay for BRCA1: Mutagenesis of the COOH-Terminal Region Reveals Critical Residues for Transcription Activation. **CANCER RESEARCH**, 60, 2411–2418, 2000.

Ho, S.; Hunt, H.; *et al*. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. **Gene**, 77(1):51-9, 1989.

IARC, disponível em <http://www.iarc.fr/> acessado em: janeiro, 2016.

IARC, World Cancer Report, 2008.

INCA, disponível em <http://www2.inca.gov.br/wps/wcm/connect/inca/portal/home?orig=inca> acessado em: janeiro, 2016.

INCA, Estimativa 2016: Incidência de câncer no Brasil.

Iversen, E.; Couch, F.; *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): 1078–1088, 2011

Jhuraney, A.; Velkova, A.; *et al*. BRCA1 Circos: a visualisation resource for functional analysis of missense variants. **J Med Genet**, 52:224–230, 2015.

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

Kakarougkas, A. & Jeggo. DNA DSB repair pathway choice: an orchestrated handover mechanism. **Br J Radiol**, 2014.

Karchin, R.; Monteiro, A.; *et al*. Functional Impact of Missense Variants in BRCA1 Predicted by Supervised Learning. **PLOS**, Volume 3, Issue 2, e26, 2007.

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

Key, T.; Verkasalo, P.; Banks, E. Epidemiology of breast câncer. **THE LANCET Oncology**, Vol 2, 2001.

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

Lee, M.; Green, R.; *et al*. Comprehensive Analysis of Missense Variations in the BRCT Domain of BRCA1 by Structural and Functional Assays. **Cancer Res**, 15; 70(12): 4880–4890, 2010.

Lux, M.; Fasching, P.; Beckmann, M. Hereditary breast and ovarian cancer: review and future perspectives. **J Mol Med**, 84: 16–28, 2006.

Mersch, J.; Jackson, M.; *et al*. Cancers Associated With BRCA1 and BRCA2 Mutations Other Than Breast and Ovarian. **Cancer**, 2014.

Mesquita, R.; Woods, N.; *et al*. Tandem BRCT Domains: DNA's Praetorian Guard. **Genes & Cancer**, vol 1, no 11, 2010.

Miki, Y.; Swensen, J.; *et al*. A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1. **SCIENCE**, VOL 266, 1994.

Millot, G.; Carvalho, M.; *et al*. A Guide for Functional Analysis of *BRCA1* Variants of Uncertain Significance (VUS). **Hum Mutat**, 33(11): 1526–1537, 2012.

Minami, A.; Nakanishi, A.; *et al*. Connection between tumor suppressor BRCA1 and PTEN in damaged DNA repair. **Frontiers in Oncology**, volume 4, article 318, 2014.

Monteiro, A. BRCA1: exploring the links to Transcription. **TIBS**, 25, OCTOBER 2000.

Monteiro, A. Common BRCA1 Variants and Transcriptional Activation. **Am J Hum Genet**, 61:761-762, 1997.

Monteiro, A.; August, A.; Hanafusa, H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. **Proc. Natl. Acad. Sci,** Vol. 93, pp. 13595–13599, 1996.

Morris, A.; MacArthur, M.; *et al*. Stereochemical Quality of Protein Structure Coordinates. **Proteins: Structure, Function and Genetics**, 12:345-364, 1992.

Morris, J.; Pangon, L.; *et al*. Genetic analysis of BRCA1 ubiquitin ligase activity and its relationship to breast cancer susceptibility. **Hum Mol Genet**, 15(4):599-606, 2006.

Moynahan, M. The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans. **Oncogene**, 21(58):8994-9007, 2002.

Paul, A. & Paul, S. The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. **Front Biosci**, 19: 605–618, 2015.

Phelan, C.; Dapic, V.; *et al*. Classification of BRCA1 missense variants of unknown clinical significance. **J Med Genet**, 42:138–146, 2005.

Plon, S.; Eccles, D.; *et al*. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. **Hum Mutat,** 29(11): 1282–1291, 2008.

PolyPhen, disponível em <http://genetics.bwh.harvard.edu/pph/data/> acessado em: janeiro, 2016.

PolyPhen-2, disponível em <http://genetics.bwh.harvard.edu/pph2/> acessado em: janeiro, 2016.

Prakash1, R.; Zhang, Y.; *et al*. Homologous recombination and human health: the roles of BRCA1, BRCA2 and associated proteins. **Cold Spring Harb Perspect Biol**, 7(4), 2015.

Pruthi, S.; Gostout, B.; Lindor, N. Identification and Management of Women With *BRCA* Mutations or Hereditary Predisposition for Breast and Ovarian Cancer. **Mayo Clin Proc***,* 85(12):1111-1120, 2010.

Ramachandran, G.; Ramakrishna, C.; Sasisekharan. Sterochemistry of Polypeptide Chain Configuration. **J. Mol. Biol.**, 7, 95-99, 1963.

Ransburgh, D.; Chiba, N.; *et al*. The effect of *BRCA1* missense mutations on homology directed recombination. **Cancer Res**, 70(3): 988–995, 2010.

RCSB PDB, disponível em <http://www.rcsb.org/pdb/home/home.do> acessado em: janeiro, 2016.

Rosen, E. & Pishvaian, M. Targeting the BRCA1/2 Tumor Suppressors. **Current Drug Targets***,* 15*,* 17-31, 2014.

Saiki, R.; Gelfand, D.; *et al*. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. **Science**, 239(4839):487-91, 1988.

Savage, K. & Harkin, P. BRCA1, a ‘complex’ protein involved in the maintenance of genomic stability. **FEBS Journal**, 282, 630–646, 2015.

Sawyer, S.; Tian, L.; *et al*. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. **Cancer Discov**, 5(2):135-42, 2015.

Scalia-Wilbur, J.; Colins, B.; *et al*. Breast Cancer Risk Assessment: Moving Beyond BRCA 1 and 2. **Semin Radiat Oncol**, 26:3-8, 2016.

Scully, R. Ganesan, S.; *et al*. Genetic Analysis of BRCA1 Function in a Defined Tumor Cell Line. **Molecular Cell**, Vol. 4, 1093–1099, 1999.

Sheng, Z.; Zhao, Y.; Huang, J. Functional Evolution of BRCT Domains from Binding DNA to Protein. **Evolutionary Bioinformatics**, 7, 87–97, 2011.

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

SIFT, disponível em <http://sift.jcvi.org/> acessado em: janeiro, 2016.

Smith, T.; Lee, M.; et al. Complete Genomic Sequence and Analysis of 117 kb of Human DNA Containing the Gene BRCAI. **GENOME RESEARCH** 6:1029-1049, 1996.

Swiss PDB Viewer, disponível em <http://spdbv.vital-it.ch/disclaim.html> acessado em: janeiro, 2016.

Vallon-Christersson, J.; Cayanan, C.; *et al*. Functional Analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. **Human Molecular Genetics**, Vol. 10, No 4, 2001.

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

Wang, L. & Di, L. BRCA1 And Estrogen/Estrogen Receptor In Breast Cancer: Where They Interact? **International Journal of Biological Sciences**, 10(5): 566-575, 2014.

WHO, disponível em <http://www.who.int/cancer/en/> acessado em: janeiro, 2016.

Williams, S & Glover, M. Structural Consequences of a Cancer-causing BRCA1-BRCT Missense Mutation. **THE JOURNAL OF BIOLOGICAL CHEMISTRY**, vol. 278, No. 4, 2630–2635, 2003.

Williams, S.; Green, R.; Glover, M. Crystal structure of the BRCT repeat region from the breast cancer-ssociated protein BRCA1.**Nature structural biology**, volume 8, number 10, 2001.

Williams, S.; Lee, M.; *et al*. Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1. **NATURE STRUCTURAL & MOLECULAR BIOLOGY**, vol 11, number 6, 2004.

Woods, N.; Baskin, R.; *et al*. Functional assays provide a robust tool for the clinical annotation of genetic variants of uncertain significance. **Genomic Medicine**, 2016.

Woods, N.; Mesquita, R.; *et al*. Charting the landscape of tandem BRCT domain-mediated protein interactions. **Science Signaling**, 5(242), 2012.

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

Worley, T.; Vallon-Christersson, J.; et al. A Naturally Occurring Allele of BRCA1 Coding for a Temperature-Sensitive Mutant Protein. **Cancer Biology & Therapy**, 1:5, 497-501, 2002.

Wu, Q.; Jubb, H.; *et al*. Phosphopeptide interactions with BRCA1 BRCT domains: More than just a motif. **Progress in Biophysics and Molecular Biology**, 117, 143e148, 2015.

Yi, Y.; Kang, H.; Bae, I. BRCA1 and Oxidative Stress. ***Cancers***, *6*, 771-795, 2014.

Zhang, J. & Powell, S. The Role of the BRCA1 Tumor Suppressor in DNA Double-Strand Break Repair. **Mol Cancer Res**, 3(10), 2005.

ANEXOS

Anexo 1

Tabela A1 – Análises *in silico* dos variantes naturais localizados na região *linker* de BRCA1. Todos os variantes naturais identificados foram submetidos aos algoritmos de predição Align GVGD, SIFT e PolyPhen. Onde: Class0 – menor risco, Class65 – maior risco e C15, C25, C35, C45 e C55 – risco intermediário progressivo; B – benigno e D – *damaging* (prejudicial); POD – possivelmente prejudicial, PRD – provavelmente prejudicial. Em vermelho, variantes previamente analisados em outros estudos.

	Perfil preditivo		
Variante	*Align GVGD*	*SIFT*	*PolyPhen*
R1737G	Class C45	D	POD
R1737T	Class C25	D	POD
R1737I	Class C35	D	B
R1737K	Class C0	B	B
R1737S	Class C35	D	PRD
G1738R	Class C65	B	POD
G1738A	Class C55	D	B
G1738V	Class C65	D	POD
G1738E	Class C65	D	B
D1739H	Class C65	D	POD
D1739N	Class C15	B	B
D1739Y	Class C65	D	PRD
D1739V	Class C65	D	B
D1739A	Class C65	D	B
D1739G	Class C65	D	B
D1739E	Class C35	B	B
V1740L	Class C25	D	B
V1740M	Class C15	D	B
V1740E	Class C65	D	POD
V1740A	Class C65	D	POD
V1740G	Class C65	D	PRD
V1741L	Class C0	D	B
V1741I	Class C0	B	B
V1741F	Class C0	D	PRD
V1741D	Class C45	B	POD
V1741A	Class C25	B	B
V1741G	Class C35	D	B
N1742Y	Class C15	D	POD
N1742H	Class C0	D	B
N1742D	Class C0	D	B
N1742I	Class C15	D	POD
N1742T	Class C0	D	B
N1742S	Class C0	D	B
N1742K	Class C15	B	POD
G1743R	Class C65	B	B
G1743A	Class C55	B	POD
G1743V	Class C65	D	PRD
G1743E	Class C65	D	POD
R1744G	Class C0	D	B
R1744T	Class C0	D	B
R1744I	Class C0	D	POD
R1744K	Class C0	D	B
R1744S	Class C0	D	B
N1745Y	Class C45	D	POD
N1745H	Class C25	D	B
N1745D	Class C0	D	B
N1745I	Class C55	D	B
N1745T	Class C25	D	B
N1745S	Class C15	D	B
N1745K	Class C35	B	POD
H1746Y	Class C65	D	PRD
H1746D	Class C65	D	POD
H1746N	Class C65	D	POD
H1746L	Class C65	D	B
H1746R	Class C25	D	B
H1746P	Class C65	D	POD
H1746Q	Class C15	D	B
Q1747E	Class C0	D	B
Q1747K	Class C0	D	B

Q1747L	Class C25	D	B
Q1747P	Class C15	D	B
Q1747R	Class C0	D	B
Q1747H	Class C0	D	POD
G1748R	Class C65	D	PRD
G1748S	Class C55	D	POD
G1748C	Class C65	D	PRD
G1748A	Class C55	D	B
G1748V	Class C65	D	POD
G1748D	Class C65	D	POD
P1749A	Class C25	D	POD
P1749T	Class C35	B	POD
P1749S	Class C65	D	POD
P1749R	Class C65	D	B
P1749L	Class C65	B	PRD
P1749Q	Class C65	B	POD
K1750Q	Class C0	D	B
K1750E	Class C0	B	B
K1750M	Class C0	D	B
K1750T	Class C0	D	B
K1750R	Class C0	D	B
K1750N	Class C0	D	B
R1751G	Class C45	D	B
R1751P	Class C35	B	B
R1751L	Class C35	D	B
R1751Q	Class C0	D	POD
A1752P	Class C0	D	PRD
A1752S	Class C0	D	B
A1752T	Class C0	D	POD
A1752G	Class C0	D	B
A1752V	Class C0	D	POD
A1752E	Class C0	D	B
R1753G	Class C65	D	B
R1753T	Class C65	D	POD
R1753I	Class C65	D	POD
R1753K	Class C25	D	B
R1753S	Class C65	D	POD
E1754Q	Class C0	D	POD
754K	Class C0	B	B
E1754V	Class C0	B	B
E1754A	Class C0	D	B
E1754G	Class C0	D	B
E1754D	Class C0	D	B

S1755T	Class C0	D	B
S1755A	Class C0	D	B
S1755P	Class C0	D	B
S1755C	Class C15	D	PRD
S1755Y	Class C0	D	POD
S1755F	Class C0	D	B
Q1756K	Class C0	D	B
Q1756E	Class C0	D	B
Q1756L	Class C0	D	B
Q1756P	Class C0	D	B
Q1756R	Class C0	D	B
Q1756H	Class C0	D	B
D1757H	Class C0	D	POD
D1757Y	Class C0	B	B
D1757N	Class C0	B	B
D1757V	Class C0	B	B
D1757A	Class C0	D	B
D1757G	Class C0	D	B
D1757E	Class C0	D	B
R1758G	Class C0	D	B
R1758T	Class C0	D	POD
R1758K	Class C0	D	B
R1758I	Class C0	D	PRD
R1758S	Class C0	D	POD
K1759Q	Class C0	B	B
K1759E	Class C0	B	B
K1759T	Class C0	D	B
K1759M	Class C0	D	POD
K1759R	Class C0	D	POD
K1759N	Class C0	D	POD