UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA

SIMONE DA COSTA E SILVA CARVALHO

**CARACTERIZAÇÃO MOLECULAR DA SÍNDROME HEREDITÁRIA
DO CÂNCER DE MAMA E/OU OVÁRIO**

Ribeirão Preto

2019

SIMONE DA COSTA E SILVA CARVALHO

CARACTERIZAÇÃO MOLECULAR DA SÍNDROME HEREDITÁRIA DO
CÂNCER DE MAMA E/OU OVÁRIO

Tese de Doutorado apresentada à Faculdade de Medicina
de Ribeirão Preto, Universidade de São Paulo, para
obtenção do título de Doutora em Ciências.

Área de Concentração: Genética

Orientador: Prof. Dr. Wilson Araújo da Silva Junior
Co-orientador: Prof. Dr. Victor Evangelista de Faria
Ferraz

Versão corrigida. A versão original encontra-se disponível tanto na Biblioteca da Unidade que aloja o
Programa, quanto na Biblioteca Digital de Teses e Dissertações da USP (BDTD)

Ribeirão Preto

2019

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Carvalho, Simone da Costa e Silva.
Caracterização Molecular Da Síndrome Hereditária Do
Câncer De Mama E/Ou Ovário / Simone da Costa e Silva
Carvalho; Orientador Silva Jr, Wilson Araujo. - Ribeirão Preto,
2019.

174 f.: il.; 30cm.

Tese (Doutorado) – Faculdade de Medicina de Ribeirão
Preto, Universidade de São Paulo, 2019.

1. BRCA1 BRCA2. 2. Reparo de DNA. 3. HBOC. 4.
Diagnóstico molecular.

FOLHA DE APROVAÇÃO

Simone da Costa e Silva Carvalho

Caracterização Molecular da Síndrome Hereditária do Câncer de Mama e/ou Ovário

Tese de Doutorado apresentada à Faculdade de
Medicina de Ribeirão Preto, Universidade de São Paulo,
para obtenção do título de Doutora em Ciências.

Área de Concentração: Genética.

Orientador: Prof. Dr. Wilson Araújo da Silva Junior

Co-orientador: Prof. Dr. Victor Evangelista de Faria
Ferraz

Aprovado em:

Banca Examinadora

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

DEDICATÓRIA

Ao Universo pela oportunidade, a mim pela coragem, aos meus pais pela possibilidade e aos demais pela continuidade.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer à Universidade de São Paulo,

Pela oportunidade e estrutura oferecida, possibilitando o meu brutal desenvolvimento profissional e pessoal. De forma conjunta, agradeço ao governo Brasileiro pela destinação de verbas à educação de forma marcante, impulsionando a qualificação em massa de muitos dos atuais profissionais. Agradeço por ter disfrutado desta época de glória do País.

Obrigada ao Departamento de Genética, em nome da Secretária Susie Nalon.

Susie você é uma pessoa incrível, que sustenta a organização do departamento. Não é atoa que seus momentos de prazer (férias) são os nossos momentos de tristeza e desespero. Você faz parte da Genética da USP de Ribeirão Preto. Obrigada por tudo que você tem feito não só por mim, mas por todos os alunos e pelo departamento. Parte das nossas vitórias são suas também, que vibra junto com a gente nos momentos de vitória e segura a mão nos momentos de angústia. Um obrigada mais que especial para você!

Agradeço aos meus orientadores Wilson Araujo e Victor Ferraz,

Por me receberem de braços abertos, me orientarem nesta longa jornada sempre com muita ética, paciência, companheirismo e entusiasmo. Agradeço pelos ricos momentos de discussões científicas e por terem aberto as portas de seus laboratórios e ambulatórios para que eu pudesse aflorar como cientista e profissional. Prof. Wilson, muito obrigada por me proporcionar a minha tão almejada experiência com técnicas de diagnóstico molecular! Um agradecimento especial ao meu eterno “The best Chefinho” Victor, pela oportunidade de vivenciar a prática clínica dos ambulatórios e me permitir a participação no processo de diagnóstico e aconselhamento genético dos pacientes envolvidos em nossas pesquisas, sempre primando pelo melhor para os pacientes.

I would like to thank my Supervisor during my Internship in Canada, Dr. Jean Yves Masson,

For being so supportive during such an important part of my life. Always willing to help whenever I needed, both in academic as in my personal life. Thanks for opening the doors of your lab to me, for leaving me work like a kid in a playground, for marking my professional life, and also for being one of the most incredible person I have ever met. I am very proud of have had spent this time of my life under your supervision. A special thanks to all your team, in the name of Amelie Rodrigue and Yan Columbe, who taught me everything in such an enthusiastically way. I can say with no doubts: Labo Jym et Ville de Quebec, Je me souviens (pour toujours)!

Agradeço à minha família, em nome dos meus pais, Francisco e Socorro,

Pelo suporte INCONDICIONAL às minhas tomadas de decisões, pelo apoio emocional, pelos paitrocínios tão importantes na construção da carreira de um cientista e por sempre acreditarem e valorizarem o meu trabalho. Tudo que tenho feito foi motivada pelo orgulho que vocês têm pela pessoa que vocês criaram e ensinaram valores tão bonitos e valiosos, que faço questão de espalhar por onde vou. Serei doutora, por vocês e para vocês!

Agradeço IMENSAMENTE aos órgãos de fomento à pesquisa: CNPq, FAEPA, Fundação Hemocentro e FAPESP,

Pela bolsa de estudos, pela disponibilidade de verbas para compra de insumos para a realização desta

pesquisa e pela ajuda financeira para participação de eventos científicos. Apesar de este ser apenas o começo da minha história, já me sinto muito orgulhosa da carreira construída até aqui e de toda a experiência que adquiri na minha área de atuação. Com as verbas disponibilizadas para a pesquisa pude retornar os resultados obtidos aos pacientes, acompanhar a evolução de famílias sob risco, como também pude compartilhar os conhecimentos adquiridos com a comunidade científica, o que me permitiu conhecer pessoalmente e discutir meus dados com muitos dos autores citados neste trabalho.

Agradeço ao meu esposo Bryan,

Pelo suporte incondicional, pela compreensão nos momentos de ausência, por sempre estar ao meu lado, sendo meu raio de sol, e por todas as vezes que com seu entusiasmo e áurea positiva, me obrigou a relaxar um pouco nos momentos em que eu mais precisava. Para mim você é a definição de companheirismo. Te amo!

Agradeço a todas as amigas maravilhosas que construí em Ribeirão Preto,

Em especial à Jéssica Praça, minha amiga e companheira de laboratório e casa, pelos anos maravilhosos que passamos dividindo nosso cotidiano, tornando-os mais leves e com certeza, INESQUECÍVEIS. Obrigada ao nosso agregado Lucas Goedert, que com certeza abrilhantou os dias e noites de vinhos, risadas e confidências no Palmeiras 602, tornando esses dias os meus momentos mais felizes durante o doutorado. Obrigada Danielle Brotto, Luíza Araújo, Camila Araújo e Eduarda Morgana, por sempre estarem ao meu lado e fazerem eu lembrar destes anos de Ribeirão Preto/São Paulo com muita saudade no coração.

Obrigada ao laboratório de Genética Molecular e Bioinformática,

Em nome dos meus amigos Isabela Ichihara, Ádamo Siena, Reginaldo Alves e Greice Molfetta, agradeço a todos do LGMB, bem como às nossas primas do Laboratório do Dr. Rodrigo Panepucci, pelos momentos de risadas intermináveis, pela companhia nos almoços no bandex, pelas cervejinhas da sexta-feira, pelos churrascos memoráveis que marcavam a conquista das nossas vitórias!! Vocês são maravilhosos, obrigada pelos anos incríveis que passamos juntos. Um obrigado mais especial ainda à Greice, que além de tudo isso, desde sempre dividiu comigo seus conhecimentos, sendo base FUNDAMENTAL para o início da minha carreira.

Agradeço ao Laboratório do Prof. Marcelo Alex de Carvalho, no INCA-RJ,

Por terem me recebido também de maneira tão simpática e acolhedora durante os quase dois meses de estágio no INCA. Foi muito gratificante conviver com você Marcelo, um profissional excelente e guerreiro, que mesmo com as dificuldades financeiras da pesquisa no nosso País, faz o seu trabalho com excelência. Obrigada por ser o profissional que é e por cuidar com tanto carinho dos seus alunos pupilos... tenho certeza que se tornarão profissionais exemplares como você! Também tenho muito orgulho de parte da minha carreira ter sido construída em seu laboratório.

Agradeço aos amigos que colaboraram com a realização deste projeto,

Sem a ajuda de vocês não teria sido possível a conclusão desta etapa. Um trabalho grande envolve sempre uma colaboração de sucesso. Muito obrigada Danielle Brotto, Reginaldo Alves, Lorena Alves, Thereza Loureiro, Clarissa Picanço, Thiago Gomes, Amélie Rodrigue e Larissa Milano pela colaboração e por estenderem a mão num momento tão importante. Este trabalho também é de vocês.

Agradeço também aos encontros sublimes da vida,

Que me apresentaram a dois seres maravilhosos: Mirela D’Arc e Franciele Busatto. Meninas, sem palavras para descrever a alegria que tenho no meu coração de a vida ter me apresentado pessoas como vocês. Mirela você é tão especial que durante minha estadia na sua casa você conheceu meus amigos mais especiais. Como pode o simples acaso explicar conexões inesperadamente tão boas? Obrigada Fran, minha guriazinha, pelo companheirismo durante o estágio em Quebec, você foi um ímã que atraiu todas as boas energias para nós nesse momento tão especial das nossas vidas. Fico feliz de ter dividido um momento tão marcante de nossas vidas com uma pessoas como vocês.

E para finalizar, meu muito obrigada ao Hemocentro de Ribeirão Preto,

Local onde nosso laboratório esteve baseado, por toda a infraestrutura oferecida, imprescindível para a realização da minha pesquisa de doutorado, e por toda a troca científica proporcionada. Meu muito obrigada!

RESUMO

CARVALHO, S. C. S. **Caracterização Molecular da Síndrome Hereditária do Câncer de Mama e/ou Ovário**. 2019. 174 f. Tese de Doutorado. Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto – São Paulo, 2019.

A Síndrome Hereditária do Câncer de Mama e/ou Ovário (HBOC) é caracterizada por famílias com história de câncer de mama e/ou ovário e associada a um padrão de herança autossômica dominante. Esta síndrome abrange cerca de 10% e 15% de todos os casos de câncer de mama e de ovário, respectivamente. Os genes *BRCA1* e *BRCA2* são genes de alta penetrância associados ao aumento de até 20x no risco para o desenvolvimento do câncer de mama e/ou ovário. Entretanto, apenas 20-30% dos casos de HBOC apresentam variantes patogênicas nestes genes e estudos têm evidenciado a contribuição de outros genes ligados às vias de reparo do DNA para o aumento de risco para HBOC. Na população brasileira, variantes em genes como *ATM*, *ATR*, *CHEK2*, *MLH1*, *MSH2*, *MSH6*, *POLQ*, *PTEN* e *TP53*, têm sido relatadas em até 7,35% dos casos estudados. Assim, o objetivo deste estudo foi realizar a pesquisa e caracterização de variantes em 21 genes relacionados à HBOC (*ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *FAM175A*, *MLH1*, *MRE11A*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* e *UIMC1*) em 95 indivíduos diagnosticados com suspeita clínica de HBOC no Sudeste brasileiro. Foram identificadas 78 variantes patogênicas ou de significado incerto com conflito de informações sobre patogenicidade na literatura, em 19 dos 21 genes analisados. Cerca de 23,4% dos pacientes apresentaram variantes patogênicas nos genes *BRCA1*, *BRCA2* e *TP53*, frequência maior que a identificada por outros estudos na população brasileira. No entanto, 83,3% das variantes identificadas são de significado incerto. A análise de associação identificou 13 destas VUS, em 8 genes, como significativamente associadas ao risco para HBOC. Os pacientes portadores destas variantes não apresentaram alterações no número de cópias para os genes *BRCA1/BRCA2*. Cinco variantes no *BRCA1* foram identificadas como sendo segregadas em haplótipo, mas este não foi associado ao risco para HBOC. Duas variantes nos genes *FAM175A* e *UIMC1* apresentaram alta frequência nos casos estudados e pacientes duplo heterozigotos apresentaram deficiência no reparo de danos causado após irradiação. Duas variantes no gene *PALB2* associadas ao risco não afetaram a interação com *BRCA1* e apresentaram atividade de reparo similar à proteína selvagem. No entanto, uma variante no *CHEK2* levou à instabilidade da proteína e diminuição nos níveis de ativação de *CHK2*. Embora não associadas ao risco, duas variantes identificadas nos genes *PMS2* e *BRIP1* foram classificadas como possivelmente patogênicas. Esta variante no gene *BRIP1* apresentou sensibilidade ao tratamento à mitomicina C, evidenciando uma deficiência na função desta proteína. Estudos como este evidenciam como outros genes envolvidos em vias de reparo do DNA podem estar contribuindo para a patogênese do HBOC. A caracterização destas variantes é imprescindível para estimar os riscos associados e assim possibilitar o acompanhamento clínico e aconselhamento genético adequados para pacientes diagnosticados clinicamente com HBOC, que, em sua maioria, permanecem sem um diagnóstico molecular elucidado.

ABSTRACT

CARVALHO, S. C. S. **Molecular Characterization of the Hereditary Breast and Ovarian Cancer Syndrome.** 2019. 174 f. Doctoral Thesis. Ribeirão Preto Medical School – University of São Paulo, Ribeirão Preto – São Paulo, 2019.

The Hereditary Breast and Ovarian Cancer Syndrome (HBOC) is characterized by families with a history of breast/ovarian cancer with an autosomal dominant inheritance pattern. This syndrome covers about 10% to 15% of all cases of breast and ovarian cancer, respectively. The *BRCA1* and *BRCA2* genes are high penetrance genes associated with an increased risk of up to 20x for breast and ovarian cancer. However, only 20-30% of HBOC cases present pathogenic mutations in those genes, and other DNA repair genes have emerged as increasing the risk for HBOC. In the Brazilian population, mutations in genes such as *ATM*, *ATR*, *CHEK2*, *MLH1*, *MSH2*, *MSH6*, *POLQ*, *PTEN* and *TP53* have been reported in up to 7.35% of the studied cases. The aim of this study was to investigate and characterize mutations in 21 HBOC-related genes (*ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *FAM175A*, *MLH1*, *MRE11A*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* and *UIMC1*) in 95 individuals clinically diagnosed with HBOC in the Southeastern Brazil. We identified 78 pathogenic mutations and variants of uncertain significance with conflicting data on pathogenicity for 19 of the 21 genes analysed. About 23.4% of the patients presented pathogenic mutations in *BRCA1*, *BRCA2* and *TP53*, a frequency higher than that identified by other studies in the Brazilian population. However, 83.3% of the identified variants are of uncertain significance. Five *BRCA1* variants were identified as being segregated in haplotype, but this was not associated with risk for HBOC. Two variants in *FAM175A* and *UIMC1* genes presented high frequency in the studied cases and when in double heterozygosity, presented deficiency in the repair of DNA damage after irradiation. The association analysis showed 13 VUS in 8 genes, as significantly associated with increased risk for HBOC, and the patients that carry those VUS did not present copy number variations for *BRCA1/BRCA2* genes. Two *PALB2* variants associated with risk did not affect the interaction with *BRCA1* and presented repair activity similar to the wild type protein. However, one variant in *CHEK2* led to protein instability and decreased levels of *CHK2* activation. Although not associated with an increased risk, two variants identified in *PMS2* and *BRIP1* were classified as possibly pathogenic. The *BRIP1* variant showed sensitivity to mitomycin C treatment, evidencing a deficiency in the protein function. Data such as this show how much other DNA repair genes may be contributing to the HBOC pathogenesis, emphasizing how essential the characterization of variants of unknown significance is for estimating the risks and thus enabling adequate clinical follow-up and genetic counseling for patients clinically diagnosed with HBOC, which, for the most part, remain without an elucidated molecular diagnosis.

LISTA DE FIGURAS

Figura 1. Principais etapas e proteínas envolvidas no processo de reparo por recombinação homóloga..	28
Figura 2. Representação esquemática das proteínas BRCA1 e BRCA2 e seus principais domínios de ligação e fosforilação por outras proteínas da via de reparo por recombinação homóloga.....	28
Figura 3. Principais etapas e proteínas envolvidas no processo de reparo de bases mal pareadas (MMR).....	32
Figura 4. Esquema de avaliação das pacientes, de acordo com as metodologias de análise utilizadas..	37
Figura 5. Esquema representativo da geração dos fragmentos de interesse a partir de DNA genômico da biblioteca sequenciada	41
Figura 6. Fluxograma de análise do manual de boas práticas do GATK	43
Figura 7. Esquema representativo da reação de MLPA	44
Figura 8. Representação esquemática do sistema duplo híbrido em mamíferos para avaliação da interação de BRCA1 e PALB2.	52
Figura 9. Curva de sobrevivência dos pacientes estudados pelo tempo de vida desde o diagnóstico em anos	60
Figura 10. Demonstrativo da qualidade geral das reads geradas no sequenciamento do painel gênico das 95 amostras de HBOC	66
Figura 11. Gráficos de qualidade gerados a partir do arquivo <i>fastq</i> de uma única amostra.....	67
Figura 12. Proporção de Transições (Ti) e Transversões (Tv) observadas nos 21 genes sequenciados para as 96 amostras sequenciadas.	68
Figura 13. Cobertura cumulativa das regiões sequenciadas para os 21 genes, para cada uma das 95 amostras de HBOC sequenciadas.	69
Figura 14. Regiões e efeitos das variantes identificadas nos 21 genes estudados.	70
Figura 15. Tipo e quantidade de variantes identificadas por gene analisado.	71
Figura 16. Representação esquemática dos genes <i>BRCA1</i> e <i>BRCA2</i> e diagnóstico molecular das variantes adicionais investigadas	73
Figura 17. Representação esquemática do gene e proteína p53 e localização da variante investigada adicionalmente p.Arg337His.	74
Figura 18. Gráfico <i>Oncoprint</i> mostrando a prevalência de variantes dentre as amostras estudadas, de acordo com o tipo e patogenicidade de cada uma delas	76
Figura 19. Porcentagem de células positivas para fosforilação de γ H2Ax antes e após irradiação com raios-X (2 Gy)	82
Figura 20. Variantes com conflito de informações de patogenicidade e análise da probabilidade de associação destas variantes com HBOC.	86

Figura 21. Representação esquemática da proteína BARD1 e variantes encontradas no estudo e associadas ao risco para HBOC	88
Figura 22. Representação esquemática da proteína CHK2 e variantes encontradas no estudo e associadas ao risco para HBOC	90
Figura 23. <i>Western blot</i> evidenciando a expressão das proteínas CHK2 mutadas p.R137Q e p.I160M em relação à proteína selvagem e endógena	91
Figura 24. <i>Western blot</i> evidenciando os níveis de expressa da proteína CHK2 fosforilada, antes, 30 e 60 minutos após irradiação.	92
Figura 25. <i>Western blot</i> evidenciando a expressão das proteínas MLH1 selvagem (wt) e mutada p.E102E em células HeLa.....	94
Figura 26. Atividade de luciferase relativa à interação PALB2-BRCA1	95
Figura 27. Atividade de luciferase relativa à interação do resíduo 317 de PALB2 mutado por variantes naturais com BRCA1 em células HEK293FT	96
Figura 28. Curva de sobrevivência a diferentes concentrações de Olaparib (inibidor de PARP) de células HeLa após silenciamento da proteína PALB2 endógena.....	98
Figura 29. Representação esquemática da proteína PMS2 e variantes encontradas no estudo e associadas ao risco para HBOC	99
Figura 30. Análise de instabilidade de microssatélites com base na amplificação de 5 marcadores mononucleotídicos em DNA extraído de tecido tumoral pareado com DNA de sangue total.	100
Figura 31. <i>Western blot</i> evidenciando a expressão das proteínas PMS2 selvagem (Wt) e mutada (V717M e D792N) em células HEK293T	101
Figura 32. Curva de sobrevivência a diferentes concentrações de Mitomicina C de células HEK293T após silenciamento da proteína BRIP1 endógena.....	102
Figura 33. Curva de sobrevivência dos pacientes de acordo com o genótipo pelo tempo de vida desde o diagnóstico em anos.....	104

LISTA DE TABELAS

Tabela 1. Concentração de DNA utilizado para a transfecção dos plasmídeos transfectados para o ensaio de interação BRCA1-PALB2 em sistema duplo híbrido em mamíferos.	53
Tabela 2. Principais informações clínicas coletadas a partir da revisão de prontuários médicos de 92 pacientes com suspeita clínica de HBOC.	62
Tabela 3. Dados quantitativos do sequenciamento, após o alinhamento e mapeamento das <i>reads</i> geradas à sequência referência.	67
Tabela 4. Frequências estimadas dos haplótipos envolvendo as cinco variantes frequentes no gene BRCA1 na série de casos estudada.	79
Tabela 5. Caracterização fenotípica de acordo com o genótipo dos pacientes estudados.	105
Tabela A1. Primers utilizados para detecção de variantes pontuais, citadas na Metodologia deste documento.	118
Tabela A2. Primers utilizados para mutagênese dirigida em plasmídeos utilizados para os testes funcionais <i>in vitro</i> e <i>in vivo</i>	119
Tabela A3. Anticorpos utilizados para relevar a superexpressão das proteínas.	120
Tabela A4. Caracterização clínica e molecular dos pacientes estudados.	121
Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.	126

LISTA DE ABREVIACÕES

Alu	Retroelementos de inserção de aproximadamente 300 nucleotídeos
BC	do inglês <i>Breast Cancer</i> = Câncer de mama
CDI	Carcinoma Ductal Invasivo
DSB	do inglês <i>Double Strand Breaks</i> = Quebras de dupla fita no DNA
DSS	do inglês <i>DNA Damage Site</i> = Sítio de danos ao DNA
Ex	Abreviação de éxon
FDA	do inglês <i>Food and Drug Administration</i> = Agência reguladora de medicamentos dos Estados Unidos
HBOC	do inglês <i>Hereditary Breast and Ovarian Cancer Syndrome</i> = Síndrome Hereditária do Câncer de Mama e Ovário
HER-2	Fator de crescimento <i>human epidermal growth factor receptor 2</i>
HGSOC	do inglês <i>High Grade Serous Ovarian Cancer</i> = Câncer seroso de ovário de alto grau
HR	do inglês <i>Homologous Repair</i> = Reparo por recombinação homóloga
LFS	do inglês <i>Li-Fraumeni Syndrome</i>
MMR	do inglês <i>Mismatch Repair</i> = Reparo de bases mal pareadas
MSI	do inglês <i>Microsatellite Instability</i> = Instabilidade de microssatélites
Não-BRCA	Pacientes que não apresentaram variantes nos genes <i>BRCA1/BRCA2</i>
NCCN	do inglês <i>National Comprehensive Cancer Network</i>
NHEJ	do inglês <i>Non-homologous DNA End-Joining</i> = reparo por junção de pontas não homólogas
OC	do inglês <i>Ovarian Cancer</i> = Câncer de Ovário
PARP	poli(ADP-ribose) polimerase
PCR	do inglês <i>Polymerase Chain Reaction</i>
RE	Receptores de estrogênio
RP	Receptores de Progesterona
SDS-PAGE	do inglês <i>SDS-polyacrylamide gel electrophoresis</i>
SNP	do inglês <i>Single Nucleotide Polymorphism</i>
TN	Tripla-negativo
VUS	do inglês <i>Variant of Unknown Significance</i> = Variante de significado incerto

LISTA DE SÍMBOLOS

°C	Graus celsius
γ H2AX	Histona H2AX fosforilada
Kb	Kilobases
kDa	Kilodaltons
μ g	Microgramas
μ L	Microlitro
μ M	Micromolar
mL	Mililitro
mM	Milimolar
min	Minutos
ng	Nanogramas
nM	Nanomolar
n°	Número
%	Porcento
ER+ PR+	Positivo para receptores de estrogênio e/ou para receptores de progesterona
rpm	Rotações por minuto
s	Segundos
pb	Tamanho em pares de bases
Ti/Tv	Taxa de transição e transversão
X^2	Teste estatístico do Qui-quadrado
V	Voltagem

SUMÁRIO

1 INTRODUÇÃO	17
1.1 OS CÂNCERES DE MAMA E OVÁRIO	17
1.2 A SÍNDROME HEREDITÁRIA DO CÂNCER DE MAMA E/OU OVÁRIO	20
1.3 AS VIAS DE REPARO DE DANOS AO DNA E O RISCO AUMENTADO PARA HBOC	26
1.3.1 A via de reparo por recombinação homóloga (HR)	26
1.3.2 A via de reparo de bases mal pareadas (MMR)	31
2 JUSTIFICATIVA	33
3 OBJETIVOS	36
3.1 OBJETIVO GERAL	36
3.2 OBJETIVOS ESPECÍFICOS	36
4 MATERIAIS E MÉTODOS	37
4.1 CARACTERIZAÇÃO DAS AMOSTRAS E ASPECTOS ÉTICOS	38
4.2 CARACTERIZAÇÃO CLÍNICA DOS PACIENTES	39
4.3 SEQUENCIAMENTO DO PAINEL GÊNICO	40
4.3.1 Preparo da biblioteca genômica para análise dos genes <i>ABRAXAS1</i> ou <i>FAM175A</i> , <i>ATM</i> , <i>ATR</i> , <i>BARD1</i> , <i>BRCA1</i> , <i>BRCA2</i> , <i>BRIP1</i> , <i>CDH1</i> , <i>CHEK2</i> , <i>MLH1</i> , <i>MRE11</i> , <i>MSH2</i> , <i>MSH6</i> , <i>NBN</i> , <i>PALB2</i> , <i>PMS2</i> , <i>PTEN</i> , <i>RAD50</i> , <i>RAD51</i> , <i>TP53</i> e <i>UIMC1</i>	40
4.3.2 Reação de Sequenciamento	41
4.3.3 Análise de dados	42
4.3.4 Anotação das variantes encontradas	43
4.4 <i>MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION</i> (MLPA)	44
4.5 <i>SCREENING</i> PARA DETECÇÃO DA VARIANTE c.-107A>T NO GENE <i>BRCA1</i>	45
4.6 <i>SCREENING</i> PARA DETECÇÃO DE INSERÇÃO Alu NO GENE <i>BRCA2</i>	45
4.7 <i>SCREENING</i> DA VARIANTE p.Arg337His NO GENE <i>TP53</i>	46
4.8 ANÁLISE DE HAPLÓTIPOS NO GENE <i>BRCA1</i> ASSOCIADOS AO RISCO PARA HBOC	47
4.9 ANÁLISE DE PATOGENICIDADE <i>in silico</i> E VALIDAÇÃO DAS VARIANTES CANDIDATAS	48
4.10 ANÁLISE DE ASSOCIAÇÃO ENTRE FENÓTIPO E GENÓTIPO	49
4.11 TESTES DE PATOGENICIDADE <i>IN VIVO</i>	49
4.11.1 Avaliação dos níveis de fosforilação de γ H2AX	50
4.11.2 Mutagênese sítio-dirigida para geração dos plamídeos mutados	51
4.11.3 Análise da interação de PALB2-BRCA1 utilizando sistema de duplo-híbrido em mamíferos	52

4.11.4 Superexpressão de proteínas mutadas em células humanas silenciadas para os genes endógenos de interesse	53
4.11.5 Verificação dos níveis de expressão proteica por <i>western blot</i>	55
4.11.6 Efeito das mutações em PALB2 na sensibilidade à inibidores de poli(ADP-ribose) polimerase (PARPi)	56
4.11.7 Níveis de fosforilação das proteínas mutadas após dano ao DNA	57
4.11.8 Efeito das mutações em BRIP1 na sensibilidade à droga Mitomicina C	58
4.12 ANÁLISE DE INSTABILIDADE DE MICROSSATÉLITES (MSI) PARA OS GENES DA VIA MMR	59
5 RESULTADOS E DISCUSSÃO	60
5.1 CARACTERIZAÇÃO CLÍNICA DOS PACIENTES	60
5.2 SEQUENCIAMENTO DO PAINEL GÊNICO	65
5.2.1 Análise da qualidade do sequenciamento	65
5.2.2 Anotação das variantes encontradas	70
5.3 CARACTERIZAÇÃO DAS VARIANTES ENCONTRADAS	71
5.4 ASSOCIAÇÃO GENÓTIPO VERSUS FENÓTIPO	102
6 CONCLUSÕES	106
REFERÊNCIAS	109
APÊNDICE	118
ANEXOS	132
1 Carta de Aprovação do estudo pelo Comitê de Ética do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto.	132
2 Manuscrito do trabalho que será submetido à revista BMC Medical Genomics.	133

1 INTRODUÇÃO

1.1 OS CÂNCERES DE MAMA E OVÁRIO

O câncer de mama constitui o tipo tumoral mais frequente em mulheres em todo o mundo (FERLAY et al., 2014). No Brasil, o câncer de mama é também o tipo de câncer mais comum no Brasil, sendo responsável pela maior causa de óbitos por câncer na população feminina sobretudo na faixa etária entre 40 e 69 anos. Com um risco estimado de 56,33 casos a cada 100 mil mulheres, a estimativa é de 59.700 novos casos de câncer de mama para o biênio 2018-2019 (INCA, 2017). Esta alta incidência pode ser explicada pelas mudanças nos hábitos reprodutivos das mulheres e famílias, bem como pelo aumento nas taxas de detecção precoce da doença por meio do eficiente rastreamento mamográfico utilizado atualmente (SIEGEL; MILLER, 2019).

Os diferentes tipos de carcinomas mamários compreendem diversas entidades biológicas, com características patológicas e implicações clínicas distintas. Portanto, conhecer e entender os diferentes subtipos torna-se uma ferramenta importante para uma decisão terapêutica mais específica e eficaz (DAI et al., 2016). Os carcinomas de mama podem ser classificados em carcinoma “*in situ*” ou invasivo/infiltrante, que são grupos de tumores bastante heterogêneos, subdivididos de acordo com os diferentes subtipos histológicos presentes. Os principais subtipos conhecidos são os carcinomas ductal, lobular, ductal/lobular, mucinoso ou colóide, tubular, medular e o papilar. Dentre estes, o carcinoma ductal constitui o subtipo mais comum com uma prevalência de 70–80% dentre aquelas lesões invasivas (KEENEY et al., 2017; MALHOTRA et al., 2010). Existem ainda os subtipos classificados de acordo com os marcadores moleculares, importantes mediadores das vias de sinalização de crescimento celular. Os principais marcadores imunohistoquímicos utilizados são os receptores de estrogênio (ER), receptores de progesterona (PR) e a proteína *human epidermal growth factor receptor 2* (HER2) e, de acordo com o status de cada um deles os tumores podem ser subdivididos em quatro principais grupos: Luminal A, Luminal B, HER-2 e Basal/triplo-negativo (DAI et al., 2016). O status destes marcadores resultam em prognósticos diferentes e determinam o tipo de terapia mais adequado para cada paciente, como por exemplo, o uso de antagonistas dos receptores de estrogênio, inibidores de aromatase e anticorpos monoclonais específicos para a proteína HER-2 (CHEANG et al., 2009; SCHNITT, 2010; VALLEJOS et al., 2010).

Os subtipos Luminal A e B são caracterizados pela positividade do tumor para os receptores ER e/ou PR e negatividade para amplificação de HER-2 ([ER+|PR+] HER-), diferindo apenas quanto à

expressão do marcador nuclear de proliferação celular Ki67 (DAI et al., 2016). O subtipo Luminal A apresenta baixa expressão de Ki67 (<14%) enquanto o subtipo Luminal B apresenta expressão Ki67 >14% e por isso está associado a um pior prognóstico, à presença de graus tumorais mais elevados e a uma maior prevalência de recorrência local, quando comparado ao subtipo Luminal A (CHEANG et al., 2009). Existe ainda um subtipo Luminal classificado como Luminal HER, caracterizado por [ER+|PR+] HER+, podendo ou não apresentar expressão de Ki67 (VODUC et al., 2017). É sabido que os hormônios esteroides são produzidos por células tumorais e que desempenham papel importante na proliferação destas células, no entanto, tumores ER+|RP+ são bastante responsivos a terapias endócrinas (SUZUKI; MIKI; OHUCHI, 2008; WIDSCHWENDTER et al., 2013). Juntos, os cânceres dos subtipos Luminal são aqueles mais frequentes, presentes em aproximadamente 70% dos casos de câncer de mama, e apresentam o fenótipo menos agressivo dentre os carcinomas invasivos (SCHNITT, 2010). O subtipo HER-2 (ER-PR-HER2+) é caracterizado pela superexpressão da proteína HER-2 que é associada ao aumento da capacidade proliferativa, supressão da apoptose e aumento na motilidade e do potencial invasivo e metastático. Isto, por sua vez está associado a um fenótipo clínico mais agressivo e diminuição na sobrevida dos pacientes (SLAMON et al., 2011). Apesar de um pior prognóstico, estes casos geralmente estão associados a um bom desfecho clínico, devido às opções terapêuticas disponíveis para estes casos, como o Trastuzumab (HERCEPTIN, Genentech Inc.), que se liga a HER-2 diminuindo a proliferação e sobrevivência de células tumorais (DAI et al., 2016).

O câncer de mama de subtipo triplo-negativo (TNBC) ou *basal-like*, é assim caracterizado por ser negativo para receptores de estrogênio e progesterona, não apresentar superexpressão ou amplificação de HER-2 (ER-PR-HER-), e apresentar positividade para citoqueratinas basais como CK5/6 e EGFR (BLOWS et al., 2010). Cerca de 80% dos casos de câncer de mama associado a variantes no gene *BRCA1* são cânceres do tipo triplo-negativo (SCHNITT, 2010). Este subtipo tem em geral, um curso mais agressivo que aqueles com positividade para os receptores hormonais e estão associados a altas taxas de metástases viscerais e do sistema nervoso central, bem como a casos de recorrência e morte precoces (AFGHAHI; TELLI; KURIAN, 2016). Este pior prognóstico pode estar associado à dificuldade de terapias alvo como é possível para os outros subtipos. Felizmente, novas drogas estão sendo disponibilizadas para pacientes com TNBC, como é o caso da droga Atezolizumab (TECENTRIQ, Genentech Inc.), que é um inibidor de PD-L1 que permite a ação de células T tumor-específicas, aumentando a sobrevida média de pacientes em até 70% (U.S. FOOD AND DRUG ADMINISTRATION, 2019). Isto demonstra que além da classificação imunohistoquímica, a caracterização molecular dos diferentes cânceres é de suma importância para o desenvolvimento de

novas terapias alvo (SCHNITT, 2010).

Muitos são os fatores de risco que contribuem para o desenvolvimento do câncer de mama e estes incluem desde perfis hormonais, endógenos ou exógenos, bem como a reprodução e obesidade, por exemplo. Entretanto, o fator de risco mais importante é a presença de histórico familiar (LALLOO; EVANS, 2012). Cerca de 5-10% de todos os casos de câncer de mama seguem um padrão de herança mendeliano autossômico dominante, sendo caracterizados como cânceres hereditários. Já uma parcela de 15-20% de todos os casos são ditos familiares e estão relacionados a mulheres que possuem dois ou mais parentes de primeiro ou segundo grau afetados por algum tipo de câncer (ECONOMOPOULOU; DIMITRIADIS; PSYRRI, 2015).

Já o câncer de ovário representa 3,6% dos tumores femininos no mundo (FERLAY et al., 2014). No Brasil constitui o oitavo tipo tumoral mais incidente em mulheres, com 6.150 casos novos estimados para até o final de 2019 (INCA, 2017). Comparado ao câncer de mama, o câncer de ovário não apresenta uma incidência tão elevada, porém possui uma alta taxa de mortalidade chegando a 62% dos casos, enquanto a mortalidade para o câncer de mama é de aproximadamente 15% na população mundial (SIEGEL; MILLER, 2019). Histologicamente, os carcinomas ovarianos podem se apresentar em cinco principais tipos: carcinoma seroso de alto grau, que é aquele mais frequente com uma prevalência de até 70% dos casos; carcinoma endometriais (10%); de células claras (10%); e menos frequentes, os carcinomas mucinoso (3%) e de baixo grau (<5%) (STEWART; WILD, 2014). Além de ser o tipo de câncer de ovário mais frequente, o câncer seroso de alto grau (HGSOC) é considerado o tumor ginecológico mais letal, sendo responsável por 70-80% dos óbitos em decorrência de câncer de ovário no mundo. A alta mortalidade está relacionada à presença de sintomas iniciais vagos e inespecíficos o que faz com que cerca de 80% dos casos sejam diagnosticados em estágios mais avançados da doença (BURGESS; KOLESAR, 2018). Além disso, os casos de HGSOC apresentam alta heterogeneidade celular, instabilidade genômica e estão associados a altas taxas de recidiva e resistência à quimioterapia (YANG et al., 2018).

Os principais fatores de risco para o câncer de ovário são a presença de histórico familiar de câncer de mama e/ou ovário - geralmente associadas a variantes em genes das vias de reparo do DNA, com cerca de 24% de todos os casos sendo do tipo hereditário (BURGESS; KOLESAR, 2018). Outros fatores tais como menopausa tardia, endometriose; nuliparidade, o uso de terapias de reposição hormonal, obesidade, e tabagismo estão associados com o desenvolvimento do câncer ovariano (STEWART; WILD, 2014).

De forma independente aos diferentes fatores de predisposição, os cânceres são doenças que se desenvolvem a partir de variantes em genes que comprometem a capacidade de divisão e proliferação celular (ROMERO-LAORDEN; CASTRO, 2017). Os cânceres mais frequentes são aqueles de início mais tardio, que afetam principalmente indivíduos acima dos 60 anos de idade. Estes cânceres, também são conhecidos como cânceres esporádicos e estão associados à alterações genéticas somáticas que levam à tumorigênese, principalmente em decorrência da ativação de oncogenes (MEANEY-DELMAN; BELLCROSS, 2013). No entanto, cerca de 5%-10% dos casos de câncer resultam de variantes germinativas herdadas que promovem a perda de função de genes supressores tumorais e defeitos em genes de reparo no DNA, estes últimos que constituem o evento genético mais comum relacionado aos cânceres hereditários (JEGGO; PEARL; CARR, 2016). Nestes casos, o alelo mutado é herdado de forma autossômica dominante e a presença de eventos somáticos levam à perda de função do alelo selvagem, dando início ao processo de tumorigênese nos tecidos onde ocorreu o segundo evento mutacional (ROMERO-LAORDEN; CASTRO, 2017). A Síndrome Hereditária do Câncer de Mama e/ou Ovário (HBOC) constitui a principal síndrome envolvendo cânceres de mama e ovário hereditário e é associada a variantes germinativas nos genes *BRCA1* e *BRCA2*, que comprometem a ativação da principal via de reparo de quebras de dupla fita no DNA por recombinação homóloga (HR) (NIELSEN; VAN OVEREEM HANSEN; SØRENSEN, 2016). A segunda síndrome hereditária mais comum é a Síndrome de Lynch, que além de casos de mama e ovário, apresenta casos de cânceres colorretal e endometrióide associados a variantes germinativas em genes da via de reparo de bases mal pareadas no DNA (MMR) (AL BAKIR; GABRA, 2014).

1.2 A SÍNDROME HEREDITÁRIA DO CÂNCER DE MAMA E/OU OVÁRIO

A Síndrome Hereditária do Câncer de Mama e/ou Ovário (HBOC – do inglês, *Hereditary Breast and Ovarian Cancer*) é caracterizada por famílias com história de câncer de mama e/ou ovário herdado de maneira autossômica dominante, compreendendo até 5-10% de todos os casos de câncer de mama e até 80% dos casos de câncer hereditário de ovário (AL BAKIR; GABRA, 2014; SILVA et al., 2014). Esta síndrome foi primeiramente descrita no século XIX quando a esposa do físico Francês Paul Broca desenvolveu câncer de mama em idade jovem e a análise familiar revelou uma forte segregação da doença com a família. Isso levou à busca por meios de predição de risco para família afetadas e, em 1971, com a proposição do modelo *two-hits* de carcinogênese (KNUDSON, 1971) foi possível a caracterização do fundo genético da síndrome (NIELSEN; VAN OVEREEM HANSEN; SØRENSEN, 2016).

As famílias portadoras desta síndrome possuem risco de 50-80% para desenvolver câncer de mama e de 30-50% para câncer de ovário durante a vida (ROY; CHUN; POWELL, 2012). Existem critérios clínicos estabelecidos para caracterização desta síndrome, que são preconizados pelo *The National Comprehensive Cancer Network*® (NCCN), que é uma aliança composta por 28 centros de estudos de câncer nos Estados Unidos. Os principais critérios clínicos para pacientes portadores de HBOC preconizados pelo NCCN são: diagnóstico da doença antes dos 45 anos (início precoce quando comparadas aos cânceres esporádicos de mama e ovário), acometimento por câncer de mama e ovário no mesmo paciente ou família, câncer de mama bilateral e, embora menos frequente, casos de câncer de mama no sexo masculino. A presença de outros tipos de cânceres nos familiares afetados como câncer de pâncreas, próstata e melanoma, também são critérios para classificação de HBOC (POUMPOURIDOU; KROUPIS, 2012). Quando o paciente apresenta pelo menos um destes critérios estabelecidos, ele já é indicado para o manejo clínico que envolve a análise de risco, investigação molecular e o aconselhamento genético. É aconselhável a investigação clínica e molecular do próprio afetado, mas pode ser realizada a investigação de um caso familiar não afetado que esteja sob risco, caso o do próprio paciente não seja possível (NCCN GUIDELINES, 2019).

No Serviço de Genética Médica do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, o Ambulatório de Genética do Câncer, chefiado pelo Prof. Dr. Victor Evangelista de Faria Ferraz, acompanha e oferece o aconselhamento genético pré e pós teste a mais de 180 pacientes sob suspeita clínica de HBOC. Dentre os pacientes assistidos, oito famílias estão sob investigação do risco para HBOC em decorrência de histórico familiar ou pessoal de câncer mama e/ou ovário.

A HBOC é também conhecida como *BRCA-Related Breast and/or Ovarian Cancer Syndrome* uma vez que está associada a variantes nos genes de susceptibilidade ao câncer de mama *BRCA1* e *BRCA2* (AL BAKIR; GABRA, 2014). O *BRCA1* (OMIM#604370) foi o primeiro gene de susceptibilidade ao câncer de mama descrito (MIKI et al., 1994), e logo em seguida foi caracterizado o gene *BRCA2* (OMIM#612555) (WOOSTER et al., 1995), o que possibilitou o uso de testes de predição de risco em família afetadas. Estes são genes de alta penetrância e a presença de variantes patogênicas em heterozigose conferem risco aumentado para o desenvolvimento do câncer de mama e ovário (POUMPOURIDOU; KROUPIS, 2012). Os genes *BRCA1* e *BRCA2* são caracterizados como supressores tumorais, que codificam grandes proteínas expressas durante as fases S e G2 do ciclo celular. Ambos têm sido associados também à diferenciação de células epiteliais mamárias, atuando como reguladores negativos do crescimento destas células, tendo sua expressão controlada por hormônios esteroides produzidos pelos ovários. Possuem importante papel em vias de reparo de

quebras de dupla fita de DNA por recombinação homóloga (HR), atuando também na regulação do ciclo celular e transcrição (DUROCHER et al., 1996; FARMER et al., 2005; TUTT et al., 2001; VENKITARAMAN et al., 2002).

Ambos os genes *BRCA1/BRCA2* estão associados a um aumento de risco de até 20 vezes para câncer de mama (BC) e ovário (OC) (HONDOW et al., 2011). No entanto, a presença de variantes em *BRCA1* está associada ao aumento de risco para BC em idade mais jovem (30-50 anos de idade), quando comparado a *BRCA2* (40-60 anos) (KOTSOPOULOS, 2018). É sabido ainda, que pacientes com variantes no gene *BRCA1* apresentam até 57% de chance de desenvolver BC e até 40% para OC até os 70 anos, enquanto que para variantes em *BRCA2* os números são um pouco menores: 49% de chance para câncer de mama e apenas 18% para câncer de ovário (COUCH; NATHANSON; OFFIT, 2014). Os riscos estimados para acometimento por OC após BC primário após 10 anos, é de 10% para portadores de variantes em *BRCA1* e de 7% quando em *BRCA2* (LYNCH; SNYDER; CASEY, 2013). De fato, variantes no gene *BRCA1* estão associados tumores com maior taxa de proliferação, cânceres mais agressivos e invasivos, geralmente de tipo triplo-negativo, apresentam graus tumorais mais elevados e um pior prognóstico quando comparadas ao *BRCA2* (LAKHANI et al., 1998; SPURDLE et al., 2014; ZHONG et al., 2015). A maioria dos casos de câncer de mama associados a variantes no *BRCA1* são casos de carcinoma ductal invasivo (74% dos casos). Casos de carcinoma medular também podem ser encontrados em casos de HBOC, com uma frequência significativa quando comparada à frequência na população geral (13% e 2% respectivamente), e estão mais relacionados a variantes no gene *BRCA2* (KEENEY et al., 2017; LAKHANI et al., 1998). Já para o câncer de ovário o subtipo relacionado a variantes nos genes *BRCA1* e *BRCA2* mais predominante é carcinoma seroso de alto grau e, cânceres primário de tuba uterina e peritoneal também fazem parte do espectro de câncer associado a variantes em *BRCA1/BRCA2* (ACOG PRACTICE BULLETIN, 2009).

Portanto, a busca por variantes nestes genes supressores tumorais possui grande significância na predição de risco para estes cânceres. Uma vez que a variante é identificada em mulheres diagnosticadas com HBOC, o risco para desenvolvimento da doença pode ser investigado em membros da família reduzindo assim a mortalidade por meio da adoção de medidas preventivas como a mastectomia e/ou ooforectomia profiláticas, além de estratégias de rastreio eficientes para a detecção da doença ainda nos estágios iniciais (PLEVRITIS et al., 2006). A realização da mastectomia profilática está associada à redução do risco de câncer de mama em até 90% em portadores de variantes nos genes *BRCA1/2* (ROMERO-LAORDEN; CASTRO, 2017). A ooforectomia também reduz o risco para câncer de ovário em até 95% e, quando realizada antes da menopausa, também há uma redução

significativa no risco de câncer de mama de até 72% em portadores de variantes em *BRCA1*, e de até 46% para portadores de variantes em *BRCA2* (LYNCH; SNYDER; CASEY, 2013).

Na base de dados do *Human Gene Mutation Database* (HGMD) já foram descritas 2977 variantes em *BRCA1*, sendo variantes *missense* o tipo mais comum. No gene *BRCA2*, 3214 variantes são descritas, com pequenas deleções sendo o tipo mais frequente neste gene. Diversos estudos utilizando sequenciamento de painel gênico em pacientes com HBOC identificaram uma prevalência de variantes patogênicas nos genes *BRCA1* e *BRCA2* em torno de 4,6% - 54% na população mundial (BUYS et al., 2017; DAROOEI et al., 2017; SCHROEDER et al., 2015; TEDALDI et al., 2017). Na América Latina, o sequenciamento de painel gênico na população Colombiana evidenciou uma frequência de 17,64% (COCK-RADA et al., 2018). Na população brasileira, o único estudo utilizando análise de painel gênico evidenciou uma frequência de 8,3% de variantes patogênicas em *BRCA1/BRCA2* (TIMOTEO et al., 2018). Demais estudos envolvendo o sequenciamento de regiões codificantes de *BRCA1* e *BRCA2* evidenciam uma frequência de variantes de 16,6% - 21,5% (CIPRIANO et al., 2019; FERNANDES et al., 2016; PALMERO et al., 2016; SILVA et al., 2014).

Os tipos mais comuns de variantes patogênicas identificadas nos genes *BRCA1* e *BRCA2* na população brasileira são variantes de ponto, em geral do tipo *nonsense*, e pequenas deleções associadas a perda no *frame* de leitura (variantes *frameshift*). Embora menos frequentes, grandes inserções e deleções também são identificadas em até 4,9% dos casos de HBOC (EWALD et al., 2016; PALMERO et al., 2018). Dentre estes grandes rearranjos genômicos, aquele mais frequente é a inserção de fragmento Alu no éxon 3 (c.156_157insAlu) do gene *BRCA2*. Os fragmentos Alu são retroelementos de aproximadamente 300 nucleotídeos e que já foram reportados como inseridos em diversos genes, incluindo *BRCA1* e *BRCA2*, e associados ao aumento de risco para câncer (FELICIO et al., 2018b; MACEDO; ALEMAR; ASHTON-PROLLA, 2019). A inserção de fragmentos Alu no éxon 3 do gene *BRCA2* foi primeiramente descrita em 2005 por Teugels e colaboradores como uma mutação fundadora proveniente da população Portuguesa em pacientes com HBOC. Esta variante é bastante comum na população brasileira devido à imigração Portuguesa durante o período de colonização e, a sua patogênese está associada à perda de parte importante do éxon 3 neste gene (FELICIO et al., 2018b). A integridade desta região é essencial para a ligação da proteína com PALB2 e RAD51, para que haja uma correta ressecção de fragmentos de DNA em áreas de quebra de dupla fita e o consequente reparo por recombinação homóloga (DUCY et al., 2019).

No gene *BRCA1*, a variante mais prevalente em diversas regiões do Brasil é a variante *frameshift* c.5266dupC (PALMERO et al., 2018). Esta variante está localizada no éxon 20 do gene *BRCA1* e a

deleção desta citosina promove o ganho de um códon de parada prematuro causando truncamento da proteína em seu domínio BRCT, importante para interação com CtIP, Abraxas e BARD1, proteínas que atuam na sinalização de danos e recrutamento de moléculas efetoras aos sítios de danos ao DNA (DA COSTA et al., 2008; GILMORE et al., 2017; ROY; CHUN; POWELL, 2012). Esta variante é também comumente encontrada em populações latino americanas, sendo bastante descrita em pacientes com HBOC na população brasileira, especialmente em casos de câncer de ovário (COTRIM et al., 2019; MAISTRO et al., 2016). É uma mutação fundadora proveniente da população Ashkenazi, bastante frequente em populações Norte Europeias (HAMEL et al., 2011), o que pode justificar a alta frequência encontrada na população brasileira, que é marcada por um forte componente ancestral europeu (PENA et al., 2011).

Entretanto, apesar da alta penetrância destes genes, apenas cerca de 20% dos casos de câncer de mama familiar têm sido de fato atribuídos a variantes nos genes *BRCA1* e *BRCA2*, enquanto 5-10% têm sido atribuídos a outros genes de susceptibilidade tais como *TP53*, *STK11*, *PTEN*, *ATM* e *CHEK2* (ALORAIFI et al., 2015). Com o avanço das tecnologias para diagnóstico genético, diversas variantes têm sido identificadas nos genes *BRCA1/BRCA2*, no entanto, apenas cerca de 8% das variantes nestes genes conferem risco aumentado para o câncer de mama. Isto evidencia que variantes em outros genes podem também aumentar o risco para o câncer e estar contribuindo para a patogênese do HBOC (AL BAKIR; GABRA, 2014; BARZAN et al., 2013). Apesar dos protocolos de manejo clínico já serem bem estabelecidos para pacientes com variantes em *BRCA1/BRCA2*, aqueles pacientes negativos para alterações nestes genes sofrem com a falta de acompanhamento clínico e aconselhamento genético adequados, mesmo apresentando características clínicas similares, forte história familiar e risco aumentado para câncer (LEE et al., 2018), o que reforça a necessidade de estudos na caracterização de outros genes relacionados ao aumento de risco para HBOC.

Assim, uma vez que a susceptibilidade genética aumentada em decorrência de variantes nos genes *BRCA1* e *BRCA2* constitui apenas uma parcela de todo o cenário do câncer de mama (POUMPOURIDOU; KROUPIS, 2012; VENKITARAMAN et al., 2002), nas últimas duas décadas vários estudos têm reunido esforços para caracterizar outros genes envolvidos na patogênese no HBOC, utilizando abordagens de análise genômica. Já foram identificadas mais de 7000 variantes associados ao aumento de risco para o câncer de mama em aproximadamente 1000 genes dos quais destacam-se o *BRCA1*, *BRCA2*, *TP53*, *PTEN*, *CHEK2*, *ATM*, *BRIPI*, *PALB2*, *PPM1D*, *CDH1*, *MLH1*, *MRE11*, *MSH2*, *MSH6*, *MUTYH*, *NBN*, *PMS1*, *PMS2*, *RAD50*, *RAD51C*, *STK11* e *BARD1* (KUUSISTO et al., 2011; ZHANG et al., 2011; LYNCH; SNYDER; CASEY, 2013; MAHDI;

NASSIRI; NASIRI, 2013). Estudos neste sentido têm sido importante para caracterizar não apenas os cânceres hereditários, como também os cânceres esporádicos. Um estudo realizado com 560 amostras de câncer de mama identificou potenciais variantes “*driver*” somáticas, em 93 genes codificadores de proteínas, dentre eles, os 10 mais mutados foram: *TP53*, *PIK3CA*, *MYC*, *CCND1*, *PTEN*, *ERBB2*, *ZNF703/FGFR1*, *GATA3*, *RBI* e *MAP3K1* (NIK-ZAINAL et al., 2016). Para o câncer de ovário, um estudo recente demonstrou que além de *BRCA1*, *BRCA2* e *TP53*, genes como *CSMD3*, *PIK3CA*, *FAT3*, *CDK12*, *KRAS*, *PTEN*, *RBI*, *EFEMP1* e *NF1* também apresentaram variantes em casos de carcinoma seroso (YANG et al., 2018).

Na população brasileira, informações de variantes em genes além de *BRCA1/BRCA2* associadas ao aumento de risco para HBOC são muito escassas, sendo variantes pontuais patogênicas em genes como *TP53* e *CHEK2* aquelas mais comumente relatadas na literatura. Uma variante tão frequente quanto as descritas anteriormente em *BRCA1* e *BRCA2* é a variante c.1010G>A (p.R337H) no gene *TP53*, que é associada ao aumento de risco para a Síndrome de Li Fraumeni e é frequentemente encontrada em pacientes com HBOC no Brasil (CURY; FERRAZ; SILVA, 2014). É uma variante fundadora originada na região Sul brasileira, que está presente no éxon 10 da proteína p53 (HAHN et al., 2018) onde está localizado o domínio de oligomerização. Este domínio é importante para a formação de tetrâmeros que promovem a transativação de p53 e facilitam a ligação com o DNA (PARK et al., 2018; PARK; WANG; HWANG, 2019). Devido ao efeito fundador, esta variante está presente em 0,3% da população do Sul e Sudeste Brasileiro (PALMERO et al., 2008). A variante c.11110delC no gene *CHEK2* é uma das variantes mais estudadas em casos de HBOC, especialmente em casos negativos para variantes nos genes *BRCA1/BRCA2* (CIPRIANO et al., 2019), e está associada ao aumento de risco de até 2x para câncer de mama em mulheres e de até 10x em homens (FELIX et al., 2014; SILVA et al., 2014). A variante está localizada no éxon 10 de *CHK2*, região que codifica o domínio catalítico da proteína, importante para autofosforilação e ativação de *CHK2* em resposta a danos ao DNA (ZANNINI; DELIA; BUSCEMI, 2014). Esta variante já foi reportada com uma prevalência de até 14,3% na população brasileira (PALMERO et al., 2016; SILVA et al., 2014). Além destas variantes, microdeleções no éxon 2 do gene *PTEN* sugeridas como polimorfismos (SILVA et al., 2014), e a variante c.-1060A>G em *POLQ* (BRANDALIZE et al., 2014) já foram descritas na população brasileira. O único estudo envolvendo sequenciamento de painel gênico em pacientes com HBOC evidenciou variantes nos genes *ATM*, *ATR*, *CDH1*, *MLH1* e *MSH6* (TIMOTEO et al., 2018).

No entanto, mesmo com uma quantidade relevante de genes já associados, muitos casos de

HBOC ainda permanecem sem um diagnóstico molecular pois variantes causais que possam ser utilizadas como marcadores moleculares tem sido pobremente caracterizadas e relacionadas a esta síndrome, ressaltando a necessidade de uma investigação mais acurada (STUCKEY; ONSTAD, 2015). Estudos recentes têm evidenciado que variantes em genes relacionados às vias de reparo do DNA, como por exemplo a via HR que tem BRCA1/BRCA2 como proteínas-chave, podem constituir a causa de HBOC em pacientes sem diagnóstico molecular definido (FELICIO et al., 2018a). De fato, o reparo de quebras de dupla fita no DNA por recombinação homóloga se dá em decorrência de interações bioquímicas de BRCA1 e BRCA2 com diversas outras proteínas envolvidas na mesma via. Portanto, acredita-se que assim como em BRCA1/BRCA2, variantes que alteram a função de qualquer uma das demais proteínas envolvidas neste processo podem comprometer o processo de reparo e contribuir para o aumento de risco para HBOC (BUISSON et al., 2017; PAUTY et al., 2017; ROY; CHUN; POWELL, 2012; STUCKEY; ONSTAD, 2015).

1.3 AS VIAS DE REPARO DE DANOS AO DNA E O RISCO AUMENTADO PARA HBOC

1.3.1 A via de reparo por recombinação homóloga (HR)

Em mamíferos, a via de reparo por recombinação homóloga é a principal via de reparo de quebras de dupla fita de DNA (DSB), causadas principalmente pelo estresse replicativo, agentes alquilantes, presença de *cross-links*, irradiação ionizante e estresse oxidativo (MEHTA; HABER, 2014). Variantes em genes da via HR já são associadas ao aumento de risco para câncer, uma vez que células deficientes em HR apresentam alta instabilidade genômica e a presença de variantes em homozigose pode ser letal nos estágios iniciais do desenvolvimento do embrião (KRAJEWSKA et al., 2015; PRAKASH et al., 2015). A segunda via responsável pelo reparo de DSB é via de junção de pontas não homólogas (NHEJ), que ocorre em todas as fases do ciclo celular. No entanto essa via é mais susceptível a erros durante o processo de reparo, permitindo inserções ou deleções na região do dano, enquanto HR utiliza a sequência intacta da cromátide irmã como molde para o reparo do DNA, mas está limitada às fases S/G2 do ciclo celular (HER et al., 2016; PRAKASH et al., 2015).

Os genes supressores tumorais *BRCA1* e *BRCA2* codificam proteínas-chave que atuam em diferentes estágios da via HR (**Figura 1**), formando macro complexos protéicos que atuam tanto na detecção de quebras de dupla fita de DNA, como também no recrutamento de outras proteínas que

possuem o papel de mediadoras e efetoras do processo de reparo (ROY; CHUN; POWELL, 2012). Diversos estudos têm apontado o papel de outras proteínas, além de *BRCA1/BRCA2* no aumento de risco para o HBOC. Genes como *ATM*, *BARD1*, *BRIP1*, *CHEK2*, *MRE11*, *NBN*, *PALB2*, *RAD51C*, *RAD50* e *TP53*, bem como *MLH1*, *MSH2*, *MSH6* e *PMS2*, genes relacionados ao reparo de bases mal pareadas no DNA, têm sido relatados como frequentemente mutados em casos de câncer de mama e ovário (ALORAIFI et al., 2015; HIROTSU et al., 2015; MINION et al., 2015; SCHUBERT et al., 2019; SUSZYNSKA et al., 2019; TEDALDI et al., 2017; TIMOTEO et al., 2018; TOSS et al., 2015). Estes genes não só apresentam uma alta frequência de variantes, como muitos apresentam penetrância alta a moderada para OC e BC, já sendo considerados para o manejo clínico (NCCN GUIDELINES, 2019).

De maneira simplificada, no processo de reparo HR, uma das primeiras modificações bioquímicas que ocorre em resposta ao dano ao DNA é ativação de ATM, que promove a ativação de outras proteínas de sinalização do dano como ATR, CHK2 e H2Ax (BUISSON et al., 2017; ZANNINI; DELIA; BUSCEMI, 2014). A fosforilação da histona γ H2Ax promove a sua ubiquitinação nas Lys63 que são reconhecidas por RAP80. Em reposta, RAP80 é ubiquitinada e reconhecida por Abraxas, que está ligada a BRCA1 por meio de seu domínio BRCT (**Figura 2**). Assim ocorre o recrutamento do complexo BRCA1-A (RAP80, Abraxas, BRCA1) aos sítios de dano ao DNA (HER et al., 2016; WANG et al., 2007). Após isso, MDC1 também é recrutado aos sítios de dano ao DNA (DSS) e sua fosforilação promove o recrutamento de RFN8 via domínio FHA. O recrutamento de RFN8 promove a ubiquitilação de histonas H1 nas regiões que circundam o DSS levando ao recrutamento de RFN168 e aumentando o sinal de ubiquitilação na região (HER et al., 2016). Em resposta aos danos ao DNA, CtIP também é fosforilado e se liga a BRCA1 também pelo domínio BRCT (**Figura 2**). Há então o recrutamento do complexo MRN nuclease (MRE11, RAD50 e NBS1) que junto com BRCA1, promovem a ressecção da extremidade 5' do sítio de dano ao DNA e, inibem NHEJ (NEPOMUCENO et al., 2017; ROY; CHUN; POWELL, 2012). Após isso, há a ligação de proteínas RPA que protegem o fragmento de DNA fita simples (ssDNA) recém resseccionado. O processo de ressecção também promove a ativação de ATR, que juntamente com ATM e CHK2, fosforilam/ativam BRCA1 e PALB2 (**Figuras 1 e 2**) (BUISSON et al., 2017). BRCA1 recruta PALB2, que por sua vez recruta BRCA2, levando a formação do complexo BRCA1-PALB2-BRCA2. Este complexo recruta RAD51 e dá início à busca pela região homóloga à sequência de ssDNA na cromátide irmã ou cromossomo homólogo intactos, formando uma estrutura denominada D-loop (BUISSON et al., 2010). As DNA polimerases δ ou η promovem a síntese de DNA na região e então o complexo com RAD51 é dissociado, e RAD52 atua permitindo a ligação dependente de síntese dos

fragmentos gerados (DUCY et al., 2019), levando à reparação da região da quebra de dupla fita de DNA.

Figura 1. Principais etapas e proteínas envolvidas no processo de reparo por recombinação homóloga. A bolinha em amarelo com a letra “P”, representa o sinal de fosforilação, relacionado à ativação da proteína. As caixas em amarelo claro na parte de cima da Figura caracterizam as proteínas envolvidas na detecção do dano (elipses em azul mais claro), proteínas mediadoras e proteínas efetoras. As proteínas em elipses de cor azul mais escuro representam os macrocomplexos proteicos com BRCA1. As caixas em bordô representam os complexos proteicos. Imagem adaptada de Roy; Chun e Powel (2012).

Figura 2. Representação esquemática das proteínas BRCA1 e BRCA2 e seus principais domínios de ligação e fosforilação por outras proteínas da via de reparo por recombinação homóloga. Ring = Domínio ring N-terminal. NLS

= sítios de localização nuclear. SCD = domínio de cluster SQ/TQ. BRCT = domínio BRCT. Repetições BRC = Domínios de repetição BRC. H = domínio helicoidal. T = domínio torre. OB = sítio de ligação a oligonucleotídeos. Imagem adaptada de Roy; Chun e Powel (2012).

Diante disso, torna-se claro que outros genes desta via de reparo podem desempenhar papel tão importante quanto *BRCA1/BRCA2* para a patogênese do HBOC. Como exemplo, o gene *ATM*, importante ativador de proteínas como BRCA1, CHK2, CtIP e p53 também tem sido frequentemente reportado como mutado, e é associado ao risco para câncer de mama apresentando baixa-moderada penetrância para HBOC (TAPIA et al., 2008; TAVERA-TAPIA et al., 2017). Por estarem envolvidos com a deficiência de HR, pacientes portadores de variantes em *ATM* e em genes como *CHEK2*, *NBN*, *PALB2* e *BRCA1/BRCA2* podem ser beneficiados com o tratamento do tumor utilizando inibidores de PARP (NIELSEN; VAN OVEREEM HANSEN; SØRENSEN, 2016). O gene *ATR* é outro importante sinalizador de danos no DNA, e também é responsável pela ativação de BRCA1, p53 e CHK1 (KOBAYASHI et al., 2013). Buisson e colaboradores (2017), evidenciaram que ATR também atua recrutando PALB2 aos DSS e promove a formação do complexo BRCA1-PALB2 após irradiação.

A proteína Abraxas, que é codificada pelo gene *FAM175A*, é responsável pelo recrutamento de BRCA1 aos DSS. Estudos têm evidenciando que quando há a perda de função desta proteína, há uma deficiência no recrutamento de BRCA1 levando ao comprometimento dos *checkpoints* de G2/M após irradiação. Em modelos *knockout*, são observados ainda, altos níveis de ressecção nos sítios de dano e, uma diminuição na sobrevida e aumento na incidência tumoral em ratos (HAMDI et al., 2016; RENAULT et al., 2016). RAP80, que é codificada pelo gene *UIMC1* também possui papel importante na sinalização dos danos ao DNA por recrutar o complexo Abraxas-BRCA1 até os sítios de dano por meio da ligação à cadeias de ubiquitina de histonas H2AX (HER et al., 2016). Células deficientes em RAP80 são hipersensíveis à radiação ionizante, também reduzem a formação de focos de BRCA1 e estão associadas à instabilidade genômica e à deficiência do reparo por recombinação homóloga (KIM; CHEN; YU, 2007; LI et al., 2017).

Após a sinalização dos danos, o complexo MRN (MRE11, RAD50 e NBS1) promove a ressecção da região onde há o dano na fita de DNA por meio do recrutamento da exonuclease Exo1 e nuclease Dna2 aos DSS (GOBBINI et al., 2018). Variantes nos genes *MRE11* e *NBN* também estão associadas à síndrome de Ataxia-Telangiectasia-like (assim como variantes em *ATM*) e à Síndrome Quebras de Nijmegen, respectivamente; e a depleção destas proteínas está associada ao aumento da susceptibilidade ao câncer, bem como levam à letalidade embrionária em decorrência de defeitos na proliferação celular e promovem o aumento do processo de ressecção em regiões de dano (PETRONI

et al., 2018).

Após isso, junto com BRCA1 e BRCA2, PALB2 interage de forma direta com RAD51 permitindo a dissociação de RPA do DNA recém resseccionado promovendo a ligação de RAD51. PALB2 também promove a formação da estrutura D-loop durante a invasão da dupla fita da cromátide irmã ou do cromossomo homólogo intactos, permitindo assim o reparo da quebra de dupla fita por homologia (BUISSON et al., 2010; PAUTY et al., 2017). Variantes em *PALB2* são causa da Anemia de Fanconi, mas recentemente têm sido relacionadas também a um aumento de risco de até 30x para câncer de mama, quando comparado a pacientes não portadores de variantes neste gene (NEPOMUCENO et al., 2017). Células deficientes em PALB2 apresentam diminuição na atividade de HR e formação de focos de RAD51, e apresentam sensibilidade a drogas inibidoras de PARP, evidenciando seu papel fundamental em HR (PAUTY et al., 2014).

Outra proteína importante nesta via de reparo é BARD1, que apresenta alta homologia com BRCA1. BARD1 forma heterodímeros com BRCA1 que funcionam como ubiquitina ligases (E3) como também atuam como mediadores diretos na recombinação homóloga por recrutar RAD51 aos sítios de danos ao DNA (ADAMOVICH et al., 2019). Variantes neste gene têm sido frequentemente relatadas em pacientes com câncer de mama/ovário hereditário (QUEZADA URBAN et al., 2018) e associadas a um risco moderado para câncer de mama (SUSZYNSKA et al., 2019; WEBER-LASSALLE et al., 2018). Estudos têm demonstrado também que variantes em *BARD1* levaram a uma deficiência em HR, e tornaram as células mais sensíveis aos danos ao DNA (ADAMOVICH et al., 2019).

Existe ainda o complexo BRCA1-B (BRCA1, BRIP1 e TOPBP1) (**Figura 1**), cuja precisa função ainda é desconhecida mas estudos tem sugerido a sua relação à ativação de *checkpoints* em resposta a erros nas forquilhas de replicação na fase S do ciclo celular, mediado por ATR e CHK1 (ROY; CHUN; POWELL, 2012). No entanto, é sabido que BRIP1 interage com BRCA1 por meio do domínio BRCT e, defeitos nesta proteína já foram relacionados à deficiência na capacidade de reparo HR e superativação dos checkpoints associada à instabilidade genômica (CANTOR et al., 2001; KUMARASWAMY; SHIEKHATTAR, 2007). Um outro estudo relata que a depleção de BRIP1 também promoveu alterações na morfogênese mamária levando um perfil neoplásico pelo aumento da proliferação celular e capacidade de invasão (DAINO et al., 2013). Variantes em homozigose neste gene estão principalmente associadas à Anemia de Fanconi, no entanto, variantes no gene *BRIP1* estão relacionadas ao aumento de risco para câncer de ovário, sendo este gene descrito como de baixa ou nenhuma penetrância para câncer de mama (EASTON et al., 2016; RAMUS et al., 2015; SEAL et al.,

2006).

1.3.2 A vida de reparo de bases mal pareadas (MMR)

A deficiência de proteínas relacionadas à via de reparo de bases mal pareadas no DNA (MMR) é causa principal da segunda Síndrome de câncer hereditário mais frequente: a Síndrome de Lynch, que além de casos de cânceres colorretal e endometrióide, apresenta em seu espectro casos de câncer de mama e de ovário (AL BAKIR; GABRA, 2014). Apesar de variantes em genes da via HR estarem mais associados ao risco para HBOC, com o advento do sequenciamento de painéis gênicos diversas variantes têm sido identificadas em genes que nem sempre refletem o quadro clássico associado a algumas síndromes (KAPOOR et al., 2015). Assim, diversos estudos têm relatado a presença de variantes nos genes *MLH1*, *MSH2*, *MSH6* e *PMS2* em pacientes com HBOC e associados ao risco para câncer de mama e de ovário (BONACHE et al., 2018; CASTÉRA et al., 2014; KAPOOR et al., 2015; MINION et al., 2015; SUSZYNSKA et al., 2019; TEDALDI et al., 2017), inclusive na população brasileira (TIMOTEO et al., 2018). Espenschied e colaboradores (2017), sugerem em seu estudo que pacientes portadores de variantes nos genes *MSH6* e *PMS2* apresentaram um fenótipo mais próximo de HBOC do que de Lynch devido à presença de casos de cânceres de mama e ovário. O mesmo grupo evidenciou ainda que pode haver sobreposição fenotípica de casos de Lynch e HBOC, como evidenciaram em sua série de casos estudada em que 36% dos casos preenchiem critérios clínicos para ambas as síndromes.

O sistema de reparo MMR desempenha papel fundamental na degradação e reconstrução de regiões no DNA contendo bases mal pareadas causadas por erros no sistema de checagem da DNA polimerase, bem como na correção de inserções e deleções causadas por *loops* em regiões de DNA repetitivo como regiões de microssatélite (JIRICNY, 2006). O complexo MutS α é responsável pela detecção do dano e se liga às bases pareadas incorretamente (**Figura 3**). MutS α é um heterodímero composto pelas proteínas MSH2 e MSH6, que interage com uma molécula de ATP promovendo a sua mudança conformacional em uma estrutura em grampo, essencial para o seu deslocamento pela fita de DNA. O complexo MutS β é formado por MSH2 e MSH3 e está mais relacionado ao reconhecimento de grandes estruturas em loop (JIRICNY, 2006; LI, 2008). Em seguida, o estado MutS α ativo permite sua interação com MutL α , um heterodímero formado por MLH1 e PMS2, que é ativado por PCNA e RFC, permitindo a clivagem e a remoção da fita recém sintetizada contendo o erro, por endo (FEN1) e exonucleases (EXO1). A síntese de uma nova fita livre de erros é realizada pelas DNA polimerases δ ou ϵ , seguida pela ligação das extremidades dos filamentos recém-sintetizados (**Figura 3**) (FRIEDHOFF; LI; GOTTHARDT, 2016; KUNKEL; ERIE, 2015).

Figura 3. Principais etapas e proteínas envolvidas no processo de reparo de bases mal pareadas (MMR). Imagem adaptada de Germano e colaboradores (2018).

O correto funcionamento de MMR é crucial para o processo de replicação de sequências repetitivas no DNA. Quando há a deficiência em MMR, genes que possuem sequências de microssatélites dentro de suas regiões codificantes podem acumular variantes e desencadear um processo tumorigênico em decorrência da instabilidade genômica gerada (JIRICNY, 2006; KOBAYASHI et al., 2013). Assim, tumores com variantes em genes como *MLH1*, *MSH2*, *MSH6* e *PMS2* geralmente apresentam microssatélites que diferem no tamanho em relação tecidos normais, causando um fenômeno denominado instabilidade de microssatélites ou MSI, que hoje já é usado como diagnóstico para síndrome de Lynch (ROMERO-LAORDEN; CASTRO, 2017). Para os cânceres colorretal, endometrióide e de estômago, até 30% dos tumores apresentam alta MSI. Já os cânceres de mama e ovário, os números são um pouco menores mas cerca de 1,7% e 3,7% dos tumores possuem alta instabilidade de microssatélites (GERMANO et al., 2018).

Com a grande quantidade de dados genômicos disponíveis, estudos recentes têm demonstrado que células que apresentam variantes em genes da via MMR e por isso apresentam alta MSI podem

ser beneficiados com novas terapias utilizando inibidores de WRN, uma proteína da família das RecQ helicases (SETON-ROGERS, 2019). De fato, recentemente o FDA aprovou a droga pembrolizumab, um inibidor de PD-1, que é indicada para tumores que apresentam alta MSI, incluindo tumores de mama (U.S. FOOD AND DRUG ADMINISTRATION, 2019).

Além de genes relacionados às vias de reparo de danos ao DNA, variantes em genes supressores tumorais como *PTEN* e *CDHI* também tem sido frequentemente relatadas em pacientes com HBOC. *CDHI* codifica a proteína de adesão E-caderina, e variantes que afetam a sua função estão associadas ao comprometimento da adesão celular, aumento da capacidade invasiva das células e por consequência, à quadros de metástase (MATEUS et al., 2009). Variantes neste gene estão mais associadas a casos de câncer gástrico, mas estudos têm caracterizado *CDHI* como um gene de penetrância moderada para câncer de mama, evidenciando uma frequência de até 4% de variantes em pacientes com HBOC (KOBAYASHI et al., 2013; MINION et al., 2015; SUSZYNSKA et al., 2019). O *PTEN* é outro gene que é frequentemente mutado em diversos tipos de cânceres, e está associado à regulação negativa da via de sinalização AKT, importante na sobrevivência celular (KOBAYASHI et al., 2013). Embora variantes em *PTEN* estejam principalmente relacionadas à Síndrome de Cowden, também são associadas a um alto risco para câncer de mama (SUSZYNSKA et al., 2019), apresentando uma frequência de variantes de até 6% em casos de câncer de mama e ovário hereditário (BONACHE et al., 2018; BUYS et al., 2017; MINION et al., 2015).

2 JUSTIFICATIVA

De acordo com estatísticas do INCA (2017), o câncer de mama constitui a maior causa de óbitos por câncer na população feminina brasileira principalmente na faixa etária de 40 e 69 anos, devido à sua alta incidência (FERLAY et al., 2015; SIEGEL; MILLER, 2019). Já o câncer de ovário, apesar de menos frequente, é um tumor de difícil diagnóstico, principalmente devido à presença de sintomas iniciais vagos e inespecíficos o que faz com que cerca de 80% dos casos sejam diagnosticados em estágio mais avançados da doença (BURGESS; KOLESAR, 2018; SIEGEL; MILLER, 2019; WHITTEMORE; HARRIS; ITNYRE, 1992).

Cerca de 10-15% desses cânceres são hereditários, o que faz do teste genético aplicado às famílias com forte história familiar e/ou pessoal de HBOC seja essencial para o conhecimento do risco

genético e para o correto manejo destes pacientes (PLEVRITIS et al., 2006). No entanto, apesar de protocolos de manejo clínico já serem bem estabelecidos para pacientes com variantes em genes de alta penetrância como *BRCA1/BRCA2*, aqueles pacientes negativos para alterações nestes genes sofrem com a falta de acompanhamento clínico e aconselhamento genético adequados, mesmo apresentando características clínicas similares, forte história familiar e risco aumentado para câncer (LEE et al., 2018), o que reforça a necessidade de estudos na caracterização de outros genes relacionados ao aumento de risco para HBOC. Estudos têm evidenciado uma taxa de diagnóstico molecular de 4,6% - 54% na população mundial quando apenas *BRCA1/BRCA2* são investigados, evidenciando o papel de outros genes, embora com menor penetrância, na patogênese do HBOC (BUYS et al., 2017; DAROOEI et al., 2017; SCHROEDER et al., 2015; TEDALDI et al., 2017).

Com a popularização das tecnologias de sequenciamento de nova geração, genes relacionados à mesma via de reparo de *BRCA1/BRCA2*, como *ATM, BARD1, BRIP1, CHEK2, MRE11, NBN, PALB2, RAD51C, RAD50* e *TP53*, bem como *MLH1, MSH2, MSH6* e *PMS2*, relacionados à via de reparo MMR, têm sido frequentemente relatados como mutados em casos de câncer de mama e ovário (ALORAIFI et al., 2015; HIROTSU et al., 2015; MINION et al., 2015; SCHUBERT et al., 2019; SUSZYNSKA et al., 2019; TEDALDI et al., 2017; TIMOTEO et al., 2018; TOSS et al., 2015). Estes genes não só apresentam uma alta frequência de variantes, como muitos deles também apresentam penetrância alta a moderada para OC e BC, já sendo considerados para o manejo clínico (NCCN GUIDELINES, 2019).

No Brasil, apenas o estudo de Timoteo e colaboradores (2018), utilizou uma abordagem de análise de painel gênico, evidenciando uma frequência de 8,3% de variantes patogênicas em *BRCA1/BRCA2* e de 3,8% em genes como *ATR, CDH1, MLH1* e *MSH6* em pacientes sob risco de HBOC. Demais estudos têm abordado apenas a análise de regiões codificantes dos genes *BRCA1/BRCA2* evidenciando uma frequência de variantes de 16,6% - 21,5%; bem como o *screening* de variantes pontuais e microdeleções em genes como *CHEK2, PTEN* e *POLQ* (CIPRIANO et al., 2019; FERNANDES et al., 2016; PALMERO et al., 2016; SILVA et al., 2014).

Portanto, propusemos neste projeto a inclusão da estratégia de investigação baseada no sequenciamento de painel gênico, incluindo *BRCA1* e *BRCA2* e outros 19 genes relacionados às vias de reparo do DNA, com o objetivo de identificar variantes causais, caracterizar variantes de significado incerto para o fenótipo, diminuir o número de casos sem diagnóstico molecular para HBOC e assim realizar o diagnóstico clínico e molecular das famílias atendidas no Ambulatório de Genética do Câncer do Hospital das Clínicas de Ribeirão Preto da FMRP/USP. Os resultados deste projeto

apresentam importante papel epidemiológico e translacional, uma vez que a utilização desta abordagem de análise molecular permitiu a caracterização de outros genes além de *BRCA1* e *BRCA2* na etiologia do HBOC na população brasileira, como também permitiu assegurar o manejo e a vigilância clínica adequados para o rastreamento das neoplasias associadas à HBOC nestes pacientes.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Identificar variantes germinativas em genes relacionados às vias de reparo do DNA e caracterizar a contribuição destas variantes na etiologia molecular da Síndrome Hereditária do Câncer de Mama e/ou Ovário (HBOC) em pacientes atendidas pelo serviço de Genética Médica do Hospital das Clínicas de Ribeirão Preto.

3.2 OBJETIVOS ESPECÍFICOS

- Caracterizar clinicamente os pacientes com suspeita da Síndrome Hereditária de Câncer de Mama e/ou Ovário;
- Identificar e caracterizar variantes germinativas nos genes *ABRAXAS1* ou *FAM175A*, *ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MRE11*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* e *UIMC1*, em pacientes com suspeita clínica de HBOC;
- Validar por sequenciamento Sanger e caracterizar funcionalmente por meio de ferramentas *in silico* as variantes encontradas;
- Identificar alterações no número de cópias dos genes *BRCA1* e *BRCA2* em pacientes que não apresentam variantes pontuais nestes genes, naqueles que não foram identificadas quaisquer outras variantes, bem como em pacientes que tiveram variantes de significado incerto associadas ao fenótipo;
- Correlacionar os genótipos encontrados com os fenótipos das pacientes;
- Realizar análises de associação das variantes selecionadas com o fenótipo utilizando bases de dados públicos populacionais.
- Caracterizar funcionalmente as variantes candidatas por meio de testes *in vivo* e *in vitro*;
- Agregar à literatura científica, informações de variantes genéticas que possam auxiliar no diagnóstico futuro de pacientes com HBOC;
- Auxiliar no aconselhamento genético pré e pós-teste mais acurado aos indivíduos indicados para o teste genético.

4 MATERIAIS E MÉTODOS

Figura 4. Esquema de avaliação das pacientes, de acordo com as metodologias de análise utilizadas. Os números presentes nas caixas cinzas retratam os tópicos correspondentes no tópico Material e Métodos.

O plano de avaliação dos pacientes está resumidamente ilustrado na **Figura 4** e está descrito de forma detalhada em cada uma das seções seguintes. Um total de 95 pacientes com suspeita clínica de HBOC de acordo com os critérios NCCN (<https://www.nccn.org/>) v.2.2015 foram caracterizados clinicamente à partir da coleta de informações de prontuários médicos. Após isso, as amostras foram analisadas quanto à presença de variantes nos genes *ABRAXAS1*, *ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MRE11*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* e *UIMC1*. As variantes encontradas foram avaliadas quanto ao seu impacto para o fenótipo utilizando-se informações disponíveis no ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), escores de predição de patogenicidade *in silico* e quanto à anotação das variantes pela ferramenta SnpEff (<http://snpeff.sourceforge.net/>). A frequência alélica das variantes classificadas como patogênicas e de

significado incerto (VUS) foi utilizada para uma análise de associação com o fenótipo quando comparadas às frequências alélicas para as mesmas variantes encontradas nas bases de dados populacionais. Após essa análise, aquelas variantes candidatas a causais (sabidamente patogênicas e de significado incerto ou conflitante associadas ao fenótipo), bem como todas as variantes identificadas nos genes *BRCA1* e *BRCA2* foram validadas por meio do sequenciamento Sanger. As amostras que apresentaram VUS associadas com HBOC, bem como amostras negativas para variantes no genes *BRCA1/BRCA2* foram investigadas quanto à presença de alterações no número de cópias (MLPA) nos genes *BRCA1* e *BRCA2*. Foram realizados ensaios funcionais de patogenicidade para algumas das variantes associadas ao HBOC na tentativa de caracterizar a contribuição destas para o fenótipo. Após os ensaios funcionais de patogenicidade, propostos como última etapa de investigação laboratorial, foram realizadas as correlações com as informações clínicas das pacientes a fim de traçar os perfis “genótipo *versus* fenótipo” para as diferentes características clínicas encontradas. Por fim foi possível fornecer o diagnóstico molecular às pacientes e família de maneira mais informativa.

4.1 CARACTERIZAÇÃO DAS AMOSTRAS E ASPECTOS ÉTICOS

Foram avaliadas amostras de 95 indivíduos com histórico pessoal de câncer, que preencheram os critérios do NCCN (*National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology* v.2.2015 aplicado à Síndrome do Câncer Hereditário de Mama e/ou Ovário, e que foram atendidos no Ambulatório de Aconselhamento Genético do Câncer (AAGC) do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (HCFMRP-USP) até o primeiro semestre de 2016. Foi utilizado como critério de exclusão possuir score de probabilidade de variante em *BRCA1/BRCA2* PennII < 10 (<https://pennmodel2.pmacs.upenn.edu/penn2/>), uma vez que o número de pacientes com suspeita clínica era superior à quantidade de pacientes que poderíamos estudar (n=95).

A maior parte das amostras utilizadas já tinham sido previamente coletadas e fazem parte do biorrepositório de amostras existente no Centro de Medicina Genômica (HCFMRP-USP). As demais amostras foram coletadas à medida que as pacientes eram atendidas pelo AAGC e também foram estocadas no mesmo biorrepositório mediante prévia autorização dos pacientes. Os participantes foram previamente esclarecidos sobre a pesquisa e consultados sobre a possibilidade de participação e as amostras somente eram coletadas após o consentimento de participação e aplicação do Termo de Consentimento Livre e Esclarecido (TCLE). A utilização de todas as amostras e a realização deste

estudo foi aprovada pelo Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto - USP, sob processo nº 2819/2016 (Anexo I).

Como grupo controle, foram utilizados dados da análise de exoma de 28 indivíduos com idade acima de 70 anos, sem histórico pessoal de câncer, que já haviam sido coletadas em parceria com o Ambulatório de Geriatria do HCFMRP-USP e que fazem parte do mesmo banco de amostras citado anteriormente. Estas amostras foram gentilmente sequenciadas no laboratório de Genética Molecular da UNICAMP, sob responsabilidade da Profa. Dra. Iscia Lopes Cendes. A utilização de idosos como grupo controle é interessante para descartar variantes encontradas que não estão relacionadas ao câncer de acometimento em idade jovem. Para algumas análises, dados populacionais obtidos dos bancos de dados 1000 Genomes Project (<http://www.internationalgenome.org/>), ExAC (<http://exac.broadinstitute.org/>), bem como o ABraOM (<http://abraom.ib.usp.br/>) - banco de dados exclusivo de idosos da população brasileira, foram utilizados como grupo controle.

As amostras de pacientes com suspeita de HBOC e do grupo controle foram analisadas para alterações nas regiões codificantes e regiões 5' e 3'UTR (do inglês, *untranslated*) dos genes *ABRAXAS1* ou *FAM175A*, *ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MRE11*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* e *UIMC1*.

4.2 CARACTERIZAÇÃO CLÍNICA DOS PACIENTES

Os 95 pacientes investigados tiveram seus prontuários médicos revisados quanto à presença de histórico familiar de câncer, subtipo histológico, tipo, grau, estadiamento e status dos receptores hormonais do tumor, idade ao diagnóstico e idade ao óbito (quando houvesse). O status dos receptores hormonais foram classificados de acordo com a ASCO (*College of American Pathologists guidelines*), exceto para os receptores de estrogênio, que foram considerados como positivos apenas aqueles casos com mais de 10% de células positivas (YI et al., 2014). Também foram coletadas informações sobre os tumores, tais como: modo de instalação do tumor (localizado – mama unilateral; ou multifocal – vários tecidos ou mama bilateral), tamanho do tumor, presença de linfonodos e de metástases à distância. Uma vez traçado o perfil de variantes para cada um dos pacientes, foi realizada a correlação com o perfil clínico na tentativa de caracterizar alguma associação. Os testes genéticos foram realizados como parte do projeto de pesquisa, no entanto, tendo em vista o alto custo da análise genética para os pacientes, os resultados foram laudados e retornados aos pacientes como parte do aconselhamento genético.

4.3 SEQUENCIAMENTO DO PAINEL GÊNICO

4.3.1 Preparo da biblioteca genômica para análise dos genes *ABRAXAS1* ou *FAM175A*, *ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MRE11*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* e *UIMC1*

Para o sequenciamento do painel gênico foram customizados oligonucleotídeos que flanqueiam as regiões *upstream* e *downstream* dos éxons de cada um dos 21 genes investigados, formando um total de 708 amplicons com uma cobertura média de 98% das regiões de interesse. A customização é feita por meio da ferramenta DesignStudio™ (Illumina® - <https://designstudio.illumina.com/>).

Para a preparação da biblioteca genômica sequenciada foi utilizado o Kit *TruSeq Custom Amplicon Library Preparation* da Illumina® para 96 amostras. O primeiro passo foi a hibridação dos oligonucleotídeos customizados ao DNA genômico. Depois foi realizada a remoção daqueles oligos que não se hibridaram ao DNA genômico e utilizando um filtro seletor de tamanho de fragmento, aqueles fragmentos menores (oligos) foram descartados. Foram realizadas três lavagens para assegurar a completa remoção destes oligos e preparar as amostras para o próximo passo que era o de extensão e ligação. Nesta etapa, a DNA polimerase estende a região alvo a partir do oligo *upstream* e se conecta à extremidade 5' *downstream* por meio de uma ligase, resultando em sequências de interesse flanqueadas por oligos, necessários à amplificação (**Figura 5B**). Após isso foi realizada a amplificação destes fragmentos de interesse, utilizando *primers* específicos contendo tanto sequências *index* ou identificadoras (i5 e i7), importantes para a identificação das sequências geradas, como adaptadores (P5 e P7), necessários para a geração de *clusters* de DNA sequenciados (**Figura 5C e 5D**). O ciclo utilizado para a amplificação foi: 95°C por 3 min, 30 ciclos de: 95°C por 30s; 66°C por 30 s; 72°C por 60 s; e 72°C por 5 min. Após a amplificação foi realizada a purificação dos produtos da PCR com *beads* magnéticas (AMPure XP beads) a fim de descartar outros produtos da reação. Foi então realizada a normalização das quantidades de DNA para assegurar uma quantidade equivalente de DNA para cada uma das bibliotecas quando todas as amostras foram combinadas. Após isso, o *pool* das bibliotecas foi quantificado por PCR em tempo real, utilizando o kit Kapa Library Quantification (KAPA Biosystems). Após a quantificação, a biblioteca foi diluída a 17 pM no buffer de hibridação do kit de preparação da biblioteca. Em seguida, as sequências foram desnaturadas e a solução foi então inoculada no cartucho 500 ciclos v2 de sequenciamento da plataforma MiSeq Illumina®.

Figura 5. Esquema representativo da geração dos fragmentos de interesse a partir de DNA genômico da biblioteca sequenciada. FONTE: Documentação Illumina® - https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_truseq/truseqampliconcancerpanel/truseq-amplicon-cancer-panel-reference-guide-15031875-02.pdf

4.3.2 Reação de Sequenciamento

A primeira etapa da reação de sequenciamento é a clusterização. Durante esta etapa, no sequenciador, as moléculas de DNA em fita simples se ligam à superfície da *flow cell* e em seguida, são amplificadas por ponte em 500 ciclos de sequenciamento, *paired-end* (2x250), formando *clusters*. Para isso, os reagentes de uso para a reação, bem como o cartucho contendo as bibliotecas previamente preparadas foram inseridos no aparelho junto com a *flow cell*. Durante a reação, imagens dos *clusters* gerados iam sendo capturadas por meio de LED e combinações de filtros específicos para cada um

dos quatro dideoxinucleotídeos marcados com fluorescência. À medida que uma *tile* da *flow cell* tinha suas imagens capturas, a *flow cell* era movida a fim de expor a próxima *tile* para captura de imagens. Este processo se repetiu em cada ciclo do sequenciamento.

4.3.3 Análise de dados

Os dados gerados pelo sequenciador MiSeq são imagens das sequências e estas são convertidas em arquivos *bcl*, um tipo de arquivo binário que contém informações básicas das bases nucleotídicas. Utilizando o software FASTQ Generation v.1.0.0 por meio da ferramenta BaseSpace (Illumina®) (<https://basespace.illumina.com>), estes arquivos foram então convertidos em arquivos *fastq* que já são arquivos que permitem a visualização de qualidade das sequências (*reads*) que foram geradas. Estas informações de qualidade das *reads* são visualizadas utilizando-se a ferramenta fastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Após isso foi realizado o pré-processamento e mapeamento destas *reads* ao genoma referência GRCh37.75/hg19 (<http://hgdownload.cse.ucsc.edu/>), utilizando o software GATK v.3.6-0 (<https://software.broadinstitute.org/gatk/>) de acordo com o manual de boas práticas (**Figura 6**). A primeira etapa é a geração de arquivos sam (*sequencing alignment and map*) por meio da ferramenta BWA mem (<http://bio-bwa.sourceforge.net/>). Com a ferramenta AddOrReplaceReadGroups da linha de comando Picard (<https://broadinstitute.github.io/picard/>) foram adicionados a estes arquivos de mapeamento as informações de identificação de cada uma das amostras, e as *reads* mapeadas foram ordenadas de acordo com as coordenadas genômicas, resultando em um arquivo .bam (binário) que foi utilizado para realizar a chamada de variantes. Normalmente é realizada também uma marcação de sequências que estão em duplicata, no entanto para o presente protocolo de análise esta etapa não foi necessária uma vez que a biblioteca foi gerada pela metodologia de *amplicons*. Foi realizada ainda uma recalibração de indels por meio do IndelRealigner (https://software.broadinstitute.org/gatk/documentation/tooldocs/current/org_broadinstitute_gatk_to_ols_walkers_indels_IndelRealigner.php), etapa importante para evitar a perda pareamento das *reads* de indels. Foi realizado um ajuste nas posições de cada uma das *reads* que continham variantes, que é a etapa de *base recalibration*, baseada em dados do dbSNP. Foi então realizada a chamada de variantes com a ferramenta Haplotype Caller (https://software.broadinstitute.org/gatk/documentation/tooldocs/current/org_broadinstitute_gatk_to_ols_walkers_haplotypecaller_HaplotypeCaller.php) que gera um arquivo .vcf (*variant call format*).

Best Practices for Germline SNPs and Indels in Whole Genomes and Exomes - June 2016

Figura 6. Fluxograma de análise do manual de boas práticas do GATK. FONTE: Website do GATK – https://software.broadinstitute.org/gatk/img/BP_workflow_3.6.png

Foram utilizados alguns filtros de qualidade e mapeamento/alinhamento das sequências, que são padrões do manual de boas práticas do GATK. Para os SNPs, os filtros utilizados foram: $QD < 2.0$; $MQ < 40.0$; $FS > 60.0$; $MQRankSum < -12.5$ e $ReadPosRankSum < -8.0$; e para os indels: $QD < 2.0$, $FS > 200.0$, $ReadPosRankSum < -20.0$ e $InbreedingCoeff < -0.8$. Estes filtros aplicados baseiam-se em escores gerados por testes estatísticos. A aplicação destes permite a normalização dos escores de qualidade, eliminação de vieses de mapeamento em relação à qual fita está sendo mapeada (*forward* ou *reverse*) e de erros de identificação de variantes encontradas no início ou fim das sequências de cada *read*, que em geral têm uma qualidade comprometida. Foram gerados escores de confiança da chamada de alelos alterados em relação ao alelo referência; escores de mapeamento *de novo* de regiões de haplótipos, bem como escores de estimação de coeficiente de endogamia.

4.3.4 Anotação das variantes encontradas

Em seguida as variantes foram anotadas utilizando a ferramenta Snpeff (<http://>

<http://snpeff.sourceforge.net/>) que agrega à lista de variantes informações como o tipo de variante, informações de predição de efeito destas variantes nos diferentes transcritos de um gene, e a frequência em bancos de dados populacionais, por meio de extensões com bancos de dados como ClinVar v.build147 (<https://www.ncbi.nlm.nih.gov/clinvar/>) e dbNSFP v.2.9 (<https://sites.google.com/site/jpopgen/dbNSFP>).

Os gráficos utilizados para análise dos dados foram gerados utilizando os pacotes GenVisR (SKIDMORE et al., 2016) e ComplexHeatmap (GU; EILS; SCHLESNER, 2016), ambos do Bioconductor (<https://bioconductor.org/packages/release/bioc/html/GenVisR.html>) em ambiente R.

4.4 MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION (MLPA)

Cada sonda de MLPA consiste em um par de oligonucleotídeos que se hibridam a regiões alvo adjacentes para que posteriormente sejam ligadas, por uma enzima ligase, formando uma única sonda. Estas sondas geradas são amplificadas simultaneamente por meio de uma reação de PCR que utiliza o mesmo par de *primers*, que são complementares a uma pequena sequência presente na extremidade de cada sonda do kit (**Figura 7**).

Figura 7. Esquema representativo da reação de MLPA. FONTE: MLPA® General Protocol - http://www.mlpa.com/WebForms/WebFormDBData.aspx?Tag=_G1U3PYAOzf2SFuaxkika4YjruAIwx3T3q8uAJJ_V-Ws.

Para isso foi utilizado o kit SALSA® MLPA® (MRC-Holland MLPA®), composto por sondas específicas para as regiões promotoras, regiões codificantes e demais éxons do gene *BRCA1* e do gene *BRCA2*. Para cada um dos genes, foram utilizadas amostras de DNA na concentração de 100 ng (5 µL de amostra de DNA a 20 ng/µL) que foram desnaturadas por 10 min a 98°C. Após isso foi realizada a reação de hibridação com uma incubação do DNA desnaturado + 1,5 µL de MLPA buffer + 1,5 µL do mix de sondas específicas do gene investigado, a 95°C por 1 min e 60°C por 8h. Depois, para cada reação foram adicionados 25 µL H₂O + 3 µL Ligase Buffer A + 3 µL Ligase Buffer B e 1 µL da enzima Ligase-65. Este mix de ligação foi incubado por 15 min a 54°C (para a reação de ligação

das sondas) e 5 min a 98°C para inativação da enzima Ligase-65. As amostras foram resfriadas a 20°C e, um mix contendo 7,5 µL de H₂O + 2 µL de SALSA PCR mix de *primers* + 0.5 µL de enzima polimerase SALSA à temperatura ambiente é adicionado à cada tubo. Foi realizada então a reação de PCR, com um programa de 35 ciclos de 30 segundos a 95°C; 30 segundos a 60°C e 60 segundos a 72°C, seguidos de 1 ciclo de 72°C por 20 min. Depois disso, 1 µL produto de PCR gerado para cada amostra foi então adicionado a 8,8 µL de Formamida + 0,2 µL de GeneScan™ 500 LIZ™ dye Size Standard (Thermo Fisher Scientific®) e foi realizada a análise de fragmento pelo ABI 3500XL Genetic Analyzer (Applied Biosystems™).

4.5 SCREENING PARA DETECÇÃO DA VARIANTE c.-107A>T NO GENE *BRCA1*

Evans e colaboradores (2018), evidenciaram que a presença da variante c.-107A>T na região 5' UTR do gene *BRCA1* era causa de câncer hereditário de mama e ovário, por levar à metilação da região promotora de *BRCA1*, promovendo a perda de expressão desta proteína. A fim de verificar a presença desta variante na amostra estudada e relacionar à perda de expressão de *BRCA1*, as 95 amostras foram sequenciadas utilizando primers específicos para o éxon 1 do gene *BRCA1* (Apêndice, Tabela A1). Para isso, foi realizada a reação de PCR utilizando 200ng de DNA de cada amostra, 2,5 µL de PCR buffer 10x (Invitrogen - Thermo Fisher Scientific®), 0,75 µL de MgCl₂ (Invitrogen - Thermo Fisher Scientific®), 2 µL de dNTPs (Invitrogen - Thermo Fisher Scientific®) a 1,25mM, 1,5 µL de primer *forward* a 2,5 nM, 1,5 µL de primer *reverse* a 2,5 nM, 0,5 µL de Taq DNA Polymerase (Invitrogen - Thermo Fisher Scientific®) completando com água *nuclease-free* (Invitrogen - Thermo Fisher Scientific®) para um volume final igual a 25 µL. A reação foi então amplificada utilizando o seguinte ciclo: 94°C por 5 min, seguido de 35 ciclos de 94°C 1 min, 61°C 1 min e 72°C 1 min, seguido de 1 ciclo de 72°C por 5 min e depois 4°C infinito. Após isso, os produtos de PCR foram enviados para sequenciamento Sanger, utilizando os primers específicos para o éxon 1 do gene *BRCA1*, e BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific®), seguindo o protocolo estabelecido pelo fabricante. As amostras foram sequenciadas no sequenciador automático XL 3500 Genetic Analyzer (Applied Biosystems - Thermo Fisher Scientific®).

4.6 SCREENING PARA DETECÇÃO DE INSERÇÃO Alu NO GENE *BRCA2*

A inserção c.156_157insAlu é uma mutação fundadora identificada na população Portuguesa e

que apresenta uma frequência relativamente alta na população Brasileira. Esta variante é caracterizada pela inserção da sequência Alu no éxon 3 do gene *BRCA2*, sendo uma variante já caracterizada como patogênica (PEIXOTO et al., 2009). Para a análise desta variante, são utilizadas duas reações de PCR, uma para amplificação do éxon 3 do gene *BRCA2* e outra específica para a região inserida, utilizando primers específicos para amplificação da sequência Alu (Apêndice, Tabela A1). Para a primeira reação de amplificação utilizando *primers* específicos para o éxon 3 do gene *BRCA2*, foram utilizados: 100ng de DNA de cada amostra, 3 µL de PCR buffer 10x (Invitrogen - Thermo Fisher Scientific®), 1,8 µL de MgCl₂ (Invitrogen - Thermo Fisher Scientific®), 2 µL de dNTPs (Invitrogen - Thermo Fisher Scientific®) a 1,25mM, 5 nL de primer *forward* a 2,5 µM, 5 µL de primer *reverse* a 2,5 nM, 0,3 µL de Taq DNA Polymerase (Invitrogen - Thermo Fisher Scientific®) completando com água *nuclease-free* (Invitrogen) para um volume final igual a 30 µL. A reação foi então amplificada utilizando o seguinte ciclo: 94°C por 3 min, seguido de 35 ciclos de 94°C 1 min, 52°C 1 min e 72°C 4 min, seguido de 1 ciclo de 72°C por 10 min e depois 4°C infinito. O fragmento esperado para a amplificação normal do éxon 3 é de 201 pb, no entanto, quando há a inserção da sequência Alu, há a amplificação de um fragmento de cerca de 500 pb uma vez que as sequências Alu possuem um tamanho aproximado de 300 nucleotídeos. Quando uma amostra apresentou a amplificação de uma banda correspondente à amplificação de Alu, foi realizado uma segunda reação de PCR utilizando *primer forward* específico para Alu (Apêndice, Tabela A1). Para isso, foram utilizados 200ng de DNA da amostra, 15 µL de JumpStart™ Taq ReadyMix (Sigma Aldrich - Merck®), 6 µL de primer *forward* a 2,5 nM, 6 µL de primer *reverse* a 2,5 nM, completando com água *nuclease-free* (Invitrogen) para um volume final igual a 30 µL. O ciclo de amplificação foi exatamente igual ao utilizado para amplificação do éxon 3. As amostras que apresentaram amplificação utilizando primers específicos para Alu foram enviadas para sequenciamento Sanger, utilizando os mesmos *primers* utilizados para amplificação do fragmento contendo Alu, e BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific®), seguindo o protocolo estabelecido pelo fabricante. As amostras foram sequenciadas no sequenciador automático XL 3500 Genetic Analyzer (Applied Biosystems - Thermo Fisher Scientific®).

4.7 SCREENING DA VARIANTE p.Arg337His NO GENE *TP53*

Variantes no gene *TP53* estão principalmente relacionadas à Síndrome de Li Fraumeni. No entanto, a variante p.Arg337His localizada no éxon 10 do gene *TP53* é uma mutação fundadora identificada na região Sul do Brasil, que já é descrita como segregada em famílias com histórico de câncer adrenocortical, carcinoma do plexo coróide, sarcomas e câncer de mama em idade jovem

(CURY; FERRAZ; SILVA, 2014; HAHN et al., 2018). Para verificar a prevalência desta variante na série de casos estudada, todas as amostras foram amplificadas utilizando primers específicos para o éxon 10 do gene *TP53* e depois enviadas para sequenciamento Sanger. Para isso, foi realizada a reação de PCR utilizando 200ng de DNA de cada amostra, 2,5 µL de PCR buffer 10x (Invitrogen - Thermo Fisher Scientific®), 0,75 µL de MgCl₂ (Invitrogen - Thermo Fisher Scientific®), 2 µL de dNTPs (Invitrogen - Thermo Fisher Scientific®) a 1,25mM, 1,5 µL de primer *forward* a 2,5 nM, 1,5 µL de primer *reverse* a 2,5 nM, 0,5 µL de Taq DNA Polymerase (Invitrogen - Thermo Fisher Scientific®) completando com água *nuclease-free* (Invitrogen - Thermo Fisher Scientific®) para um volume final igual a 25 µL. A reação foi então amplificada utilizando o seguinte ciclo: 94°C por 5 min, seguido de 35 ciclos de 94°C 1 min, 64°C 1 min e 72°C 1 min, seguido de 1 ciclo de 72°C por 5 min e depois 4°C infinito. Após isso, os produtos de PCR foram enviados para sequenciamento Sanger, utilizando os primers específicos para o éxon 10 do gene *TP53*, e BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific®), seguindo o protocolo estabelecido pelo fabricante. As amostras foram sequenciadas no sequenciador automático XL 3500 Genetic Analyzer (Applied Biosystems - Thermo Fisher Scientific®)

4.8 ANÁLISE DE HAPLÓTIPOS NO GENE *BRCA1* ASSOCIADOS AO RISCO PARA HBOC

Para a análise de haplótipos, amostras de 136 idosos foram genotipadas para as 5 variantes de alta frequência no gene *BRCA1*. Para a genotipagem, foram utilizadas cinco sondas TaqMan (Applied Biosystems - Thermo Fisher Scientific®) para as variantes rs1799966 (assay ID: C_2615208_20), rs8176318 (assay ID: AHX1AK8), rs16941 (assay ID: C_2287888_10), rs799917 (assay ID: C_2287943_10) e rs16942 (C_2287889_20). O mix para a amplificação utilizando as sondas específicas foi realizada utilizando 2 µL de DNA a 5ng/µL, 5 µL de TaqMan master mix (Applied Biosystems - Thermo Fisher Scientific®), e 0,25 µL (200 nM) de cada uma das sonda, completando com água *nuclease-free* (Invitrogen - Thermo Fisher Scientific®) para um volume final igual a 10 µL. O mix foi feito para cada uma das sondas, para cada uma das 136 amostras, utilizando placas de 96 poços. O ciclo utilizado para a amplificação foi: 95°C por 10 min, 40 ciclos de 92°C por 15 s e 60°C por 1 min. A leitura das placas foi realizada utilizando o 7500 Real-Time PCR Systems (Applied Biosystems - Thermo Fisher Scientific®) e os resultados foram analisados utilizando o software do fabricante.

Após obtenção dos resultados da genotipagem, foi realizada a análise de frequências dos

possíveis haplótipos, baseado na estimativa máxima de verossimilhança, utilizando o pacote Haplo.stats (SINNELL, SCHAID, 2018), em ambiente R. Este programa também calcula a associação dos haplótipos estimados com a doença, considerando valor de $p < 0.05$ como estatisticamente significativo.

4.9 ANÁLISE DE PATOGENICIDADE *in silico* E VALIDAÇÃO DAS VARIANTES CANDIDATAS

As informações de efeito e patogenicidade foram plotadas na anotação com o SnpEff, como citado anteriormente. Essas informações são extraídas pelo dbNSFP v.2.9, que compila algoritmos de predição de patogenicidade *in silico* como o UMD-Predictor, Sift, PolyPhen-2 e Mutation Taster para as variantes já anotadas. Para uma caracterização *in silico* mais acurada, foram utilizadas ainda as ferramentas de predição de patogenicidade CADD (<https://cadd.gs.washington.edu/>) e AlignGVGD (http://agvgd.hci.utah.edu/agvgd_input.php). Além disso, a ferramenta SnpEff ainda dá uma predição de impacto na escala de alto, moderado ou baixo, que se baseia no tipo de variante encontrada. Por exemplo, variantes *frameshift*, de ganho de códon de parada, que afetam sítios de contatos com outras proteínas e que afetam as duas bases importantes para o reconhecimento dos sítios de *splicing* são assumidas como tendo um alto impacto na proteína, causando perda ou truncamento de sua função. Já variantes como *missense* ou deleções/inserções *inframe* podem manter algumas propriedades funcionais da proteína e, portanto, são consideradas de impacto moderado. Existem ainda as variantes consideradas de baixo impacto, que são aquelas menos prováveis de causar algum efeito na funcionalidade da proteína como aquelas variantes sinônimas ou intrônicas situadas próximas às regiões de *splicing*; e aquelas de impacto dito modificador, que engloba as variantes em regiões não codificantes ou em genes que não codificam proteínas e que são de difícil predição ou não há qualquer evidência de impacto (CINGOLANI, 2017).

As variantes foram então filtradas a fim de selecionar apenas aquelas presentes em regiões codificantes (*missense*, *nonsense*, *frameshift*, inserções e deleções) e em sítios de *splicing*. Foram consideradas “Patogênicas” aquelas variantes assim caracterizadas no ClinVar; como “Possivelmente patogênicas” aquelas classificadas como de impacto alto pelo SnpEff, preditas como patogênicas por pelo menos 5 das 6 ferramentas *in silico* utilizadas, e classificadas como VUS pelo ClinVar. Algumas variantes foram classificadas como “Provavelmente benignas” quando se tratavam de variantes de alto impacto (SnpEff) mas mesmo assim eram classificadas como benignas pelo ClinVar e, “Benignas” quando assim eram classificadas pelo ClinVar e pelas ferramentas de predição *in silico*. As variantes candidatas a associadas ao fenótipo, bem como aquelas já classificadas como patogênicas foram

sequenciadas pela metodologia de Sanger para validação dos resultados encontrados pelo método de NGS. A validação foi realizada de acordo com as recomendações do kit BigDye® Terminator v3.1 Cycle Sequencing (Applied Biosystems™) para sequenciamento pelo ABI 3500XL 3500xL Genetic Analyzer (Applied Biosystems™).

4.10 ANÁLISE DE ASSOCIAÇÃO ENTRE FENÓTIPO E GENÓTIPO

Foram realizados testes de associação das variantes identificadas com o risco para HBOC. Para isso foram consultadas as frequências alélicas de cada uma das variantes na série de casos estudada, no grupo controle, bem como em bancos de dados populacionais. Foi utilizado o banco de dados da população brasileira ABraOm que é alimentado por dados do sequenciamento do exoma de 609 indivíduos acima de 60 anos. Quando a frequência das variantes era igual a zero nesta base de dados, foram utilizadas então as frequências alélicas das populações Europeia – não finlandesa, Latina, Americana, Africana e aquelas categorizadas como “outras”, disponíveis no 1000 Genomes ou ExAC. Foi realizada análise de odds ratio (OR) utilizando o teste exato de Fisher, e os valores de “p” foram obtidos utilizando o teste do X^2 de Pearson. Quando a frequência alélica foi igual a zero em todos os bancos de dados populacionais não foi possível calcular os valores de OR.

Para as análises de associação dos dados clínicos com os achados moleculares, foi utilizado teste do X^2 de Pearson. Para as curvas de sobrevivência, foram utilizadas curvas de sobrevivência Kaplan-Meier, utilizando o teste de log-rank. Foi também utilizada a regressão de Cox para análise de sobrevivência relativa a riscos proporcionais.

4.11 TESTES DE PATOGENICIDADE *IN VIVO*

Os testes funcionais para a análise de variantes de significado desconhecido ou duvidoso têm contribuído para a classificação destas variantes. Estes ensaios não estão diretamente relacionados ao entendimento dos processos de desenvolvimento do câncer, mas sim nas modificações funcionais ou estruturais nas proteínas, provendo evidências sobre a patogenicidade das variantes encontradas (MILLOT et al., 2011). Assim os testes funcionais foram realizados para as VUS que permanecerem com patogenicidade duvidosa, e para variantes candidatas a variantes causais significativamente associadas ao fenótipo.

4.11.1 Avaliação dos níveis de fosforilação de γ H2AX

Foi realizado o ensaio de reparo de quebras de dupla fita de DNA, cujo principal mecanismo de reparo envolvido é o de reparo por recombinação homóloga. O ensaio se baseia na quantificação de quebras na dupla fita de DNA pela marcação de H2AX, histona que substitui a histona H2A convencional em alguns nucleossomos e pode ser rapidamente fosforilada na serina 139 na região carboxi-terminal em resposta ao dano de DNA (WATTERS et al., 2009). Para isso, utilizando anticorpos específicos que detectam a histona H2AX fosforilada, é possível quantificar o nível de quebras de DNA após o tratamento com radiação ionizante. Para isso, leucócitos de pacientes portadores das variantes a serem testados foram extraídos utilizando Ficoll-Paque PLUS (GE Healthcare Life Sciences) de acordo com as recomendações do fabricante. Foram plaqueados 3×10^5 por poço numa placa de 96 poços, sendo utilizados três poços para cada condição (triplicata experimental). As células permaneceram em cultura por 3 dias em meio RPMI (Gibco - Thermo Fisher Scientific®) acrescido de 10% de soro bovino fetal (Gibco - Thermo Fisher Scientific®), fitohemaglutinina PHA-M 1% (Gibco - Thermo Fisher Scientific®) e interleucina-2 IL-2 30U/mL (Thermo Fisher Scientific®) em estufa 37°C CO₂ 5%. No terceiro dia as células foram submetidas à irradiação (raios X) a uma dose de 2 Gy (Energia = 6 MV, Taxa de dose = 400cGy/min) utilizando o acelerador linear UNIQUE (Varian). As células foram coletadas logo após a irradiação (tempo zero), 2h, 4h e 6h após a irradiação: o conteúdo de cada poço foi transferido para um tubo de 1,5 mL, os poços foram lavados com 150 μ L de PBS 1x e em seguida tripsinizados com 30 μ L de TrypLE (Gibco - Thermo Fisher Scientific®) por 5 minutos em estufa 37°C CO₂ 5%. O conteúdo dos poços foi transferido para seus respectivos tubos e as amostras foram centrifugadas a 1500 rpm por 3 minutos. As células foram lavadas com 1 mL de PBS 1x, centrifugadas a 1500 rpm por 3 minutos, o sobrenadante foi retirado e as células foram fixadas com solução de metanol 80% em PBS 1x a -20°C por pelo menos 24h. Após a fixação as células foram centrifugadas a 1800 rpm por 3 minutos e em seguida foram bloqueadas com solução BSA 2% em PBS 1x por 1h com agitação. Após o bloqueio, as células foram centrifugadas a 2200 rpm por 3 minutos, foi retirado o sobrenadante e as células foram incubadas com o anticorpo primário anti- γ H2AX (ABCAM, ab26350) 1:100, *overnight* a 4°C em agitador. No dia seguinte, foi adicionado 1 mL de BSA 2% em PBS 1x em cada tudo e as amostras foram centrifugadas a 2200 rpm por 3 minutos. Foi descartado o sobrenadante e as células foram incubadas com o anticorpo secundário IgG anti-mouse conjugado com Alexafluor 488 (ABCAM, ab150113) por 1h em agitador. Após isso, foi adicionado 1 mL de BSA 2% em PBS 1x em cada tudo e as amostras foram centrifugadas a 2200 rpm por 3 minutos. Foi descartado o sobrenadante e as células foram por citometria de fluxo (FACSCanto) e as células marcadas com γ H2AX fosforilada

foram quantificados. Os resultados representam a média $\pm$ SD de dois experimentos independentes, realizados em triplicata.

4.11.2 Mutagênese sítio-dirigida para geração dos plamídeos mutados

A estratégia de mutagênese foi utilizada para geração *in vitro* das variantes estudadas, a partir de vetores específicos contendo a sequência codificante de cada um dos genes estudados. Esta técnica se baseia em reações de PCR para a geração de uma sequência alvo encerrando a variante desejada. Para o teste de interação PALB2 e BRCA1, utilizando sistema duplo híbrido em mamíferos foi utilizado o kit QuickChange Site-Directed Mutagenesis System (Stratagene - Agilent Technologies) e Q5® Site-Directed Mutagenesis Kit (New England Biolabs).

Inicialmente, foi utilizada uma reação de PCR utilizando como molde o plasmídeo pVP16-PALB2 NT WT (Parte N-terminal da sequência selvagem de *PALB2*) e *primers* específicos para as variantes p.Lys18Arg e aquelas no códon 317 (Apêndice, Tabela A2). Para a reação foram utilizados 50 ng do plasmídeo molde, 10 μ L de master mix da enzima PrimeSTAR® Max (Clontech®), 1 μ L do *primer forward* a 10pM, 1 μ L do *primer reverse* a 10pM, completando o volume para 20 μ L com água nuclease-free (Invitrogen - Thermo Fisher Scientific®). Depois era realizada a reação de PCR utilizando o seguinte ciclo: 98°C por 10 s, 30 ciclos de 98°C por 10 s, 55°C 5 s e 72°C 40 s, seguidos de 72°C por 40 s e 4°C infinitos. Depois disso foi realizada a digestão do produto de PCR com a enzima DpnI (New England Biolabs) que atua seletivamente em uma sequência específica de DNA quando metilada (5'-Gm6ATC-3'), para eliminar o plasmídeo molde susceptíveis a esse tipo de digestão. Os produtos das reações de digestão foram, então, submetidos a rotinas de transformação e seleção de bactérias competentes *E. coli* DH5' α utilizando os métodos de eletroporação (AUSUBEL et al, 2002, NEUMANN et al., 1982). Todas as construções geradas foram confirmadas por sequenciamento pela metodologia de Sanger.

Para análise de sensibilidade a inibidores de PARP para as variantes p.Lys18Arg e p.Thr317Pro em *PALB2*, bem como para as variantes nos genes *BRIP1*, *CHEK2*, *MLH1* e *PMS2*, foram realizadas reações de mutagênese dirigida utilizando o kit Q5 site-directed mutagenesis protocol (New England Biolabs), de acordo com as recomendações do fabricante e utilizando *primers* e plasmídeos específicos para cada uma das variantes, em cada um dos genes analisados (Apêndice, Tabela A2). Após isso os produtos de PCR eram tratados com a reação KLD (*Kinase, Ligase and DpnI*) para remoção do plasmídeos molde metilado e depois foram submetidos a rotinas de transformação por choque térmico,

utilizando bactérias *E. coli* 5'α competentes (New England Biolabs) de acordo com as recomendações do fabricante. Todas as construções geradas foram confirmadas por sequenciamento pela metodologia de Sanger.

4.11.3 Análise da interação de PALB2-BRCA1 utilizando sistema de duplo-híbrido em mamíferos

Esse ensaio é baseado na ativação da transcrição da Luciferase (utilizada como gene repórter) por meio da interação entre as proteínas GAL4 ligada ao domínio BRCT de BRCA1 (isca), região de interação com PALB2 e, VP16 ligada à região N-terminal (aminoácidos 1 ao 319) de PALB2 (presa), que é a região de interação com BRCA1. Em condições normais, a proteína GAL4 interage com um domínio de ligação ao DNA (**Figura 8**), e quando há a interação de BRCA1 com PALB2, que está fusionada ao ativador transcrricional VP16, ocorre a transcrição de luciferase (**Figura 8**). Caso a variante testada interfira na interação de BRCA1 com PALB2, não haverá a ativação transcrricional de Luciferase (**Figura 8**). Desta forma, a interação entre as proteínas BRCA1-PALB2 foi avaliada pela atividade enzimática do produto do gene repórter.

Figura 8. Representação esquemática do sistema duplo híbrido em mamíferos para avaliação da interação de BRCA1 e PALB2. Adaptada de Rodrigue et al., 2019 (ainda não publicado)

Para isso, células HEK293FT foram plaqueadas em placa de 96 poços (3.000 células/poço) em DMEM 1x (Corning) acrescido de 10% de soro bovino fetal (Gibco - Thermo Fisher Scientific®) e 1% de Penicilina e Streptomicina (Gibco - Thermo Fisher Scientific®) (meio DMEM completo) e depois foram colocadas em estufa a 37°C, em atmosfera de 5% de CO₂. Após 24h, foi realizada a transfecção dos plasmídeos (**Tabela 1**), utilizando o reagente PEI 25K (3 µL de PEI:1 µg de DNA). Para isso os plasmídeos foram misturados ao PEI e com meio de cultura sem soro e sem antibiótico durante 30 minutos e em seguida, essa mistura foi adicionada a cada uma dos poços correspondentes da placa de 96 poços. Foram transfectados os plasmídeos contendo as variantes de interesse em PALB2, bem como as variantes p.L21A (SY; HUEN; CHEN, 2009) e p.L35P (FOO et al., 2017), utilizadas como controle

negativo de interação. Foram transfectados também o plasmídeo contendo a sequência da proteína luciferase repórter de *Photinus pyralis* com um promotor sob o controle de operadores de GAL4 (pG5Luc), e o plasmídeo de controle interno de transfecção (pGR-TK), que contém a região condificante da Luciferase de *Renilla reniformis* controlado pelo promotor HSV TK, constitutivamente ativo em células humanas (**Tabela 1**). As placas permaneceram na estufa a 37°C e após 3h o meio foi substituído por DMEM completo. Após 24h da transfecção, as células foram tratadas com o tampão de lise e uma alíquota do lisado celular foi transferida para um microtubo de 1,5mL para a leitura do ensaio no luminômetro GloMax (Promega). Foi realizada a leitura da atividade de luciferase utilizando o kit Dual-Luciferase® Reporter Assay System (Promega), seguindo as recomendações do fabricante. A atividade de luciferase foi expressa em porcentagem normalizada em relação a atividade média de PALB2 WT (selvagem). Os resultados representam a média $\pm$ SD de três experimentos independentes, realizados em quadruplicata.

Tabela 1. Concentração de DNA utilizado para a transfecção dos plasmídeos transfectados para o ensaio de interação BRCA1-PALB2 em sistema duplo híbrido em mamíferos.

Construção plasmidial	Variante	Quantidade de DNA transfectado
pG5luc		50ng/poço
pGR-TK		2ng/poço
pCDNA3-GAL4-BRCA1 CT		100ng/poço
pVP16-PALB2 NT WT		
pVP16-PALB2 NT WT	p.L21A	
pVP16-PALB2 NT WT	p.L35P	
pVP16-PALB2 NT WT	p.K18R	
pVP16-PALB2 NT WT	p.T317P	
pVP16-PALB2 NT WT	p.T317S	60ng/poço
pVP16-PALB2 NT WT	p.T317A	
pVP16-PALB2 NT WT	p.T317K	
pVP16-PALB2 NT WT	p.T317I	
pVP16-PALB2 NT WT	p.T317R	

pG5luc = plasmídeo contendo a sequência da proteína luciferase repórter de *Photinus pyralis* com um promotor sob o controle de operadores de GAL4. pGR-TK = plasmídeo de controle interno de transfecção contendo a região condificante da Luciferase de *Renilla reniformis* controlado pelo promotor HSV TK, constitutivamente ativo em células humanas. As variantes p.L21P e p.L35P foram utilizadas como controle negativo da interação BRCA1-PALB2.

4.11.4 Superexpressão de proteínas mutadas em células humanas silenciadas para os genes endógenos de interesse

Para ensaios de superexpressão das proteínas BRIP1 e PMS2, foram utilizadas células HEK293T mantidas em meio DMEM completo. Para os ensaios com a proteína BRIP1, $3,5 \times 10^5$ células HEK293T foram plaqueadas por poço numa placa de 6 poços em meio DMEM completo. Após 24h horas, o meio destas células foi substituído por OPTI-MEM (Gibco - Thermo Fisher Scientific®), acrescido de 10% de soro bovino fetal (Gibco - Thermo Fisher Scientific®) e as células foram transfectadas com 800 ng do plasmídeo pIKO (que confere resistência à puromicina) e com os plasmídeos contendo as sequências de interesse para BRIP1 (Apêndice, Tabela A2). Após 6h, o meio foi substituído por DMEM completo. Na manhã do dia seguinte as células foram tratadas com puromicina a $3 \mu\text{g/mL}$ e permaneceram em estufa a 37°C por 7 dias. Após a semana de tratamento $3,5 \times 10^5$ de células HEK293T expressando as proteínas mutadas de interesse foram plaqueadas em uma placa de 6 poços em meio DMEM completo. Após 24h horas, o meio destas células foi substituído por OPTI-MEM (Gibco - Thermo Fisher Scientific®), acrescido de 10% de soro bovino fetal (Gibco - Thermo Fisher Scientific®). As células foram então transfectadas com siRNA (pequeno RNA de interferência) controle (5'-UUCGAACGUGUCACGUCAA-3') ou específicos para a proteína BRIP1 (5'-GUACAGUACCCACCUUAUUU-3') endógena, a 50nM utilizando Lipofectamine RNAiMAX (Invitrogen - Thermo Fisher Scientific®) em meio OPTI-MEM sem soro e sem antibiótico. Na manhã do dia seguinte, as células foram novamente transfectadas com os siRNA e 24h depois as células foram ressuspensas em meio DMEM completo e quantificadas utilizando contador TC20 (Bio-Rad) e as células foram plaqueadas para o teste de sensibilidade à Mitomicina C ou coletadas para verificação da expressão por *western blot*.

Para as células transfectadas com plasmídeos contendo as sequências mutadas de PMS2, para cada condição $3,5 \times 10^5$ células HEK293T foram plaqueadas por poço numa placa de 6 poços em meio DMEM completo. Após 24h horas, o meio destas células foi substituído por OPTI-MEM (Gibco - Thermo Fisher Scientific®), acrescido de 10% de soro bovino fetal (Gibco - Thermo Fisher Scientific®). As células foram então transfectadas com siRNA (pequeno RNA de interferência) controle (5'-UUCGAACGUGUCACGUCAA-3') ou específico para a proteína PMS2 (GAACAGAAUUACAGGAAGUUU) endógena, a 50nM utilizando Lipofectamine RNAiMAX (Invitrogen - Thermo Fisher Scientific®) em meio OPTI-MEM sem soro e sem antibiótico. Na manhã do dia seguinte, as células foram novamente transfectadas com os siRNAs e ao final da tarde as foram então transfectadas com com $1,5 \mu\text{g}$ de plasmídeos de PMS2, usando Lipofectamine 2000 (Invitrogen - Thermo Fisher Scientific®) e OPTI-MEM sem soro e sem antibiótico. Na manhã do dia seguinte, o meio dos poços foi aspirado e as células foram ressuspensas em meio DMEM completo e quantificadas utilizando contador TC20 (Bio-Rad) e as células foram plaqueadas para verificação da

expressão por *western blot*.

Para os ensaios com as proteínas CHK2, MLH1 e PALB2, foram utilizadas células HeLa mantidas em meio DMEM completo. Para cada condição dos ensaios com estas proteínas, $2,4 \times 10^5$ células HeLa foram plaqueadas por poço numa placa de 6 poços em meio DMEM completo. Após 6h horas, o meio destas células foi substituído por OPTI-MEM (Gibco - Thermo Fisher Scientific®), acrescido de 10% de soro bovino fetal (Gibco - Thermo Fisher Scientific®). As células foram então transfectadas com siRNA (pequeno RNA de interferência) controle ou específicos para as proteínas CHK2 (5'-GAACCUGAGGACCAAGAAC-3'), MLH1 (5'-CAGUGUACGCAGCCUAUUU-3') e PALB2 (5'-CUUAGAAGAGGACCUUAUU-3') endógenas, a 50nM utilizando Lipofectamine RNAiMAX (Invitrogen - Thermo Fisher Scientific®) em meio OPTI-MEM sem soro e sem antibiótico. Na manhã do dia seguinte, as células transfectadas para CHK2 e MLH1 foram novamente transfectadas com os siRNAs. Após 48h (para CHK2 e MLH1) ou após 24h (para PALB2) da primeira transfecção, as células foram então transfectadas com 800 ng de plasmídeos contendo as sequências das proteínas testadas, usando Lipofectamine 2000 (Invitrogen – Thermo Fisher Scientific®) e OPTI-MEM sem soro e sem antibiótico. Aproximadamente 5h após a transfecção dos plasmídeos, o meio foi substituído por DMEM completo. Após 24h da transfecção dos plasmídeos, o meio dos poços foi aspirado e as células foram ressuspensas e quantificadas utilizando contador TC20 (Bio-Rad) e as células eram plaqueadas para testes específicos ou coletadas para verificação da expressão por *western blot*.

4.11.5 Verificação dos níveis de expressão proteica por *western blot*

As células transfectadas com os plasmídeos de interesse eram peletadas e estocadas em freezer -80°C até o momento da extração das proteínas. No dia da extração, o *pellet* era ressuspensado em 80-100 µL de buffer IP500 (50 mM de Tris-HCl pH 7,5 + 500 mM de NaCl + 0,5% de NP40) gelado, contendo inibidores de protease (1 mM PMSF + Aprotine 1:500 e Leupeptine 1:1000) e fosfatase (5 mM NaF + 1mM Na₃VO₄). Após isso os tubos eram sonificados no aparelho Bioruptor (Diagenode), a uma alta intensidade por 10 minutos 30s *on* e 30s *off*. Após isso as amostras eram centrifugadas por 10 minutos a 13000 rpm ou à velocidade máxima, a 4°C. Após isso, o sobrenadante era utilizado para a quantificação do extrato total de proteínas utilizando o método colorimétrico descrito por Bradford (1976). As proteínas eram então diluídas a uma concentração de 2 µg/µL em solução de Tris-HCl 0,5M (pH 6.8), SDS 4.4% (p/v), glicerol 20% (v/v), azul de bromofenol diluídos em água e acrescidos de

DDT 1M e 2-Mercaptoetanol 2M. Após esta etapa as proteínas poderiam ser estocadas em -20°C.

No dia da corrida do *western blot* (WB), as proteínas eram desnaturadas a 95°C por 5 minutos e após isso eram inoculadas em pocinhos de um gel de poliacrilamida 7-10% e corriam a 150V por 1h e 10 minutos, seguindo protocolo de SDS-PAGE (*SDS-polyacrylamide gel electrophoresis*) (LAEMMLI, 1970). Após isso o gel foi transferido para uma membrana de nitrocelulose Hybond PVDF (*Polyvinylidene difluoride*) (GE Healthcare) numa reação de transferência utilizando sistema Bio-Rad, em uma corrida de 1h a 100V. Após isso a membrana foi bloqueada com BSA (Sigma-Aldrich) a 5% em TBST (*Tris-Buffered Saline* + Tween 20; NaCl 136 mM, KCl 3 mM, Tris-base 25 mM; pH 7,4 + Tween 20 0,1% v/v) para proteínas fosforiladas; ou em leite em pó a 5% em PBST (*Phosphate-buffered saline* + Tween 20; NaCl 137 mM, KCl 2,7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1,8 mM + Tween 20 0,05%) para as demais proteínas. Em seguida, a membrana foi incubada *overnight* à 4°C com os anticorpos primários (Apêndice, Tabela A3) diluídos em TBST ou PBST suplementado com BSA ou leite desnatado a 5%. No dia seguinte, as membranas eram lavadas por 3 vezes com TBST ou PBST e eram então incubada por 1h com os anticorpos secundários *anti-mouse* ou *anti-rabbit* (Apêndice, Tabela A3), também diluídos em TBST ou PBST. Ao término da incubação, a membrana era lavada 3 vezes com o tampão adequado e então era realizada a revelação utilizando o Kit Amersham ECL Western Blotting Analysis System (GE Healthcare), seguida pela exposição da membrana ao filme radiográfico Amersham Hyperfilm ECL (GE Healthcare), segundo as recomendações do fabricante.

Quando necessário, as proteínas eram quantificadas utilizando o programa Image J (DAVARINEJAD, 2015). A concentração das proteínas era normalizada pela concentração da proteína endógena (tubulina ou vinculina, dependendo do tamanho da proteína alvo).

4.11.6 Efeito das mutações em PALB2 na sensibilidade à inibidores de poli(ADP-ribose) polimerase (PARPi)

A inibição de poli(ADP-ribose) polimerase (PARP) leva a uma persistência de lesões ao DNA de fita simples, o que leva ao colapso da forquilha de replicação do DNA e permite a permanência de DSBs que seriam normalmente reparadas por recombinação homóloga (GUNDERSON; MOORE, 2015). Estudos têm evidenciado que células deficientes em BRCA1/BRCA2 ou PALB2 estão associadas a um déficit severo no processo de HR o que leva à sensibilização das células mutadas ao dano no DNA após tratamento com inibidores de PARP (BUISSON et al., 2017; TKÁČ et al., 2016). Portanto, para caracterizar se variantes no gene *PALB2* estavam relacionadas ao comprometimento do

processo de reparo, foi analisada a sobrevivência celular após tratamento com Olaparib® (Selleckchem) em células transfectadas com os plasmídeos selvagens, células transfectadas com o vetor vazio e em células contendo as proteínas mutadas.

Para isso, células HeLa silenciadas para o gene *PALB2* endógeno foram transfectadas com plasmídeos contendo as sequências selvagem e mutadas de *PALB2* resistentes aos siRNAs, conforme descrito na seção 4.11.4. Após 24h da transfecção dos plasmídeos, o meio dos poços foi aspirado e as células foram tripsinizadas com Trypsin-EDTA 1X (Gibco – Thermo Fisher Scientific®) e quantificadas utilizando contador TC20 (Bio-Rad). Após isso, 3000 células foram plaqueadas por poço numa placa escura de 96 poços (Corning 3603) em DMEM completo, totalizando três poços para cada condição. As células que sobraram foram congeladas para extração de proteínas para análise de expressão proteica por *western blot*. Passadas mais ou menos 6 horas do plaqueamento das células, cinco diferentes concentrações de Olaparib (diluídas em DMSO) foram adicionadas às triplicatas de cada condição: apenas DMSO, Olaparib a 0,3125 μ M, 0,625 μ M, 1,25 μ M e 2,5 μ M. As células permaneceram na incubadora a 37°C, em atmosfera de 5% de CO₂. Após 3 dias de tratamento, 10 μ g/mL de Hoechst 33342 (Invitrogen – Thermo Fisher Scientific®) em DMEM foram adicionados a cada um dos poços e as placas permaneceram na estufa a 37°C por 1h. Após isso as placas foram levadas para aquisição de imagens, utilizando aumento de 4x no equipamento Cytation 5 Cell Imaging Multi-Mode Reader (BioTek Instruments), por meio da quantificação de núcleos corados com Hoechst por meio do programa Gen5 Data Analysis Software v3.03 (BioTek Instruments). A viabilidade celular foi expressa em porcentagem de sobrevivência ao tratamento pelo Olaparib e DMSO. Os resultados representam a média $\pm$ SD de três experimentos independentes, realizados em triplicata.

4.11.7 Níveis de fosforilação das proteínas mutadas após dano ao DNA

O gene *CHEK2* codifica uma serina/treonina quinase que é ativada por ATM em resposta às DSBs, sendo fosforilada no resíduo Thr68 o que resulta na sua dimerização, auto-fosforilação e ativação, e na transdução do sinal às proteínas mediadoras e efetoras do reparo por recombinação homóloga por meio da fosforilação de outras proteínas como p53 e BRCA1 (ECONOMOPOULOU; DIMITRIADIS; PSYRRI, 2015; MAGNI et al., 2015).

Baseado nisso, foi verificado o nível de expressão da proteína CHK2 fosforilada após indução de danos ao DNA por irradiação. Para isso, células HeLa silenciadas para o gene *CHEK2* endógeno foram transfectadas com plasmídeos contendo as sequências selvagem e mutadas de *PALB2*

resistentes aos siRNAs, conforme descrito na seção 4.11.4. Após 24h da transfecção dos plasmídeos, as placas foram irradiadas a 5 Gy, utilizando o aparelho CellRad Compact Cabinet X-ray Cell Irradiator (Faxitron). Após a irradiação, o meio dos poços foi aspirado e as células foram tripsinizadas com Trypsin-EDTA 1X (Gibco – Thermo Fisher Scientific®), lavas com PBS 1x e então peletadas para análise da expressão por *western blot*. O nível de expressão das proteínas foram quantificados utilizando o programa Image J (DAVARINEJAD, 2015) e os resultados representam a média $\pm$ SD de três experimentos independentes.

4.11.8 Efeito das mutações em BRIP1 na sensibilidade à droga Mitomicina C

A mitomicina C (MMC) é um agente quimioterapêutico que atua principalmente na alquilação do DNA após a sua redução enzimática, promovendo a formação de *cross-links* entre as fitas de DNA, provocando instabilidade genômica (PAZ et al., 2012). E estudos já têm demonstrado que proteínas FANC associadas à Anemia de Fanconi (FA) são responsáveis pelo reparo de *cross-links* entre as fitas de DNA e que a presença de variantes em *FANCD1* ou *BRIP1* estão associadas à uma maior sensibilidade a MMC e apresentam instabilidade genômica (LITMAN et al., 2005; ZOU et al., 2013).

Para verificar se variantes identificadas em *BRIP1* estavam associadas ao comprometimento da função da proteína, foi realizado teste de sensibilidade das proteínas mutadas ao tratamento com MMC. Para isso, foram utilizadas células HEK293T expressando as proteínas mutadas e silenciadas para o gene *BRIP1* endógeno conforme descrito na seção 4.11.4. Após 24h da segunda transfecção dos siRNAs, 1500 células foram plaqueadas por poço numa placa escura de 96 poços (Corning 3603) em DMEM completo, totalizando três poços para cada condição. As células que sobraram foram congeladas para extração de proteínas para análise de expressão proteica por *western blot*. Passadas mais ou menos 6 horas do plaqueamento das células, cinco diferentes concentrações de MMC (diluídas em DMEM completo) foram adicionadas às triplicatas de cada condição: apenas DMEM, MMC a 4 ng/mL, 8 ng/mL, 12 ng/mL e 16 ng/mL. As células foram mantidas na incubadora a 37°C, em atmosfera de 5% de CO₂. Após 4 dias de tratamento, as células foram tratadas com Cell Counting Kit 8 (WST-8 / CCK8) (Abcam), de acordo com recomendações do fabricante e as placas permaneceram na estufa a 37°C por 3h. Após isso as placas foram levadas para leitura utilizando o equipamento Cytation 5 Cell Imaging Multi-Mode Reader (BioTek Instruments), por meio da leitura de absorbância pelo programa Gen5 Data Analysis Software v3.03 (BioTek Instruments). A viabilidade celular foi

expressa em porcentagem de sobrevivência ao tratamento com MMC. Os resultados representam a média $\pm$ SD de três experimentos independentes, realizados em triplicata.

4.12 ANÁLISE DE INSTABILIDADE DE MICROSSATÉLITES (MSI) PARA OS GENES DA VIA MMR

Avaliar a estabilidade de marcadores de microsatélite se tornou uma importante ferramenta de avaliação do sistema de reparo MMR, uma vez que esta constitui a principal via de reparo de danos ao DNA em regiões de microsatélites (ETO et al., 2016). Embora a deficiência em genes da via MMR de reparo ao DNA estejam mais associadas à Síndrome de Lynch, cerca de 1,7% e 3,7% dos tumores de mama e ovário, respectivamente, também apresentam alta instabilidade de microsatélites (GERMANO et al., 2018). Para análise de MSI nas amostras estudadas, blocos de parafina contendo amostras de tumor das pacientes que apresentaram variantes VUS nos genes MLH1 e PMS2 tiveram seu DNA extraído utilizando o Kit Maxwell® 16 FFPE Tissue LEV (Promega®) e o equipamento Maxwell Rapid Sample Concentrator (Promega®), seguindo o protocolo do fornecedor. As amostras de DNA obtidas foram quantificadas e avaliadas quanto à pureza por meio do espectrofotômetro NanoDrop (Thermo Fisher Scientific®).

O status de MSI dos tumores foi obtido por análise do padrão de amplificação dos marcadores moleculares mononucleotídicos NR-27, NS-21, NR-24, BAT-25 e BAT-26 por PCR multiplex utilizando primers marcados com fluorescência, de acordo com protocolo descrito por Buhard e colaboradores (2004), comparando aos perfis de amplificação no DNA de tecido normal (sangue) já coletado e utilizado para as demais análises. Os *primers* utilizados para a amplificação dos marcadores está descrito na Tabela A1, no Apêndice deste documento. A análise dos *amplicons* foi realizada em eletroforese capilar utilizando o equipamento ABI Prism 3130 Genetic Analyzer (Applied Biosystems - Thermo Fisher Scientific®) e os resultados foram interpretados utilizando-se o software GeneMarker versão 1.85 (SoftGenetics®). Marcadores microsatélites são considerados instáveis quando é observada, em amplificação a partir de DNA tumoral, a presença de alelos não visualizados para o respectivo marcador amplificado a partir de DNA extraído de sangue periférico (DNA não tumoral).

5 RESULTADOS E DISCUSSÃO

5.1 CARACTERIZAÇÃO CLÍNICA DOS PACIENTES

Foram coletadas informações clínicas dos pacientes com suspeita clínica de HBOC a partir da revisão de prontuários médicos em papel, eletrônicos e em registros familiares coletados pelo serviço de Genética Médica do Hospital das Clínicas de Ribeirão Preto. Alguns pacientes não puderam ser melhor caracterizados pela falta de informações nos arquivos revisados. Para dois pacientes não foi possível identificar a idade ao diagnóstico da doença até o presente momento, portanto na **Tabela 2**, somente 92 dos 95 pacientes estudados puderam ser caracterizados. Como mostrado na **Tabela 2**, houve uma prevalência de indivíduos do sexo feminino havendo apenas um indivíduo do sexo masculino; este mesmo indivíduo era o mesmo portador do câncer tipo gástrico e foi incluído no estudo pelo forte histórico familiar de câncer de mama e ovário em idade jovem. A idade de diagnóstico da doença em todos os indivíduos variou de 22 a 72 anos e a idade atual de 31 a 76 anos. Dos 93 pacientes caracterizados 14 vieram a óbito, sendo que destes, 13 indivíduos que apresentaram metástase à distância. A mediana de tempo de vida desde o diagnóstico da doença para estes indivíduos foi de 4 anos (1 a 12 anos) (**Figura 9**).

Figura 9. Curva de sobrevivência dos pacientes estudados pelo tempo de vida desde o diagnóstico em anos. Na linha azul, cada traço preto na vertical indica os pacientes censurados, ou seja, aqueles que ainda não vieram a óbito e o tempo marcado por cada traço representa a sobrevivência destes pacientes em anos. Há pacientes acompanhados a mais tempo e que o tempo de diagnóstico da doença desde o diagnóstico até 2019 é de 28 anos, assim como há pacientes que iniciaram o acompanhamento clínico em 2016 e o tempo de doença é de apenas 3 anos.

A maioria dos pacientes possuem histórico familiar de câncer, que é uma das variáveis utilizadas como critério de inclusão dos indivíduos para investigação clínica/molecular de HBOC pelo *National Comprehensive Cancer Network (NCCN) Guidelines*. A quantidade de casos positivos para história familiar entre os indivíduos mais jovens foi menor (58,8%) do que a observada para os pacientes diagnosticados nas demais faixas etárias (85,6%-100%) (**Tabela 2**). A presença de história familiar é um fator de risco importante para pacientes com suspeita clínica de HBOC, no entanto o acometimento em idade jovem por si só já é um fator de risco significativo e suficiente para a inclusão destes pacientes na investigação clínica/molecular de HBOC, mesmo na ausência de história familiar. Ainda assim, para todos os grupos de idade ao diagnóstico o valor de probabilidade de associação à presença de histórico familiar foi significativo ($p=0.009089$), que permaneceu significativo ($p=0.003216$) mesmo quando excluído o grupo com tamanho amostral muito pequeno (61+) que poderia desviar a análise. A Síndrome do Câncer Hereditário de Mama e Ovário compreende principalmente indivíduos que apresentam cânceres de mama e/ou ovário, no entanto outros tipos podem ser encontrados e ser incluídos como critérios clínicos para HBOC, como cânceres de próstata e pâncreas, por exemplo, desde que acompanhados de uma história familiar que preencha alguns dos critérios do NCCN. Nos pacientes avaliados, 84 casos eram de cânceres de mama, sendo o tipo Carcinoma Ductal Invasivo (CDI) aquele mais frequente (83,2% dos casos de mama), corroborando os dados de (TIMOTEO et al., 2018) na população brasileira. De fato, o CDI constitui o tipo de câncer mais comum dentre os casos familiares/hereditários (KEENEY et al., 2017; LINDOR et al., 2008). Apenas 6 pacientes apresentaram câncer de ovário, sendo que um deles apresentou história pessoal de câncer de mama e de ovário (**Tabela 2**). Foi observado ainda um caso de câncer de endométrio e outro de câncer gástrico, mas ambos apresentam forte histórico familiar de cânceres de mama e ovário em idade jovem em parentes de primeiro e segundo grau, respectivamente.

Tabela 2. Principais informações clínicas coletadas a partir da revisão de prontuários médicos de 92 pacientes com suspeita clínica de HBOC.

Variável	Idade ao Diagnóstico*								p-valor§	
	20-30		31-50		51-60		61+			
	n=17	%	n=64	%	n=8	%	n=3	%		
Gênero										
Homem	-		1	1,60%	-		-		-	
Mulher	17	100%	63	98,40%	8	100%	3	100%		
Idade (mediana)										-
Nº de óbitos										0.0927
Sobrevida em anos	2	11,8%	9	14,1%	1	12,5%	2	66,7%		
(mediana)	6,5		4		6		2,5			
Histórico familiar										0.009089 0.003216
Sim	10	58,8%	55	85,6%	8	100%	3	100%		
Não	6	35,3%	5	8,0%	-		-			
NI	1	5,9%	4	6,4%	-		-			
Localização de tumor										-
Mama	15	88,7%	61	95,2%	6	75,0%	2	66,7%		
Ovário	2	11,8%	2	3,2%	1	12,5%	1	33,3%		
Endométrio	-		-		1	12,5%	-			
Estômago	-		1	1,6%	-		-			
Distribuição do tumor										0.03308 0.01502
Localizado	11	64,6%	53	82,8%	5	75,0%	2	66,7%		
Mama bilateral	2	11,8%	8	12,5%	2	25,0%	-			
Multifocal	2	11,8%	-		1		1	33,3%		
NI	2	11,8%	3	4,7%	-		-			
Subtipo CA de mama										0.5534
Luminal	6	40,0%	21	34,4%	-		1	50%		
Luminal B-like	3	20,0%	11	18,0%	2	33,3%	-			
HER2	3	20,0%	5	8,2%	2	33,3%	-			
TN	3	20,0%	19	31,1%	1	16,7%	1	50%		
NI	-		5	8,3%	1	16,7%	-			
Grau do tumor										0.5794
1	2	11,8%	6	9,4%	2	25,0%	-			
2	5	29,4%	30	46,9%	1	12,5%	1	33,3%		

Continua.

Tabela 2. Principais informações clínicas coletadas a partir da revisão de prontuários médicos de 92 pacientes com suspeita clínica de HBOC.

Continuação.

Variável	Idade ao Diagnóstico*								p-valor§
	20-30		31-50		51-60		61+		
	n=17	%	n=64	%	n=8	%	n=3	%	
Grau do tumor									
3	4	23,5%	17	26,6%	3	37,5%	1	33,3%	
NI	6	35,3%	11	17,2%	2	25,0%	1	33,3%	
Linfonodos									
Ausência	7	41,2%	21	32,8%	2	25,0%	-		
Presença	6	35,3%	32	50,0%	4	50,0%	2	66,7%	0.9411
NI	4	23,5%	11	17,2%	2	25,0%	1	33,3%	
Metástase à distância									
M0	11	64,7%	42	65,6%	7	87,5%	-		
M1	2	11,8%	12	18,7%	1	12,5%	3	66,7%	0.01195
NI	4	23,5%	10	15,6%	-		-		

NI corresponde a uma variável não-informada. *As informações estão ranqueadas de acordo com a idade ao diagnóstico do tumor. [§]Valor de probabilidade em relação às faixas de idade ao diagnóstico calculado utilizando-se o Teste do X² de Pearson.

Nos casos de câncer de mama (n=84), não foram identificadas associações significativas entre os diferentes subtipos tumorais e a idade dos pacientes ao diagnóstico. O subtipos mais frequentes foram o Luminal, com 28 pacientes apresentando positividade para os receptores de estrogênio (RE) e/ou progesterona (RP), e negativos para a amplificação da proteína *Human Epidermal growth factor Receptor-type 2* (HER-2); e o subtipo triplo-negativo (TN), com 24 pacientes apresentando negatividade para qualquer um destes três marcadores moleculares. O câncer de mama do tipo triplo-negativo é o subtipo mais frequente em indivíduos com câncer de mama em idade jovem, bem como em indivíduos com ancestralidade Africana e os judeus Ashkenazi (mutações fundadoras em *BRCA1/BRCA2*). Este subtipo tem em geral, um curso mais agressivo que aqueles com positividade para os receptores hormonais e estão associados a altas taxas de metástases viscerais e do sistema nervoso central, bem como a casos de recorrência e morte precoces (AFGHAHI; TELLI; KURIAN, 2016). Dos 14 pacientes que vieram a óbito, 13 apresentaram metástase à distância ($p=7.85e^{-12}$). No entanto, não foi identificada qualquer associação entre os subtipos de cânceres de mama e a ocorrência de óbitos ou metástases à distância ou linfonodal, entretanto, quando observado o grau do tumor, observou-se que cerca de 48% dos casos de tumores triplo negativos estavam associados a tumores de grau III ($p=0.02191$), aquele de pior prognóstico, de células pouco diferenciadas e rápida progressão (ARPINO et al., 2016).

O grau do tumor, no entanto, não teve nenhuma correlação positiva com a idade dos pacientes ao diagnóstico. Entretanto, quando consideramos a coorte total, observa-se que houve uma prevalência de cânceres mais agressivos, aqueles de graus II e III. Houve associação significativa ($p=0.02297$) da presença de metástases à distância e o acometimento de nódulos linfáticos axilares ou torácicos. Quanto à distribuição do tumor, embora a literatura relate uma alta frequência de casos de acometimento de mama bilateral nos casos de HBOC (POUMPOURIDOU; KROUPIS, 2012), os tumores de instalação unilateral foram aqueles mais frequentes na série de casos estudada ($p=0.01502$), mesmo após excluir aqueles grupos com tamanho amostral muito pequeno. Houveram 4 casos de instalação multifocal, em que um deles era o paciente que apresentou câncer de mama e ovário e os demais eram pacientes que apresentaram, além do câncer de mama, outros tipos de câncer como de tireoide e renal.

5.2 SEQUENCIAMENTO DO PAINEL GÊNICO

5.2.1 Análise da qualidade do sequenciamento

Foram sequenciadas um total de 96 amostras que era o total suportado pelo kit de sequenciamento para uma cobertura desejada de 500x. Destas, 95 amostras foram de pacientes com suspeita clínica de HBOC previamente triados e 1 amostra foi fornecida pelo kit de sequenciamento utilizado, para controle da reação. Os primeiros dados não binários gerados pelo sequenciamento foram os arquivos *fastq* que continham informações sobre a qualidade das sequências (*reads*) geradas. Por meio da ferramenta fastQC, foram gerados arquivos de extensão .html que possuem vários gráficos que demonstram a qualidade das sequências geradas para as duas *reads* (*read 1* e *read 2* do sequenciamento *paired-end*) de cada uma das amostras sequenciadas, gerando assim, um total de 192 arquivos (96x2). Um dos gráficos gerados é o da qualidade de cada uma das bases nitrogenadas lidas nas *reads*. A escala da qualidade utilizada é o PHRED *score* o qual para os sequenciamentos da plataforma Illumina, os valores desta escala variam de 0 a 40, sendo aqueles valores de 28 a 40 correspondentes às bases chamadas com uma alta qualidade, de 20 a 27 aquelas bases chamadas com uma qualidade razoável, e aqueles bases com valores abaixo de 19 correspondem às bases classificadas como de baixa qualidade. As **Figuras 10A e 10B** correspondem aos gráficos de qualidade de todas as bases chamadas para a *read 1* e para a *read 2*, respectivamente, de uma única amostra dentre as 96 que foram sequenciadas. De forma geral, todas as amostras apresentaram um padrão de qualidade semelhante ao da amostra que está representada na **Figura 10**, com todas as bases chamadas nas *reads 1* possuindo PHRED score acima de 28, enquanto que nas *reads 2* sempre apresentaram uma queda na qualidade ao final da sequência. No entanto, esse padrão observado nas *reads 2* não comprometeram a qualidade geral das sequências geradas, uma vez que a maioria das sequências apresentaram um PHRED score médio acima de 30, como pode ser observado nas **Figuras 10C e 10D**.

Apenas uma das amostras de HBOC não apresentou uma boa qualidade durante a chamada de bases, comprometendo a qualidade das todas as sequências geradas para esta amostra, como pode ser observado na **Figura 11** a presença de um padrão de qualidade completamente distinto daquele observado para as demais amostras sequenciadas. Foi checada a qualidade do DNA desta amostra e estava semelhante às demais amostras sequenciadas. Olhando os dados brutos do sequenciamento foi observada uma quantidade bastante reduzida de *reads* e *clusters* gerados para esta amostra, cerca de

apenas 2% em relação ao que foi gerado para cada uma das demais amostras. Por conta disso, essa amostra foi desconsiderada para as posteriores análises.

Figura 10. Demonstrativo da qualidade geral das reads geradas no sequenciamento do painel gênico das 95 amostras de HBOC. Imagens geradas a partir do arquivo fastq de uma única amostra que evidencia valores de qualidade de cada base lida para cada posição (A e B) e a qualidade média das sequências geradas para esta amostra (C e D).

Figura 11. Gráficos de qualidade gerados a partir do arquivo *fastq* de uma única amostra que evidencia uma qualidade comprometida para cada base lida para cada posição (A e B) e a qualidade média em torno de 22 das sequências geradas para esta amostra (C e D).

Após isso foram gerados os arquivos SAM/BAM com o alinhamento e mapeamento das *reads* à sequência genoma referência GRCh37.75. As *reads* que foram mapeadas e checadas pela ferramenta Qualimap (<http://qualimap.bioinfo.cipf.es/>), que também gera gráficos e tabelas que possibilitam uma visão geral dos dados gerados. O sequenciamento *paired-end* das 96 amostras utilizadas gerou um total de 26.299.408 *reads* e destas, cerca de 97,03% foram mapeadas à sequência referência, com uma cobertura vertical média de 336,65 *reads*, o que evidencia uma qualidade geral bastante satisfatória (Tabela 3).

Tabela 3. Dados quantitativos do sequenciamento, após o alinhamento e mapeamento das *reads* geradas à sequência referência.

	Nº de <i>reads</i>	%
Total de <i>reads</i> geradas	26.299.408	100%
Total de <i>reads</i> mapeadas	25.519.888	97,03%
Média de cobertura das amostras	336,65	

Outra informação importante é a quantidade de transições (Ti) e transversões (Tv) observadas para as variantes identificadas. Se as substituições ocorrem de forma randômica no genoma, a proporção Ti/Tv é em torno de 0,5 uma vez que é sabido que há a possibilidade de acontecerem duas transições para quatro transversões. No entanto, para dados de sequenciamento, que envolvem principalmente regiões codificantes, essa taxa está em torno de 2,0 a 3,0 (WANG et al., 2014). A **Figura 12** mostra a porcentagem de transições (Ti) e transversões (Tv) para cada uma das 96 amostras sequenciadas e observamos que de maneira geral a taxa de transição é aproximadamente duas vezes maior que a taxa de transversão, exceto para a amostra 4229, aquela mesma amostra que evidenciou baixos *scores* de qualidade (**Figura 11**). A proporção Ti/Tv para esta amostra foi cerca de 0,88 enquanto a mesma taxa varia de 2,0 a 3,0 para as demais amostras, o que confirma que a taxa de Ti/Tv pode ser utilizada como parâmetro de qualidade nos sequenciamentos de nova geração (GUO et al., 2013).

Figura 12. Proporção de Transições (Ti) e Transversões (Tv) observadas nos 21 genes sequenciados para as 96 amostras sequenciadas.

Utilizando os histogramas do número de regiões genômicas encontradas para cada valor de cobertura, gerados pelo SnpEff, foi obtido o gráfico de cobertura cumulativa das regiões para cada uma das 95 amostras de HBOC, por meio pacote GenVisR (SKIDMORE et al., 2016) em ambiente R. Com estes dados, pode-se observar que pelo menos 75% de todas as regiões sequenciadas tiveram uma cobertura entre 500x e 1000x (cores variando de roxo a amarelo) (**Figura 13**) para a maioria das amostras analisadas. Apenas a amostra 4229 apresentou baixos valores de cobertura para todas as regiões sequenciadas. Esta amostra foi desconsiderada para as etapas de análise seguintes.

Figura 13. Cobertura cumulativa das regiões sequenciadas para os 21 genes, para cada uma das 95 amostras de HBOC sequenciadas.

5.2.2 Anotação das variantes encontradas

Após o alinhamento e mapeamento das variantes encontradas, foi realizada a anotação com o SnpEff. Esta ferramenta cria um arquivo adicional que mostra as estatísticas gerais do sequenciamento. Foram identificadas 667 variantes de ponto e pequenas inserções/deleções. Mesmo tratando-se de um sequenciamento de regiões codificantes, cerca de 63,4% das variantes foram identificadas em regiões intrônicas tais como as regiões intergênicas e extremidades 5' e 3' não traduzidas dos genes (UTR). Já em regiões codificantes foram identificadas cerca de 32,28% das variantes, dentre elas, variantes sinônimas, *missense*, *frameshift*, variantes de ganho de códon de parada (stop_gain) inserções e deleções. Uma pequena parcela de variantes foi identificada como afetando regiões de sítios de *splicing* (4,19%) (**Figura 14**).

Figura 14. Regiões e efeitos das variantes identificadas nos 21 genes estudados.

Quando observamos o tipo de variante identificada por gene (**Figura 15**), os genes que apresentaram a maior quantidade de variantes foram *ATM*, *ATR*, *BRCA1*, *BRCA2* e *NBN*, que juntos apresentaram o correspondente a cerca de 44,6% do total de variantes identificadas no estudo. No entanto, como houve uma alta prevalência de variantes intrônicas, não necessariamente estes foram os mesmos genes mais afetados por variantes consideradas de impacto para o fenótipo. Os genes *BRCA1* e *BRCA2* foram os únicos a apresentar variantes do tipo *frameshift*, aquelas de maior impacto ao fenótipo.

Figura 15. Tipo e quantidade de variantes identificadas por gene analisado.

5.3 CARACTERIZAÇÃO DAS VARIANTES ENCONTRADAS

Durante a análise dos dados foi observado que dentre as variantes não havíamos identificado algumas variantes importantes e já caracterizadas como associadas ao risco para HBOC como: c.156_157insAlu no gene *BRCA2* (FELICIO et al., 2018b; PEIXOTO et al., 2009) e p.Arg337His (rs121912664) no gene *TP53* (CURY; FERRAZ; SILVA, 2014), ambas bastante caracterizadas na população brasileira. Os elementos Alu são retroelementos de aproximadamente 300 nucleotídeos que formam rearranjos genômicos já descritos em genes associados ao risco para câncer como *APC*, *ATM*, *CHEK2*, *BRCA1* e *BRCA2* (QIAN et al., 2017). A inserção de fragmentos Alu no éxon 3 do gene *BRCA2* foi primeiramente descrita em 2005 por Teugels e colaboradores como uma variante fundadora da população Portuguesa em pacientes com HBOC, associada à perda de parte importante do éxon 3, região importante para a ligação com *PALB2* e *RAD51* (DUCY et al., 2019; FELICIO et al., 2018b). Já a variante p.Arg337His no gene *TP53* está associada ao aumento de risco para a Síndrome de Li Fraumeni. É uma mutação fundadora da região Sul do Brasil, que já é descrita como segregada em famílias com histórico de câncer adrenocortical, carcinoma do plexo coróide, sarcomas e câncer de mama em idade jovem (HAHN et al., 2018). Esta variante está presente no éxon 10 do gene *TP53*, onde está localizado o domínio de oligomerização de p53. Este domínio é importante para a formação de tetrâmeros que promovem a transativação de p53 e facilitam a ligação com o DNA. Além dos

estudos de segregação, estudos funcionais têm demonstrado uma deficiência na capacidade de oligomerização e na atividade transcricional de p53 (PARK et al., 2018; PARK; WANG; HWANG, 2019). Além destas duas variantes, mais recentemente a variante c.-107A>T (rs799905) no gene *BRCA1* foi relacionada ao aumento de risco para HBOC por estar associada à metilação da região promotora de *BRCA1* (EVANS et al., 2018). As regiões destas três variantes foram cobertas pelo sequenciamento do painel gênico, no entanto, as variantes não foram chamadas após a análise dos dados. Foi então realizado o *screening* destas três variantes de ponto por sequenciamento Sanger utilizando *primers* específicos para as regiões das variantes, para as 95 amostras estudadas.

Na **Figura 16** temos a representação esquemática dos genes *BRCA1* e *BRCA2* em que podemos observar os éxons (caixas pretas) que foram analisados pelo sequenciamento do painel gênico e as variantes adicionais que foram analisadas nestes dois genes. Para o *screening* da variante c.-107A>T, foi realizado sequenciamento Sanger a partir da amplificação do Éxon 1 não codificante de *BRCA1* (**Figura 16A**). Nenhum paciente apresentou a variante c.-107A>T, sendo descartado o silenciamento epigenético da região promotora de *BRCA1* associado a esta variante. Já para investigação da inserção Alu no gene *BRCA2*, foi realizado um primeiro *screening* com *primers* específicos para o éxon 3, seguida de PCR específico para amplificação da sequência Alu. Para uma das 95 amostras estudadas, foi observada uma banda adicional de aproximadamente 579 pb após amplificação do éxon 3, que seria equivalente à amplificação do fragmento Alu inserido em heterozigose nesta região. Foi então realizada uma nova reação de PCR para esta amostra, utilizando *primers* específicos para Alu e foi observada a amplificação da banda correspondente à inserção (**Figura 16B1**). Este fragmento foi sequenciado utilizando *primers* específicos para a sequência Alu e foi confirmada a presença da inserção c.156_157insAlu nesta paciente (**Figura 16B2**). Esta paciente não apresentou nenhuma outra variante de significado incerto, sendo possível fechar o diagnóstico molecular de HBOC para este caso.

Figura 16. Representação esquemática dos genes *BRCA1* e *BRCA2* e diagnóstico molecular das variantes adicionais investigadas. (A) Representação esquemática do gene *BRCA1* evidenciando os seus 23 éxons, dos quais 22 são codificantes (caixas em cor preta) e, éxon 1 e parte dos éxons 2 e 23 que são não codificantes (caixas em cor cinza). As regiões codificantes foram investigadas pelo sequenciamento do painel gênico e a variante c.-107A>T, localizada no éxon 1, foi investigada adicionalmente por sequenciamento Sanger. Todos os pacientes apresentaram a sequência selvagem do éxon 1. (B) Representação esquemática do gene *BRCA2* evidenciando os seus 27 éxons, dos quais 26 são codificantes (caixas em cor preta) e, éxon 1 e parte dos éxons 2 e 27 que são não codificantes (caixas em cor cinza). As regiões codificantes foram investigadas pelo sequenciamento do painel gênico e a inserção c.156_157insAlu no éxon 3 foi investigada adicionalmente por amplificação do éxon 3 e validação por sequenciamento Sanger (B1 e B2). Foram realizadas duas reações de PCR independentes para amplificação do éxon 3 do gene *BRCA2* e do fragmento Alu. (B1) Poço 1 - marcador de peso molecular Generuler 1Kb. Poços 2 a 4 - PCR específica para amplificação do éxon 3 do gene *BRCA2* (201 pb). O poço “NTC” corresponde à reação controle livre de DNA e o poço “CTRL Ex 3”, a amplificação de uma amostra negativa para inserção Alu. No poço “4132 Ex 3”, amplificação de uma banda adicional resultante da inserção de um fragmento em heterozigose para a amostra 4132. Foi realizada então uma segunda reação de PCR utilizando primer *forward* específico para a sequência Alu resultando na amplificação de uma banda sugestiva da presença da sequência Alu na amostra positiva 4132 (Poço “4132 Alu”). (B2) Eletroferograma do sequenciamento do amplicon identificado no poço “4132 Alu”, confirmando a inserção da sequência Alu no éxon 3 do gene *BRCA2* para a amostra 4132.

Na **Figura 17** estão representadas a estrutura do gene e da proteína p53, evidenciando o domínio de oligomerização que engloba os éxons 9 e 10, região em que está localizada a variante p.Arg337His. Para o *screening* desta variante nas 95 amostras estudadas, o éxon 10 foi amplificado e sequenciado utilizando *primers* específicos para esta região. Um total de cinco pacientes apresentaram a variante patogênica p.Arg337His (**Figura 17**).

Figura 17. Representação esquemática do gene e proteína p53 e localização da variante investigada adicionalmente p.Arg337His. O esquema do gene *TP53* evidencia os 11 éxons presentes dos quais 10 são codificantes (caixas em cor preta) e, o éxon 1 e parte dos éxons 2 e 11 que são não codificantes (caixas em cor cinza). Já no esquema da proteína, temos representados os domínios TAD = domínio de transativação, que engloba um dos sítios de exportação nuclear; Domínio rico em prolina, ou PR; Domínio de ligação ao DNA; Domínio de oligomerização, que engloba outro sítio de exportação nuclear e; dois domínios NLS = sítios de localização nuclear. Eletroferograma evidenciando a troca em heterozigose de uma Guanina por uma Adenina no início do ÉXON10, no domínio de oligomerização, que promove a troca de uma Arginina por uma Histidina no aminoácido 337. Esta variante foi identificada em 5 das 95 amostras estudadas. Estruturas dos genes e proteína adaptadas de Dufour et al. (2013).

Quanto ao efeito, o caráter patogênico destas variantes pontuais investigadas separadamente já é conhecido. No entanto, para muitas das variantes identificadas pela análise de painel gênico o significado clínico permanece desconhecido. Assim, para melhor caracterizar as demais variantes identificadas neste estudo, foi realizada uma filtragem daquelas variantes de potencial impacto ao fenótipo. Para isso, foram selecionadas apenas variantes presentes em regiões codificantes, ou em regiões de *splicing* caracterizadas como patogênicas. Aproximadamente 11,7% (78) das variantes identificadas no sequenciamento foram então selecionadas e classificadas quanto à patogenicidade baseando-se em dois principais aspectos: 1. Caracterização do efeito das variantes pelo ClinVar, combinado à 2. Predição *in silico* de patogenicidade utilizando seis diferentes ferramentas de classificação. Esta classificação visou evitar ao máximo a exclusão ou perda de informações daquelas variantes de significado clínico desconhecido ou duvidoso, e para isso as variantes foram categorizadas em seis diferentes sub-classes: 1. Patogênicas quando o ClinVar assim já as caracterizava; 2. Possivelmente patogênica quando o ClinVar as classificava como variantes de significado incerto (VUS), ou não as classificava, mas todas as ferramentas de predição apresentavam scores sugestivos de patogenicidade; 3. Benignas quando o ClinVar e as ferramentas de predição corroboravam a classificação de benigna; 4. Possivelmente benigna quando o ClinVar assim as classificava, mas a variante era de uma classe considerada de alto impacto (*Nonsense* ou *frameshift*) e por isso predita como patogênica; 5. Conflito de informação quando o ClinVar as classificava como benignas mas algumas das ferramentas de predição as classificavam como patogênicas e, 6. Variantes de significado incerto (VUS) quando não existiam dados sobre aquela variante no ClinVar e nem as ferramentas podiam prever a sua patogenicidade (uma vez que algumas ferramentas *in silico* só funcionam para predição de variantes em regiões codificantes). Após essa classificação, foi construído um gráfico em ambiente R, utilizando o pacote ComplexHeatmap (**Figura 18**), que classifica as variantes filtradas de acordo com o gene, a frequência e a patogenicidade. No gráfico também estão plotados os resultados do *screening* das duas mutações pontuais descritas anteriormente.

Figura 18. Gráfico *Oncoprint* mostrando a prevalência de variantes dentro as amostras estudadas, de acordo com o tipo e patogenicidade de cada uma delas. O efeito de patogenicidade “Conflito de informações” refere-se àquelas variantes com conflito de resultados pelas ferramentas de predição e o ClinVar e “VUS” refere-se às variantes as quais as ferramentas de predição não conseguiram gerar scores de patogenicidade e/ou aquelas que não se tem informação pelo ClinVar. A nomenclatura está de acordo com a classificação utilizada pelo ClinVar. Os indivíduos foram classificados em grupos de idade ao diagnóstico; subdivididos de acordo com o subtipo molecular do tumor (TN= Triplo-negativo ou também conhecido como basal-like; Lum= Luminal, quando há positividade para os receptores de estrogênio e negatividade para HER2; LumHER= Luminal do positivo para os três marcadores; HER2= quando há positividade para o fator de crescimento HER2 e negatividade para receptores de estrogênio e progesterona; RP= positividade apenas para receptores de progesterona; NI= Não informado) sendo indicados os casos que não constituem casos de mama: Ovário, estômago e Endométrio; e ainda, de acordo com o grau do tumor.

A maior parte das variantes encontradas são do tipo *missense* e de significado clínico duvidoso/desconhecido. Cerca de 11,25% (9) das variantes selecionadas são variantes previamente descritas como patogênicas (*frameshift*, *stop gain*, *insertion* e *missense*) nos genes *BRCA1*, *BRCA2* e *TP53*. A prevalência destas variantes nos pacientes estudados foi de 23,4% (22) sendo a variante *frameshift* p.Gln1756Profs*74, c.5266dupC (rs80357906), no gene *BRCA1* (ENSP00000350283.3) aquela mais frequente, presente em 11 dos 22 pacientes (11,7% dos casos) que apresentaram variantes patogênicas (**Figura 18**). Esta variante foi também a mais frequente no estudo de Timoteo e colaboradores (2018) na população brasileira, embora com uma frequência de apenas 2,5%. Esta variante é comumente encontrada em populações latino americanas, sendo bastante descrita em pacientes com HBOC na população brasileira, especialmente em casos de câncer de ovário (MAISTRO et al., 2016; MAURO et al., 2019), embora tenha sido encontrada exclusivamente em casos de câncer de mama na série de casos estudada. Esta é uma variante fundadora identificada na população Ashkenazi, bastante frequente em populações Norte Europeias afetadas por câncer de mama (HAMEL et al., 2011), o que explica sua alta frequência em nossa coorte que é marcada por um forte componente ancestral europeu (PENA et al., 2011). Esta variante não foi identificada no grupo controle.

Para uma melhor análise e interpretação das variantes, as amostras foram então agrupadas quanto ao genótipo, onde foram considerados três grupos principais: amostras que apresentaram variantes patogênicas nos genes *BRCA1/BRCA2*; amostras que apresentaram variantes de significado incerto ou benignas nos genes *BRCA1/BRCA2*; e amostras que não apresentaram quaisquer variantes nestes dois genes (não-BRCA). Levando em consideração os genes *BRCA1* e *BRCA2*, a prevalência de variantes patogênicas nestes genes foi de 18,1% (**Figura 18**). Não foi identificada nenhuma variante não descrita nestes genes.

Além da mutação p.Gln1756Profs*74 no gene *BRCA1*, foram identificadas ainda duas outras variantes patogênicas: a variante *nonsense* (mutação de ganho de códon de parada) p.Trp1836Ter, c.5507G>A (rs80356962), e a variante p.Gln1111Asnfs*5, c.3331_3334delCAAG (rs80357701). No gene *BRCA2*, foram identificadas as variantes p.Ala938Profs*21, c.2808_2811delACAA (rs80359351), p.Tyr3009Serfs*7, c.9026_9030delATCAT (rs80359741) e p.Arg3128Ter, c.9382C>T (rs80359212). Todas as variantes estavam em heterozigose e nenhuma outra variante além de p.Gln1756Profs*74 no gene *BRCA1* foi identificada em mais de um paciente.

O gene *TP53* foi o terceiro gene que apresentou maior frequência de variantes patogênicas

(5,3%). Além da variante p.Arg337His, para uma das cinco pacientes que apresentou esta variante foi observada dupla heterozigose com a variante patogênica já descrita p.Arg273His, c.818G>A (rs28934576). Este paciente foi o único caso de óbito dentre os indivíduos portadores de variantes neste gene. Uma das pacientes que apresentou a variante p.Arg337His não apresentou nenhuma variante nos genes *BRCA1* e *BRCA2*. Variantes no gene *TP53* são causas da síndrome de Li-Fraumeni (LFS), caracterizada pela presença de sarcomas de partes moles, câncer de mama pré-menopausa, tumores cerebrais, carcinomas adrenocorticais e leucemia, com risco aumentado para tumores primários múltiplos e acometimento na infância (CURY; FERRAZ; SILVA, 2014). O câncer de mama em idade jovem é um dos tumores mais frequentes em pacientes com LFS e por isso pacientes com HBOC são frequentemente investigados quanto à presença variantes no gene *TP53* (CURY; FERRAZ; SILVA, 2014; NANDIKOLLA; VENUGOPAL; ANAMPA, 2017).

Apesar de variantes patogênicas nos genes *BRCA1* e *BRCA2* já terem seu papel bastante caracterizado no aumento de risco para HBOC, variantes em outros genes como *PMS2* e *BRIP1* também foram consideradas possivelmente patogênicas uma vez que são classificadas como VUS mas a maioria das ferramentas (5/6) de predição as classificaram como patogênicas. A variante p.Asp792Asn (rs587781265) foi identificada em heterozigose no gene *PMS2* (ENST00000265849) em um paciente do sexo masculino que apresentou câncer gástrico difuso e veio a óbito. Esse paciente teve uma sobrevida de 3 anos após o diagnóstico da doença, e não apresentou nenhuma outra variante de potencial impacto nos demais genes investigados. Embora a presença de variantes no gene *PMS2* esteja mais associada a casos de síndrome de Lynch, a prevalência de variantes neste gene pode chegar a até 18% em casos de HBOC (KAPOOR et al., 2015; MINION et al., 2015). A variante p.Asp792Asn já foi descrita por Arora e colaboradores (2017), como uma variante de funcionalidade moderada uma vez que foi observado um aumento da viabilidade celular e no recrutamento de H2Ax após tratamento com drogas indutoras de danos ao DNA, indicando uma deficiência moderada no sistema de reparo MMR. Já a variante p.Gln740His (rs45589637) no gene *BRIP1* (ENST00000259008) já foi descrita em casos de cânceres de mama e ovário, mas não possui nenhuma informação descrita sobre a sua funcionalidade (JALKH et al., 2017; LI et al., 2016). Esta variante foi identificada em um paciente com câncer de mama, no entanto, variantes no gene *BRIP1* estão relacionadas ao aumento de risco para câncer de ovário, sendo este gene descrito como de baixa ou nenhuma penetrância para câncer de mama (EASTON et al., 2016; RAMUS et al., 2015; SEAL et al., 2006).

Cerca de 69% dos pacientes estudados apresentaram variantes benignas ou de significado

clínico duvidoso nos genes *BRCA1* e/ou *BRCA2*. Neste grupo de pacientes, a frequência de variantes *missense* com conflito de informações sobre patogenicidade foi bastante alta, sendo os genes *BRCA1*, *FAM175A* ou *ABRAXAS1*, *ATM*, *BRCA2* e *UIMC1* aqueles mais mutados.

A alta frequência de variantes observada no gene *BRCA1* foi, na verdade, decorrente da alta prevalência de quatro variantes que pareciam estar segregando de forma dependente: p.Pro871Leu (rs799917), p.Glu1038Gly (rs16941) e p.Lys1183Arg (rs16942) no éxon 11, e p.Ser1613Gly (rs1799966) no éxon 16. As quatro variantes são descritas como benignas pelo ClinVar, no entanto, na condição de estarem segregando juntas, poderiam contribuir para o aumento de risco para HBOC. Pensando nisso, a base de dados do HapMap (<https://www.genome.gov/10001688/international-hapmap-project>) foi consultada e foi observado que a variante c.*421G>T (rs8176318), localizada na região 3'UTR do *BRCA1*, também estava presente na maioria dos indivíduos em que essas quatro variantes também estavam. A variante rs8176318 já foi associada com diminuição da expressão de *BRCA1* e com risco aumentado para câncer (DORAIRAJ et al., 2014; PELLETIER et al., 2011), entretanto, também é descrita como benigna pelo ClinVar. Todas as 95 amostras de HBOC foram então genotipadas para esta variante a fim de caracterizar se as variantes estavam em desequilíbrio de ligação na amostra estudada. Para aumentar o poder estatístico da análise de haplótipos, outras 24 amostras adicionais de HBOC que não foram sequenciadas foram genotipadas juntamente com o grupo controle composto por 136 idosos sem história pessoal de câncer. De fato as variantes estavam em desequilíbrio de ligação, sendo estimados quatro possíveis haplótipos com frequência maior que 1% para as variantes (**Tabela 4**), sendo o haplótipo 1 aquele que apresentou maior frequência na amostra estudada. No entanto, este haplótipo não foi associado ao risco aumentado para HBOC uma vez que foi significativamente mais frequente na população controle que na série de casos estudada (OR=0.64, 95% CI 0.42-0.99, p= 0.020125).

Tabela 4. Frequências estimadas dos haplótipos envolvendo as cinco variantes frequentes no gene *BRCA1* na série de casos estudada.

Hp	p.Pro871Leu (CCG→CTG)	p.Glu1038Gly (GAA→GGA)	p.Lys1183Arg (AAA→AGA)	p.Ser1613Gly (AGT→GGT)	c.*421G>T (G/T)	Freq. Hap	Controle (n=136)	HBOC (n=119)	p.valor
1	Leu	Gly	Arg	Gly	T	0.248	0.292	0.199	0.020
2	Leu	Glu	Lys	Ser	G	0.136	0.129	0.143	0.633
3	Pro	Glu	Lys	Ser	G	0.546	0.533	0.563	0.532
4	Leu	Gly	Lys	Gly	T	0.028	0.0167	0.038	0.172

Hp = Haplótipos estimados. Freq. Hap. = frequência estimada para os haplótipos. p-valor = valor estatístico calculado pelo software haplo stats para estimação dos haplótipos.

Bodian e colaboradores (2014), também relataram a alta frequência das quatro variantes p.Pro871Leu, p.Glu1038Gly, p.Lys1183Arg, p.Ser1613Gly em sua série de casos estudada. Após uma análise da composição étnica das amostras estudadas, o grupo evidenciou que a alta frequência destas variantes era ancestralidade dependente, uma vez que as amostras que apresentaram p.Pro871Leu exibiram forte componente ancestral africano e europeu, e as demais variantes um forte componente central asiático. De fato, isto pode explicar a alta frequência encontrada na população brasileira, para ambos os grupos caso e controle, que é caracterizada pela alta miscigenação étnica com a presença de um forte componente Africano e Europeu. Desta forma, acreditamos que a presença do haplótipo 1 provavelmente não está associada ao aumento da susceptibilidade ao câncer.

Cerca de 66% dos pacientes apresentaram variantes *missense* com dados conflitantes de patogenicidade no gene *FAM175A*. A variante mais frequente foi a p.Ala348Thr (c.1042G>A) que apresentou uma frequência alélica de 0,4 na população estudada, sendo identificada tanto em heterozigose como em homozigose. Um dos 28 pacientes sequenciados para o grupo controle, também apresentou esta variante. Trata-se de uma variante não descrita no ClinVar e que possui conflito de informações de patogenicidade dentre as ferramentas de predição *in silico* (Conflito de informações). É descrita pelo dbSNP e trata-se de um polimorfismo com MAF de 0,42 na população mundial (1000 Genomes Projects) e de 0,34 na população brasileira (ABraOM), e já é relatada em outros estudos como apresentando uma alta prevalência nos grupos controle (SOLYOM et al., 2013). FAM175A é uma das três proteínas que compõem a família das Abraxas, cuja interação com o domínio BRCT no gene *BRCA1* atua como um adaptador para a ligação das demais proteínas do complexo BRCA1-A, importante na sinalização do dano ao DNA e imprescindível para um correto reparo de quebra de dupla fita e consequente manutenção da estabilidade genômica (HER et al., 2016). Estudos têm demonstrado que animais e células que tiveram a proteína depletada, apresentaram uma maior sensibilidade à radiação ionizante, comprometimento do checkpoint de G2/M por impedir o recrutamento de BRCA1 aos sítios de danos ao DNA, como também levaram ao aumento na ligação de RPA e RAD51, indicativo do aumento dos níveis de ressecção de pontas de DNA o que leva à instabilidade genômica; sendo ainda associado ao desenvolvimento tumoral espontâneo em camundongos (CASTILLO et al., 2014; RENAULT et al., 2016).

Com uma frequência alélica um pouco menor na população estudada (0,10), a variante p.Pro435Leu (c.1304C>T) no gene *UIMCI*, também não descrita pelo ClinVar, foi outra VUS bastante frequente nos casos estudados. Assim como a variante no *FAM175A*, esta variante possui conflito de informações de patogenicidade e apresenta um MAF relativamente alto nas populações

brasileira (0,12) e mundial (0,24). UIMC1 ou RAP80 também é uma das proteínas do complexo BRCA1-A, e desempenha papel importante no recrutamento de BRCA1 até os sítios de danos ao DNA por meio da ligação à cadeias de ubiquitina de histonas H2AX (HER et al., 2016). Estudos já têm demonstrado que células deficientes em RAP80 são hipersensíveis à radiação ionizante, reduzem a formação de focos de BRCA1 e estão associadas à instabilidade genômica e à deficiência do reparo por recombinação homóloga (KIM; CHEN; YU, 2007; LI et al., 2017).

Diante disso, uma vez que não existem estudos que comprovem ou não a relação destas duas variantes com a doença, realizamos a análise de ativação da histona fosforilada γ H2Ax, importante marcador da sinalização de danos ao DNA (PODHORECKA; SKLADANOWSKI; BOZKO, 2010), após indução de danos por irradiação (2 Gy) em linfócitos isolados de pacientes. Para isso, foram utilizadas células de pacientes que apresentaram p.Ala348Thr no *FAM175A* e p.Pro435Leu no *UIMC1*, em dupla heterozigose, e que foram negativos para variantes nos demais genes analisados, bem como para alterações no número de cópias nos genes *BRCA1* e *BRCA2*.

Apesar da fosforilação da histona γ H2Ax ser um evento que ocorre antes do recrutamento de Abraxas1, RAP80 e BRCA1 aos sítios de dano ao DNA (LI et al., 2017), esta é uma importante ferramenta para análise da cinética de reparo ao DNA. Se as proteínas Abraxas1, RAP80 e BRCA1 estão funcionais, serão capazes de sinalizar os danos ao DNA às moléculas efetoras que realizarão o processo de reparo o que por sua vez refletirá na diminuição dos níveis de γ H2Ax algumas horas após a indução do dano. Isto é confirmado ao ver a cinética de reparo de células que possuem a proteína BRCA1 truncada por uma variante do tipo *frameshift* (controle positivo), que são incapazes de retornar aos níveis basais de fosforilação de γ H2Ax após 6h quando comparada a células controle (**Figura 19**). Nas amostras do controle negativo (paciente sem histórico pessoal/familiar de câncer, barras em verde) é observado um aumento dos níveis de γ H2Ax fosforilada logo após a irradiação, apresentando um aumento de mais de 46% da marcação 4h após a irradiação. No entanto, estes danos são completamente reparados após 6h, com as células retornando ao estado basal de ativação de γ H2Ax. Já as células portadoras das variantes estudadas apresentaram níveis basais de γ H2Ax fosforilada muito similares aos do controle positivo, atingindo um pico de fosforilação 2h após a irradiação apresentando cerca de 72% das células marcadas para γ H2Ax. De forma similar ao controle positivo, as células testadas não recuperaram os níveis basais de fosforilação de γ H2Ax após 6h da irradiação. No entanto, não foram evidenciadas diferenças significativas das amostras em relação ao controle negativo. Apesar de ter sido observada uma variabilidade muito grande nos níveis de fosforilação das amostras do controle positivo e das amostras testadas, estes resultados podem

indicar o comprometimento da detecção de danos ao DNA pelo sistema de reparo em pacientes portadoras destas VUS. No entanto, a avaliação individualizada de ambas as variantes é essencial para melhor caracterizar a contribuição de cada uma delas para o fenótipo.

Figura 19. Porcentagem de células positivas para fosforilação de γ H2Ax antes e após irradiação com raios-X (2 Gy). Cada barra representa as médias de duas replicatas biológicas de triplicatas experimentais. CTRL = controle. não-IR são células não irradiadas. Foi realizada análise estatística utilizando teste ANOVA e teste de múltiplas comparações de Dunnetts e Tukey. Não foram identificadas diferenças significativas entre as amostras.

Outro gene com alta prevalência de variantes dentre as amostras foi o gene *ATM* (29%). Cerca de 16,8% das amostras apresentaram a variante p.Asp1853Asn, c.5557G>A (rs1801516), caracterizada como benigna pelo ClinVar. Estudos sugerem que esta variante não está associada ao aumento de risco para HBOC (JIN et al., 2011) e de fato, um dos 28 pacientes do grupo controle apresentaram esta variante.

Foi identificada também a variante p.Pro932Thr (c.2794C>A) no gene *ATR*, que é considerada possivelmente patogênica. Trata-se de uma variante nova e, portanto, não é descrita pela base de dados do ClinVar. Na análise *in silico*, todas as ferramentas de predição de patogenicidade a classificaram como patogênica. No entanto, a mesma paciente que apresentou esta variante também apresentou a VUS p.Pro435Leu no gene *UIMC1*. Esta paciente apresentou ainda a variante sinônima de significado incerto p.Glu102Glu, c.306G>A (rs63751665) no gene *MLH1* (ENST00000231790). Esta variante também é classificada como patogênica por todas as ferramentas de predição consultadas e é caracterizada na base de dados do ClinVar, porém não existem estudos caracterizando ou mesmo reportando essa variante. Portanto, é difícil associar qualquer informação ou caracterizar a variante nova p.Pro932Thr (c.2794C>A) no gene *ATR*, uma vez que a paciente também apresentou

outras variantes que podem estar relacionadas ao fenótipo, bem como a ausência na população controle impossibilita a realização de estudos de associação com o risco para HBOC.

Neste grupo de pacientes também foi observada uma alta frequência de variantes nos genes envolvidos com a via MMR de reparo de DNA. Dentre eles, os genes *PMS2* e *MSH2* foram aqueles que apresentaram uma maior frequência de variantes, com cerca de 19% e 10% dos pacientes estudados apresentando variantes nestes genes, respectivamente. Diversos estudos corroboram nossos dados, evidenciando os genes *PMS2* e *MSH2* como frequentemente mutados em casos de HBOC (BONACHE et al., 2018; CASTÉRA et al., 2014; KAPOOR et al., 2015). Variantes nos genes MMR estão mais associados a casos de câncer de ovário, em sua maioria relacionados à síndrome de Lynch, síndrome que junto com HBOC correspondem àquelas de herança autossômica dominante mais comuns entre as mulheres com cânceres ginecológicos (RASHID et al., 2016). No entanto, dos pacientes que apresentaram variantes em *PMS2* e *MSH2*, 23 eram pacientes com câncer de mama, um caso de câncer de endométrio e outro de câncer gástrico, com nenhum caso apresentando câncer de ovário. Apesar de alguns casos estudados terem apresentado câncer de endométrio e gástrico, os mesmos preencheram critérios NCCN para HBOC e não de Lynch. De fato, Espenschied e colaboradores (2017), sugerem em seu estudo que pacientes portadores de variantes nos genes *MSH6* e *PMS2* apresentaram um fenótipo mais próximo de HBOC do que de Lynch devido a presença de casos de cânceres de mama e ovário. O mesmo grupo evidenciou ainda que pode haver sobreposição fenotípica de casos de Lynch e HBOC, como por exemplo, em sua série de casos estudada em que 36% dos casos preenchiam critérios clínicos para ambas as síndromes.

Na série de casos estudada, aproximadamente 12,8% (n=12) das amostras não apresentaram quaisquer variantes nos genes *BRCA1* e *BRCA2*. Nestes pacientes foram identificadas variantes em 10 dos 21 genes estudados: *ATM*, *ATR*, *BARD1*, *CDH1*, *FAM175A*, *MLH1*, *MSH6*, *PMS2*, *TP53* e *UIMC1*. Para confirmar que os genes *BRCA1/BRCA2* não estavam relacionados ao fenótipo nestes casos, foi realizada também a pesquisa de alterações no número de cópias para estes dois genes pela técnica de MLPA e nenhuma amostra apresentou alterações nos genes *BRCA1* e *BRCA2*. Foi dentre estes pacientes que foi identificada a variante p.Arg337His no gene *TP53*, descrita anteriormente.

Com a análise do painel gênico, foi possível identificar variantes patogênicas e estabelecer o diagnóstico molecular para 23,4% dos pacientes estudados. O único estudo envolvendo a análise de painel gênico em outros pacientes com HBOC na população brasileira evidenciou uma frequência de variantes patogênicas em apenas 12,1% dos casos estudados, embora tenha identificado outros genes mutados como *ATM*, *ATR*, *MLH1*, *MSH2* e *MSH6* (TIMOTEO et al., 2018). Na América Latina,

outro estudo descreveu variantes patogênicas nos genes *BRCA1/BRCA2*, *PALB2*, *ATM*, *MSH2* em 22,4% dos pacientes estudados, frequência similar à encontrada em nosso estudo (COCK-RADA et al., 2017). Demais estudos em outras populações também descrevem frequências similares ou mais baixas de pacientes afetados por variantes patogênicas nestes e demais genes como *BRIP1*, *CHEK2*, *MLH1*, *MSH6* e *PTEN* (BONACHE et al., 2018; QUEZADA URBAN et al., 2018; TEDALDI et al., 2017). Na série de casos estudada foram identificadas variantes consideradas possivelmente patogênicas nos genes *PMS2* e *ATR*, no entanto, variantes sabidamente patogênicas em casos negativos para variantes em *BRCA1/BRCA2* foram identificadas em apenas 4% dos pacientes, que apresentaram variantes patogênicas no gene *TP53*. Esta frequência, no entanto, é similar à encontrada em outros genes por Timoteo e colaboradores (2018), e Cock-Rada e colaboradores (2017), em seus estudo na população Brasileira e Colombiana, respectivamente. Na população Espanhola, Bonache e colaboradores (2018), descrevem 8% dos pacientes carregando variantes patogênicas em outros genes que não *BRCA1/BRCA2*.

Apesar de o número de casos com diagnóstico molecular ter corroborado as frequências identificadas em outras populações, muitas são as variantes identificadas sem um claro significado clínico. A maior parte das variantes identificadas neste estudo (83,3%) constituem variantes com conflito de informações de patogenicidade – de acordo com o ClinVar e ferramentas de predição de patogenicidade *in silico* – o que dificulta a interpretação destas variantes dentro de um contexto clínico. Na tentativa de melhor caracterizar estas variantes, foi realizado um estudo de associação com o risco para HBOC utilizando as frequências alélicas encontradas nos pacientes estudados em comparação com as frequências alélicas das mesmas variantes descritas em bancos de dados populacionais. Foi utilizado principalmente o banco de dados do Arquivo Brasileiro Online de Mutações (ABraOM), que sequenciou o exoma de 609 indivíduos com idade acima de 60 anos. O uso de um banco de dados da mesma população que aquela estudada é de extrema importância uma vez que a população brasileira possui um alto nível de miscigenação, o que pode ser um viés para as análises de associação. No entanto, quando as frequências alélicas na população brasileira foram iguais a zero, foi utilizada a frequência alélica das variantes em populações Europeia – não finlandesa, Latina, Americana, Africana e naquelas categorizadas como “outras”, relatadas pelos bancos de dados ExAC (<http://exac.broadinstitute.org/>), 1000 Genomes Project (<http://www.internationalgenome.org/>) e gnomAD (<https://gnomad.broadinstitute.org/>). Das 69 variantes de significado incerto, 13 variantes em 8 genes foram significativamente associadas com risco para HBOC quando comparadas às frequências alélicas disponíveis nos bancos de dados populacionais (**Figura 20**). Para todas as variantes associadas ao fenótipo, foi realizada análise de

alterações no número de cópias para os genes *BRCA1* e *BRCA2* pela técnica de MLPA e nenhum paciente apresentou alterações nestes genes. Nenhuma destas variantes foram identificadas no grupo controle.

Figura 20. Variantes com conflito de informações de patogenicidade e análise da probabilidade de associação destas variantes com HBOC. A análise de associação foi realizada comparando as frequências alélicas encontradas nos pacientes estudados e comparadas às frequências alélicas nas bases de dados populacionais * na população brasileira (bases de dados populacionais ABraOM) ou mundial (1000Genomes e ExAC, levando em consideração as seguintes populações: Europeia – não finlandesa, Latina, Americana, Africana e aquelas categorizadas como “outras”). \$ B = Benigna; LB = Provavelmente benigna; US = significado incerto; P = Patogênica. Os valores de *odds ratio* (OR) foram calculados utilizando o teste exato de Fisher, e os valores de “p” foram obtidos utilizando o teste do χ^2 de Pearson. A ausência de valores de OR acontece quando a frequência alélica é igual a zero nos bancos de dados populacionais. As variantes marcadas em vermelho são aquelas que foram significativamente associadas ao HBOC.

Dos 94 pacientes estudados, 15 apresentaram variantes associadas ao risco para HBOC. Destes, três são pacientes que vieram o óbito: uma que apresentou a variante p.Ala617Thr no gene *CDH1*, outra que apresentou p.Ile160Met no gene *CHEK2*, ambos casos de câncer de mama, e outro que apresentou a variante p.Asp792Asn no gene *PMS2*, considerada possivelmente patogênica, em um paciente que teve câncer gástrico. Dois dos casos de óbito (variantes em *CDH1* e *PMS2*) são casos que não apresentaram quaisquer variantes nos genes *BRCA1/2*, assim como outros dois pacientes que também apresentaram p.Ala617Thr no gene *CDH1*, e outro que apresentou a variante p.Asn255Ser no gene *BARD1*. Um paciente, que apresentou a variante p.Leu239Gln no gene *BARD1*, também apresentou a variante patogênica p.Trp1836Ter no gene *BRCA1*. Os outros 10 pacientes apresentaram variantes de significado incerto nos genes *BRCA1/2*. Exceto para a variante p.Arg137Gln no gene *CHEK2*, caracterizada como provavelmente benigna, todas as variantes associadas ao fenótipo são caracterizadas como variantes de significado incerto ou com dados conflitantes de patogenicidade pelo ClinVar.

Diante destes dados, ensaios funcionais têm sido aplicados no sentido de validar a contribuição destas variantes que foram correlacionadas com o HBOC na patogênese da doença. Para algumas destas variantes, entretanto, não foi possível realizar a avaliação funcional uma vez que muitos dos ensaios ainda permanecem em fase de padronização. Foi realizada a tentativa de estudos de segregação da doença na família, que infelizmente não foi possível pela falta de indivíduos afetados vivos ou por dificuldades no recrutamento dos familiares. Das variantes que foram associadas ao fenótipo, com exceção das variantes p.Ala617Thr no gene *CDH1* e p.Lys18Arg no gene *PALB2*, todas apresentaram frequência alélica igual a zero na população brasileira, sendo usadas as frequências alélicas na população mundial para estimar as probabilidades de associação.

O gene *BARD1* foi aquele que apresentou maior número de variantes associadas ao fenótipo: p.Leu239Gln, c.716T>A (rs200359745), p.Asn255Ser, c.764A>G (rs138904906) e p.Lys423Arg, c.1268A>G (rs749383704). Variantes neste gene têm sido frequentemente relatadas em pacientes com câncer de mama/ovário hereditário (QUEZADA URBAN et al., 2018) e associadas a um risco moderado para câncer de mama (SUSZYNSKA et al., 2019; WEBER-LASSALLE et al., 2018). As variantes p.Asn255Ser e p.Lys423Arg são descritas no ClinVar mas não apresentam nenhum estudo relatando sua frequência na população e nem dados referentes à sua caracterização funcional. Já a variante p.Leu239Gln, já foi descrita em um estudo na população norte americana, como variante de significado incerto identificada em um dos 488 casos estudados (TUNG et al., 2016). A proteína BARD1 possui três principais domínios: um domínio RING na região N-terminal, um domínio com

três fragmentos de repetição Ankyrin (ANK) e dois domínios BRCT na sua região C-terminal, e é no domínio RING onde ocorre a ligação com BRCA1, formando heterodímeros que são recrutados e permanecem nos sítios de danos ao DNA (CIMMINO; FORMICOLA; CAPASSO, 2017). As variantes p.Leu239Gln e p.Asn255Ser estão localizadas entre os domínios RING e ANK (**Figura 21**), sendo classificadas como possivelmente patogênicas por duas das seis ferramentas de predição utilizadas. Já a variante p.Lys423Arg está localizada na região inicial do domínio ANK (**Figura 21**), região importante para a ligação de BARD1 com p53 e ativação da apoptose (FEKI et al., 2005). Apesar deste domínio não estar relacionado com reparo de DNA, um estudo recente caracterizou 4 variantes em regiões próximas à observada neste estudo (aminoácidos 460 a 560), que apresentaram deficiência no reparo por recombinação homóloga, evidenciando este domínio como importante para o reparo de danos ao DNA (ADAMOVICH et al., 2019). Três das seis ferramentas de predição *in silico* caracterizaram essa variante como patogênica. No entanto, diante da escassez de informações sobre estas variantes, estudos funcionais se fazem necessários para caracterização destas VUS em *BARD1* na patogênese do HBOC.

Figura 21. Representação esquemática da proteína BARD1 e variantes encontradas no estudo e associadas ao risco para HBOC. As caixas brancas representam as duas variantes identificadas em *BARD1* e a localização de cada uma delas. As caixas em diferentes cores representam os três diferentes domínios da proteína: RING *domain* em amarelo e verde; Ankyrin (ANK) repeat *domain* em azul e lilás e, BRCT *domain* em tons de rosa. Embaixo, escala que vai de 0 a 800 representando a localização das variantes e dos domínios da proteína que possui 777 aminoácidos.

A variante p.Met2775Arg (rs80359073) no gene *BRCA2* foi outra variante identificada como associada ao risco para HBOC. Esta variante está localizada na região C-terminal de *BRCA2*, em um domínio de ligação com DNA de fita simples, cuja interação é importante para disponibilizar moléculas de RAD51 – também ligadas à *BRCA2*, na região de quebras no DNA para um correto reparo por recombinação homóloga (VON NICOLAI et al., 2016). Essa variante é caracterizada como variante de significado incerto e já foi descrita em pacientes com câncer de próstata (MAIER et al.,

2014). A metionina da posição 2775 não é um aminoácido conservado evolutivamente e a arginina nesta mesma posição faz parte da cadeia de aminoácidos de espécies de murinos (YANG et al., 2002), no entanto, quatro ferramentas de predição de patogenicidade *in silico* classificam esta variante como possivelmente patogênica. Isso sugere que a integridade desta região é necessária para um correto reparo de quebras de dupla fita no DNA.

Dentre as variantes associadas ao fenótipo, a variante p.Ala617Thr no gene *CDHI* foi aquela que apresentou frequência alélica mais alta, sendo significativamente mais frequente na série de casos estudada que na população brasileira. Variantes no gene *CDHI* são mais associadas a casos de câncer gástrico difuso hereditário, bem como a casos de câncer de mama do tipo lobular (STUEBS et al., 2018). De fato, uma das pacientes tem história de parente de segundo grau com câncer gástrico, porém todos os casos índices apresentaram casos de câncer de mama do tipo Carcinoma Ductal Invasivo, preenchendo critério clínicos para HBOC. Diferentemente dos outros genes analisados, *CDHI* codifica a proteína de adesão E-caderina, e variantes que afetam a função desta proteína estão associadas ao comprometimento da adesão celular, aumento da capacidade invasiva das células e por consequência, à quadros de metástase (MATEUS et al., 2009). De fato, os três casos apresentaram metástases linfonodais, com uma paciente apresentando também metástase à distância acompanhada de óbito. Esta variante afeta um motivo conservado presente no domínio de ligação de íons cálcio que são necessários para conferir estabilidade à proteína E-caderina em sua forma ativa (SHAPIRO et al., 1995). Foi primeiramente descrita na população brasileira em uma paciente de HBOC por Lajus e Sales (2015), que sugerem que esta variante pode estar relacionada à formação de complexos celulares instáveis. De fato, duas ferramentas de predição *in silico* classificaram esta variante como patogênica. Timoteo e colaboradores (2018) classificam esta variante como patogênica. Entretanto, Suriano e colaboradores (2003), realizaram testes funcionais para avaliar a capacidade de proliferação, migração e morfologia de células expressando a proteína mutada para p.Ala617Thr e observaram um fenótipo muito próximo do selvagem para todos os ensaios realizados. Assim, acredita-se que esta variante, por si só, pode não estar relacionada ao aumento de risco para HBOC.

Outro gene que apresentou variantes associadas ao risco para HBOC foi *CHEK2*, gene importante na sinalização de danos ao DNA. CHK2 é ativada pela fosforilação do resíduo Thr68 do seu domínio SCD por ATM, o que permite a propagação do sinal de danos ao DNA ao fosforilar e ativar outras proteínas efetoras como BRCA1 (CRAIG; HUPP, 2004). As duas variantes p.Arg137Gln e p.Ile160Met significativamente associadas ao fenótipo estão localizadas no domínio FHA (*Forkhead associated*) de CHK2 (**Figura 22**). Em resposta à fosforilação por ATM, o domínio catalítico KD

(*Kinase Domain*) de CHK2 é autofosforilado promovendo a ligação do domínio SCD (*SQ/TQ Cluster Domain*) fosforilado ao FHA de outra proteína CHK2, formando dímeros que se dissociam formando monômeros em seu estado ativo (ZANNINI; DELIA; BUSCEMI, 2014).

Figura 22. Representação esquemática da proteína CHK2 e variantes encontradas no estudo e associadas ao risco para HBOC. As caixas brancas representam as duas variantes identificadas em CHK2 e a localização de cada uma delas. As caixas em diferentes cores representam diferentes domínios da proteína: FHA *domain* em verde musgo e laranja, onde estão localizadas as duas variantes; *Protein Kinase* (KD) *domain* ou Domínio catalítico, em azul e verde e, em rosa, sítio ativo de fosforilação nos resíduos Ser (serina) e Thr (treonina). Embaixo, escala que vai de 0 a 400 representando a localização das variantes e dos domínios da proteína que possui 543 aminoácidos.

A variante p.Arg137Gln (p.R137Q) já foi descrita em outros estudos de pacientes com câncer de mama e/ou ovário (ROEB; HIGGINS; KING, 2012) e já é descrita no ClinVar como possivelmente benigna. No entanto, é predita como possivelmente patogênica por duas ferramentas de predição *in silico*. Estudos funcionais evidenciaram que p.R137Q não alterou a estabilidade da proteína e apresentou uma resposta a danos ao DNA e atividade catalítica normais (BELL et al., 2007; DESRICHARD et al., 2011; SODHA et al., 2006). Já a variante p.Ile160Met, possui conflito de informações de patogenicidade no ClinVar e é classificada como patogênica por quatro das seis ferramentas de predição *in silico*. Também já foi descrita em pacientes com câncer hereditário de mama (DFAULT et al., 2004; MOHAMAD et al., 2015) e Roeb e colaboradores (2012), caracterizaram esta variante como apresentando atividade parcial de reparo aos danos ao DNA. Para ambas as variantes foi possível realizar avaliação funcional durante o estágio no exterior no laboratório de Estabilidade Genômica, sob orientação do Prof. Dr. Jean-Yves Masson. Foi induzida a expressão das proteínas mutadas em células HeLa silenciadas para *CHEK2* endógeno e foi observado o nível de ativação de CHK2, relativo à expressão da proteína fosforilada, após indução de danos ao DNA causados por irradiação (utilizando uma dose de 5 Gy). Primeiramente foi observado que a proteína mutada p.I160M apresenta certa instabilidade uma vez que a expressão foi reduzida em relação à

expressão da proteína selvagem, bem como em relação a p.R137Q (**Figura 23**). Mesmo após aumentar a quantidade de proteína p.I160M em até 5x, não foi possível atingir os níveis de expressão da proteína selvagem.

Figura 23. Western blot evidenciando a expressão das proteínas CHK2 mutadas p.R137Q e p.I160M em relação à proteína selvagem e endógena. siCTRL corresponde à expressão da proteína CHK2 endógena (61 kDa); VV= Vetor pcDNA3 vazio; wt= *wild type* ou proteína selvagem Vinculina= proteína endógena (116 kDa).

Analisando a ativação de CHK2 por meio dos níveis de fosforilação da proteína no resíduo Thr68, foi observado que 30 minutos após a irradiação, a proteína selvagem é fosforilada em resposta ao dano ao DNA, apresentando diminuição dos sinais de fosforilação 60 minutos após a irradiação. A proteína mutada p.R137Q apresentou níveis de fosforilação similares aos da proteína selvagem, no entanto, os níveis de fosforilação de p.I160M foram significativamente mais baixos ($p < 0,01$) quando comparados à proteína selvagem (**Figura 24**). Isto sugere que a variante p.I160M, localizada em um domínio importante do gene *CHEK2*, pode estar relacionada ao fenótipo por alterar a estabilidade bem como a ativação da proteína CHK2, comprometendo a sinalização e o reparo de danos ao DNA. A

paciente que apresentou a variante p.I160M apresentou ainda outras variantes nos genes *BRCA1* e *PMS2*, também caracterizadas como benignas pelo ClinVar e não associadas ao fenótipo, e ainda a VUS p.Ala348Thr no gene *FAM175A*. Esta paciente apresentou câncer de mama bilateral sincrônico com diagnóstico aos 34 anos e que veio a óbito dois anos depois após apresentar quadros de metástase.

Figura 24. Western blot evidenciando os níveis de expressa da proteína CHK2 fosforilada, antes, 30 e 60 minutos após irradiação. pCHEK2 (Thr68) corresponde à marcação da proteína fosforilada pelo anticorpo que reconhece a fosforilação no resíduo Thr68 da proteína. Expressão das proteínas mutadas e selvagem e presença de sinais de fosforilação mais fracos quando as células não são expostas à Irradiação. IR = Irradiação (dose de 5 Gy). As células foram coletadas antes e 30 e 60 após a irradiação (5 Gy). Vinculin = Vinculina, a proteína endógena utilizada. EV = Empty vector (vetor vazio). **a e **b = significância estatística com $p < 0.01$ quando comparado à expressão de pCHEK2 na proteína selvagem 30' e 60' após irradiação, respectivamente. Os valores de significância foram obtidos utilizando teste ANOVA e teste de múltiplas comparações de Dunnetts.

O gene *FAM175A* também apresentou uma variante significativamente associada ao fenótipo. A variante mais frequente neste gene foi a p.Ala348Thr, que possui frequência alélica elevada nas populações controle o que sugere que esta variante seja apenas um polimorfismo e que, portanto, não está relacionada ao risco para a doença. Apesar disso, células que expressam esta variante apresentaram deficiência na cinética de reparo após danos ao DNA por radiação ionizante (**Figura 19**). Já a variante p.Arg163Ser, c.489G>T (rs535462791), presente em apenas um paciente da série de casos estudada foi significativamente associada ao fenótipo ($p = 2.2 \times 10^{-16}$) (**Figura 20**). Esta variante é

descrita no dbSNP mas não é descrita no ClinVar. É predita como patogênica por cinco das 6 ferramentas de predição utilizadas nesse estudo, uma vez que está localizada no domínio Pad1 (MPN) na região N-terminal do gene *ABRAXAS1*, região de ligação de RAP80 e outras proteínas importantes na sinalização de danos ao DNA (WU et al., 2016). RAP80 media o recrutamento de BRCA1 aos sítios de danos ao DNA por meio de ligação com BRCA1, mediada por ABRAXAS, desta forma, acredita-se que variantes neste domínio de ligação em ABRAXAS podem impedir a localização de RAP80 e outras proteínas de ligação necessárias para um correto reparo destas lesões (RENAULT et al., 2016; WANG et al., 2009). No entanto, na ausência de caracterização funcional ou estudos de segregação da variante com a doença na família, fica difícil acessar a real contribuição desta variante para o fenótipo.

A variante sinônima p.Glu102Glu, identificada no gene *MLH1* também foi significativamente associada ao HBOC ($p=2.2 \times 10^{-16}$) (**Figura 20**). Esta variante é descrita como VUS pelo ClinVar devido à falta de estudos que caracterizem o papel desta variante na doença. Trata-se de uma variante localizada no final do éxon 3, que envolve um sítio de *splicing*. É classificada como patogênica por todas as ferramentas *in silico* que realizam predição de patogenicidade para variantes sinônimas (CADD, UMD predictor e mutation taster). A paciente que apresentou esta variante foi um caso de Carcinoma Ductal Invasivo com diagnóstico aos 40 anos, que também apresentou a variante p.Val213Met também no gene *MLH1*, predita como benigna pelo ClinVar, bem como a VUS p.Pro435Leu no gene *UIMCI* e a variante não descrita p.Pro932Thr no gene *ATR*, predita como patogênica pelas ferramentas de predição *in silico*.

Durante o estágio no exterior, foi realizado um teste piloto para verificar a estabilidade da proteína na presença da variante p.Glu102Glu. Plasmídeos contendo as sequências selvagem e mutada de *MLH1* foram transfectados em células HeLa silenciadas para a proteína *MLH1* endógena. Foi observada uma expressão equivalente entre as proteínas mutada e selvagem (**Figura 25**), o que sugere que a variante não influencia na estabilidade da proteína expressa. No entanto, demais estudos são necessários a fim de melhor caracterizar esta variante no gene *MLH1*.

Figura 25. Western blot evidenciando a expressão das proteínas MLH1 selvagem (wt) e mutada p.E102E em células HeLa. siCTRL corresponde à expressão da proteína MLH1 endógena (80-90 kDa); VV= Vetor pcDNA3 vazio; wt= *wild type* ou proteína selvagem; MLH1= expressão da proteína MLH1, utilizando anti-MLH1. V5 = expressão da proteína exógena transfectada, contendo a tag V5, utilizando anti-V5. Tubulina= proteína endógena (50-55 kDa), utilizando anti-alpha tubulina.

Um outro gene que apresentou variantes associadas ao fenótipo foi o *PALB2*. Variantes bialélicas neste gene são causas conhecidas da anemia da Fanconi do tipo N e, mais recentemente, variantes em heterozigose foram descritas como associadas ao risco moderado para câncer de mama (DUCY et al., 2019; NEPOMUCENO et al., 2017). A variante p.Arg18Lys, c.53A>G (rs138789658), possui conflito de informações de patogenicidade no ClinVar, com estudos que a classificam como benigna, provavelmente benigna ou de significado incerto. Quanto às predições de patogenicidade, três das seis ferramentas de predição *in silico* a classificaram como patogênica. Esta variante está localizada no domínio *coiled coil* de PALB2 que é importante para interação com BRCA1, no entanto, estudos já demonstraram que esta variante não afeta a interação com BRCA1, mas que está associada à redução da capacidade de reparo por HR, o que sugere que a variante pode exercer algum efeito na conformação da proteína (FOO et al., 2017). Esta variante foi encontrada em dois pacientes da série de casos estudada, que apresentaram câncer de mama unilateral com diagnóstico aos 33 e 42 anos, e que não apresentaram variantes patogênicas em nenhum dos demais genes investigados. Uma das pacientes um caso de câncer de mama triplo negativo, mas nenhum dos casos apresentou metástases linfonodais ou à distância. Já a variante p.Thr317Pro, c.949A>C (rs587780223), classificada como VUS pelo ClinVar, foi identificada em um paciente que também apresentou câncer de mama unilateral do tipo triplo-negativo e com presença de metástases linfonodais. Diferente de p.Arg18Lys, esta variante não possui nenhum tipo de descrição ou caracterização na literatura. Está presente na região vizinha ao domínio de ligação ao DNA de PALB2 e é caracterizada como possivelmente patogênica por apenas duas das seis ferramentas de predição *in silico*.

A fim de caracterizar funcionalmente estas variantes, durante estágio no Departamento de

Pesquisa Clínica do INCA, sob supervisão do Prof Dr. Marcelo Alex de Carvalho, investigamos a capacidade de ligação das proteínas mutadas com BRCA1, uma vez que uma das variantes está localizada no domínio *coiled coil* de ligação com BRCA1, e a outra está localizada nas proximidades. Para isso foi utilizando um sistema de duplo híbrido em mamíferos e como observado na **Figura 26**, nenhuma das duas variantes afeta a capacidade de ligação de PALB2 com BRCA1, uma vez que ambas apresentaram capacidade de ligação muito similar ao *PALB2 wild type* (Wt). Os dados referentes à variante p.Lys18Arg foram gentilmente cedidos pelo Prof. Dr. Marcelo, uma vez que eles já estavam trabalhando na caracterização funcional desta e de outras VUS em *PALB2*.

Figura 26. Atividade de luciferase relativa à interação PALB2-BRCA1. Os valores absolutos de atividade de luciferase foram normalizados aos valores de luciferase de *Renilla reniformis* (pGR-TK utilizado como controle de transfecção). Os valores são expressos em %, em relação à atividade de PALB2 Wt (selvagem). Foi utilizado o vetor vazio pVP16 (VV) para demonstrar a especificidade do ensaio como também para verificar a atividade intrínseca do domínio BRCT presente no constructo de BRCA1. Foram utilizadas duas variantes já caracterizadas como patogênicas, como controles negativos de interação, p.L21A e p.L35P (Sy et al., 2009; Foo et al., 2017). *Western blot* evidenciando a ausência da expressão de BRCA1 exógeno em células não transfectadas (NT), e expressão de BRCA1 ligado à proteína GAL4 nas células transfectadas com o vetor vazio (VV), utilizando Anti-GAL4. Abaixo, ausência da expressão de PALB2 exógeno em células não transfectadas (NT), e expressão nas células transfectadas com as construções de PALB2 ligada à proteína VP16 (Utilizando Anti-VP16). Foi utilizada a análise de variância (one-way ANOVA) e o pós-teste de Dunnett (Dunnett's post hoc). Ambas as variantes de PALB2 testadas apresentaram capacidade de ligação com BRCA1 similar a PALB2 Wt, não apresentando nenhuma diferença significativa.

Foi ainda avaliado se qualquer outra variante de ocorrência natural no resíduo 317 além de p.Thr317Pro poderia interferir na capacidade de ligação com BRCA1, a fim de caracterizar se a

integridade da região circunvizinha à região *coiled coil* também era importante para interação com BRCA1. Para isso, foi realizada mutagênese dirigida para outras cinco possíveis trocas não sinônimas de aminoácidos no códon 317 e foi observado que quando há a troca de uma Treonina para uma Arginina (R) ou para uma Alanina (A) a capacidade de ligação com BRCA1 é diminuída significativamente, evidenciando que apesar de não estar localizado no domínio *Coiled coil* de PALB2, a integridade das regiões circunvizinhas também é importante para interação com BRCA1 (Figura 27).

Figura 27. Atividade de luciferase relativa à interação do resíduo 317 de PALB2 mutado por variantes naturais com BRCA1 em células HEK293FT. Os valores absolutos de atividade de luciferase foram normalizados aos valores de luciferase de *Renilla reniformis* (pGR-TK utilizado como controle de transfecção). Os valores são expressos em %, em relação à atividade de PALB2 Wt = proteína selvagem. Foi utilizado o vetor vazio pVP16 (VV) para demonstrar a especificidade do ensaio como também para verificar a atividade intrínseca do domínio BRCT presente no construto de BRCA1. Foram utilizadas duas variantes já caracterizadas como patogênicas, como controles negativos de interação, p.L21A e p.L35P (Sy et al., 2009; Foo et al., 2017). *Western blot* evidenciando a ausência da expressão de PALB2 exógeno em células não transfectadas, e expressão nas células transfectadas com as construções de PALB2 ligada à proteína VP16 (Utilizando Anti-VP16). Foi utilizada a análise de variância (one-way ANOVA) e o pós-teste de Dunnett (Dunnett's post hoc). As variantes p.T317A e p.T317R apresentaram interação reduzida com BRCA1 ($p < 0.05$ e $p < 0.01$, respectivamente) em relação à proteína exógena Wt.

Uma vez que o domínio BRCT de BRCA1 possui uma atividade intrínseca de ativação transcricional (MONTEIRO; AUGUST; HANAFUSA, 2002), apresentando até 60% da atividade relativa de luciferase em células transfectadas com o vetor vazio (vide VV = vetor vazio nas Figuras

26 e 27), os resultados do ensaio de duplo híbrido podem apresentar um certo viés. A fim de melhor caracterizar estas variantes foi realizado também um ensaio de sobrevivência ao Olaparib® durante o estágio no exterior, a fim de avaliar a funcionalidade do sistema de reparo em células transfectadas com as variantes p.K18R e p.T317P. Olaparib® é uma droga que têm sido usada para o tratamento de tumores deficientes no sistemas de reparo do DNA, uma vez que inibe proteínas da família PARP que são importantes na manutenção do sistema de reparo de quebras de fita simples no DNA (GUNDERSON; MOORE, 2015). Com a inibição destas proteínas, a persistência de quebras de fita simples leva ao colapso da forquilha de replicação, gerando quebras de dupla fita. Assim, células com deficiência no sistema de reparo por recombinação homóloga (HR), principal via de reparo de quebras de dupla fita, tornam-se sensíveis ao tratamento com Olaparib®, por um mecanismo conhecido como letalidade sintética (BUISSON et al., 2017). Desta forma, avaliar a sobrevivência de células mutadas a drogas inibidoras de PARP se torna uma ferramenta importante na avaliação da integridade do sistema de reparo (DUCY et al., 2019). Ambas as variantes testadas apresentaram resistência ao tratamento com Olaparib®, de forma semelhante à proteína selvagem (WT) (**Figura 28**) o que sugere que nenhuma das duas variantes estejam relacionadas ao comprometimento da via de reparo por recombinação homóloga. Os dados referentes à variante p.Lys18Arg foram gentilmente cedidos pelo Prof. Dr. Jean-Yves Masson, uma vez que eles já estavam trabalhando na caracterização funcional desta e de outras VUS em *PALB2*. Os resultados funcionais são sugestivos de que ambas as variantes não estejam relacionadas ao estabelecimento da doença nestas pacientes, entretanto, estudos de segregação das variantes com a família ainda são necessários para descartar a relação destas variantes com o fenótipo.

Figura 28. Curva de sobrevivência a diferentes concentrações de Olaparib (inibidor de PARP) de células HeLa após silenciamento da proteína PALB2 endógena. Células HeLa foram transfectadas com siRNA para PALB2 endógeno e transfectadas com constructos de vetor YFP vazio (VV), PALB2 *wild type* (Wt - selvagem) e PALB2 mutada (p.K18R e p.T317P). As células transfectadas com siRNA controle e com o vetor YFP vazio (VV) foram utilizadas como controles. *Western blot* evidenciando a expressão de PALB2 endógeno (~150 kDa), e exógeno nas células transfectadas com as construções de PALB2 ligada à proteína fluorescente YFP. Tubulina (50-55 kDa) foi utilizada como proteína endógena. Para ambas as variantes foi observada sobrevivência ao tratamento com Olaparib similar àquela observada para a proteína PALB2 wt (selvagem).

O gene *PMS2* também apresentou duas variantes significativamente associadas ao HBOC (**Figura 29**). Ambas as variantes estão localizadas no domínio MutL na região C-terminal de *PMS2* (**Figura 29**), que juntamente com sua região N-terminal, compõem os sítios de ligação com MLH1 para formação dos heterodímeros do complexo MutL α , importante para incisão da fita de DNA permitindo a excisão do fragmento mutado (DE BARROS et al., 2018). A variante p.Val717Met, c.2149G>A (rs201671325), associada com p=0.000421, possui conflitos de informações de patogenicidade no ClinVar sendo caracterizada como benigna, possivelmente benigna e VUS. No entanto, das ferramentas de predição de patogenicidade utilizadas, apenas o AlignGVGD não a classifica como possivelmente patogênica. Estudos funcionais demonstraram que a proteína *PMS2* mutada para esta variante apresentou-se estável e foi proficiente em MMR, mas foi identificada instabilidade de microssatélites no tecido tumoral, e portanto não foi descartada a possibilidade de

p.Val717Met constituir um alelo hipomórfico (GONZÁLEZ-ACOSTA et al., 2017). Já a variante p.Asp792Asn, c.2374G>A (rs587781265) ($p=2.2e^{-16}$), foi identificada em um paciente com câncer de estômago que veio a óbito. Esta variante já foi descrita nos tópicos acima, sendo relacionada a um comprometimento moderado do sistema de reparo MMR (ARORA et al., 2017).

Figura 29. Representação esquemática da proteína PMS2 e variantes encontradas no estudo e associadas ao risco para HBOC. As caixas brancas representam as duas variantes identificadas em *PMS2* e a localização de cada uma delas. As caixas em diferentes cores representam diferentes domínios da proteína: domínio histidina quinase em verde e laranja, o domínio transdutor que contém o domínio ribossomal S5, em tons de azul e lilás e, domínio MutL na região C-terminal em rosa, onde ocorre a dimerização com MLH1 e onde foram identificadas as duas variantes. Embaixo, escala que vai de 0 a 750 representando a localização das variantes e dos domínios da proteína que possui 862 aminoácidos.

Para melhor caracterizar as variantes identificadas neste gene, foi então realizada a análise de instabilidade de microssatélites (MSI) nas amostras que apresentaram estas variantes. Avaliar a estabilidade de marcadores de microssatélite se tornou uma importante ferramenta de avaliação do sistema de reparo MMR, uma vez que esta constitui a principal via de reparo de danos ao DNA em regiões de microssatélites (ETO et al., 2016). Como a deficiência em genes do MMR estão mais associadas a casos de Síndrome de Lynch, já existem protocolos de investigação de MSI estabelecidos para os cânceres colorretal e de endométrio, que se baseiam principalmente no padrão de amplificação de 5 marcadores mono ou dinucleotídicos em regiões de microssatélites (BUHARD et al., 2004; UMAR et al., 2004). No entanto, estudos já demonstraram que a avaliação de MSI com marcadores já caracterizados para os demais tipos tumorais também é informativa em casos de câncer de mama, embora a frequência de casos de instabilidade seja menor (CONDORELLI et al., 2019; IMYANITOV et al., 2000; SEITZ et al., 2003). Assim, para uma das amostras (3462) que apresentou a variante p.Val717Met, foi possível realizar a pesquisa de instabilidade de microssatélites utilizando DNA extraído de sangue periférico, em comparação ao perfil de amplificação de microssatélites em DNA extraído de tecido tumoral (**Figura 30**). A mesma análise foi possível para duas outras variantes que não foram associadas ao fenótipo: a variante p.Asp784Asn, c.2350G>A (rs143340522) no gene *PMS2*

que apresentou um valor de associação próximo de significativo ($p=0.05498$), e a variante p.Ala834Thr, c.2500G>A (rs63750757) no gene *MSH2* (ENST00000233146), com valor de $p=0.1277$. Para nenhuma das amostras foi observada MSI, o que sugere que estas variantes não estão associadas à instabilidade genômica nestes pacientes. A paciente (4264) que apresentou a variante p.Asp784Asn no gene *PMS2* que é descrita em casos de síndrome de Lynch (YURGELUN et al., 2015), foi uma daquelas que apresentou a variante patogênica p.Arg337His no gene *TP53* que já constitui a causa molecular de HBOC nesta paciente. As pacientes que apresentaram as outras duas variantes testadas também apresentaram variantes de significado incerto em outros genes como *BRCA1*, *FAM175A* e *UIMC1*, mas permanecem sem um diagnóstico molecular definido.

Figura 30. Análise de instabilidade de microssatélites com base na amplificação de 5 marcadores mononucleotídicos em DNA extraído de tecido tumoral pareado com DNA de sangue total. Cada cor representa um marcador. Os picos em cor laranja representam o marcador de tamanho dos picos utilizado. O paciente não apresenta instabilidade de microssatélites uma vez que o padrão de amplificação foi o mesmo dentre os tecidos analisados, para todos os marcadores utilizados.

Para ambas as variantes associadas, foi realizado um teste piloto para avaliar a expressão das proteínas mutadas e foi observado que as duas variantes parecem não interferir na estabilidade da proteína, uma vez que apresentaram padrão de expressão similar ao da proteína selvagem (**Figura 31**).

Figura 31. Western blot evidenciando a expressão das proteínas PMS2 selvagem (Wt) e mutada (V717M e D792N) em células HEK293T. siCTRL corresponde à expressão da proteína PMS2 endógena (100 kDa); VV= Vetor pcDNA3 vazio; Wt= *wild type* ou proteína selvagem; PMS2= expressão da proteína PMS2, utilizando anti-PMS2. V5 = expressão da proteína exógena transfectada, contendo a tag V5, utilizando anti-V5. Tubulina= proteína endógena (50-55 kDa), utilizando anti-alpha tubulina.

Além dos genes descritos acima, o gene *BRIP1* também apresentou variantes com dados conflitantes de patogenicidade mas que não foram associadas ao risco para HBOC. Das três variantes descritas neste gene, a variante p.Arg173Cys (rs4988345) apresentou frequência alélica maior no grupo HBOC do que na população controle, embora esta associação não tenha sido significativa ($p=0.07966$). Esta variante possui conflito de informações de patogenicidade no ClinVar, sendo caracterizada tanto como benigna quanto como VUS. Alguns estudos reportam esta variante como provavelmente benigna por relatarem a presença no grupo controle, bem como a ausência de associação com o risco para outros cânceres (KOTE-JARAI et al., 2009; LEWIS et al., 2005). Apesar disso, estudos relatam que variantes em *BRIP1* apresentam baixa ou nenhuma penetrância para câncer de mama (EASTON et al., 2016; SEAL et al., 2006). No entanto, esta variante está presente no domínio de localização nuclear da proteína BRIP1, região conservada e essencial para a correta localização da proteína durante o processo de reparo de DNA. De fato, Lei e Vorechovsky (2003), demonstraram que o sinal de células expressando a proteína mutada para p.Arg173Cys foi significativamente reduzido no núcleo quando comparado a proteína selvagem. Já a variante p.Gln740His (rs45589637), abordada nas seções anteriores, já foi descrita em casos de cânceres de mama e ovário, mas não possui nenhuma informação descrita sobre a sua funcionalidade (JALKH et al., 2017; LI et al., 2016), apesar de ser classificada como potencialmente patogênica por várias ferramentas de predição *in silico*. Pensando nisso, foram realizados experimentos de sensibilidade da proteína BRIP1 mutada para estas duas variantes, à mitomicina C (MMC). Esta droga é um quimioterápico responsável pela a formação de *cross-links* entre as fitas de DNA, provocando instabilidade genômica (PAZ et al., 2012) e a presença de variantes em *FANCF* ou *BRIP1*, proteína

responsável pelo reparo de lesões do tipo *cross-link* estão associadas à uma maior sensibilidade a MMC e à instabilidade genômica (LITMAN et al., 2005; ZOU et al., 2013). No ensaio piloto, a proteína mutada para a variante p.Arg173Cys que apresentou alta frequência no grupo HBOC quando comparada às populações controle, foi observada uma sensibilidade ao tratamento com mitomicina C similar ao apresentado pelas células transfectadas com a proteína selvagem. Já a proteína mutada para a variante p.Gln740His, que foi classificada como possivelmente patogênica, foi discretamente mais sensível ao tratamento com MMC, apresentando uma curva mais próxima àquela das células deficientes em BRIP1 (**Figura 32**). Apesar da alta variância, estes dados sugerem que a variante p.Gln740His possa estar relacionada ao comprometimento da função de BRIP1, no entanto, testes mais elaborados são mandatórios para a confirmação deste fenótipo.

Figura 32. Curva de sobrevivência a diferentes concentrações de Mitomicina C de células HEK293T após silenciamento da proteína BRIP1 endógena. Células HEK293T foram transfectadas com siRNA para BRIP1 endógena e transfectadas com constructos de vetor pcDNA3 vazio (VV), BRIP1 *wild type* (Wt - selvagem) e BRIP1 mutada (p.R173C e p.Q740H). As células transfectadas com siRNA controle e com o vetor pcDNA3 vazio (VV) foram utilizadas como controles.

Diante destes dados, ressaltamos que estudos funcionais adicionais, bem como os estudos de segregação destas variantes com as famílias afetadas, são necessários para a caracterização desta e demais variantes de significado clínico duvidoso.

5.4 ASSOCIAÇÃO GENÓTIPO VERSUS FENÓTIPO

As Tabelas A4 e A5, no Apêndice deste documento apresentam a compilação das características clínicas e moleculares de todos os pacientes estudados. O grupo de pacientes com variantes patogênicas nos genes *BRCA1* e *BRCA2* apresentaram uma idade mediana de 34 anos no momento do diagnóstico (idade 3 anos mais jovem que a dos pacientes que não apresentaram variantes patogênicas nestes genes), e todos constituíam casos de câncer de mama com graus tumorais mais elevados e cerca de 52,9% eram casos de câncer de mama triplo-negativo (**Tabela 5**).

Como descrito nas seções anteriores, a maioria dos casos estudados foram casos de câncer de mama, com apenas 6 pacientes apresentando câncer de ovário (**Tabela 5**) (sendo um caso que apresentou câncer de mama e de ovário). Nenhum destes casos apresentou variantes patogênicas nos genes *BRCA1* e *BRCA2*, no entanto, três foram casos que vieram óbito em decorrência de metástases à distância. Dos três pacientes que vieram a óbito dois apresentaram o haplótipo 1. Foi então realizada a análise de associação da presença do haplótipo 1 com acometimento por metástase, grau tumoral, óbito e subtipo tumoral e também não foi identificada nenhuma associação significativa.

Foi observada também uma prevalência geral de casos apresentando tumores de graus mais agressivos (III e II) ($p=0.03686$) (**Tabela 5**). No entanto, para o subgrupo de pacientes com variantes patogênicas nos genes *BRCA1/BRCA2*, foi observado que cerca de 90% dos pacientes com variantes patogênicas em *BRCA1* apresentaram tumores de grau III, enquanto cerca de 80% daqueles pacientes apresentando variantes patogênicas no gene *BRCA2* apresentavam tumores de graus I e II (**Tabela 5**). Apesar do grande número de casos com falta de informação (NI) nesta análise, já é descrito na literatura que variantes no gene *BRCA1* estão associados a cânceres mais agressivos, geralmente de tipo triplo-negativo, apresentam graus tumorais mais elevados e um pior prognóstico quando comparadas ao *BRCA2* (SPURDLE et al., 2014; ZHONG et al., 2015). De fato, dos 9 pacientes com variantes patogênicas em *BRCA1/BRCA2* do tipo triplo negativo, 7 apresentaram variantes em *BRCA1*. Curiosamente, aqueles pacientes que não apresentaram variantes ou alterações no número de cópias nos genes *BRCA1/BRCA2* também apresentaram uma maioria de casos de tumores de graus II e III (**Tabela 5**), o que sugere que a presença das variantes identificadas nos outros genes, também podem estar contribuindo para o risco de HBOC.

A presença de metástase foi fortemente correlacionada com a ocorrência de óbitos ($p=7.85e^{-12}$) com 13 dos 14 pacientes vindo a óbito após metástase à distância. No entanto, não foi observada diferença significativa na sobrevida de pacientes em relação ao genótipo (**Figura 33**). Curiosamente,

o número de óbitos foi maior para o grupo de pacientes que apresentaram variantes benignas ou de significado incerto nos genes *BRCA1/BRCA2*, grupo com maior número de indivíduos. Apenas um caso de óbito apresentou variante patogênica no gene *BRCA1* e dois casos foram de pacientes não-BRCA (**Tabela 5**). Aqueles pacientes portadores de variantes benignas ou de significado incerto também apresentaram uma menor sobrevida em relação ao paciente que era portador da variante patogênica no gene *BRCA1* (**Figura 33**). De fato, o efeito cumulativo de alelos de baixa penetrância nos genes *BRCA1/BRCA2*, assim como variantes raras em genes de menor penetrância seguindo um modelo poligênico pode estar relacionado ao aumento de risco para a doença (ECONOMOPOULOU; DIMITRIADIS; PSYRRI, 2015; GRACIA-AZNAREZ et al., 2013; SAWYER et al., 2012; SMYTH et al., 2015), o que reforça a necessidade de uma melhor caracterização destas variantes.

Figura 33. Curva de sobrevivência dos pacientes de acordo com o genótipo pelo tempo de vida desde o diagnóstico em anos. Cada traço cinza na vertical indica os pacientes censurados, ou seja, aqueles que ainda não vieram a óbito e o tempo marcado por cada traço representa a sobrevivência destes pacientes em anos. Há pacientes acompanhados a mais tempo e que o tempo de diagnóstico da doença desde o diagnóstico até 2019 é de 28 anos, assim como há pacientes que iniciaram o acompanhamento clínico em 2016 e o tempo de doença é de apenas 4 anos. Não foram encontradas diferenças estatísticas de sobrevida entre os genótipos.

Tabela 5. Caracterização fenotípica de acordo com o genótipo dos pacientes estudados.

Variável	Genótipo						p-valor&
	Patogênica BRCA*		Benignas e VUS BRCA ‡		não-BRCA		
	n=17	%	n=65	%	n=12	%	
Gênero							
Homem			1	1,5			
Mulher	17	18,1	64	98,5	12	100	
Idade ao diagnóstico (mediana)							
	24 - 57 (34)		22 - 72 (37)		31 - 47 (36,5)		
Óbito							
	1	5,9	11	16,9	2	16,6	0.0927
Sobrevivência em anos (mediana)		8		3		8	
História familiar							
Presente	14	82,3	52	80	10	83,3	0.294
Ausente	3	17,7	16,9	16,9	2	16,7	
NI			3,2	3,1			
Localização do tumor							
Mama	17	100,0	57	87,7	12	100	0.6034
Ovário			6	9,3			
Endométrio			1	1,5			
Estômago			1	1,5			
Distribuição tumoral							
Unilateral ou localizado	12	70,6	48	73,8	10	83,3	0.2376
Bilateral (mama)	5	29,4	6	9,3	1	8,3	
Tumores multicêntricos			5	7,7			
NI			6	9,3	1	8,3	
Subtipo (Câncer de mama)							
Luminal	4	23,5	20	30,8	5	41,7	0.4425
Luminal HER	2	11,8	11	16,9	3	25	
HER2	2	11,8	7	10,8	1	8,3	
TN	9	52,9	13	20	1	8,3	
RP			1	1,5			
NI			13	20	2	16,7	
Grau do tumor							
I	1	5,9	7	10,8	1	8,3	0.03686
II	3	17,6	29	44,6	5	41,7	
III	11	64,7	11	16,9	4	33,3	
NI	2	11,8	18	27,7	2	16,7	

Continua.

Tabela 5. Caracterização fenotípica de acordo com o genótipo dos pacientes estudados.
 Continuação.

Variável	Genótipo						p-valor ^{&}
	Patogênica BRCA*		Benignas e VUS BRCA [‡]		não-BRCA		
	n=17	%	n=65	%	n=12	%	
Linfonodo							0.1984
Presente	7	41,2	31	47,7	7	58,3	
Ausente	8	47,1	16	24,6	3	25	
NI	2	11,8	18	27,7	2	16,7	
Metástase à distância							0.1964
Presente	1	5,9	38	58,5	7	58,3	
Ausente	15	88,2	14	21,5	3	25	
NI	1	5,9	13	20	2	16,7	

^{*}Variantes previamente caracterizadas na literatura como patogênicas (ClinVar). [‡]Pacientes com variantes benignas ou de significado incerto nos genes *BRCA1/2* e demais genes. NI corresponde a uma variável não-informada. ^{*}As informações estão ranqueadas de acordo com a idade ao diagnóstico do tumor. [§]Valor de probabilidade em relação ao status mutacional de *BRCA1/BRCA2* calculado utilizando-se o Teste do Chi-quadrado. NI = não informado; HER2 = tumores apresentando superexpressão da proteína HER2; TN = Triplo negativo; RP = Receptor de progesterona.

6 CONCLUSÕES

Este constitui o terceiro estudo caracterizando a população brasileira com HBOC e abordando a análise de painel de genes de reparo em pacientes com suspeita clínica de HBOC. Estudos como este são de extrema importância epidemiológica, por caracterizar e acrescentar à literatura dados de variantes em genes de risco alto a moderado nesta população. Para muitas das variantes identificadas, este estudo constitui o primeiro relato na população brasileira, trazendo também dados inéditos da associação de variantes ao aumento de risco para HBOC. No entanto, muitas das variantes identificadas continuam sem um significado clínico definido, mas a descrição destas variantes de significado incerto concentra sua importância para estudos futuros que utilizarão estes dados para caracterizá-las quanto ao risco para HBOC. Estudos de segregação das variantes candidatas a causais com a doença na família são de suma importância para validação dos dados encontrados nos estudos funcionais, bem como no estudo de associação e são uma perspectiva futura deste projeto.

Na série de casos estudada, cerca de 23% dos pacientes tiveram o diagnóstico molecular elucidado. Apesar de este ser um trabalho de pesquisa, os resultados provenientes da análise do painel foram retornados aos pacientes do Ambulatório de Genética do Câncer do Hospital das Clínicas – FMRP/USP, permitindo-lhes o estudo de risco mais acurado; a pesquisa das variantes bem como a assistência e aconselhamento genético para as famílias. Para algumas pacientes, o diagnóstico molecular possibilitou a realização de cirurgias profiláticas, reduzindo o risco de recorrência em até 90%. Destes pacientes, cerca de 17% apresentaram variantes de alta penetrância nos genes *BRCA1* e *BRCA2*, frequência maior que a apontada pelo estudo de Timoteo e colaboradores (2018), na população brasileira. Outros 5% apresentaram variantes patogênicas em *TP53*. Cerca de 18% dos pacientes também apresentaram variantes candidatas a causais, uma frequência maior que a descrita por Timoteo et al (2018); sendo destes 2,1% com variantes possivelmente patogênicas em *PMS2* e *BRIP1* e, 15,9% apresentando variantes significativamente associadas ao risco para HBOC em genes já associados a um risco alto a moderado para câncer de mama e ovário (*BARD1*, *BRCA2*, *CDH1*, *CHEK2*, *FAM175A*, *MLH1*, *PALB2* e *PMS2*). Estudos funcionais sugeriram variantes nos genes *CHEK2*, *BRIP1*, *FAM175A* e *UIMC1* como associadas ao comprometimento da função da proteína e demais estudos são necessários para melhor caracterizar estas variantes na patogênese do HBOC.

Estudos como este evidenciam como outros genes envolvidos em vias de reparo do DNA podem estar contribuindo para a patogênese do HBOC e como a caracterização destas variantes é

imprescindível para estimar estes riscos e assim possibilitar o acompanhamento clínico e aconselhamento genético adequados para pacientes diagnosticados clinicamente com HBOC, que em sua maioria, permanecem sem um diagnóstico clínico elucidado.

Apesar do extenso trabalho para a caracterização das variantes identificadas nestes pacientes, muito ainda pode ser elucidado nestes casos, como a busca por alterações no número de cópias para estes genes, bem como a análise de regiões não codificantes que também podem estar relacionadas ao aumento de risco para os cânceres de mama e ovário. Como perspectivas futuras, testes funcionais adicionais e estudos de segregação serão aplicados a fim de melhor caracterizar as variantes associadas ao fenótipo neste estudo.

REFERÊNCIAS

- ACOG practice bulletin No. 103: Hereditary breast and ovarian cancer syndrome. Obstetrics and Gynecology**, 2009.
- ADAMOVICH, A. I. et al. Functional analysis of BARD1 missense variants in homology-directed repair and damage sensitivity. **PLOS Genetics**, v. 15, n. 3, p. 1–21, 2019.
- AFGHAHI, A.; TELLI, M. L.; KURIAN, A. W. Genetics of triple-negative breast cancer: Implications for patient care. **Current Problems in Cancer**, v. 40, n. 2–4, p. 130–140, 2016.
- AL BAKIR, M.; GABRA, H. The molecular genetics of hereditary and sporadic ovarian cancer: implications for the future. **British medical bulletin**, v. 112, n. 1, p. 57–69, 2014.
- ALORAIFI, F. et al. Gene analysis techniques and susceptibility gene discovery in non-BRCA1/BRCA2 familial breast cancer. **Surgical Oncology**, v. 24, n. 2, p. 100–109, 2015.
- ARORA, S. et al. Functional analysis of rare variants in mismatch repair proteins augments results from computation-based predictive methods. **Cancer Biology & Therapy**, v. 18, n. 7, p. 519–533, 3 jul. 2017.
- ARPINO, G. et al. Tumor characteristics and prognosis in familial breast cancer. **BMC Cancer**, 2016.
- BARZAN, D. et al. Comparison of genetic variation of breast cancer susceptibility genes in Chinese and German populations. **European journal of human genetics : EJHG**, v. 21, n. 11, p. 1286–92, nov. 2013.
- BELL, D. W. et al. Genetic and functional analysis of CHEK2 (CHK2) variants in multiethnic cohorts. **International Journal of Cancer**, 2007.
- BLOWS, F. M. et al. Subtyping of Breast Cancer by Immunohistochemistry to Investigate a Relationship between Subtype and Short and Long Term Survival : A Collaborative Analysis of Data for 10 , 159 Cases from 12 Studies. v. 7, n. 5, 2010.
- BODIAN, D. L. et al. Germline Variation in Cancer-Susceptibility Genes in a Healthy, Ancestrally Diverse Cohort: Implications for Individual Genome Sequencing. **PLoS ONE**, v. 9, n. 4, p. e94554, 11 abr. 2014.
- BONACHE, S. et al. Multigene panel testing beyond BRCA1/2 in breast/ovarian cancer Spanish families and clinical actionability of findings. **Journal of Cancer Research and Clinical Oncology**, v. 144, n. 12, p. 2495–2513, 10 dez. 2018.
- BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, 1976.
- BRANDALIZE, A. P. C. et al. A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. **BMC cancer**, v. 14, p. 850, 2014.
- BUHARD, O. et al. Quasimonomorphic Mononucleotide Repeats for High-Level Microsatellite Instability Analysis. **Disease Markers**, v. 20, n. 4–5, p. 251–257, 2004.
- BUISSON, R. et al. Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination. **Nature structural & molecular biology**, v. 17, n. 10, p. 1247–54, 2010.
- BUISSON, R. et al. Coupling of Homologous Recombination and the Checkpoint by ATR. **Molecular Cell**, v. 65, n. 2, p. 336–346, 2017.
- BURGESS, B. T.; KOLESAR, J. M. Defective DNA repair in hereditary ovarian cancers: Implications for therapy. **American Journal of Health-System Pharmacy**, v. 75, n. 21, p. 1697–1707, 1 nov. 2018.
- BUYS, S. S. et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. **Cancer**, p. 1–10, 2017.
- CANTOR, S. B. et al. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. **Cell**, 2001.
- CASTÉRA, L. et al. Next-generation sequencing for the diagnosis of hereditary breast and ovarian

cancer using genomic capture targeting multiple candidate genes. **European journal of human genetics : EJHG**, n. October 2013, p. 1305–1313, 19 fev. 2014.

CASTILLO, A. et al. The BRCA1-interacting protein Abraxas is required for genomic stability and tumor suppression. **Cell Reports**, 2014.

CHEANG, M. C. U. et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. **Journal of the National Cancer Institute**, v. 101, n. 10, p. 736–750, 2009.

CIMMINO, F.; FORMICOLA, D.; CAPASSO, M. Dualistic Role of BARD1 in Cancer. **Genes**, v. 8, n. 12, p. 375, 8 dez. 2017.

CIPRIANO, N. M. et al. Mutation screening of TP53, CHEK2 and BRCA genes in patients at high risk for hereditary breast and ovarian cancer (HBOC) in Brazil. **Breast Cancer**, v. 26, n. 3, p. 397–405, 2019.

COCK-RADA, A. M. et al. A multi-gene panel study in hereditary breast and ovarian cancer in Colombia. **Familial Cancer**, 2018.

CONDORELLI, R. et al. Genomic alterations in breast cancer: level of evidence for actionability according to ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). **Annals of Oncology**, v. 30, n. 3, p. 365–373, 1 mar. 2019.

COUCH, F. J.; NATHANSON, K. L.; OFFIT, K. Two Decades After BRCA: Setting Paradigms in Personalized Cancer Care and Prevention. **Science**, v. 343, n. 6178, p. 1466–1470, 28 mar. 2014.

CRAIG, A. L.; HUPP, T. R. The regulation of CHK2 in human cancer. **Oncogene**, v. 23, n. 52, p. 8411–8418, 2004.

CURY, N. M.; FERRAZ, V. E. F.; SILVA, W. A. TP53 p.R337H prevalence in a series of Brazilian hereditary breast cancer families. **Hereditary Cancer in Clinical Practice**, v. 12, n. 1, p. 8, 13 dez. 2014.

DA COSTA, E. C. B. et al. Founder effect of the BRCA1 5382insC mutation in Brazilian patients with hereditary breast ovary cancer syndrome. **Cancer Genetics and Cytogenetics**, 2008.

DAI, X. et al. Cancer hallmarks, biomarkers and breast cancer molecular subtypes. **Journal of Cancer**, v. 7, n. 10, p. 1281–1294, 2016.

DAINO, K. et al. Loss of the BRCA1-Interacting Helicase BRIP1 Results in Abnormal Mammary Acinar Morphogenesis. **PLoS ONE**, v. 8, n. 9, p. e74013, 6 set. 2013.

DAROOEI, M. et al. Pedigree and BRCA gene analysis in breast cancer patients to identify hereditary breast and ovarian cancer syndrome to prevent morbidity and mortality of disease in Indian population. **Tumor Biology**, v. 39, n. 2, p. 101042831769430, 2017.

DAVARINEJAD, H. Quantifications of Western Blots with ImageJ. **University of York**, 2015.

DE BARROS, A. C. et al. DNA mismatch repair proteins MLH1 and PMS2 can be imported to the nucleus by a classical nuclear import pathway. **Biochimie**, 2018.

DESRICHARD, A. et al. CHEK2 contribution to hereditary breast cancer in non-BRCA families. **Breast Cancer Research**, 2011.

DORAIRAJ, J. J. et al. A germline mutation in the BRCA1 3'UTR predicts Stage IV breast cancer. **BMC Cancer**, 2014.

DUCY, M. et al. The Tumor Suppressor PALB2: Inside Out. **Trends in Biochemical Sciences**, v. 44, n. 3, p. 226–240, 2019.

DUFAULT, M. R. et al. Limited relevance of the CHEK2 gene in hereditary breast cancer. **International Journal of Cancer**, 2004.

DUFOUR, A. et al. Inactivation of TP53 correlates with disease progression and low miR-34a expression in previously treated chronic lymphocytic leukemia patients. **Blood**, v. 121, n. 18, p. 3650–3657, 2 maio 2013.

DUROCHER, F. et al. Mutation analysis of the BRCA1 gene in 23 families with cases of cancer of the breast. p. 814–819, 1996.

EASTON, D. F. et al. No evidence that protein truncating variants in BRIP1 are associated with breast cancer risk: Implications for gene panel testing. **Journal of Medical Genetics**, 2016.

ECONOMOPOULOU, P.; DIMITRIADIS, G.; PSYRRI, A. Beyond BRCA: New hereditary breast cancer susceptibility genes. **Cancer Treat Rev**, v. 41, n. 1, p. 1–8, 2015.

ESPENSCHIED, C. R. et al. Multigene Panel Testing Provides a New Perspective on Lynch Syndrome. **J Clin Oncol**, v. 35, n. 22, p. 2568–2575, 2017.

ETO, T. et al. Modal variety of microsatellite instability in human endometrial carcinomas. **Journal of Cancer Research and Clinical Oncology**, 2016.

EVANS, D. G. R. et al. A Dominantly Inherited 5' UTR Variant Causing Methylation-Associated Silencing of BRCA1 as a Cause of Breast and Ovarian Cancer. **American Journal of Human Genetics**, v. 103, n. 2, p. 213–220, 2018.

EWALD, I. P. et al. BRCA1 and BRCA2 rearrangements in Brazilian individuals with hereditary breast and ovarian cancer syndrome. **Genetics and Molecular Biology**, v. 39, n. 2, p. 223–231, 2016.

FARMER, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. v. 239, n. 1991, p. 236–239, 2005.

FEKI, A. et al. BARD1 induces apoptosis by catalysing phosphorylation of p53 by DNA-damage response kinase. **Oncogene**, 2005.

FELICIO, P. S. et al. Genetic alterations detected by comparative genomic hybridization in BRCAX breast and ovarian cancers of Brazilian population. **Oncotarget**, v. 9, n. 44, 8 jun. 2018a.

FELICIO, P. S. et al. Screening and characterization of BRCA2 c.156_157insAlu in Brazil: Results from 1380 individuals from the South and Southeast. **Cancer Genetics**, v. 228–229, p. 93–97, dez. 2018b.

FELIX, G. E. et al. Germline mutations in BRCA1, BRCA2, CHEK2 and TP53 in patients at high-risk for HBOC: characterizing a Northeast Brazilian Population. **Human genome variation**, v. 1, n. May, p. 14012, 2014.

FERLAY, J. et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. **Int J Cancer**, v. 136, n. 5, p. E359-86, 2014.

FERLAY, J. et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. **International journal of cancer**, 2015.

FERNANDES, G. C. et al. Prevalence of *BRCA1/BRCA2* mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. **Oncotarget**, v. 7, n. 49, p. 80465–80481, 2016.

FOO, T. K. et al. Compromised BRCA1-PALB2 interaction is associated with breast cancer risk. **Oncogene**, n. January, p. 1–10, 2017.

FRIEDHOFF, P.; LI, P.; GOTTHARDT, J. **Protein-protein interactions in DNA mismatch repairDNA Repair**, 2016.

GERMANO, G. et al. The clinical impact of the genomic landscape of mismatch repair-deficient cancers. **Cancer Discovery**, v. 8, n. 12, p. 1518–1528, 2018.

GILMORE, B. L. et al. Molecular Analysis of BRCA1 in Human Breast Cancer Cells under Oxidative Stress. **Scientific Reports**, 2017.

GOBBINI, E. et al. The MRX complex regulates Exo1 resection activity by altering DNA end structure. **The EMBO Journal**, 2018.

GONZÁLEZ-ACOSTA, M. et al. Elucidating the clinical significance of two PMS2 missense variants coexisting in a family fulfilling hereditary cancer criteria. **Familial Cancer**, 2017.

GRACIA-AZNAREZ, F. J. et al. Whole Exome Sequencing Suggests Much of Non-BRCA1/BRCA2 Familial Breast Cancer Is Due to Moderate and Low Penetrance Susceptibility Alleles. **PLoS ONE**, v. 8, n. 2, 2013.

GU, Z.; EILS, R.; SCHLESNER, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. **Bioinformatics**, v. 32, n. 18, p. 2847–2849, 2016.

GUNDERSON, C. C.; MOORE, K. N. Olaparib: An oral PARP-1 and PARP-2 inhibitor with promising activity in ovarian cancer. **Future Oncology**, 2015.

GUO, Y. et al. Three-stage quality control strategies for DNA re-sequencing data. **Briefings in Bioinformatics**, v. 15, n. 6, p. 879–889, 2013.

HAHN, E. C. et al. TP53 p.Arg337His germline mutation prevalence in Southern Brazil: Further evidence for mutation testing in young breast cancer patients. **PLoS ONE**, 2018.

HAMDI, Y. et al. Association of breast cancer risk with genetic variants showing differential allelic expression: Identification of a novel breast cancer susceptibility locus at 4q21. **Oncotarget**; Vol 7, No 49, v. 7, n. 49, p. 80140–80163, 2016.

HAMEL, N. et al. On the origin and diffusion of BRCA1 c.5266dupC (5382insC) in European populations. **European Journal of Human Genetics**, 2011.

HER, J. et al. Factors forming the BRCA1-A complex orchestrate BRCA1 recruitment to the sites of DNA damage. **Acta Biochimica et Biophysica Sinica**, v. 48, n. 7, p. 658–664, jul. 2016.

HIROTSU, Y. et al. Multigene panel analysis identified germline mutations of DNA repair genes in breast and ovarian cancer. **Molecular Genetics & Genomic Medicine**, v. 3, n. 5, p. 459–466, 2015.

HONDOW, H. L. et al. A high-throughput protocol for mutation scanning of the BRCA1 and BRCA2 genes. **BMC Cancer**, v. 11, n. 1, p. 265, 24 dez. 2011.

IMYANITOV, E. N. et al. Evidence for microsatellite instability in bilateral breast carcinomas. **Cancer Letters**, 2000.

INCA, I. N. DE C. J. A. G. DA S. **Estimativa 2018. Incidência de câncer no Brasil**. [s.l: s.n.].

JALKH, N. et al. Next-generation sequencing in familial breast cancer patients from Lebanon. **BMC Medical Genomics**, 2017.

JEGGO, P. A.; PEARL, L. H.; CARR, A. M. **DNA repair, genome stability and cancer: A historical perspective** *Nature Reviews Cancer*, 2016.

JIN, B. et al. Are polymorphisms of the ataxia telangiectasia mutated gene associated with breast cancer risk? **Breast Cancer Research and Treatment**, v. 128, n. 1, p. 293–295, 16 jul. 2011.

JIRICNY, J. The multifaceted mismatch-repair system. **Nature Reviews Molecular Cell Biology**, v. 7, n. 5, p. 335–346, 12 maio 2006.

KAPOOR, N. S. et al. Multigene Panel Testing Detects Equal Rates of Pathogenic BRCA1/2 Mutations and has a Higher Diagnostic Yield Compared to Limited BRCA1/2 Analysis Alone in Patients at Risk for Hereditary Breast Cancer. **Annals of surgical oncology**, 2015.

KEENEY, M. G. et al. Non- BRCA familial breast cancer: review of reported pathology and molecular findings. **Pathology**, p. 1–8, 2017.

KIM, H.; CHEN, J.; YU, X. Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. **Science**, 2007.

KNUDSON, A. G. Mutation and cancer: statistical study of retinoblastoma. **Proceedings of the National Academy of Sciences of the United States of America**, v. 68, n. 4, p. 820–3, abr. 1971.

KOBAYASHI, H. et al. Hereditary breast and ovarian cancer susceptibility genes (Review). **Oncology Reports**, v. 30, n. 3, p. 1019–1029, 2013.

KOTE-JARAI, Z. et al. A recurrent truncating germline mutation in the BRIP1/FANCI gene and susceptibility to prostate cancer. **British Journal of Cancer**, v. 100, n. 2, p. 426–430, 2009.

KOTSOUPOULOS, J. BRCA Mutations and Breast Cancer Prevention. **Cancers**, 2018.

KRAJEWSKA, M. et al. Regulators of homologous recombination repair as novel targets for cancer treatment. **Frontiers in genetics**, v. 6, n. March, p. 96, 2015.

KUMARASWAMY, E.; SHIEKHATTAR, R. Activation of BRCA1/BRCA2-Associated Helicase BACH1 Is Required for Timely Progression through S Phase. **Molecular and Cellular Biology**, v. 27, n. 19, p. 6733–6741, 2007.

KUNKEL, T. A.; ERIE, D. A. Eukaryotic Mismatch Repair in Relation to DNA Replication. **Annual Review of Genetics**, 2015.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, 1970.

LAJUS, T. B. P.; SALES, R. M. D. CDH1 germ-line missense mutation identified by multigene

sequencing in a family with no history of diffuse gastric cancer. **Gene**, v. 568, n. 2, p. 215–219, set. 2015.

LAKHANI, S. R. et al. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. **Journal of the National Cancer Institute**, 1998.

LALLOO, F.; EVANS, D. G. Familial Breast Cancer. **Clinical Genetics**, v. 82, n. 2, p. 105–114, 2012.

LEE, J. Y. et al. BRCA1/2-negative, high-risk breast cancers (BRCAx) for Asian women: genetic susceptibility loci and their potential impacts. **Scientific Reports**, v. 8, n. 1, p. 1–11, 2018.

LEI, H.; VORECHOVSKY, I. BACH1 517C>T transition impairs protein translocation to nucleus: A role in breast cancer susceptibility? **International Journal of Cancer**, v. 104, n. 3, p. 389–391, 10 abr. 2003.

LEWIS, A. G. et al. Mutation analysis of FANCD2, BRIP1/BACH1, LMO4 and SFN in familial breast cancer. **Breast cancer research : BCR**, v. 7, n. 6, p. R1005–R1016, 2005.

LI, G.-M. Mechanisms and functions of DNA mismatch repair. **Cell Research**, v. 18, n. 1, p. 85–98, 2008.

LI, J. et al. Targeted massively parallel sequencing of a panel of putative breast cancer susceptibility genes in a large cohort of multiple-case breast and ovarian cancer families. **Journal of Medical Genetics**, 2016.

LI, Y. et al. USP13 regulates the RAP80-BRCA1 complex dependent DNA damage response. **Nature Communications**, 2017.

LINDOR, N. M. et al. Concise Handbook of Familial Cancer Susceptibility Syndromes - Second Edition. **JNCI Monographs**, v. 2008, n. 38, p. 3–93, 2008.

LITMAN, R. et al. BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCD1. **Cancer Cell**, v. 8, n. 3, p. 255–265, 2005.

LYNCH, H. T.; SNYDER, C.; CASEY, M. J. Hereditary ovarian and breast cancer: what have we learned? **Annals of oncology : official journal of the European Society for Medical Oncology / ESMO**, v. 24 Suppl 8, n. Supplement 8, p. viii83–viii95, nov. 2013.

MACEDO, G. S.; ALEMAR, B.; ASHTON-PROLLA, P. Reviewing the characteristics of BRCA and PALB2-related cancers in the precision medicine era. **Genetics and Molecular Biology**, 2019.

MAGNI, M. et al. CCAR2/DBC1 is required for Chk2-dependent KAP1 phosphorylation and repair of DNA damage. **Oncotarget**, v. 6, n. 19, p. 17817–17831, 2015.

MAIER, C. et al. Subgroups of familial and aggressive prostate cancer with considerable frequencies of BRCA2 mutations. **Prostate**, 2014.

MAISTRO, S. et al. Germline mutations in BRCA1 and BRCA2 in epithelial ovarian cancer patients in Brazil. **BMC Cancer**, v. 16, n. 1, p. 934, 3 dez. 2016.

MALHOTRA, G. K. et al. Histological, molecular and functional subtypes of breast cancers. **Cancer Biology and Therapy**, v. 10, n. 10, p. 955–960, 2010.

MATEUS, A. R. et al. E-cadherin mutations and cell motility: a genotype-phenotype correlation. **Experimental cell research**, v. 315, n. 8, p. 1393–402, 2009.

MAURO, C. DE C. et al. Prevalence of BRCA1 and BRCA2 pathogenic and likely pathogenic variants in non-selected ovarian carcinoma patients in Brazil. **BMC Cancer**, v. 19, n. 1, p. 1–9, 2019.

MEANEY-DELMAN, D.; BELLCROSS, C. A. **Hereditary breast/ovarian cancer syndrome. A primer for obstetricians/gynecologists** *Obstetrics and Gynecology Clinics of North America*, 2013.

MEHTA, A.; HABER, J. E. Sources of DNA Double-Strand Breaks and Models of Rec. **Cold Spring Harbor Perspectives in Biology**, v. 6, p. 1–19, 2014.

MIKI, Y. et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. **Science (New York, N.Y.)**, v. 266, n. 5182, p. 66–71, 1994.

MILLOT, G. A. et al. Assessment of human nter and cter BRCA1 mutations using growth and

localization assays in yeast. **Human Mutation**, v. 32, n. 12, p. 1470–1480, 2011.

MINION, L. E. et al. Hereditary predisposition to ovarian cancer, looking beyond BRCA1/BRCA2. **Gynecologic Oncology**, v. 137, n. 1, p. 86–92, 2015.

MOHAMAD, S. et al. Low prevalence of CHEK2 gene mutations in multiethnic cohorts of breast cancer patients in Malaysia. **PLoS ONE**, 2015.

MONTEIRO, A. N. A.; AUGUST, A.; HANAFUSA, H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. **Proceedings of the National Academy of Sciences**, 2002.

NANDIKOLLA, A. G.; VENUGOPAL, S.; ANAMPA, J. Breast Cancer - Targets and Therapy. **Breast Cancer - Targets and Therapy**, v. 9, p. 207–215, 2017.

NEPOMUCENO, T. C. et al. The role of PALB2 in the DNA damage response and cancer predisposition. **International Journal of Molecular Sciences**, v. 18, n. 9, p. 1–20, 2017.

NEUMANN, E. et al. Gene transfer into mouse lyoma cells by electroporation in high electric fields. **The EMBO journal**, 1982.

NIELSEN, F. C.; VAN OVEREEM HANSEN, T.; SØRENSEN, C. S. Hereditary breast and ovarian cancer: new genes in confined pathways. **Nature Reviews Cancer**, v. 16, n. 9, p. 599–612, 2016.

NIK-ZAINAL, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. **Nature**, p. 1–20, 2016.

OUHTIT, A.; GUPTA, I.; SHAIKH, Z. BRIP1, a potential candidate gene in development of non-BRCA1/2 breast cancer. **Frontiers in bioscience (Elite edition)**, v. 8, n. 1, p. 289–98, 2016.

PALMERO, E. I. et al. Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. **Cancer Letters**, 2008.

PALMERO, E. I. et al. Screening for germline BRCA1, BRCA2, TP53 and CHEK2 mutations in families at-risk for hereditary breast cancer identified in a population-based study from Southern Brazil. **Genetics and Molecular Biology**, v. 39, n. 2, p. 210–222, 2016.

PALMERO, E. I. et al. The germline mutational landscape of BRCA1 and BRCA2 in Brazil. **Scientific Reports**, v. 8, n. 1, p. 9188, 15 dez. 2018.

PARK, J.-H. et al. Mouse Homolog of the Human TP53 R337H Mutation Reveals Its Role in Tumorigenesis. **Cancer Research**, v. 78, n. 18, p. 5375–5383, 15 set. 2018.

PARK, J.-H.; WANG, P.-Y.; HWANG, P. M. Modeling the prevalent germline *TP53* R337H mutation in mouse. **Oncotarget**, 2019.

PAUTY, J. et al. Exploring the roles of PALB2 at the crossroads of DNA repair and cancer. **The Biochemical journal**, v. 460, n. 3, p. 331–42, 2014.

PAUTY, J. et al. Cancer-causing mutations in the tumor suppressor PALB2 reveal a novel cancer mechanism using a hidden nuclear export signal in the WD40 repeat motif. **Nucleic Acids Research**, v. 45, n. 5, p. 2644–2657, 2017.

PAZ, M. M. et al. A new mechanism of action for the anticancer drug Mitomycin C: Mechanism-based inhibition of thioredoxin reductase. **Chemical Research in Toxicology**, v. 25, n. 7, p. 1502–1511, 2012.

PEIXOTO, A. et al. The c.156-157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. **Breast Cancer Research and Treatment**, v. 114, n. 1, p. 31–38, 2009.

PELLETIER, C. et al. Rare BRCA1 haplotypes including 3'UTR SNPs associated with breast cancer risk. **Cell Cycle**, 2011.

PENA, S. D. J. et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. **PLoS ONE**, 2011.

PETRONI, M. et al. MRE11 inhibition highlights a replication stress-dependent vulnerability of MYCN-driven tumors. **Cell Death and Disease**, 2018.

PLEVRITIS, S. K. et al. Cost-effectiveness of Screening BRCA1/2 Mutation Carriers With Breast Magnetic Resonance Imaging Sylvania. **JAMA**, v. 295, p. 2374–2384, 2006.

PODHORECKA, M.; SKLADANOWSKI, A.; BOZKO, P. H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy. **Journal of Nucleic Acids**, v. 2010, p. 1–9, 2010.

POUMPOURIDOU, N.; KROUPIS, C. Hereditary breast cancer: Beyond BRCA genetic analysis; PALB2 emerges. **Clinical Chemistry and Laboratory Medicine**, v. 50, n. 3, p. 423–434, 2012.

PRAKASH, R. et al. Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. **Cold Spring Harbor Perspectives in Biology**, v. 7, n. 4, p. a016600, 1 abr. 2015.

QIAN, Y. et al. Identification of pathogenic retrotransposon insertions in cancer predisposition genes. **Cancer Genetics**, 2017.

QUEZADA URBAN, R. et al. Comprehensive Analysis of Germline Variants in Mexican Patients with Hereditary Breast and Ovarian Cancer Susceptibility. **Cancers**, v. 10, n. 10, p. 361, 27 set. 2018.

RAMUS, S. J. et al. Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer. **Journal of the National Cancer Institute**, 2015.

RASHID, M. U. et al. A novel deleterious c.2656G>T MSH2 germline mutation in a Pakistani family with a phenotypic overlap of hereditary breast and ovarian cancer and Lynch syndrome. **Hereditary Cancer in Clinical Practice**, v. 14, n. 1, p. 1–6, 2016.

RENAULT, A.-L. et al. ABRAXAS (FAM175A) and Breast Cancer Susceptibility: No Evidence of Association in the Breast Cancer Family Registry. **PloS one**, v. 11, n. 6, p. e0156820, 2016.

ROEB, W.; HIGGINS, J.; KING, M. C. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. **Human Molecular Genetics**, 2012.

ROMERO-LAORDEN, N.; CASTRO, E. Inherited mutations in DNA repair genes and cancer risk. **Current Problems in Cancer**, v. 41, n. 4, p. 251–264, jul. 2017.

ROY, R.; CHUN, J.; POWELL, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. **Nature reviews. Cancer**, v. 12, n. 1, p. 68–78, 2012.

SAWYER, S. et al. A role for common genomic variants in the assessment of familial breast cancer. **Journal of Clinical Oncology**, 2012.

SCHNITT, S. J. Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy. **Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc**, v. 23 Suppl 2, n. S2, p. S60–S64, 2010.

SCHROEDER, C. et al. HBOC multi-gene panel testing: comparison of two sequencing centers. **Breast Cancer Research and Treatment**, v. 152, n. 1, p. 129–136, 2015.

SCHUBERT, S. et al. The identification of pathogenic variants in BRCA1/2 negative, high risk, hereditary breast and/or ovarian cancer patients: High frequency of FANCM pathogenic variants. **International Journal of Cancer**, 2019.

SEAL, S. et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. **Nature Genetics**, v. 38, n. 11, p. 1239–1241, 2006.

SEITZ, S. et al. Identification of microsatellite instability and mismatch repair gene mutations in breast cancer cell lines. **Genes Chromosomes and Cancer**, 2003.

SETON-ROGERS, S. Prioritizing synthetic lethal targets with functional genomics. **Nature Reviews Cancer**, v. 19, n. June, p. 2019, 2019.

SHAPIRO, L. et al. Structural basis of cell-cell adhesion by cadherins. **Nature**, 1995.

SIEGEL, R. L.; MILLER, K. D. Cancer Statistics , 2019. v. 69, n. 1, p. 7–34, 2019.

SILVA, F. C. et al. Hereditary breast and ovarian cancer: assessment of point mutations and copy number variations in Brazilian patients. **BMC medical genetics**, v. 15, p. 55, jan. 2014.

SKIDMORE, Z. L. et al. GenVisR: Genomic Visualizations in R. **Bioinformatics**, v. 32, n. 19, p. 3012–3014, 2016.

SLAMON, D. et al. Adjuvant Trastuzumab in HER2-Positive Breast Cancer. **New Engl J Med**, v. 375, n. 14, p. 683–693, 2011.

SMYTH, C. et al. Quantifying the cumulative effect of low-penetrance genetic variants on breast

cancer risk. **Molecular Genetics and Genomic Medicine**, 2015.

SODHA, N. et al. Rare germ line CHEK2 variants identified in breast cancer families encode proteins that show impaired activation. **Cancer Research**, 2006.

SOLYOM, S. et al. NIH Public Access. v. 4, n. 122, 2013.

SPURDLE, A. B. et al. Refined histopathological predictors of BRCA1 and BRCA2 mutation status: A large-scale analysis of breast cancer characteristics from the BCAC, CIMBA, and ENIGMA consortia. **Breast Cancer Research**, 2014.

STEWART, B.; WILD, C. World Cancer Report 2014. **International Agency for Research on Cancer**, 2014.

STUCKEY, A. R.; ONSTAD, M. A. Hereditary breast cancer: an update on risk assessment and genetic testing in 2015. **American Journal of Obstetrics and Gynecology**, v. 213, n. 2, p. 161–165, 2015.

STUEBS, F. et al. CDH1 mutation screen in a BRCA1/2-negative familial breast-/ovarian cancer cohort. **Archives of Gynecology and Obstetrics**, v. 297, n. 1, p. 147–152, 9 jan. 2018.

SURIANO, G. E-cadherin germline missense mutations and cell phenotype: evidence for the independence of cell invasion on the motile capabilities of the cells. **Human Molecular Genetics**, v. 12, n. 22, p. 3007–3016, 18 set. 2003.

SUSZYNSKA, M. et al. Large-scale meta-analysis of mutations identified in panels of breast/ovarian cancer-related genes — Providing evidence of cancer predisposition genes. **Gynecologic Oncology**, n. xxxx, p. 1–11, 2019.

SUZUKI, T.; MIKI, A. Y.; OHUCHI, A. N. Intratumoral estrogen production in breast carcinoma : significance of aromatase. p. 270–277, 2008.

SY, S. M. H.; HUEN, M. S. Y.; CHEN, J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. **Proceedings of the National Academy of Sciences**, 2009.

TAPIA, T. et al. ATM allelic variants associated to hereditary breast cancer in 94 Chilean women: Susceptibility or ethnic influences? **Breast Cancer Research and Treatment**, v. 107, n. 2, p. 281–288, 2008.

TAVERA-TAPIA, A. et al. Almost 2% of Spanish breast cancer families are associated to germline pathogenic mutations in the ATM gene. **Breast Cancer Research and Treatment**, v. 161, n. 3, p. 597–604, 2017.

TEDALDI, G. et al. Multiple-gene panel analysis in a case series of 255 women with hereditary breast and ovarian cancer. **Oncotarget**, 2017.

TEUGELS, E. et al. De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. **Human mutation**, 2005.

TIMOTEO, A. R. DE S. et al. A portrait of germline mutation in Brazilian at-risk for hereditary breast cancer. **Breast Cancer Research and Treatment**, v. 172, n. 3, p. 637–646, 2018.

TKÁČ, J. et al. HELB Is a Feedback Inhibitor of DNA End Resection. **Molecular Cell**, v. 61, n. 3, p. 405–418, 2016.

TOSS, A. et al. Hereditary Ovarian Cancer: Not Only BRCA 1 and 2 Genes. **BioMed research international**, v. 2015, n. Figure 1, p. 341723, 2015.

TUNG, N. et al. Frequency of germline mutations in 25 cancer susceptibility genes in a sequential series of patients with breast cancer. **Journal of Clinical Oncology**, 2016.

TUTT, A. et al. Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. v. 20, n. 17, 2001.

UMAR, A. et al. Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability. **J Natl Cancer Inst.**, 2004.

VENKITARAMAN, A. R. et al. Cancer Susceptibility and the Functions of BRCA1 and BRCA2. v. 108, p. 171–182, 2002.

VODUC, K. D. et al. Breast Cancer Subtypes and the Risk of Local and Regional Relapse.

JOURNAL OF CLINICAL ONCOLOGY, v. 28, n. 10, 2017.

VON NICOLAI, C. et al. A second DNA binding site in human BRCA2 promotes homologous recombination. **Nature Communications**, 2016.

WANG, B. et al. Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. **Science**, 2007.

WANG, B. et al. NBA1, a new player in the Brcal A complex, is required for DNA damage resistance and checkpoint control. **Genes and Development**, 2009.

WANG, J. et al. Genome measures used for quality control are dependent on gene function and ancestry. **Bioinformatics**, v. 31, n. 3, p. 318–323, 2014.

WATTERS, G. P. et al. H2AX phosphorylation as a genotoxicity endpoint. **Mutation Research - Genetic Toxicology and Environmental Mutagenesis**, v. 679, n. 1–2, p. 50–58, 2009.

WEBER-LASSALLE, N. et al. Germline loss-of-function variants in the BARD1 gene are associated with familial breast cancer. **Information und Programm: www.senologiekongress.de**, v. 15, p. 1–6, 2018.

WHITTEMORE, A. S.; HARRIS, R.; ITNYRE, J. Characteristics relating to ovarian cancer risk: collaborative analysis of 12 {US} case-control studies. {II}. {Invasive} epithelial ovarian cancers in white women. {Collaborative} {Ovarian} {Cancer} {Group}. **American Journal of Epidemiology**, v. 136, n. 10, p. 1184–1203, 1992.

WIDSCHWENDTER, M. et al. The sex hormone system in carriers of BRCA1 / 2 mutations : a case-control study. **Lancet Oncology**, v. 14, n. 12, p. 1226–1232, 2013.

WOOSTER, R. et al. Identification of the breast cancer susceptibility gene BRCA2. **Nature**, v. 378, n. 6559, p. 789–92, 1995.

WU, Q. et al. Structure of BRCA1-BRCT/Abraxas Complex Reveals Phosphorylation-Dependent BRCT Dimerization at DNA Damage Sites. **Molecular Cell**, v. 61, n. 3, p. 434–448, 2016.

YANG, H. et al. BRCA2 function in DNA binding and recombination from a BRCA1-DSS1-ssDNA structure. **Science**, v. 297, n. September, p. 1837–1847, 2002.

YANG, S. Y. C. et al. Landscape of genomic alterations in high-grade serous ovarian cancer from exceptional long- and short-term survivors. **Genome medicine**, 2018.

YI, M. et al. Which threshold for ER positivity? a retrospective study based on 9639 patients. **Annals of oncology : official journal of the European Society for Medical Oncology / ESMO**, 2014.

YURGELUN, M. B. et al. Identification of a Variety of Mutations in Cancer Predisposition Genes in Patients with Suspected Lynch Syndrome. **Gastroenterology**, 2015.

ZANNINI, L.; DELIA, D.; BUSCEMI, G. CHK2 kinase in the DNA damage response and beyond. **Journal of Molecular Cell Biology**, v. 6, n. 6, p. 442–457, 2014.

ZHONG, Q. et al. Effects of BRCA1- And BRCA2-related mutations on ovarian and breast cancer survival: A meta-analysis. **Clinical Cancer Research**, 2015.

ZOU, J. et al. FancJ regulates interstrand crosslinker induced centrosome amplification through the activation of polo-like kinase 1. **Biology Open**, v. 2, n. 10, p. 1022–1031, 2013.

APÊNDICE

Tabela A1. Primers utilizados para detecção de variantes pontuais, citadas na Metodologia deste documento.

Gene	Amplicon	Seqüência do <i>Primer</i> (5' - 3')	Tamanho do amplicon (pb)	T.A. (°C)
<i>BRCA1</i>	ÉXON 1 F	GGCAGACTGGGTGGCCAAT	493	64
	ÉXON 1 R	ACTCAGTGCCCCCTTCCTGATC		
<i>BRCA2</i>	ÉXON 3 F	GTCAGTGGTTAAACTAAGGTGGGA	201	52
	ÉXON 3 R	GAAGCCAGCTGATTATAAGATGGTT		
	ALU F	GACACCATCCCGGCTGAAA	~500	52
	ÉXON 3 R	CCCCAGTCTACCATATTGCAT		
<i>TP53</i>	ÉXON 10 F	TCTCCCCCTCCTCTGTTG	197	59
	ÉXON 10 R	GATCCTATGGCTTTCCAACCT		

Marcadores utilizados para análise de microssatélites (MSI)

<i>Inhibitor of apoptosis protein-1</i>	NR-27 26 Adeninas 5'UTR	AACCATGCTTGCAAACCACT CGATAATACTAGCAATGACC	90	56
	<i>SLC7A8</i>	NR-21 25 Timinas 5'UTR	GAGTCGCTGGCACAGTTCTA CTGGTCACTCGCGTTTACAA	
<i>Zinc finger 2</i>	NR-24 24 Timinas 3'UTR	GCTGAATTTTACCTCCTGAC ATTGTGCCATTGCATTCCAA	129	
	<i>c-kit</i>	BAT-25 21 Timinas INTRON 16	TACCAGGTGGCAAAGGGCA TCTGCATTTTAACTATGGCTC	
<i>hMSH2</i>	BAT-26 27 Adeninas INTRON 5	CTGCGGTAATCAAGTTTTTAG AACCATTCAACATTTTAAACCC	182	

pb = pares de bases. T.A. = temperatura de *annealing*. F = *Forward*. R = *Reverse*. Para os marcadores de microssatélite: NR27 e BAT-25 foram marcados com VIC (emite sinal verde no espectro visível); NR-21 e BAT-26 foram marcados com FAM (emite sinal azul no espectro visível) e NR-24 foi marcado com NED (emite sinal amarelo no espectro visível, que aparece na cor preta, quando consultado nos eletroferogramas). Fonte: Adaptado de Buhard e colab. (2004).

Tabela A2. Primers utilizados para mutagênese dirigida em plasmídeos utilizados para os testes funcionais *in vitro* e *in vivo*.

Gene	Plasmídeo	Variante	Primer	Seqüência do Primer (5' - 3')	T.A. (°C)	Método de mutagênese
PALB2	VP16-PALB2 NT	p.Lys18Arg	Forward	GCTAATTTCTCCCTTAAC TTTTC		Prime Star
		c.53A>G	Reverse	GGAAAAGTTAAGGGAGAAATTAG		
	VP16-PALB2 NT	p.Thr317Pro	Forward	GTAAAAGTGGCCAACTGCCCCAAGT TCTAGGA		Prime Star
		c.949A>C	Reverse	GGGCAGTTGGCCACTTTTACTTATAG CTTT		
	VP16-PALB2 NT	p.Thr317Ser	Forward	GTAAAAGTGGCCAACTGCCCTCAAGT TCTAGGA		Prime Star
		c.949A>T	Reverse	GGAAAAGTTAAGGGAGAAATTAG		
	VP16-PALB2 NT	p.Thr317Ala	Forward	GTAAAAGTGGCCAACTGCCCCGCAAG TTCTAGGA	55	Prime Star
		c.949A>G	Reverse	GGAAAAGTTAAGGGAGAAATTAG		
	VP16-PALB2 NT	Thr317Lys	Forward	GTAAAAGTGGCCAACTGCCCCAAAAG TTCTAGGA		Prime Star
		c.950C>A	Reverse	GGAAAAGTTAAGGGAGAAATTAG		
	VP16-PALB2 NT	p.Thr317Ile	Forward	GTAAAAGTGGCCAACTGCCCATAAAGT TCTAGGA		Prime Star
		c.950C>T	Reverse	GGAAAAGTTAAGGGAGAAATTAG		
	VP16-PALB2 NT	p.Thr317Arg	Forward	GTAAAAGTGGCCAACTGCCCAGAAG TTCTAGGA		Prime Star
		c.950C>G	Reverse	GGAAAAGTTAAGGGAGAAATTAG		
PALB2	peYFP-c1-PALB2	p.Lys18Arg	Forward	GAAAAGTTAAGGGAGAAACTTGC	58	Q5
		c.53A>G	Reverse	CTTCTCCTCACAGCTGAG		
	peYFP-c1-PALB2	p.Thr317Pro	Forward	CCAAGTGGCCcCAAGTTCTAATTTAG	60	Q5
		c.949A>C	Reverse	CCACTTTTACTTATAGCTTTATTAC		
BRIP1	pcDNA3-3xFLAG-BRIP1	p.Arg173Cys	Forward	GATTAGAAAAtGTCATTGCTTTG	56	Q5
		c.517C>T	Reverse	TGCTGTGTAGTTTCTAAG		
	pcDNA3-3xFLAG-BRIP1	p.Gln740His	Forward	AATTACTGCAtGTGTACTATGAC	56	Q5
		c.2220G>T	Reverse	CATCAAAATTTGTTTTTTCTCC		
CHEK2	pcDNA3-Flag-CHEK2	p.Arg137Gln	Forward	GATAAATACCaAACATACAGCAAG	56	Q5
		c.410G>A	Reverse	TGTTCTTTTCAGCAGTGG		
	pcDNA3-Flag-CHEK2	p.Ile160Met	Forward	TTGCATACATgGAAGATCACAG	57	Q5
		c.480A>G	Reverse	TGTAAGAGTTTTTAGGACC		
MLH1	pcDNA3-V5-MLH1	p.Glu102Glu	Forward	TTGAGAGTGAaGCTTTGGCCA	66	Q5
		c.306G>A	Reverse	AGCCATAGGTAGAAATACTGGC		
PMS2	pcDNA3-V5-PMS2	p.Val717Met	Forward	GCAGCACACCaTGCTCCAGGG	68	Q5
		c.2149G>A	Reverse	TGCAGCATCTCGAAGTTATACTTCTC		
	pcDNA3-V5-PMS2	p.Asp792Asn	Forward	CATGCTGAGCaACAGCCCTGG	65	Q5
		c.2374G>A	Reverse	AAGATCAGTTCATCGACGTCC		

T.A. = temperatura de *Annealing*.

Tabela A3. Anticorpos utilizados para relevar a superexpressão das proteínas.

Anticorpo	Proteína	Diluição	Diluyente	Origem	Marca
B1310	BRIP1	1µL:2000µL	PBST 0,05%	<i>Rabbit</i>	Sigma
67595	CHK2	1µL:5000µL	TBST 0,1%	<i>Rabbit</i>	Novus
2661S (Thr68)	pCHK2	1µL:1000µL	TBST 0,1%	<i>Rabbit</i>	Cell signaling
Sc-510	GAL4	1µL:500µL	TBST 0,1%	<i>Mouse</i>	Santa Cruz
67381	MLH1	1µL:5000µL	TBST 0,1%	<i>Rabbit</i>	Novus
-	PALB2	1µL:5000µL	PBST 0,05%	<i>Rabbit</i>	Produção independente
67069	PMS2	1µL:5000µL	TBST 0,1%	<i>Rabbit</i>	Novus
Ab7291	Tubulina	1µL:200.000µL	PBST 0,05%	<i>Mouse</i>	Abcam
V9131	Vinculina	1µL:100.000µL	PBST 0,05%	<i>Mouse</i>	Sigma
R960-25	V5	1µL:5000µL	PBST 0,05%	<i>Mouse</i>	Invitrogen
Sc-7545	VP16	1µL:50µL	TBST 0,1%	<i>Mouse</i>	Santa Cruz
Sc-2005	<i>IgG Anti-Mouse</i>	1µL:10.000µL	PBST 0,05%	-	Santa Cruz
Ni01	<i>IgG Anti-Rabbit</i>	1µL:10.000µL	PBST 0,05%	-	Calbiochem

Tabela A4. Caracterização clínica e molecular dos pacientes estudados.

Amostras	Genótipo	Grupo de idade ao diagnóstico	Idade ao diagnóstico	Óbito	Localização do tumor	Tipo do tumor	Distribuição do tumor	História familiar	Subtipo molecular	Grau do tumor	Estadiamento linfonodal	Metástase à distância
2697	BRCA VUS	NI	NI	NÃO	Mama	Carcinoma Ductal Invasivo	NI	SIM	NI	NI	NI	NI
2699	BRCA VUS	31-50	36	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	NI	NI	NI	NI
2723	BRCA1 patogênica	31-50	31	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	TN	III	N0	M0
2724	BRCA VUS	20-30	29	NÃO	Mama	Carcinoma Ductal Invasivo tipo Lobular (?)	U	SIM	Lum	NI	N0	M0
2726	BRCA VUS	20-30	26	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	LumHER	II	N2	M0
2742	BRCA VUS	31-50	41	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	Lum	I	N1	M0
2750	BRCA VUS	31-50	35	SIM	Mama	Carcinoma Ductal Invasivo	B	SIM	TN	II	N2	M1 - PELE, ABDOME E MO
2753	BRCA VUS	31-50	40	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	HER2	II	NI	NI
2754	não-BRCA	31-50	41	SIM	Mama	Carcinoma Ductal Invasivo	U	SIM	HER2	III	N2	M1 - PULMONAR, ÓSSEA E SNC
2775	BRCA VUS	51-60	58	NÃO	Mama	Carcinoma Ductal Infiltrante	U	SIM	HER2	NI	N1	M0
2779	BRCA VUS	31-50	37	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	HER2	III	N0	M0
2781	BRCA VUS	51-60	52	SIM	Ovário	Carcinoma Seroso Papilífero de Ovário	U	SIM	NI	NI	NI	M1 - FÍGADO E PLEURAL
2784	BRCA VUS	31-50	37	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	NI	N2	M1 - ÓSSEA
2785	BRCA1 patogênica	31-50	42	NÃO	Mama	Carcinoma Ductal "In situ" Invasivo	B	SIM	LumHER	III	N0	M0
2801	BRCA VUS	20-30	30	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	LumHER	II	N0	M0
2812	BRCA1 patogênica	20-30	25	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	HER2	III	N1	M0
2815	BRCA VUS	20-30	25	NÃO	Mama	Carcinoma Ductal	M	SIM	NI	NI	NI	NI
2963	não-BRCA	31-50	47	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	NI	NI	NI	NI
2972	BRCA VUS	31-50	45	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	II	N1	M0
2977	BRCA VUS	20-30	30	SIM	Mama e Ovário	Carcinoma Ductal Invasivo e Teratoma Cístico Maduro de Ovário	M	SIM	Lum	NI	NI	M1 - HEPÁTICA E ÓSSEA
2995	BRCA VUS	31-50	35	NÃO	Mama	Carcinoma Ductal Invasivo	U	NI	HER2	NI	N2	M0
3002	BRCA VUS	31-50	32	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	II	N0	M0
3051	BRCA1 patogênica	20-30	30	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	HER2	III	NI	M0

Continua.

Tabela A4. Caracterização clínica e molecular dos pacientes estudados.

Continuação.

Amostras	Genótipo	Grupo de idade ao diagnóstico	Idade ao diagnóstico	Óbito	Localização do tumor	Tipo do tumor	Distribuição do tumor	História familiar	Subtipo molecular	Grau do tumor	Estadiamento linfonodal	Metástase à distância
3056	BRCA VUS	51-60	55	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	III	N0	M0
3078	BRCA VUS	31-50	41	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	II	N0	M0
3083	BRCA VUS	20-30	22	NÃO	Ovário	Tumor Seroso Borderline de Ovário Padrão Cistadenofibromatoso Papilífero	B	NÃO	NI	NI	NI	NI
3097	BRCA VUS	31-50	34	SIM	Mama	Carcinoma Ductal Invasivo	B	SIM	Lum	II	N0	M1 - PULMONAR, HEPÁTICA E ÓSSEA
3113	BRCA VUS	31-50	33	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	II	N1	M0
3114	BRCA VUS	61+	64	NÃO	Mama	Carcinoma Ductal Invasivo	M	SIM	Lum	II	N2	M1
3115	BRCA VUS	31-50	35	NÃO	Mama	Carcinoma Ductal Invasivo	U	NI	LumHER	II	N2	M1 - FÍGADO
3116	BRCA VUS	51-60	52	NÃO	Mama	Carcinoma Ductal Invasivo	M	SIM	LumHER	II	N2	M0
3132	BRCA1 patogênica	31-50	33	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N0	M0
3141	BRCA1 patogênica	31-50	31	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N1	M0
3155	BRCA1 patogênica	31-50	47	NÃO	Mama	Carcionma Ductal Invasivo	U	SIM	Lum	II	N0	M0
3166	BRCA VUS	NI	NI	NÃO	Mama	Carcinoma Ductal (?) "In situ"	NI	NI	NI	NI	NI	NI
3187	BRCA VUS	31-50	43	NÃO	Ovário	Adenocarcioma de Ovário	U	SIM	NI	II	NI	PERITONEAL (?)
3227	TP53 patogênica	20-30	25	SIM	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	II	N2	M1 - CÉREBRO, FÍGADO, PULMONAR E ÓSSEA
3441	BRCA VUS	51-60	60	NÃO	Endométrio	Adenocacnioma endometrióide de endométrio	U	SIM	NI	I	NI	M0
3462	BRCA VUS	20-30	26	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	I	N3	M0
3617	BRCA VUS	20-30	29	NÃO	Mama	Carcinoma Ductal Invasivo	NI	NÃO	Lum	NI	NI	NI
3634	não-BRCA	31-50	45	NÃO	Mama	Carcinoma Ductal "In situ"	NI	SIM	NI	NI	NI	NI
3634	BRCA1 patogênica	31-50	45	NÃO	Mama	Carcinoma Ductal "In situ"	NI	SIM	NI	NI	NI	NI
3639	BRCA VUS	20-30	27	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	III	NI	M0
3649	BRCA VUS	31-50	41	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	II	N1	M0

Continua.

Tabela A4. Caracterização clínica e molecular dos pacientes estudados.

Continuação.

Amostras	Genótipo	Grupo de idade ao diagnóstico	Idade ao diagnóstico	Óbito	Localização do tumor	Tipo do tumor	Distribuição do tumor	História familiar	Subtipo molecular	Grau do tumor	Estadiamento linfonodal	Metástase à distância
3650	BRCA VUS	31-50	35	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	II	N1	M0
3651	BRCA VUS	31-50	32	NÃO	Mama	Carcinoma Ductal Invasivo	NI	SIM	RP	II	N0	M0
3662	TP53 patogênica	31-50	39	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	III	N1A	M0
3664	BRCA VUS	20-30	28	NÃO	Mama	Carcinoma Ductal Invasivo	NI	SIM	HER2	NI	NI	NI
3671	BRCA VUS	31-50	32	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	HER2	II	N1C	M0
3703	BRCA VUS	31-50	33	NÃO	Ovário	Cistoadenocarcinoma Mucinoso de Ovário e Tumor Mucinoso borderline	U	SIM	NI	II	NI	NI
3706	não-BRCA	31-50	34	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	Lum	II	N2	M1 - FÍGADO, OSSEA
3716	não-BRCA	31-50	39	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	II	N1	M0
3722	BRCA1 patogênica	31-50	39	NÃO	Mama	Carcinoma Ductal tipo Basal da Mama	U	SIM	TN	NI	N1	M0
3728	BRCA1 patogênica	31-50	34	SIM	Mama	Carcinoma Ductal "In situ"	U	SIM	TN	NI	NX	M1 - PULMÃO
3761	não-BRCA	31-50	36	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	II	N0	M0
3766	não-BRCA	31-50	34	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	III	N3A	M0
3772	não-BRCA	31-50	37	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	Lum	II	N1A	M0
3782	BRCA VUS	31-50	47	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	I	N0	M0
3802	BRCA VUS	31-50	46	SIM	Estômago	Câncer Gástrico Difuso - Adenocarcinoma de corpo gástrico	U	SIM	NI	NI	N0	M0
3806	BRCA VUS	31-50	42	SIM	Mama	Carcinoma Ductal Invasivo	B	SIM	LumHER	II	N1	MX - PULMONAR E SNC
3842	BRCA VUS	31-50	37	NÃO	Mama	Carcinoma Ductal Invasivo	NI	SIM	LumHER	II	N1	M0
3876	BRCA VUS	31-50	37	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	II	N1	M0
3897	BRCA VUS	31-50	42	NÃO	Mama	Papiloma Intraductal	U	SIM	NI	NI	NI	NI
3920	BRCA VUS	20-30	30	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	Lum	I	NI	M0
4016	BRCA VUS	31-50	32	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	III	NI	NI
4020	BRCA VUS	31-50	40	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	II	N1A	M0
4037	BRCA VUS	31-50	33	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	TN	NI	N0	M0
4063	BRCA VUS	31-50	49	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	Lum	II	N1	M0

Continua.

Tabela A4. Caracterização clínica e molecular dos pacientes estudados.

Continuação.

Amostras	Genótipo	Grupo de idade ao diagnóstico	Idade ao diagnóstico	Óbito	Localização do tumor	Tipo do tumor	Distribuição do tumor	História familiar	Subtipo molecular	Grau do tumor	Estadiamento linfonodal	Metástase à distância
4094	BRCA1 patogênica	31-50	35	NÃO	Mama	Carcinoma Medular Atípico e Carcinoma Ductal Invasivo	U	SIM	TN	III	N0	M0
4122	BRCA VUS	51-60	52	NÃO	Mama	Carcinoma Lobular Invasivo	U	SIM	NI	I	N0	M0
4132	BRCA1 patogênica	31-50	35	NÃO	Mama	Carcinoma Ductal Invasivo	U	NI	Lum	I	N0	M0
4133	BRCA VUS	31-50	40	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	II	N0	M0
4135	BRCA1 patogênica	31-50	40	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N1	M0
4137	BRCA2 patogênica	31-50	31	NÃO	Mama	Carcinoma Ductal Invasivo	U	NI	TN	I	N0	M0
4138	BRCA VUS	31-50	47	SIM	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N0	MX - PULMONAR
4139	BRCA VUS	31-50	34	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N1	M0
4144	BRCA VUS	20-30	27	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N2A	M0
4145	não-BRCA	31-50	31	SIM	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	III	N1	MX - ÓSSEA E FIGADO
4146	BRCA VUS	31-50	37	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	II	NI	NI
4147	BRCA2 patogênica	20-30	24	NÃO	Mama	Carcinoma Ductal Invasivo	U	NI	Lum	II	N2a	M0
4161	BRCA VUS	61+	66	SIM	Ovário	Carcinoma seroso de alto grau de provável primário de tuba uterina	M	SIM	NI	NI	NI	MX - PLEURAL E PERITONEAL
4166	BRCA VUS	31-50	45	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N1	M0
4173	não-BRCA	31-50	31	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N1	M0
4177	BRCA VUS	20-30	27	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	TN	II	N0	M0
4186	BRCA2 patogênica	51-60	57	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	TN	III	N2	M0
4214	BRCA VUS	31-50	35	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	NI	N0	M0
4220	BRCA VUS	51-60	53	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	HER2	III	N2a	M0
4226	BRCA VUS	31-50	34	SIM	Mama	Carcinoma Ductal Invasivo	U	SIM	Lum	I	N1	M1 - COLUNA, ÓSSEA E SNC
4228	BRCA VUS	31-50	41	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	II	N1	M0
4250	BRCA VUS	31-50	35	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	III	N1	M1 - ÓSSEA, HEPÁTICA E PULMONAR
4259	não-BRCA	31-50	33	NÃO	Mama	Carcinoma Ductal Invasivo	U	NÃO	Lum	II	N0	M0

Continua.

Tabela A3. Caracterização clínica e molecular dos pacientes estudados.

Continuação.

Amostras	Genótipo	Grupo de idade ao diagnóstico	Idade ao diagnóstico	Óbito	Localização do tumor	Tipo do tumor	Distribuição do tumor	História familiar	Subtipo molecular	Grau do tumor	Estadiamento linfonodal	Metástase à distância
4262	BRCA2 patogênica	31-50	36	NÃO	Mama	Carcinoma Ductal Invasivo	B	SIM	LumHER	II	NI	NI
4264	não-BRCA	31-50	41	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	LumHER	I	N0	M0
4268	BRCA VUS	61+	72	SIM	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	III	N1	M1 - SNC
4279	BRCA VUS	31-50	39	NÃO	Mama	Carcinoma Ductal Invasivo	U	SIM	TN	II	N1	M0

NI = Não informado; B = Câncer de mama bilateral; U = Câncer de mama unilateral; M = Câncer multifocal, ou seja, ocorrência em mais de um sítio. Na coluna subtipo molecular, a classificação é feita apenas para tumores de mama, onde: TN = Triplo negativo; Lum = Luminal A ou Luminal B; LumHER = Luminal HER. HER2 = positivo apenas para HER2; RP = positivo apenas para receptor de progesterona.

Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.

Gene	Nome da variante		dbSNP ID	ClinVar	Predição In silico						Amostras
	DNA codificante	Proteína			CADD	Align GVG	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
Genes HR											
ATM	c.1541G>A	p.Gly514Asp	rs2235000	Benign/Likely Benign	25.7	Class C65	Polymorphism	Tolerated	Probably damaging	Disease Causing	3664; 4146
	c.1636C>G	p.Leu546Val	rs2227924	Benign/Likely Benign	11.58	Class C25	Polymorphism	Damaging	Possibly damaging	Polymorphism	3617; 3634
	c.1810C>T	p.Pro604Ser	rs2227922	Benign/Likely Benign	23.3	Class C65	Probably polymorphism	Tolerated	Possibly damaging	Disease Causing	2775
	c.2442C>A	p.Asp814Glu	rs3218695	Benign	15.88	Class C35	Polymorphism	Tolerated	Benign	Polymorphism	2753; 2784
	c.2572T>C	p.Phe858Leu	rs1800056	Conflicting information on pathogenicity	13.50	Class C15	Polymorphism	Damaging	Possibly damaging	Polymorphism	4268
	c.4258C>T	p.Leu1420Phe	rs1800058	Conflicting information on pathogenicity	15.47	Class C15	Polymorphism	Tolerated	Benign	Disease Causing	3650
	c.5557G>A	p.Asp1853Asn	rs1801516	Benign/Likely Benign	23.2	Class C15	Polymorphism	Tolerated	Benign	Polymorphism	2699; 2724; 2775; 3002; 3132 (homoz); 3141; 3166; 3187; 3728 (homoz); 4063; 4133; 4135; 4137; 4138 (homoz); 4147; 4226 (homoz)
	c.6995T>C	p.Leu2332Pro	rs4988111	Benign/Likely Benign	15.87	Class C65	Polymorphism	Tolerated	Benign	Polymorphism	3617; 3634
	c.7740A>C	p.Arg2580Ser	rs199915459	VUS	15.65	Class C65	Pathogenic	Tolerated	Benign	Disease Causing	3671
	c.5558A>T	p.Asp1853Val	rs1801673	Conflicting information on pathogenicity	24.2	Class C65	Pathogenic	Damaging	Possibly damaging	Disease Causing	4186; 4264
ATR	c.2794C>A	p.Pro932Thr	-	-	27.0	Class C35	Pathogenic	Damaging	Probably damaging	Disease Causing	4020
	c.7300C>G	p.Pro2434Ala	rs33972295	Benign/Likely Benign	23.5	Class C25	Polymorphism	Damaging	Probably damaging	Disease Causing	4228
	c.946G>A	p.Val316Ile	rs28897764	Benign/Likely Benign	18.61	Class C25	Probably polymorphism	Tolerated	Benign	Disease Causing	2726; 3116; 3671; 3703; 4228
BARD1	c.-83C>T	-	rs71579840	-	8.485	-	-	-	-	-	3002
	c.1972C>T	p.Arg658Cys	rs3738888	Benign/Likely Benign	26.5	Class C65	Probably Pathogenic	Damaging	Probably damaging	Disease Causing	3671
	c.1268A>G	p.Lys423Arg	rs749383704	VUS	21.8	Class C25	Probably Pathogenic	Tolerated	Benign	Disease Causing	2995
	c.764A>G	p.Asn255Ser	rs138904906	VUS	16.75	Class C45	Polymorphism	Tolerated	Probably damaging	Polymorphism	3716
	c.716T>A	p.Leu239Gln	rs200359745	VUS	1.061	Class C65	Probably Pathogenic	Tolerated	Benign	Polymorphism	3051

Continua.

Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.
Continuação.

Gene	Nome da variante		dbSNP ID	ClinVar	Predição In silico						Amostras
	DNA codificante	Proteína			CADD	Align GVGD	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
Genes HR											
BRCA1	c.*421G>T	-	rs8176318	Benign	4.78	-	-	-	-	-	2697 (homoz); 2699; 2742; 2750; 2753; 2779; 2801; 2815; 2972; 2977; 3002; 3056; 3078; 3097; 3114; 3115; 3116; 3132; 3141; 3166; 3227; 3462; 3617; 3671; 3703; 3728; 3806; 3842; 3897; 4016; 4020; 4135; 4138; 4146; 4147; 4161; 4166; 4177; 4220; 4226; 4228; 4279
	c.3119G>A	p.Ser1040Asn	rs4986852	Benign	14.86	Class C45	Polymorphism	Tolerated	Probably damaging	Polymorphism	2699; 2785; 2995; 3002; 3114; 3876; 4020; 4037; 4132
	c.5019G>A	p.Met1673Ile	rs1799967	Benign	22.0	Class C0	Probably polymorphism	Tolerated	Benign	Disease Causing	3897
	c.4598G>T	p.Ser1533Ile	rs1800744	Benign	16.14	Class C65	Polymorphism	Damaging	Possibly damaging	Polymorphism	3113
	c.1648A>C	p.Asn550His	rs56012641	Benign	17.67	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	4132
	c.1067A>G	p.Gln356Arg	rs1799950	Benign	17.80	Class C35	Polymorphism	Damaging	Probably damaging	Polymorphism	2724; 2775; 3187; 3703; 4133; 4139
	c.2077G>A	p.Asp693Asn	rs4986850	Benign	15.84	Class C15	Polymorphism	Damaging	Benign	Polymorphism	2815; 2977; 3097; 3116; 3671; 4122; 4220
	c.5507G>A	p.Trp1836Ter	rs80356962	Pathogenic	44	-	Pathogenic	-	-	Disease Causing	3051
	c.5329dupC	p.Gln1756Profs*74	rs397507247	Pathogenic	35	-	-	-	-	-	2812; 3132; 3141; 3155; 3639; 3722; 3728; 4093; 4135; 4137; 4186
	c.3331_3334delCAAG	p.Gln1111Asnfs*5	rs80357701	Pathogenic	23.7	-	-	-	-	-	2723
	c.2612C>T	p.Pro871Leu	rs799917	Benign	17.97	Class C65	-	Tolerated	Benign	Polymorphism	2697 (homoz); 2699; 2724; 2726; 2742 (homoz); 2750; 2753; 2779 (homoz); 2801; 2812; 2815; 2972; 2977; 3002; 3056 (homoz); 3078 (homoz); 3083; 3097; 3114; 3115 (homoz); 3116 (homoz); 3132; 3141; 3166; 3227; 3441(homoz); 3462; 3617(homoz); 3650(homoz); 3651; 3664(homoz); 3671; 3703; 3728; 3782; 3802; 3806; 3842(homoz); 3897; 3920; 4016; 4020; 4037; 4063; 4093; 4122; 4135; 4138; 4139; 4144; 4146(homoz); 4147; 4161; 4166 (homoz); 4177; 4186 (homoz); 4214; 4220; 4226 (homoz); 4228; 4250 (homoz); 4262; 4268 (homoz); 4279
	c.3548A>G	p.Lys1183Arg	rs16942	Benign		Class C25	-	Tolerated	Benign	Polymorphism	2697 (homoz); 2699; 2742; 2750; 2753; 2801; 2815; 2972; 2977; 3002; 3056; 3078 (homoz); 3097; 3114; 3116; 3141; 3166; 3462; 3617 (homoz); 3651; 3703; 3806; 3842 (homoz); 3897; 4016; 4020; 4122; 4135; 4138; 4146; 4161; 4166; 4262; 4268 (homoz); 4279
	c.4900A>G	p.Ser1613Gly	rs1799966	Benign		Class C55	Polymorphism	Damaging	Possibly damaging	Polymorphism	2697; 2699; 2742; 2750; 2753; 2779; 2784; 2801; 2815; 2972; 2977; 3002; 3056; 3078 (homoz); 3097; 3114; 3115; 3116; 3132; 3141; 3166; 3227; 3462; 3617 (homoz); 3651; 3671; 3703; 3728; 3782; 3806; 3842 (homoz); 3897; 4016; 4020; 4122; 4135; 4138; 4146; 4147; 4161; 4166; 4177; 4220; 4226; 4228; 4262; 4268 (homoz); 4279

Continua.

Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.
Continuação.

Gene	Nome da variante		dbSNP ID	ClinVar	Predição In silico						Amostras
	DNA codificante	Proteína			CADD	Align GVGD	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
Genes HR											
BRCA1	c.536A>G	p.Tyr179Cys	rs56187033	Benign	24.7	Class C65	Pathogenic	Damaging	Probably damaging	Disease Causing	4132
	c.591C>T	p.Cys197Cys	rs1799965	VUS	14.63	-	Probably Pathogenic	-	-	-	4063
	c.3113A>G	p.Glu1038Gly	rs16941	Benign	22.2	Class C65	Polymorphism	Damaging	Possibly damaging	Polymorphism	2697; 2699; 2742; 2750; 2753; 2779; 2784; 2801; 2815; 2972; 2977; 3002; 3056; 3078 (homoz); 3097; 3114; 3115; 3116; 3132; 3141; 3166; 3227; 3462; 3617 (homoz); 3651; 3671; 3703; 3728; 3782; 3806; 3842; 3897; 4016; 4020; 4122; 4135; 4138; 4146; 4147; 4161; 4166; 4220; 4226; 4228; 4262; 4268 (homoz); 4279
BRCA2	c.156_157insAlu	-	-	Pathogenic	-	-	-	-	-	-	-
	c.811G>A	p.Gly271Arg	rs786204274	Conflicting information on pathogenicity	20.6	Class C65	Probably Pathogenic	Damaging	Benign	Polymorphism	2977
	c.3869G>A	p.Cys1290Tyr	rs41293485	Benign	14.39	Class C65	Probably Pathogenic	Tolerated	Benign	Polymorphism	3662
	c.4258G>T	p.Asp1420Tyr	rs28897727	Benign	15.81	Class C65	Polymorphism	Damaging	Benign	Polymorphism	3649
	c.6100C>T	p.Arg2034Cys	rs1799954	Benign	20.4	Class C65	Polymorphism	Damaging	Possibly damaging	Polymorphism	3441; 3617; 4279
	c.8149G>T	p.Ala2717Ser	rs28897747	Benign	15.38	Class C65	Polymorphism	Tolerated	Possibly damaging	Polymorphism	3002
	c.8850G>T	p.Lys2950Asn	rs28897754	Conflicting information on pathogenicity	22.4	Class C65	Probably polymorphism	Damaging	Probably damaging	Disease Causing	2781
	c.865A>C	p.Asn289His	rs766173	Benign	17.23	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	2995; 3051; 3056; 3651; 3722; 4166
	c.2808_2811delACAA	p.Ala938Profs*21	rs80359351	Pathogenic	24.3	-	-	-	-	-	4147
	c.8851G>A	p.Ala2951Thr	rs11571769	Benign	26.2	Class C55	Polymorphism	Damaging	Probably damaging	Disease Causing	3051
	c.9026_9030delATCAT	p.Tyr3009Serfs*7	rs80359741	Pathogenic	-	-	-	-	-	-	2785
	c.9382C>T	p.Arg3128Ter	rs80359212	Pathogenic	48	-	Pathogenic	-	-	Disease Causing	4262
	c.9976A>T	p.Lys3326Ter	rs11571833	Benign	36	-	Pathogenic	-	-	Disease Causing	3650
c.8324T>G	p.Met2775Arg	rs80359073	Conflicting information on pathogenicity	24.1	Class C65	Pathogenic	Tolerated	Benign	Disease Causing	3116	

Continua.

Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.
Continuação.

Gene	Nome da variante		dbSNP ID	ClinVar	Predição In silico						Amostras
	DNA codificante	Proteína			CADD	Align GVGD	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
Genes HR											
BRIP1	c.3693A>G	p.Ile1231Met	rs876659290	VUS	13.29	Class C0	Polymorphism	Damaging	Benign	Polymorphism	3132; 3728
	c.517C>T	p.Arg173Cys	rs4988345	Benign/Likely Benign	24.9	Class C65	Pathogenic	Damaging	Probably damaging	Disease Causing	4122; 4173
	c.2220G>T	p.Gln740His	rs45589637	VUS	12.53	Class C15	Probably Pathogenic	Damaging	Probably damaging	Disease Causing	3078
CHEK2	c.410G>A	p.Arg137Gln	rs368570187	Likely Benign	16.54	Class C35	Probably polymorphism	Tolerated	Benign	Disease Causing	3116
	c.480A>G	p.Ile160Met	rs575910805	Conflicting information on pathogenicity	22.6	Class C0	Probably polymorphism	Damaging	Probably damaging	Disease Causing	3097
FAM175A	c.489G>T	p.Arg163Ser	rs535462791	-	22.3	Class C65	Pathogenic	Tolerated	Possibly damaging	Disease Causing	3187
	c.1042G>A	p.Ala348Thr	rs12642536	-	13.44	Class C55	Polymorphism	Damaging	Possibly damaging	Polymorphism	2697; 2699; 2723; 2724; 2726; 2742; 2750; 2754 (homoz); 2775; 2779; 2781 (homoz); 2784 (homoz); 2812; 2972; 2977; 3002; 3078; 3097; 3113 (homoz); 3114 (homoz); 3115; 3132 (homoz); 3141 (homoz); 3155 (homoz); 3166 (homoz); 3441; 3462; 3617; 3634; 3639; 3649 (homoz); 3651; 3662; 3706; 3716; 3722; 3728 (homoz); 3761; 3772; 3806; 3876; 3897; 3920; 4016; 4037; 4063; 4093; 4132 (homoz); 4137; 4144; 4145; 4147; 4161; 4166; 4177; 4226 (homoz); 4228; 4250; 4259; 4264; 4279
MRE11A	c.1011C>G	p.Ser337Arg	rs115244417	Benign	21.9	Class C65	Probably polymorphism	Tolerated	Benign	Disease Causing	4135
	c.2101A>G	p.Met701Val	rs1805362	Benign/Likely Benign	16.49	Class C15	Polymorphism	Damaging	Benign	Disease Causing	3650
NBN	c.202T>G	p.Leu68Val	rs1200599843	VUS	15.95	Class C25	Polymorphism	Tolerated	Benign	Disease Causing	2785
	c.797C>T	p.Pro266Leu	rs769420	Benign	24.6	Class C65	Polymorphism	Damaging	Probably damaging	Disease Causing	3078
PALB2	c.2794G>A	p.Val932Met	rs45624036	Benign/Likely Benign	25.3	Class C15	Polymorphism	Tolerated	Probably damaging	Disease Causing	3842
	c.53A>G	p.Lys18Arg	rs138789658	Conflicting information on pathogenicity	24.3	Class C25	Polymorphism	Damaging	Probably damaging	Polymorphism	3897; 4037
	c.949A>C	p.Thr317Pro	rs587780223	VUS	4.012	Class C35	Polymorphism	Tolerated	Possibly damaging	Polymorphism	4139

Continua.

Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.
Continuação.

Gene	Nome da variante		dbSNP ID	ClinVar	Predição In silico						Amostras
	DNA codificante	Proteína			CADD	Align GVGD	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
Genes HR											
<i>RAD51</i>	c.164C>T	p.Ala55Val	rs145617142	VUS	23.6	Class C55	Probably polymorphism	Tolerated	Possibly damaging	Disease Causing	3116
<i>UIMC1</i>	c.43C>T	p.Arg15Trp	rs13167812	-	22.5	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	2750; 4250
	c.1304C>T	p.Pro435Leu	rs3733876	-	23.8	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	2699; 2724; 2754; 2815; 3056; 3083; 3132 (homoz); 3634; 3649; 3671; 3761; 3842; 3920; 4020; 4063 (homoz); 4138; 4173; 4214
Genes MMR											
<i>MLH1</i>	c.306G>A	p.Glu102Glu	rs63751665	VUS	22.4	-	Pathogenic	-	-	Disease Causing	4020
	c.637G>A	p.Val213Met	rs2308317	Benign	23.9	Class C15	Polymorphism	Tolerated	Possibly damaging	Disease Causing	4020
	c.1217G>A	p.Ser406Asn	rs41294980	Benign	13.09	Class C45	Polymorphism	Tolerated	Possibly damaging	Polymorphism	2963
	c.2146G>A	p.Val716Met	rs35831931	Benign	24.4	Class C15	Probably polymorphism	Damaging	Probably damaging	Disease Causing	2754
<i>MSH2</i>	c.2500G>A	p.Ala834Thr	rs63750757	Likely Benign	33.00	Class C55	Pathogenic	Damaging	Probably damaging	Disease Causing	4214
	c.380A>G	p.Asn127Ser	rs17217772	Benign	22.7	Class C45	Probably polymorphism	Damaging	Possibly damaging	Disease Causing	3650; 3920; 4146; 4228
	c.965G>A	p.Gly322Asp	rs4987188	Benign	23.0	Class C65	Probably polymorphism	Tolerated	Possibly damaging	Disease Causing	2815; 3113; 3441; 4264
<i>MSH6</i>	c.1186C>G	p.Leu396Val	rs2020908	Benign	16.97	Class C25	Polymorphism	Tolerated	Possibly damaging	Disease Causing	2699; 2754
	c.2633T>C	p.Val878Ala	rs2020912	Benign	10.23	Class C55	Polymorphism	Tolerated	Benign	Disease Causing	4147
<i>PMS2</i>	c.59G>A	p.Arg20Gln	rs10254120	Benign	16.65	Class C35	Polymorphism	Tolerated	Possibly damaging	Polymorphism	4146; 2963; 3097; 3116; 3722; 3782; 3806; 4137; 4145; 4220; 4262
	c.2374G>A	p.Asp792Asn	rs587781265	VUS	29.8	Class C15	Probably Pathogenic	Damaging	Probably damaging	Disease Causing	3802
	c.2350G>A	p.Asp784Asn	rs143340522	VUS	27.8	Class C15	Polymorphism	Damaging	Probably damaging	Disease Causing	4264
	c.2149G>A	p.Val717Met	rs201671325	Conflicting information on pathogenicity	20.7	Class C15	Probably polymorphism	Damaging	Probably damaging	Disease Causing	3116; 3462
	c.1866G>A	p.Met622Ile	rs1805324	Benign	18.96	Class C0	Polymorphism	Tolerated	Possibly damaging	Disease Causing	3772

Continua.

Tabela A5. Variantes de potencial impacto ao fenótipo identificadas neste estudo e caracterizadas quanto à sua patogenicidade.
 Continuação.

Gene	Nome da variante		dbSNP ID	ClinVar	Predição In silico						Amostras
	DNA codificante	Proteína			CADD	Align GVGD	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
Genes HR											
<i>PMS2</i>	c.1688G>T	p.Arg563Leu	rs63750668	Benign/Likely Benign	10.65	Class C65	Polymorphism	Tolerated	Possibly damaging	Polymorphism	2972; 3227; 4016
Outros genes											
<i>CDH1</i>	c.1849G>A	p.Ala617Thr	rs33935154	Conflicting information on pathogenicity	15.13	Class C55	Polymorphism	Tolerated	Benign	Disease Causing	3664; 4145; 4166
<i>TP53</i>	c.1010G>A	p.Arg337His	rs121912664	Pathogenic	22.6	Class C25	-		Probably damaging	Disease Causing	2699; 3056; 3227; 3662; 4264
	c.818G>A	p.Arg273His	rs28934576	Pathogenic/Likely Pathogenic	24.0	Class C25	Pathogenic	Damaging	Probably damaging	Disease Causing	3227
	c.1688G>T	p.Arg563Leu	rs63750668	Benign/Likely Benign	10.65	Class C65	Polymorphism	Tolerated	Possibly damaging	Polymorphism	2972; 3227; 4016

ANEXOS

- 1 Carta de Aprovação do estudo pelo Comitê de Ética do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto.

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

FMRP-USP
RIBEIRÃO PRETO

Ribeirão Preto, 14 de abril de 2016.

Ofício nº 1291/2016
CEP/MGV

PROCESSO HCRP nº 2839/2016

Prezados Pesquisadores,

O trabalho intitulado **“CARACTERIZAÇÃO MOLECULAR DA SÍNDROME HEREDITÁRIA DO CÂNCER DE MAMA E/OU OVÁRIO”**, foi analisado pelo Comitê de Ética em Pesquisa, em sua 425ª Reunião Ordinária realizada em 11/04/2016, e enquadrado na categoria: **APROVADO**, bem como a Criação do Biorrepositório, Termos de Consentimento Livre e Esclarecido e Termos de Consentimento para Guarda de Material Biológico Versão 2 - modificada em 21 de março de 2016.

De acordo com Carta Circular nº 003/2011/CONEP/CNS, datada de 21/03/2011, o sujeito de pesquisa ou seu representante, quando for o caso, deverá rubricar todas as folhas do Termo de Consentimento Livre e Esclarecido – TCLE – apondo sua assinatura na última do referido Termo; o pesquisador responsável deverá da mesma forma, rubricar todas as folhas do Termo de Consentimento Livre e Esclarecido – TCLE – apondo sua assinatura na última página do referido Termo.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (ICH-GCP), bem como a Resolução nº 466/2012 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente.

DRª MARCIA GUIMARÃES VILLANOVA
Coordenadora do Comitê de Ética em
Pesquisa do HCRP e da FMRP-USP

Ilustríssimos Senhores
SIMONE DA COSTA E SILVA CARVALHO
PROF. DR. WILSON ARAÚJO DA SILVA JUNIOR (Orientador)
Depto. de Genética da FMRP-USP

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO
Campus Universitário – Monte Alegre
14048-900 Ribeirão Preto SP

Comitê de Ética em Pesquisa do HCRP e FMRP-USP
FWA-00002733; IRB-00002186 e
Registro Plataforma Brasil /CONEP nº 5440
(016) 3602-2228
cep@hcrp.usp.br

www.hcrp.usp.br

Title:

Germline variants in DNA repair genes associated with Hereditary Breast and Ovarian Cancer Syndrome: Analysis of a 21 gene panel in the Brazilian population.

Authors:

Simone da Costa e Silva Carvalho 1,2,3, Nathalia Moreno Cury 1,3, Danielle Barbosa Brotto 1,3, Luiza Ferreira de Araujo 1,3, Reginaldo Cruz Alves Rosa 1,2, Lorena Alves Texeira 4, Jéssica Rodrigues Praça 3, Adriana Aparecida Marques 3, Kamila Chagas Peronni 3, Greice Andreotti Molfetta 1,2, Julio Cesar Moriguti 4, Dirce Maria Carraro 5, Edenir Inêz Palmero 6, Patricia Ashton-Prolla 7, Victor Evangelista de Faria Ferraz 1,2,8, Wilson Araujo da Silva Junior 1,2,3.

Institutions:

1) Department of Genetics, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil; 2) Center for Medical Genomics at University Hospital of the Ribeirão Preto Medical School of University of São Paulo, Ribeirão Preto, SP, Brazil; 3) Regional Blood Center at University Hospital of the Ribeirão Preto Medical School of University of São Paulo, Ribeirão Preto, SP, Brazil; 4) Division of Internal Medicine and Geriatrics, Department of Internal Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil. 5) International Research Center/CIPE, AC Camargo Cancer Center, Sao Paulo, SP, Brazil. 6) Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, SP, Brazil. 7) Laboratório de Medicina Genômica, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil. 8) Department of Medical Genetics of University Hospital of the Ribeirão Preto Medical School of University of São Paulo, Ribeirão Preto, SP, Brazil.

*corresponding author:

Tel.: +55 16 3315-3293

e-mail address: wilsonjr@usp.br

Abstract

Background:

The Hereditary Breast and Ovarian Cancer Syndrome (HBOC) occurs in families with a history of breast/ovarian cancer presenting an autosomal dominant inheritance pattern. This syndrome covers about 10% to 15% of all cases of breast and ovarian cancer. *BRCA1* and *BRCA2* are high penetrance genes associated with an increased risk of up to 20-fold for breast and ovarian cancer. However, only 20-30% of HBOC cases present pathogenic variants in those genes, and other DNA repair genes have emerged as increasing the risk for HBOC. In the Brazilian population, variants in *ATM*, *ATR*, *CHEK2*, *MLH1*, *MSH2*, *MSH6*, *POLQ*, *PTEN*

and *TP53* genes have been reported in up to 7.35% of the studied cases. Here we screened and characterized variants in 21 DNA repair genes in individuals clinically diagnosed with HBOC.

Methods:

We systematically analyzed 708 amplicons encompassing the coding and flanking regions of 21 genes related to DNA repair pathways (*ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *FAM175A*, *MLH1*, *MRE11A*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* and *UIMC1*). A total of 95 individuals clinically diagnosed with HBOC in the Southeastern Brazil were sequenced and 25 samples were evaluated for CNVs in *BRCA1* and *BRCA2* genes. Identified SNVs were assessed in terms of population allele frequency and its functional effects were predicted through *in silico* algorithms.

Results:

We identified 80 variants in 19 of our 21 DNA repair gene panel. About 23.4% of the patients presented pathogenic variants in *BRCA1*, *BRCA2* and *TP53*, a frequency higher than that identified among previous studies performed in the Brazilian population. We identified a novel variant in *ATR* gene which was scored as pathogenic by all *in silico* prediction tools. The association analysis revealed 13 missense variants in *BARD1*, *BRCA2*, *CHEK2*, *CDH1*, *FAM175A*, *MLH1*, *PALB2* and *PMS2* genes, as significantly associated with increased relative risk to HBOC, and the patients carrying those variants did not present large insertions or deletions in *BRCA1/BRCA2* genes. We also identified variants characterized as possibly pathogenic in *PMS2* and *BRIP1*.

Conclusions:

This study constitutes the third report of a multi-gene analysis in the Brazilian population, and also addresses the first report of many germline variants associated with HBOC in Brazil. Although further functional analyses are necessary to better characterize the contribution of those variants to the phenotype, those findings would improve the risk estimation and clinical follow-up of patients clinically diagnosed with HBOC, in the future.

Keywords: HBOC, DNA repair genes, multi-gene panel screening, next-generation sequencing, molecular diagnosis, *BRCA1*, *BRCA2*.

Background

Hereditary Breast and Ovarian Cancer Syndrome (HBOC) occurs in families with history of breast and ovarian cancers with an autosomal dominant inheritance pattern. It comprehends about 5-10% of all breast cancer (BC) cases and up to 80% of all ovarian cancers (OC) [1, 2], and the affected families present a 50-80% increase in lifetime risk to BC and 30-50% to OC [3]. The National Comprehensive Cancer Network (NCCN) [4] is an alliance that predict the guidelines used for detection, prevention, as well as for adoption of strategies for risk reduction for HBOC affected families. The clinical diagnosis is mainly based on the following criteria, according to NCCN: patients diagnosed with BC before 45 years or with invasive

OC at any age, personal or familial recurrence of BC or OC, bilateral BC, and presence of male BC. Furthermore, patients under risk to HBOC may also present pancreatic and prostate cancers [4]. Due to it, in order to help demystifying the association of HBOC only to BC and OC risk in women [5], recently it has been proposed to change the name of HBOC to King Syndrome, in honor to Mary-Claire King who first described the locus associated with hereditary breast and ovarian cancers [6].

During the 1990's, germline variants in the breast cancer susceptibility genes *BRCA1* and *BRCA2* were first described as increasing risk to HBOC [7, 8]. Variants in *BRCA1* are associated with earlier-onset BC (30-50 years), when compared to *BRCA2* variants that increase the BC risk mainly for individuals of 40-60 years old [9]. The BC and OC risk rates also vary between *BRCA1* and *BRCA2* genes, with *BRCA1* carriers presenting a risk of up to 57% for BC and 40% for OC, while for *BRCA2* carriers the risk is slightly lower, 49% and 18% for BC and OC, respectively [10].

Molecular diagnosis is a very important step on the clinical management of HBOC patients since it allows the family risk assessment, mortality reduction as well as allows the adoption of prophylactic measures as preventive mastectomy and/or oophorectomy, reducing the risk in up to 95% in *BRCA1/BRCA2* carriers [11–13]. However, despite the high penetrance and the high frequency of variants found in *BRCA1/2* genes, only about 20% of hereditary BC and OC have been attributed to the presence of pathogenic variants in those genes, meanwhile about 5-10% have been associated with other susceptibility genes as *TP53*, *STK11*, *PTEN*, *ATM* e *CHEK2* [14]. Studies have demonstrated molecular diagnosis rates of about 4,6% - 54% when only *BRCA1/2* are screened, which evidences the association of other less penetrant genes to HBOC pathogenesis [15–18]. Even though the protocols for clinical management are well established for *BRCA1/2* carries, patients tested negative for pathogenic *BRCA1/BRCA2* variants lack the proper clinical follow-up and genetic counselling even presenting similar clinical characteristics and BC/OC increased risk [19]. It reinforces the need of not only description but also the characterization of other genes associated with HBOC risk.

With the popularization of the next-generation sequencing technologies (NGS), genes encoding proteins that work in the homologous DNA repair pathway (HR), as well as mismatch repair (MMR) pathway, have been frequently reported as mutated in hereditary BC and OC cases [16, 20–27]. Most genes are not only frequently mutated but they have also been taken into account by NCCN guidelines to the clinical management of patients under risk since they are associated with a high to moderate penetrance to BC and OC [4].

However, in the Brazilian population, besides *BRCA1* and *BRCA2*, the characterization of other DNA repair genes related to HBOC susceptibility is still in infancy. The main available data encompasses the screening of hotspots variants and microdeletions in *CHEK2*, *PTEN*, *POLQ* and *TP53* genes [2, 28–31], and to date, only two studies using NGS technology are available in Brazil. Recently, the screening of whole exome in Brazilian patients negative for *BRCA1/BRCA2* pathogenic variants revealed other genes such as *ATM* and *BARD1* carrying pathogenic variants [27]. Another study using a multi-gene screening showed a prevalence of 9.8% of patients carrying *BRCA1/BRCA2* pathogenic variants and 4.5% carrying pathogenic variants in *ATR*, *CDH1*, *MLH1* and *MSH6* genes [25].

In this study, we screened 95 samples of patients clinically diagnosed with HBOC using a multi-gene panel sequencing both flanking and coding regions of *BRCA1*, *BRCA2* and other 19 DNA repair genes. Also, 25 samples were tested for *BRCA1/BRCA2* copy number variations (CNVs). The molecular screening was performed to identify causal germline variants and characterize variants of unknown/uncertain significance (VUS) in order to improve the molecular diagnosis. Our data report a global analysis of 21 DNA repair genes to the HBOC etiology, which are contributing to the epidemiology of HBOC in Brazil.

Methods

Patient samples and clinical data

Individuals referred to the Cancer Genetics Counseling Service of the University Hospital of the Ribeirão Preto Medical School of the University of São Paulo (HCFMRP-USP, Ribeirão Preto – Brazil) for cancer risk assessment from 2008-2016 were evaluated. A total of 95 unrelated subjects with clinical diagnosis of HBOC, according to NCCN Clinical Practice Guidelines in Oncology v.2.2015 [4], presenting a cumulative risk of *BRCA1* and *BRCA2* variants lower than 10%, using PennII model (<https://pennmodel2.pmacs.upenn.edu/penn2/>), and a personal history of cancer were eligible for further investigation.

The clinical and pathologic data was abstracted from medical records of the HCFMRP-USP and included personal and family cancer histories, cancer histology, stage, and receptor status. The College of American Pathologists (CAP) guidelines were used to define progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) positivity, but for estrogen receptors we used the 10% threshold for positivity [32].

Samples of 28 elderly people (over 70 years old) negative for personal history of cancer, were used as control group and had their whole exome sequenced by the Molecular Genetics Laboratory of UNICAMP (Campinas, SP), headed by Dr. Iscia Lopes Cendes, who kindly provided the results.

Genomic DNA of both case and control samples were extracted from whole blood using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI). The samples were part of the Center for Medical Genomics Biorepository (HCFMRP-USP) and were used for this analyses only after approval by the Ethics Research Committee of the HCFMRP-USP (n. 2819/2016).

The genetic test results from this analysis were returned to study participants, helping the clinical decision making when suitable.

Multi-gene panel screening

We used a TruSeq Custom Amplicon Library Preparation Kit (Illumina, San Diego, CA) for the enrichment of coding and flanking regions of 21 DNA repair genes (*ATM*, *ATR*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *FAM175A*, *MLH1*, *MRE11*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51*, *TP53* and *UIMC1*). A total of 708 amplicons for a 98% mean coverage were custom designed using the Illumina Design Studio (Illumina, San

Diego, CA). Paired-end sequencing was performed on a MiSeq equipment (Illumina, San Diego, CA), using the MiSeq sequencing kit v2 (2x250) (Illumina, San Diego, CA). The bcl files were converted into fastq using the FASTQ Generation v.1.0.0 software, available on BaseSpace (Illumina, San Diego, CA). The mapping and variant calling were performed using BWA mem and Haplotype Caller, respectively, following the GATK v.3.6-0 (<https://software.broadinstitute.org/gatk/>) best practices guidelines for germline single nucleotide polymorphisms (SNPs) and insertion/deletions (indels) detection, using the GRCh37.75/hg19 as reference genome (<http://hgdownload.cse.ucsc.edu/>). We used Snpeff for variant annotation (<http://http://snpeff.sourceforge.net/>).

The graphics to represent the sequencing data were built using the Bioconductor GenVisR [33] and ComplexHeatmap [34] packages on R environment.

Variants classification and prioritization

All variants were classified according to recommendations of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) [35] using the VarSome variant search engine [36]. To a more accurate variant characterization, we also assessed the ClinVar classification (<https://www.ncbi.nlm.nih.gov/clinvar/>), and the pathogenicity scores of the 6 following *in silico* prediction tools: CADD [37], AlignGVGD [38], UMD-Predictor [39], SIFT [40], PolyPhen [41] and MutationTaster [42].

In order to prioritize a smaller number of variants for further characterization, we refined the whole set of variants in favor of remain with those classified as pathogenic according to ACMG/AMP criteria, as well as, with all VUS and benign variants (according to VarSome and ClinVar) present both in coding and splicing regions, if they were predicted as damaging/pathogenic by the *in silico* prediction tools. We decided to maintain the benign variants in this set of prioritized variants in order to avoid disregard variants of potential effect to the phenotype, once ClinVar and VarSome classifications are not always supported by strong evidences (segregational and functional data). Thereat, sometimes we refer those variants as presenting conflicting data on pathogenicity.

Sanger validation

All samples that presented pathogenic variants, as well as all those significantly associated with relative risk to HBOC were submitted to Sanger validation. Briefly, 100 ng of whole blood DNA of individuals carrying those variants was submitted to PCR amplification performed with Taq DNA polymerase (Promega, Madison, WI). The amplification products were sequenced in both directions using BigDye Terminator v3.1 (Life Technologies, Carlsbad, CA) and specific primers for each region, in the ABI 3500XL Genetic Analyzer (Life Technologies, Carlsbad, CA), according to manufacturer's instructions. Sequencing data were analyzed with the Geneious R7 software v7.1 using the GRCh37/hg19 sequence as reference. Primer sequences are available under request.

Analysis of CNVs in *BRCA1* and *BRCA2* genes

To exclude the presence of large insertions/deletions in *BRCA1/BRCA2* genes that might not have been detected by NGS, we performed the Multiplex Ligation-dependent Probe Amplification (MLPA) analysis for patients that did not present any variants on *BRCA1/BRCA2* (n=12) after the multi-gene panel screening, as well as for those patients carrying variants that were significantly associated with relative risk to HBOC (n=15). In order that, we used the P087-*BRCA1* and P090-*BRCA2* kits (MRC-Holand, Amsterdam, NH) according to the manufacturer's recommendations. Briefly, the DNA from patients and control samples were pre-heated to 98°C, and then the salt solution and probe mix were added to the DNA. After the ligation of annealed nucleotides, the targeted genes were amplified using polymerase chain reaction (PCR). PCR products were separated using the ABI3500XL Genetic Analyzer (Applied Biosystems, Foster City, CA), and the fragments were analysed using the Coffalyser software v.140701.0000 (MRC-Holand, Amsterdam, NH).

Screening for the c.156_157insAlu variant in *BRCA2*

All 95 HBOC samples were screened for the variant c.156_157insAlu on *BRCA2* gene, which was not detected by the multi-gene panel analysis. We performed two rounds of PCR: a first PCR reaction for *BRCA2* exon 3 amplification (forward primer: GTCACTGGTTAAACTAAGGTGGGA and reverse primer: GAAGCCAGCTGATTATAAGATGGTT), and a second PCR specific for Alu fragment amplification (forward primer: GACACCATCCCGGCTGAAA, reverse primer: CCCCAGTCTACCATATTGCAT). The cycling conditions were 94°C for 3 min, 35 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 4 min, and a final extension of 72°C for 10 min. For the sample that presented a fragment amplification bigger than that expected for *BRCA2* exon 3 amplification (around 200pb), the specific Alu PCR was performed using the same cycling conditions applied for *BRCA2* exon 3 amplification. The PCR product were then sequenced in both directions using BigDye Terminator v3.1 (Life Technologies, Carlsbad, CA,) and Alu specific primers in the ABI 3500XL Genetic Analyzer (Life Technologies, Carlsbad, CA), according to manufacturer's instructions.

Haplotype analysis for high frequency *BRCA1* benign variants

We performed a haplotype analysis in order to assess if five high frequency *BRCA1* variants (c.*421G>T, p.Pro871Leu, p.Glu1038Gly, p.Lys1183Arg, and p.Ser1613Gly) were segregating together and were associated to HBOC risk. Based on previous results of our group, that also found these *BRCA1* variants presenting a high frequency in a small HBOC cohort (unpublished data), we joined the two HBOC cohorts (total n=119) and also genotyped 108 additional control samples for the five *BRCA1* SNVs (total n=136) to perform a more accurate statistical analysis.

Additionally, in order to assess the frequency of those five *BRCA1* SNVs in other Brazilian populations, we genotyped 94 HBOC versus 94 control samples from Porto Alegre Clinics

Hospital (Porto Alegre, RS, Brazil); 171 HBOC versus 185 control samples from A.C. Camargo Cancer Center (São Paulo, SP, Brazil), and also 72 HBOC versus 72 control samples from Barretos Cancer Hospital (Barretos, SP, Brazil). We then performed the haplotype analysis.

We applied a TaqMan Allele Discrimination assay (Applied Biosystems, Foster City, CA), using designed probes and primers specific to each *BRCA1* variant: c.*421G>T (assay ID: AHX1AK8), p.Pro871Leu (assay ID: C__2287943_10), p.Glu1038Gly (assay ID: C_2287888_10), p.Lys1183Arg (C__2287889_20), and p.Ser1613Gly (assay ID: C_2615208_20). For each reaction, we used 2 µL of each sample at 5ng/µL, 5 µL of TaqMan master mix (Applied Biosystems, Foster City, CA), and 0.25 µL (200 nM) of each probe, to a final volume of 10 µL, placed in 96-well PCR plates. The cycling conditions were 95°C for 10 min, 40 cycles of 92°C for 15 s and 60°C for 1 min, and 60°C for 1 min, and a final extension of 72°C for 10 min. The amplification was performed using the 7500 Real-Time PCR Systems (Applied Biosystems, Foster City, CA) and the results were analysed using the manufacturer's software.

After that, we assessed the haplotype frequency estimation for all samples using the haplo.stats package version 1.7.9 (<https://cran.r-project.org/web/packages/haplo.stats/index.html>), on R environment. The haplo.stats analysis also estimates the association among haplotypes and the disease, considering p value < 0.05 as statically significant.

Risk association analysis and statistical tests

For the risk association analysis we used the allele frequencies found in our cohort, compared to the allele frequencies of the same variants available in the AbraOM public database which includes the exome sequencing data of 609 elderly Brazilians [43]. We decided to use public databases instead of the allele frequencies on the control group, due to low number of individuals sequenced. When the allele frequencies on AbraOM was zero, we used the European non-Finnish, Latin, American, African and frequencies available on 1000 Genomes [44] or ExAC [45] databases. We performed an odds ratio (OR) analysis applying the Fisher's exact test. The p-values were assessed using the Pearson's X^2 test. For the variants that presented allele frequency equal to zero in the public databases, wasn't possible to reach OR values.

For assessing the clinical and molecular correlations, we applied Pearson's X^2 test.

For those two analysis we used the R commander [46] tools on R environment and considered results as statistically significant at a p-value of 0.05 or less.

For the survival (Kaplan Meier) analysis, we used Logrank test for trend and Mantel-Cox, as recommended by GraphPad Prism 8.1.2. We also assessed the results for the Gehan-Breslow-Wilcoxon test.

Results

Patients clinical characterization

Most of patients (n=84) were diagnosed with breast cancer, showing a prevalence of 82.4% (n=80) of Invasive Ductal Carcinoma (IDC) (Additional files: Table S1). The Luminal and Triple-negative (TN) were the most frequent molecular subtypes, presenting a frequency of 33.3% and 28.6% of BC cases, respectively. In general, most of the patients (n=65) presented tumors of intermediate to high grades (II and III), independently of the age of diagnosis. Only 6.3% (6) were diagnosed with ovarian cancer, of which half of cases were Serous Ovarian Cancer (Table 1, and Additional files: Table S1). One patient presented with diffuse gastric cancer and another, endometrial adenocarcinoma, and both presented with strong history of breast and ovarian cancers in their families. Only one case presented with both asynchronous BC and OC. Most of the cases (85.3%) were diagnosed between 22 and 49 years, and 13.6% (n=13) deceased due to distant metastasis occurrence (Table 1).

Multi-gene panel screening

We identified 667 single nucleotide variants (SNVs) and small insertions/deletions in 94 out of 95 samples screened for variants in their coding and flanking regions of 21 DNA repair genes. One sample were excluded due to a general low quality in the base calling. We then prioritized variants filtering them according to the following criteria: 1 – Variants classified as pathogenic according to ACMG/AMP criteria, and 2 - VUS and benign variants present both in coding and splicing regions, and predicted as damaging/pathogenic by the *in silico* prediction tools. This filtering aimed to select the possible candidate variants without losing variants of unknown significance (VUS) that wasn't characterized yet but may exert some effect to the phenotype. We filtered 82 variants in 19 genes with *RAD50* and *PTEN* presenting no possible candidate variants. Considering those prioritized variants, about 81% of the patients presented variants in *BRCA1* gene, although genes as *FAM175A*, *ATM*, *BRCA2* and *UIMC1* also emerged as presenting a high frequency of variants in our cohort.

Figure 1 shows the most prevalent variants detected in the studied samples. About 11.2% (n=9) were frameshift, stop gain, insertion or missense variants, previously described as pathogenic in *BRCA1*, *BRCA2* and *TP53* genes, with a prevalence of 23.4% (n=22). The most prevalent pathogenic variant was the frameshift p.Gln1756Profs*74 (c.5266dupC) in *BRCA1* (ENSP00000350283.3) gene, present in half cases which beared *BRCA1* mutations (n=11), followed by the variant p.Arg337His (c.1010G>A) in *TP53* (ENST00000269305.8), found in other 5 patients. Our results are also introduce the first report of two known pathogenic variants in the Brazilian population: the p.Tyr3009Serfs*7 (c.9026_9030delATCAT) on *BRCA2*, and p.Arg273His (c.818G>A) in *TP53*. None of those pathogenic variants were found in the control group.

Considering *BRCA1* and *BRCA2* genes, we also found five benign variants in *BRCA1* gene, as presenting a high frequency in our cohort: the 3'UTR c.*421G>T, p.Pro871Leu (c.2612C>T), p.Glu1038Gly (c.3113A>G), p.Lys1183Arg (c.3548A>G), and p.Ser1613Gly (c.4900A>G). Based on previous results of our group that also found those variants in a high frequency in a small HBOC cohort (unpublished data), we sought to investigate whether those variants were segregating together and if they were associated to an increased HBOC risk. Haplotype

analysis by Haplo.Stats program identified 5 haplotypes with frequencies above 1% (Table 3). Haplotype 2, with all five SNVs, was the second most frequent haplotype found (24.8%) in our study. However, this haplotype was significantly more frequent in the control group ($p=0.020$), being not associated to an increased HBOC risk.

To further investigate if there is any correlation between *BRCA1* haplotypes and HBOC risk, we performed the haplotype analysis using HBOC and control samples from other three cancer centers in Brazil: Porto Alegre Clinics Hospital (HPOA), A.C. Camargo Cancer Center (ACC) and Barretos Cancer Hospital (HCB). Haplotype analysis results were similar for all three centers. The Haplotype 2 (Table 3) were not significant in the other three centers (Haplotype in red, Additional files: Table S2), but Haplotype 3, which encompasses only the p.Pro871Leu SNV, showed a significant difference between HBOC and control groups in the three other cancer centers ($p=0.027$; $p=0.007$; $p=0.026$ respectively) (Haplotype in bold, Additional files: Table S2), but also with a higher frequency in the control group, suggesting no correlation with an increased risk to HBOC Syndrome.

Table 3. Haplotype estimation for five high frequency *BRCA1* benign variants found on this cohort.

Hp	p.Pro871Leu (CCG→CTG)	p.Glu1038Gly (GAA→GGA)	p.Lys1183Arg (AAA→AGA)	p.Ser1613Gly (AGT→GGT)	c.*421G>T (G→T)	Hap. freq.	Control (n=136)	HBOC (n=119)	p-value
1	Pro	Glu	Lys	Ser	G	0.546	0.533	0.563	0.532
2	Leu	Gly	Arg	Gly	T	0.248	0.292	0.199	0.020
3	Leu	Glu	Lys	Ser	G	0.136	0.129	0.143	0.633
4	Pro	Glu	Lys	Ser	T	0.028	0.017	0.038	0.172
5	Leu	Gly	Lys	Gly	T	0.028	0.017	0.038	0.172
6	Leu	Gly	Lys	Ser	T	0.008	0.007	0.007	NA
7	Leu	Gly	Lys	Gly	G	0.004	0.004	0.424	NA
8	Leu	Glu	Arg	Gly	G	0.003	0.003	0.000000002	NA
9	Leu	Gly	Arg	Gly	G	0.002	0.00000002	0.004	NA

Hp = estimated haplotypes. Hap. freq. = haplotype frequency. p-value = haplotype score statistic p-value calculated by Haplo.stats. NA = when the haplotype score statistic p-value could not be calculated.

About 12.8% (n=12) of the patients did not present any variants in the *BRCA1/BRCA2* genes (Table 2). Most cases (76.6%) presented missense VUS or benign missense variants according to VarSome and ClinVar, that were qualified as being pathogenic by the *in silico* prediction tools, which may unable the clinical interpretation and risk estimation during the genetic counselling for carriers. The association study with those variants identified 8 genes carrying 13 variants as significantly associated with an increased risk to HBOC when compared to the allele frequencies described in public databases. Genes as *BARD1*, *CHEK2*, *PALB2* and *PMS2* presented more than one variant associated with risk (Fig. 2).

The prevalence of variants associated with HBOC was about 16% (n=15), and most of them (n=13) were present in double heterozygosis variants with conflicting data on pathogenicity in *BRCA1/BRCA2*. *BARD1*, *CHEK2*, *PALB2* and *PMS2* presented more than one variant associated with risk (Fig. 3), and the variant p.Ala617Thr (c.1849G>A) in *CDH1* gene presented the highest allele frequency (AF=0.01595745). One patient presented a pathogenic variant in *BRCA1* in double heterozygosis with one *BARD1* prioritized variant (Table 2).

The variants associated with risk in *BARD1*, *CHEK2*, *FAM175A*, *PALB2* and *PMS2* genes were not found in the elderly samples used as control, and this study constitutes the first report of those variants in the Brazilian population.

All patients carrying variants associated with an increased risk, as well those who didn't present any *BRCA1/BRCA2* variants tested negative for *BRCA1/BRCA2* CNVs.

Clinical characteristics of germline variants-carriers

The prevalence of pathogenic variants in *BRCA1* and *BRCA2* was about 18% (n=17), with only four patients presenting *BRCA2* pathogenic variants. We observed that 90% of carriers of *BRCA1* pathogenic variants presented with high grade tumors (grade III) while about 80% of *BRCA2* carriers presented with tumors with grades I and II. Additionally, most of *BRCA1*-variant carriers were diagnosed with triple negative BC (Fig. 1). The non-*BRCA1/2* group also presented high frequency of intermediate to high grades tumors (grades II and III) (Fig. 1, Table 1) which may suggest that other genes are associated to moderately-poorly differentiated tumors like it is known for *BRCA1/BRCA2*-carriers [47]. The presence of metastasis was strongly correlated with death ($p=7.85e-12$) once 13 out of 14 patients that deceased presented distant metastasis.

A total of 12 individuals (12.8%) did not present any variants or CNVs in *BRCA1/BRCA2* and were grouped as non-*BRCA1/2* patients. This group presented variants in *ATM*, *ATR*, *BARD1*, *CDH1*, *FAM175A*, *MLH1*, *MSH6*, *PMS2*, *TP53* and *UIMC1* genes. All non-*BRCA1/2* patients were BC cases, showing a median age at diagnosis of 36.5 years and a median survival of 8 years (Table 1). However, we did not observe any association of death with the genotype of the patients. Surprisingly, the patients that presented pathogenic variants in *BRCA1/BRCA2* showed a trend toward better survival with most of cases that died being the ones that presented VUS, benign or no variants in *BRCA1/2* genes (Fig. 4).

Figure 4. Survival of patients after clinical diagnosis according to the genotype regarding the presence of *BRCA1/BRCA2* variants. The grey lines represent the censored data (cases that did not deceased) and the time of follow-up after clinical diagnosis, since we studied patients diagnosed with cancer 28 years ago and some diagnosed 4 years ago. Conflicting data on pathogenicity refers to VUS and benign variants that were predicted as pathogenic by the *in silico* tools. We did not find any significant difference between the genotypes.

Discussion

Genes as *BRCA1*, *BRCA2* and *TP53* presented pathogenic variants in 23.4% (n=22) of the investigated cases. The only study with a multi-gene analysis in Brazil has shown genes as *BRCA1*, *BRCA2*, *ATM*, *ATR*, *MLH1*, *MSH2* and *MSH6* carrying pathogenic variants but with a much lower frequency (9.5%) [25].

The most prevalent variant was the frameshift p.Gln1756Profs*74 (c.5266dupC) in *BRCA1*, found in 11.7% of patients. This variant was also described in the study of Timoteo et al. (2018) [25], but with a frequency of only 3%. This variant is commonly found in South American populations, being well described in Brazil, especially in ovarian cancer cases [48, 49], although it has been found only in breast cancer cases in our cohort. It is a founder Ashkenazi Jewish variant and it is very common among North European BC affected populations [50]. It may explain the high frequency found in the Southeast of Brazil, which is marked by a strong European ancestry [51].

Four patients presented variants in *BRCA2* genes: p.Ala938Profs*21; p.Tyr3009Serfs*7; p.Arg3128Ter and, the third most common variant in Brazilian population, c.156_157insAlu. The Alu retroelements are fragments of approximately 300 nucleotides that are reported as inserted in many genes such as *BRCA1* and *BRCA2* and are related to an increased cancer risk [52, 53]. The Alu insertion in *BRCA2* exon 3 was first reported by Teugels et al (2005) [54] as a Portuguese founder variant in HBOC patients, and due to the Portuguese immigration during Brazil colonization, this variant is frequently found in Brazilian populations [52]. The pathogenicity of this insertion is attributed to the exon 3 skipping, which causes the loss of the PALB2 and RAD51 binding region, essential to HR repair [55].

Five patients also presented the pathogenic variant p.Arg337His in *TP53* gene. This is a founder variant of South Brazil, known as segregating in families with sarcomas, adrenocortical and choroid plexus carcinomas, and breast cancer at early onset [31, 56]. It is located in the oligomerization domain of p53 and besides segregation studies, it has been shown that this variant is associated with a decreased oligomerization and transcriptional activities of p53 [57, 58].

However about 76.6% of the cases presented VUS and variants with conflicting data on pathogenicity in *BRCA1/BRCA2* as well as in other investigated genes based on data from VarSome, ClinVar or pathogenicity tools here employed. In this group we found one patient carrying the previously undescribed variant p.Pro932Thr (c.2794C>A) on *ATR* gene, which is predicted as pathogenic/possibly pathogenic by all *in silico* tools used in this study. This patient also presented variants in other genes as *BRCA1*, *UIMC1* and *MLH1*, but tested negative for *BRCA1/BRCA2* CNVs. It is a case of unilateral BC with lymph node metastasis diagnosed with 40 years old and with a 4 year survival after diagnosis.

For those cases who did not present any pathogenic variant we observed a high frequency of the five *BRCA1* benign variants: the 3'UTR c.*421G>T, p.Pro871Leu (c.2612C>T), p.Glu1038Gly (c.3113A>G), p.Lys1183Arg (c.3548A>G) and p.Ser1613Gly (c.4900A>G). As we can see on Table 3, these variants were segregating together, and constituted the second most frequent haplotype found in this study. Despite this, the haplotype containing the five SNVs was significantly more frequent in controls (29.2%) when compared to HBOC cases (19.9%)

($p=0.020$), which suggests that these variants are not associated with an increased risk to HBOC. Indeed, four of these variants were previously described as presenting a high frequency in a healthy cohort in an ethnic dependent manner, with p.Pro871Leu presenting high African and European ancestry, and p.Glu1038Gly, p.Lys1183Arg, and p.Ser1613Gly, associated to the Central Asiatic ethnic component [59]. It may explain the high frequency of these variants in the studied population.

The genes *FAM175A*, *UIMC1* and *ATM* also presented a high frequency of missense variants in our cohort. About 66% of the patients carry the variant p.Ala348Thr (c.1042G>A) in *FAM175A*, which is not characterized by ClinVar but is predicted as pathogenic by 3 *in silico* tools. The allele frequency for this variant was 0.4 in our cohort and population databases describe p.Ala348Thr with a MAF=0.34 in Brazil [43] and MAF=0.42 worldwide [60]. One individual from the control group also presented this variant, corroborating other studies [61], which according to ACMG/AMP criteria this is a benign variant. The p.Pro435Leu (c.1304C>T) in *UIMC1* is another VUS not described on ClinVar that presented a high allele frequency (0.10) in our HBOC cases. It also has an high MAF in the population databases (0.12 [43] and 0.24 [60]). Together with Abraxas, RAP80 is part of the BRCA1-A complex which is important for recruiting BRCA1 to double-strand break (DSB) sites [62] and studies have shown that truncating variants in both proteins are associated with increased irradiation sensitivity, deficient BRCA1 recruitment to DSB sites and genomic instability [63–66]. Three patients that carried only these two variants were evaluated for *BRCA1/BRCA2* CNVs and all tested negative. Due to their high allele frequency, those variants are classified as benign by the ACMG/AMP, however, a more accurate characterization is mandatory to address a clinical significance for those variants, once both are not characterized yet and we cannot discard its contribution to risk following a polygenic inheritance pattern, for example. Another gene that presented high frequency of variants was *ATM* (Fig. 1). About 16.8% out of the patients that presented variants in *ATM* carried the variant p.Asp1853Asn (c.5557G>A), characterized as benign by ClinVar and VarSome. Studies with this variant had shown that it is not associated with an increased risk to HBOC [67] and corroborating it, we found this variant in one of the 28 control samples.

We also observed a high frequency of missense variants in MMR genes, especially for *PMS2* and *MSH2* which were mutated in 19% and 10% of the cases, respectively (Fig. 1). Despite truncating variants in those genes constitute the cause of Lynch Syndrome (LS), it is common to find an overlap between HBOC and LS cases since both syndromes are well known by predisposing to BC and OC [68]. Many studies have reported MMR genes as associated with an increased risk to HBOC [69–71] and indeed, they have been taken into account by NCCN guidelines to the clinical management of patients under risk for hereditary BC and OC [4, 72].

However, most patients (76.6%) carry missense VUS or variants presenting conflicting data on pathogenicity. The association analysis based on Brazilian [43] and worldwide public databases [60] revealed 13 variants in *BARD1*, *CDH1*, *CHEK2*, *FAM175A*, *MLH1*, *PALB2* and *PMS2* genes as associated with HBOC, with a prevalence of 15.9% (Fig. 2). The variant p.Ala617Thr (c.1849G>A) in *CDH1* gene was the most frequent among the studied cases. Differently to the other genes, *CDH1* encodes the adhesion protein E-cadherin and variants

in this gene are associated with defects on cell adhesion, an increase in the invasive activity and, consequently, to metastasis [73]. *CDH1* truncating variants are associated with risk to gastric diffuse cancer and in fact, one patient presented familial history of gastric cancer, however, all three cases presented BC and filled NCCN criteria to HBOC. This variant has been previously described in the Brazilian population as pathogenic [25, 74] but functional assays with cells expressing the mutated protein have shown wild type morphology and normal proliferation and migration activities [75], which suggests this variant may not lead to protein truncation.

The *BARD1* was the gene that presented more variants associated with HBOC risk. *BARD1* form heterodimers with *BRCA1* playing important role as both E3 ubiquitin ligase as homologous repair mediators by recruiting RAD51 to DSB sites [76].

Variants in this genes have been associated with a deficiency on HR and increased sensitivity to DNA damage, classifying *BARD1* as a gene of moderate penetrance to BC and OC [24, 76–78]. All three associated variants are described as VUS on ClinVar, but p.Asn255Ser (c.764A>G) and p.Lys423Arg (c.1268A>G) lack studies characterizing its effects on protein functions. Indeed, this is the first study reporting both variants in a HBOC cohort from Brazil. The third variant p.Leu239Gln (c.716T>A) has been described in the North American population and was also characterized as a VUS [79]. Despite being predicted as likely benign by VarSome, p.Leu239Gln and p.Asn255Ser are predicted as pathogenic by 2 out of 6 *in silico* tools and are located between the RING and ANK *BARD1* domains (Fig. 3A). RING is the region of *BRCA1* binding and it is important for heterodimers formation [80]. p.Leu239Gln was found in double heterozygosis with the pathogenic variant p.Trp1836Ter in *BRCA1*, but p.Asn255Ser was found in a non-*BRCA1/2* BC patient. Regarding p.Lys423Arg variant, it is located in ANK domain which plays important role on apoptosis activation due to p53 binding [81]. Despite ANK is not related to the DNA repair process, the evaluation of variants located between amino acids 460-560 have shown an HR deficiency demonstrating that this domain is also important to a correct DNA repair [76]. In fact, three *in silico* tools classified this variant as pathogenic, however, only functional or segregation analyses are required to confirm the suggested pathogenic effect of those variants.

The role of *BRCA1/BRCA2* genes to the HBOC pathogenesis is already well characterized. The VUS p.Met2775Arg (c.8324T>G) in *BRCA2* was found in one BC patient in double heterozygosis with other associated variants as p.Arg137Gln in *CHEK2* and p.Val717Met in *PMS2*. p.Met2775Arg have been described in prostate cancer cases and is characterized as possibly pathogenic by 4 *in silico* prediction tools despite this variant does not affect an conserved residue [82, 83]. It is located in the C-terminal of *BRCA2* proteins which is important for single strand DNA binding as well as for delivering RAD51 molecules to DSB sites allowing a correct homologous repair [84]. It indicated that the integrity of this region is essential to a correct HR. Taking into account that this patient presented three variants significantly associated with HBOC, we suggest this genotype may play an additive effect on breast cancer risk in this case.

CHEK2 gene also presented two variants associated with risk (Fig 3B). Chk2 play an important role on signalling the DNA damage through phosphorylating effector proteins as *BRCA1* [85]. Both variants p.Arg137Gln and p.Ile160Met are located on FHA domain (Fig.

3B), which after Chk2 phosphorylation and KD domain activation, binds to SCD domains of other Chk2 activated protein, forming dimers that convert into active monomers, signalling the DNA damage [86]. p.Arg137Gln and p.Ile160Met are predicted as pathogenic/possibly pathogenic by two and four *in silico* tools, respectively. However, functional analyses have shown that p.Arg137Gln is not associated with protein instability and HR deficiency [87–89] which corroborates its likely benign classification by VarSome and ClinVar. On the other hand, p.Ile160Met is a VUS that has been related to a moderate HR deficiency [90], and in fact, carriers of p.Ile160Met variant presented a worst clinical condition, presenting bilateral BC and dying after pulmonary, bone and hepatic metastases in this study. Due to the localization and the clinical features, we suggest that p.Ile160Met may play a role in the risk to HBOC.

Besides presenting with the most frequent variant found in this cohort, *FAM175A* also presented the p.Arg163Ser (c.489G>T) variant as significantly associated with HBOC relative risk (Fig. 2). It is a VUS according to VarSome, which is not described by ClinVar but is characterized as pathogenic by 5 of 6 prediction tools. p.Arg163Ser is located in the Pad1 domain in the N-terminal region of FAM175A, an important RAP80 and other signalling proteins binding domain [91]. Both proteins are mandatory for BRCA1 recruitment to DSB sites and variants affecting that region of FAM175A may affect the correct DSBs signalling [63, 92].

The synonymous variant p.Glu102Glu (c.306G>A) in *MLH1* is predicted as likely benign by VarSome, and is characterized as VUS by ClinVar but was associated with HBOC risk (Fig. 2). It affects a splicing region in the end of *MLH1* exon 3. Due to it, p.Glu102Glu is predicted as pathogenic by all *in silico* tools that return pathogenicity scores for synonymous variants (CADD, UMD predictor and mutation taster). The patient carrying this variant is a BC case that also presented other benign variants in *MLH1* and *BRCA1*, a VUS in *UIMC1*, as well as the novel variant p.Pro932Thr in *ATR*. As previously described, truncating variants on MMR proteins are known by increasing the risk for both BC and OC [69–71]. However, there is no further evidences of the deleteriousness of this variant.

Regarding *PALB2* gene, two N-terminal variants were found as associated with HBOC risk. Despite *PALB2* biallelic mutations are associated with Fanconi Anemia, heterozygous variants are known to confer a moderate risk to BC [55, 93]. According to VarSome, p.Arg18Lys (c.53A>G) is a VUS which also presents conflicting interpretations of pathogenicity by ClinVar, and is predicted as pathogenic by 3 *in silico* tools. It is localized in the PALB2 coiled coil domain (Fig. 3C), the BRCA1 binding region, but studies have shown that this variant do not affect the PALB2-BRCA1 interaction although it promotes a reduction on HR activity [94]. This variant was found in two BC patients, with one case being a triple-negative subtype (TNBC) (Table 2, and Additional files: Table S1). The p.Thr317Pro (c.949A>C) is a VUS found in a TNBC case that presented lymph nodes metastasis. It is located nearby the DBD domain, which important for PALB2 DNA binding [55] (Fig. 3C), but differently of p.Arg18Lys, there is no report of this variant in other studies, and it is characterized as possibly pathogenic by two prediction tools.

The last risk-associated gene was *PMS2*, which presented two C-terminal variants located in the MutL domain that together with the N-terminal region constitute the MLH1 binding

region (Fig. 3D). This region is important for MutL α heterodimers formation, necessary to the correct mismatched DNA fragment excision [95]. The p.Val717Met (c.2149G>A) is a VUS that present conflicting information of pathogenicity by ClinVar database and only AlignGVGD do not predict it as pathogenic. Functional assay have demonstrated a protein stability and MMR proficiency, however the samples carrying this variant presented microsatellite instability [96]. The p.Asp792Asn (c.2374G>A) variant was identified in a gastric diffuse cancer case that deceased three year after the diagnosis. It has been described as presenting a moderate decrease in mismatch repair activity [97], which corroborates our analysis association. Due to it, we suggest that those variants may be related to increased risk to HBOC, but segregation studies and functional characterization are mandatory to access the contribution of those variants to HBOC etiology.

Conclusions

Our study comprises the third multi-gene screening in HBOC patients in the Brazilian population, showing a higher frequency of pathogenic variants than previously reported [25]. In addition, our work brings the first report of pathogenic variants on *BRCA2* and *TP53* genes in the Brazilian population, and also depicts the first report of the novel *ATR* missense variant p.Pro932Thr (c.2794C>A). This study also brings a descriptive characterization of variants found in HBOC patients, evidencing about 16% of patients carrying variants significantly associated with HBOC risk, and constituting the first report of missense variants in *BARD1*, *BRCA2*, *CHEK2*, *FAM175A*, *PALB2* and *PMS2* in Brazil. Besides segregation analyses and functional characterization are mandatory to confirm the deleteriousness of the variants described here, these results bring insights on the contribution of other genes to HBOC pathogenesis. Our data also aggregates epidemiologic information about the prevalence of germline variants in DNA repair genes on the Brazilian population, which together with further characterization will may help guide the clinical decision and risk assessment for patient under increased risk to HBOC in the future.

Declarations

Abbreviations

ACMG/AMP = American College of Medical Genetics and Genomics and the Association for Molecular Pathology; ASCO = College of American Pathologists; BC = Breast Cancer; DSB = Double Strand Breaks; ER = Estrogen Receptors; HBOC = Hereditary Breast and Ovarian Cancer Syndrome; HER2 = human epidermal growth factor receptor 2; IDC = Invasive Ductal Carcinoma; indels = insertion/deletions; IR = Irradiation; NCCN = National Comprehensive Cancer Network; NGS = Next Generation Sequencing; OC = Ovarian Cancer; PR = Progesterone Receptors; SNP = Single Nucleotide Polymorphisms; SNVs = Single Nucleotide Variants; TN = Triple-negative; VUS = Variants of Unknown Significance.

Acknowledgments

We are very thankful to all family members and patients who participated in this study. We would like to thank the Cancer Genetic Counseling Service of the General Hospital of the Ribeirão Preto Medical School of University of São Paulo (HCFMRP-USP) for all the support regarding the patients care. We also thank Dr. Iscia Lopes Cendes by providing the sequencing data for the elderly samples used as control group. We appreciate the collaboration of Dr. Jean-Yves Masson who kindly revised and gave special comments regarding the manuscript. We are very thankful to the Regional Blood Center of Ribeirão Preto (FUNDHERP) which offered all the infrastructure needed to perform this study. Our special thanks to the National Council for Scientific and Technological Development (CNPq), which supported the Doctoral scholarship for the first author and made it possible.

Funding

This work was supported by the Government of Brazil through the National Council for Scientific and Technological Development (CNPq) (S.C.S.C Project 141974/2015-0) and the São Paulo Research Foundation (FAPESP) (W.A.S-J grant #2013/08135-2).

Availability of data and materials

The publicly available datasets analyzed during the current study are available in the AbRaOM [43], 1000 genomes [44] and ExAC [45] databases. The authors declare that all relevant data are included in the article and its additional materials files, and that it is also available from the corresponding author by request.

Authors' contributions

SCSC, LAT and VEFF, performed the clinical characterization. SCSC, DBB, RCAR, AAM, KCP, and GAM, performed the multi-gene panel sequencing, PCR standardization, Sanger sequencing validation and MLPA. JCM provided the control samples. SCSC, JRP and RCAR performed the bioinformatics analysis. SCSC, NMC, and LFA worked on the statistical and haplotype analyses. EIP, DMC, PAP provided the samples for haplotype analyses of *BRCA1* SNVs. VEFF and WASJ provided HBOC DNA samples. SCSC, GAM, RCAR, VEFF, and WASJ wrote the manuscript. SCSC had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

Ethics approval and consent to participate

A written informed consent was obtained from all patients and controls, and the ethical approval was granted by the Ethics Research Committee of the General Hospital of the Ribeirão Preto Medical School-USP (n. 2819/2016).

Consent for publication

Not applicable.

Competing interests

Not applicable.

Figure Legends

Figure 1. Molecular and clinical spectrum of prioritized variants found in 94 HBOC samples screened for variants in 21 DNA repair genes.

The graph shows the frequency of variants found per gene, and the effect of each variant according to VarSome. The samples were also classified according to the age at diagnosis, molecular subtype and tumor grade. In molecular subtype, TN= Triple-negative subtype; Lum= both Luminal A and Luminal B subtypes, when presenting positivity to estrogen and/or progesterone receptors and lack HER2 expression; LumHER= Luminal positive for all three markers; HER2= when the HER2 protein is overexpressed with negative estrogen and progesterone receptors; PR= positivity to only progesterone receptors; NI= Not-informed. For the molecular subtypes we also sign the cases that are not BC cases: Ovarian, Stomach and Endometrium. The numbers in the bottom represent the samples' code.

Figure 2. Association analysis of 72 prioritized variants with conflicting data on pathogenicity to HBOC risk.

The risk association analyses were performed comparing the allele frequencies found in our cohort to frequencies found in public databases (*) AbraOM, ExAC and 1000 Genomes. In ClinVar status (\$), B = Benign; LB = Likely Benign; US = Uncertain Significance; P = Pathogenic. The association was done using Fisher's exact test, and the p-values were assessed using the Pearson's χ^2 test. The lack of allele frequencies in the databases unabled the odds ratios (OR) estimation. The variants in red are the ones significantly associated with HBOC risk.

Figure 3. Schematic representation of BARD1, CHK2, PALB2 and PMS2 proteins, and the variants associated with increased risk to HBOC. (A) Linear representation of BARD1 protein depicting the RING, Ankyrin (ANK), and BRCT domain boundaries [98], and the three variants found in that gene;(B) CHK2 depicting the SQ/TQ cluster domain (SCD), forkhead-associated domain (FHA), and the kinase domain (KD) [99], showing the localization of the two variants found in that gene; (C)PALB2 protein with its main domains depicted: coiled coil, ChAM, MRG15-binding domain I and II (MBD I and II), WD40 repeats domain, and the nuclear exportation site (NES) [55], showing the variants found as significantly associated HBOC risk; and (D) PMS2 with its ATP and MLH1 binding domains, and its endonuclease activity site [100], depicting the variants identified in that gene. The graphs were built using the lolliplot function of the GenVisR package, on R environment, and were adapted by the authors.

Figure 4. Survival of patients after clinical diagnosis according to the genotype regarding

the presence of BRCA1/BRCA2 variants. The grey lines represent the censored data (cases that did not die) and the time of follow-up after clinical diagnosis, since we studied patients diagnosed with cancer 28 years ago and some diagnosed 4 years ago. Conflicting data on pathogenicity refers to VUS and benign variants that were predicted as pathogenic by the in silico tools. We did not find any significant difference between the genotypes.

Table legends

Table 1. Phenotypic and genotypic characterization of the cohort according to BRCA mutational status.

* Variants previously characterized as pathogenic (ClinVar). † Patients carrying benign or variants of unknown significance on *BRCA1/BRCA2* genes. ‡ The association between the genotypes and the clinical characteristics were calculated using the Pearson's χ^2 test. HER2 = when the HER2 protein is overexpressed; TN = Triple-negative; PR = Positive for progesterone receptors. NI = Not-informed.

Table 2. Variants found in the cohort and its pathogenicity prediction.

Table 3. Haplotype estimation for five high frequency BRCA1 SNVs found in this cohort.

Hp = estimated haplotypes. Hap. freq. = haplotype frequency. p-value = haplotype score statistic p-value calculated by Haplo.stats. NA = when the haplotype score statistic p-value could not be calculated.

Additional files

Table S1. Clinical characterization of the HBOC cohort.

NI = Not-informed; CNS = Central Nervous System. The tumor size, lymph node staging and the metastasis status are reported according to MOC Brazil guidelines for tumor staging (<https://mocbrasil.com/>).

Table S2. Haplotype estimation for five high frequency BRCA1 SNVs in three different Cancer Centers in Brazil.

Here we show only haplotypes with frequencies higher than 1%. In red, the haplotype found as significantly more frequent in the control group in our cohort. In bold, the haplotype that was significantly more frequent in all three other Brazilian Cancer Centers. HPOA = Hospital da Clínicas de Porto Alegre, Porto Alegre, RS, Brazil; ACC = A.C. Camargo Cancer Center, São Paulo, SP, Brazil; HCB = Barretos Cancer Hospital, Barretos, SP, Brazil. Hp = estimated haplotypes. Hap. freq. = haplotype frequency. p-value = haplotype score statistic p-value calculated by Haplo.stats. NA = when the haplotype score statistic p-value could not be calculated.

References

1. Al Bakir M, Gabra H. The molecular genetics of hereditary and sporadic ovarian cancer: implications for the future. *Br Med Bull*. 2014;112:57–69. doi:10.1093/bmb/ldu034.
2. Silva FC, Lisboa BC, Figueiredo MC, Torrezan GT, Santos EM, Krepschi AC, et al. Hereditary breast and ovarian cancer: assessment of point mutations and copy number variations in Brazilian patients. *BMC Med Genet*. 2014;15:55. doi:10.1186/1471-2350-15-55.
3. Roy R, Chun J, Powell SN. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat Rev Cancer*. 2012;12:68–78. doi:10.1038/nrc3181.
4. Clinical N, Guidelines P, Guidelines N. Genetic / Familial High-Risk Assessment : Breast and Ovarian. 2015.
5. Pritchard CC. New name for breast-cancer syndrome could help to save lives. *Nature*. 2019.
6. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* (80-). 1990.
7. Miki Y, Swensen J, Shattuck-eidens D, Futreal PA, Tavtigian S, Liu Q, et al. A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1 Published by : American Association for the Advancement of Science Stable URL : <http://www.jstor.org/stable/2884716>. *Science* (80-). 1994;266:66–71.
8. Wooster, G B, J L, S S, S S, J M, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature*. 1995;378:789–92.
9. Kotsopoulos J. BRCA Mutations and Breast Cancer Prevention. *Cancers* (Basel). 2018.
10. Couch FJ, Nathanson KL, Offit K. Two Decades After BRCA: Setting Paradigms in Personalized Cancer Care and Prevention. *Science* (80-). 2014;343:1466–70. doi:10.1126/science.1251827.
11. Plevritis SK, Kurian AW, Sigal BM, Daniel BL, Ikeda DM, Stockdale FE, et al. Cost-effectiveness of Screening BRCA1/2 Mutation Carriers With Breast Magnetic Resonance Imaging Sylvia. *JAMA*. 2006;295:2374–84.
12. Romero-Laorden N, Castro E. Inherited mutations in DNA repair genes and cancer risk. *Curr Probl Cancer*. 2017;41:251–64. doi:10.1016/j.currproblcancer.2017.02.009.
13. Lynch HT, Snyder C, Casey MJ. Hereditary ovarian and breast cancer: what have we learned? *Ann Oncol*. 2013;24 Suppl 8 Supplement 8:viii83–95. doi:10.1093/annonc/mdt313.
14. Aloraifi F, Boland MR, Green AJ, Geraghty JG. Gene analysis techniques and susceptibility gene discovery in non-BRCA1/BRCA2 familial breast cancer. *Surg Oncol*. 2015;24:100–9.
15. Darooei M, Poornima S, Salma BU, Iyer GR, Pujar AN, Annapurna S, et al. Pedigree and BRCA gene analysis in breast cancer patients to identify hereditary breast and ovarian cancer syndrome to prevent morbidity and mortality of disease in Indian population. *Tumor Biol*. 2017;39:101042831769430. doi:10.1177/1010428317694303.
16. Tedaldi G, Tebaldi M, Zampiga V, Danesi R, Arcangeli V, Ravegnani M, et al. Multiple-gene panel analysis in a case series of 255 women with hereditary breast and ovarian cancer. *Oncotarget*. 2017. doi:10.18632/oncotarget.16791.
17. Buys SS, Sandbach JF, Gammon A, Patel G, Kidd J, Brown KL, et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. *Cancer*. 2017;:1–10. doi:10.1002/cncr.30498.

18. Schroeder C, Faust U, Sturm M, Hackmann K, Grundmann K, Harmuth F, et al. HBOC multi-gene panel testing: comparison of two sequencing centers. *Breast Cancer Res Treat.* 2015;152:129–36.
19. Lee JY, Kim J, Kim SW, Park SK, Ahn SH, Lee MH, et al. BRCA1/2-negative, high-risk breast cancers (BRCAX) for Asian women: genetic susceptibility loci and their potential impacts. *Sci Rep.* 2018;8:1–11.
20. Aloraifi F, Boland MR, Green AJ, Geraghty JG. Gene analysis techniques and susceptibility gene discovery in non-BRCA1/BRCA2 familial breast cancer. *Surg Oncol.* 2015;24:100–9. doi:10.1016/j.suronc.2015.04.003.
21. Hirotsu Y, Nakagomi H, Sakamoto I, Amemiya K, Oyama T, Mochizuki H, et al. Multigene panel analysis identified germline mutations of DNA repair genes in breast and ovarian cancer. *Mol Genet Genomic Med.* 2015;3:459–66. doi:10.1002/mgg3.157.
22. Minion LE, Dolinsky JS, Chase DM, Dunlop CL, Chao EC, Monk BJ. Hereditary predisposition to ovarian cancer, looking beyond BRCA1/BRCA2. *Gynecol Oncol.* 2015;137:86–92. doi:10.1016/j.ygyno.2015.01.537.
23. Schubert S, van Luttikhuisen JL, Auber B, Schmidt G, Hofmann W, Penkert J, et al. The identification of pathogenic variants in BRCA1/2 negative, high risk, hereditary breast and/or ovarian cancer patients: High frequency of FANCM pathogenic variants. *Int J Cancer.* 2019.
24. Suszynska M, Klonowska K, Jasinska AJ, Kozlowski P. Large-scale meta-analysis of mutations identified in panels of breast/ovarian cancer-related genes — Providing evidence of cancer predisposition genes. *Gynecol Oncol.* 2019; xxxx:1–11.
25. Timoteo AR de S, Gonçalves AÉMM, Sales LAP, Albuquerque BM, de Souza JES, de Moura PCP, et al. A portrait of germline mutation in Brazilian at-risk for hereditary breast cancer. *Breast Cancer Res Treat.* 2018;172:637–46. doi:10.1007/s10549-018-4938-0.
26. Toss A, Tomasello C, Razzaboni E, Contu G, Grandi G, Cagnacci A, et al. Hereditary Ovarian Cancer: Not Only BRCA 1 and 2 Genes. *Biomed Res Int.* 2015;2015 Figure 1:341723. doi:10.1155/2015/341723.
27. Torrezan GT, Fernanda G, Almeida SR De. Complex Landscape of Germline Variants in Brazilian Patients With Hereditary and Early Onset Breast Cancer. 2018;9 May:1–11.
28. Fernandes GC, Michelli RAD, Galvão HCR, Paula AE, Pereira R, Andrade CE, et al. Prevalence of BRCA1/BRCA2 mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. *Oncotarget.* 2016;7:80465–81.
29. Cipriano NM, de Brito AM, de Oliveira ES, de Faria FC, Lemos S, Rodrigues AN, et al. Mutation screening of TP53, CHEK2 and BRCA genes in patients at high risk for hereditary breast and ovarian cancer (HBOC) in Brazil. *Breast Cancer.* 2019;26:397–405. doi:10.1007/s12282-018-00938-z.
30. Palmero EI, Alemar B, Scheller-Faccini L, Hainaut P, Moreira-Filho CA, Ewald IP, et al. Screening for germline BRCA1, BRCA2, TP53 and CHEK2 mutations in families at-risk for hereditary breast cancer identified in a population-based study from Southern Brazil. *Genet Mol Biol.* 2016;39:210–22.
31. Cury NM, Ferraz VEF, Silva WA. TP53 p.R337H prevalence in a series of Brazilian

- hereditary breast cancer families. *Hered Cancer Clin Pract*. 2014;12:8. doi:10.1186/1897-4287-12-8.
32. Yi M, Huo L, Koenig KB, Mittendorf EA, Meric-Bernstam F, Kuerer HM, et al. Which threshold for ER positivity? a retrospective study based on 9639 patients. *Ann Oncol*. 2014.
 33. Skidmore ZL, Wagner AH, Lesurf R, Campbell KM, Kunisaki J, Griffith OL, et al. GenVisR: Genomic Visualizations in R. *Bioinformatics*. 2016;32:3012–4.
 34. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32:2847–9.
 35. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015.
 36. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. VarSome: the human genomic variant search engine. *Bioinformatics*. 2019.
 37. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014.
 38. Tavtigian S V., Byrnes GB, Goldgar DE, Thomas A. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. *Hum Mutat*. 2008.
 39. Frédéric MY, Lalande M, Boileau C, Hamroun D, Claustres M, Bérout C, et al. UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity - Application to four genes: FBN1, FBN2, TGFB1, and TGFB2. *Hum Mutat*. 2009.
 40. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003.
 41. Adzhubei I, Schmidt S, Peshkin L, Ramensky V, Gerasimova A, Bork P, et al. PolyPhen-2 : prediction of functional effects of human nsSNPs. *Nat Methods*. 2010.
 42. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014.
 43. Naslavsky MS, Yamamoto GL, de Almeida TF, Ezquina SAM, Sunaga DY, Pho N, et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum Mutat*. 2017.
 44. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526:68–74. doi:10.1038/nature15393.
 45. Exome Aggregate Consortium. ExAC Browser. Online. 2016.
 46. Fox J. The R Commander: A Basic-Statistics Graphical User Interface to R. *J Stat Softw*. 2015.
 47. van der Groep P, van der Wall E, van Diest PJ. Pathology of hereditary breast cancer. *Cell Oncol*. 2011;34:71–88. doi:10.1007/s13402-011-0010-3.
 48. Mauro C de C, Jbili R, Carraro DM, de Queiroz Soares DC, da Costa AABA, de Brot L, et al. Prevalence of BRCA1 and BRCA2 pathogenic and likely pathogenic variants in non-selected ovarian carcinoma patients in Brazil. *BMC Cancer*. 2019;19:1–9.

49. Maistro S, Teixeira N, Encinas G, Katayama MLH, Niewiadonski VDT, Cabral LG, et al. Germline mutations in BRCA1 and BRCA2 in epithelial ovarian cancer patients in Brazil. *BMC Cancer*. 2016;16:934. doi:10.1186/s12885-016-2966-x.
50. Hamel N, Feng BJ, Foretova L, Stoppa-Lyonnet D, Narod SA, Imyanitov E, et al. On the origin and diffusion of BRCA1 c.5266dupC (5382insC) in European populations. *Eur J Hum Genet*. 2011.
51. Pena SDJ, di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy F de SG, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One*. 2011.
52. Felicio PS, Alemar B, Coelho AS, Berardinelli GN, Melendez ME, Lengert AVH, et al. Screening and characterization of BRCA2 c.156_157insAlu in Brazil: Results from 1380 individuals from the South and Southeast. *Cancer Genet*. 2018;228–229:93–7. doi:10.1016/j.cancergen.2018.09.001.
53. Macedo GS, Alemar B, Ashton-Prolla P. Reviewing the characteristics of BRCA and PALB2-related cancers in the precision medicine era. *Genet Mol Biol*. 2019.
54. Teugels E, De Brakeleer S, Goelen G, Lissens W, Sermijn E, De Grève J. De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Hum Mutat*. 2005.
55. Ducy M, Sesma-Sanz L, Guitton-Sert L, Lashgari A, Gao Y, Brahiti N, et al. The Tumor Suppressor PALB2: Inside Out. *Trends Biochem Sci*. 2019;44:226–40.
56. Hahn EC, Bittar CM, Vianna FSL, Netto CBO, Biazús JV, Cericatto R, et al. TP53 p.Arg337His germline mutation prevalence in Southern Brazil: Further evidence for mutation testing in young breast cancer patients. *PLoS One*. 2018.
57. Park J-H, Wang P-Y, Hwang PM. Modeling the prevalent germline TP53 R337H mutation in mouse. *Oncotarget*. 2019.
58. Park J-H, Li J, Starost MF, Liu C, Zhuang J, Chen J, et al. Mouse Homolog of the Human TP53 R337H Mutation Reveals Its Role in Tumorigenesis. *Cancer Res*. 2018;78:5375–83. doi:10.1158/0008-5472.CAN-18-0016.
59. Bodian DL, McCutcheon JN, Kothiyal P, Huddleston KC, Iyer RK, Vockley JG, et al. Germline variation in cancer-susceptibility genes in a healthy, ancestrally diverse cohort: Implications for individual genome sequencing. *PLoS One*. 2014.
60. 1000 Genomes Project Consortium. An integrated map of genetic variation. *Nature*. 2012.
61. Solyom S, Aressy B, Pyrkäs K, Patterson-fortin J, Mannermaa A, Greenberg RA, et al. NIH Public Access. 2013;4.
62. Her J, Soo Lee N, Kim Y, Kim H. Factors forming the BRCA1-A complex orchestrate BRCA1 recruitment to the sites of DNA damage. *Acta Biochim Biophys Sin (Shanghai)*. 2016;48:658–64. doi:10.1093/abbs/gmw047.
63. Renault A-L, Lesueur F, Coulombe Y, Gobeil S, Soucy P, Hamdi Y, et al. ABRAXAS (FAM175A) and Breast Cancer Susceptibility: No Evidence of Association in the Breast Cancer Family Registry. *PLoS One*. 2016;11:e0156820. doi:10.1371/journal.pone.0156820.
64. Castillo A, Paul A, Sun B, Huang TH, Wang Y, Yazinski SA, et al. The BRCA1-interacting protein Abraxas is required for genomic stability and tumor suppression. *Cell*

Rep. 2014.

65. Kim H, Chen J, Yu X. Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science* (80-). 2007.
66. Li Y, Luo K, Yin Y, Wu C, Deng M, Li L, et al. USP13 regulates the RAP80-BRCA1 complex dependent DNA damage response. *Nat Commun*. 2017.
67. Jin B, Jiang F, Liu W, Chen N, Ding Z. Are polymorphisms of the ataxia telangiectasia mutated gene associated with breast cancer risk? *Breast Cancer Res Treat*. 2011;128:293–5. doi:10.1007/s10549-011-1385-6.
68. Espenschied CR, Laduca H, Li S, McFarland R, Gau C-L, Hampel H. Multigene Panel Testing Provides a New Perspective on Lynch Syndrome. *J Clin Oncol*. 2017;35:2568–75.
69. Castéra L, Krieger S, Rousselin A, Legros A, Baumann J-J, Bruet O, et al. Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. *Eur J Hum Genet*. 2014; October 2013:1305–13. doi:10.1038/ejhg.2014.16.
70. Kapoor NS, Curcio LD, Blakemore CA, Bremner AK, McFarland RE, West JG, et al. Multigene Panel Testing Detects Equal Rates of Pathogenic BRCA1/2 Mutations and has a Higher Diagnostic Yield Compared to Limited BRCA1/2 Analysis Alone in Patients at Risk for Hereditary Breast Cancer. *Ann Surg Oncol*. 2015. doi:10.1245/s10434-015-4754-2.
71. Bonache S, Esteban I, Moles-Fernández A, Tenés A, Duran-Lozano L, Montalban G, et al. Multigene panel testing beyond BRCA1/2 in breast/ovarian cancer Spanish families and clinical actionability of findings. *J Cancer Res Clin Oncol*. 2018;144:2495–513. doi:10.1007/s00432-018-2763-9.
72. Rashid MU, Naeemi H, Muhammad N, Loya A, Yusuf MA, Lubiński J, et al. A novel deleterious c.2656G>T MSH2 germline mutation in a Pakistani family with a phenotypic overlap of hereditary breast and ovarian cancer and Lynch syndrome. *Hered Cancer Clin Pract*. 2016;14:1–6. doi:10.1186/s13053-016-0056-3.
73. Mateus AR, Simões-Correia J, Figueiredo J, Heindl S, Alves CC, Suriano G, et al. E-cadherin mutations and cell motility: a genotype-phenotype correlation. *Exp Cell Res*. 2009;315:1393–402.
74. Lajus TBP, Sales RMD. CDH1 germ-line missense mutation identified by multigene sequencing in a family with no history of diffuse gastric cancer. *Gene*. 2015;568:215–9. doi:10.1016/j.gene.2015.05.035.
75. Suriano G. E-cadherin germline missense mutations and cell phenotype: evidence for the independence of cell invasion on the motile capabilities of the cells. *Hum Mol Genet*. 2003;12:3007–16. doi:10.1093/hmg/ddg316.
76. Adamovich AI, Tapahsama B, Wingo M, Kathryn D, Jie N, Fernanda Martins R, et al. Functional analysis of BARD1 missense variants in homology-directed repair and damage sensitivity. *PLOS Genet*. 2019;15:1–21.
77. Quezada Urban R, Díaz Velásquez C, Gitler R, Rojo Castillo M, Sirota Toporek M, Figueroa Morales A, et al. Comprehensive Analysis of Germline Variants in Mexican Patients with Hereditary Breast and Ovarian Cancer Susceptibility. *Cancers (Basel)*. 2018;10:361. doi:10.3390/cancers10100361.
78. Weber-Lassalle N, Borde J, Weber-Lassalle K, Klaschik K, Neidhardt G, Ernst C, et al.

- Germline loss-of-function variants in the BARD1 gene are associated with familial breast cancer. Inf und Program www.senologiekongress.de. 2018;15:1–6.
79. Tung N, Lin NU, Kidd J, Allen BA, Singh N, Wenstrup RJ, et al. Frequency of germline mutations in 25 cancer susceptibility genes in a sequential series of patients with breast cancer. *J Clin Oncol*. 2016.
 80. Cimmino F, Formicola D, Capasso M. Dualistic Role of BARD1 in Cancer. *Genes (Basel)*. 2017;8:375. doi:10.3390/genes8120375.
 81. Feki A, Jefford CE, Berardi P, Wu JY, Cartier L, Krause KH, et al. BARD1 induces apoptosis by catalysing phosphorylation of p53 by DNA-damage response kinase. *Oncogene*. 2005.
 82. Maier C, Herkommer K, Luedeke M, Rinckleb A, Schrader M, Vogel W. Subgroups of familial and aggressive prostate cancer with considerable frequencies of BRCA2 mutations. *Prostate*. 2014.
 83. Yang H, Jeffrey PD, Miller J, Kinnucan E, Sun Y, Thoma NH, et al. BRCA2 function in DNA binding and recombination from a BRCA1-DSS1-ssDNA structure. *Science* (80-). 2002;297 September:1837–47.
 84. Von Nicolai C, Ehlén Å, Martin C, Zhang X, Carreira A. A second DNA binding site in human BRCA2 promotes homologous recombination. *Nat Commun*. 2016.
 85. Craig AL, Hupp TR. The regulation of CHK2 in human cancer. *Oncogene*. 2004;23:8411–8.
 86. Zannini L, Delia D, Buscemi G. CHK2 kinase in the DNA damage response and beyond. *J Mol Cell Biol*. 2014;6:442–57.
 87. Bell DW, Kim SH, Godwin AK, Schiripo TA, Harris PL, Haserlat SM, et al. Genetic and functional analysis of CHEK2 (CHK2) variants in multiethnic cohorts. *Int J Cancer*. 2007.
 88. Desrichard A, Bidet Y, Uhrhammer N, Bignon YJ. CHEK2 contribution to hereditary breast cancer in non-BRCA families. *Breast Cancer Res*. 2011.
 89. Sodha N, Mantoni TS, Tavtigian S V., Eeles R, Garrett MD. Rare germ line CHEK2 variants identified in breast cancer families encode proteins that show impaired activation. *Cancer Res*. 2006.
 90. Roeb W, Higgins J, King MC. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. *Hum Mol Genet*. 2012.
 91. Wu Q, Paul A, Su D, Mehmood S, Foo TK, Ochi T, et al. Structure of BRCA1-BRCT/Abraxas Complex Reveals Phosphorylation-Dependent BRCT Dimerization at DNA Damage Sites. *Mol Cell*. 2016;61:434–48.
 92. Wang B, Hurov K, Hofmann K, Elledge SJ. NBA1, a new player in the Brcal A complex, is required for DNA damage resistance and checkpoint control. *Genes Dev*. 2009.
 93. Nepomuceno TC, De Gregoriis G, de Oliveira FMB, Suarez-Kurtz G, Monteiro AN, Carvalho MA. The role of PALB2 in the DNA damage response and cancer predisposition. *Int J Mol Sci*. 2017;18:1–20.
 94. Foo TK, Tischkowitz M, Simhadri S, Boshari T, Zayed N, Burke KA, et al. Compromised BRCA1-PALB2 interaction is associated with breast cancer risk. *Oncogene*. 2017; January:1–10. doi:10.1038/onc.2017.46.
 95. de Barros AC, Takeda AAS, Dreyer TR, Velazquez-Campoy A, Kobe B, Fontes MRM.

DNA mismatch repair proteins MLH1 and PMS2 can be imported to the nucleus by a classical nuclear import pathway. *Biochimie*. 2018.

96. González-Acosta M, del Valle J, Navarro M, Thompson BA, Iglesias S, Sanjuan X, et al. Elucidating the clinical significance of two PMS2 missense variants coexisting in a family fulfilling hereditary cancer criteria. *Fam Cancer*. 2017.

97. Arora S, Huwe PJ, Sikder R, Shah M, Browne AJ, Lesh R, et al. Functional analysis of rare variants in mismatch repair proteins augments results from computation-based predictive methods. *Cancer Biol Ther*. 2017;18:519–33. doi:10.1080/15384047.2017.1326439.

98. Karppinen SM, Heikkinen K, Rapakko K, Winqvist R. Mutation screening of the BARD1 gene: evidence for involvement of the Cys557Ser allele in hereditary susceptibility to breast cancer. *J Med Genet*. 2004.

99. Cai Z, Chehab NH, Pavletich NP. Structure and Activation Mechanism of the CHK2 DNA Damage Checkpoint Kinase. *Mol Cell*. 2009.

100. Borrás E, Chang K, Pande M, Cuddy A, Bosch JL, Bannon SA, et al. In silico systems biology analysis of variants of uncertain significance in lynch syndrome supports the prioritization of functional molecular validation. In: *Cancer Prevention Research*. 2017.

Gene_mutation	rs	HBOC (n=94) (Allele frequency)	Populational databases * (Allele frequency)	ClinVar status \$	P-val
ATM p.Arg2580Ser	rs199915459	0.00531915	0.00082100	Uncertain significance	0.1277
ATM p.Asp1853Asn	rs1801516	0.10638298	0.09359600	Benign/Likely benign	0.5784
ATM p.Asp1853Val	rs1801673	0.01006383	0.00656800	Benign/Likely benign	0.5365
ATM p.Asp814Glu	rs3218695	0.01006383	0.00492600	Benign	0.3325
ATM p.Gly514Asp	rs2235000	0.01006383	0.01888300	Benign/Likely benign	0.4259
ATM p.Leu1420Phe	rs1800058	0.00531915	0.00903100	Conflicting interpretations of pathogenicity (B, LB, US)	0.6066
ATM p.Leu2332Pro	rs4988111	0.01006383	0.00985200	Benign/Likely benign	0.9195
ATM p.Leu546Val	rs2227924	0.01006383	0.01395700	Benign/Likely benign	0.7137
ATM p.Phe858Leu	rs1800056	0.00531915	0.00821000	Conflicting interpretations of pathogenicity (B, LB, US)	0.6754
ATM p.Pro604Ser	rs2227922	0.00531915	0.00821000	Benign/Likely benign	0.6754
ATR p.Pro2434Ala	rs33972295	0.00531915	0.00328400	Benign/Likely benign	0.6626
ATR p.Pro932Thr	-	0.00531915	NA	-	NA
ATR p.Val316Ile	rs28897764	0.02659574	0.01888300	Benign/Likely benign	0.4811
BARD1 c.-83C>T	rs71579840	0.00531915	0.02006700	Likely benign	0.5864
BARD1 p.Arg658Cys	rs3738888	0.00531915	0.00246300	Conflicting interpretations of pathogenicity (B, LB, US)	0.4938
BARD1 p.Asn255Ser	rs138904906	0.00531915	0.00004589	Uncertain significance	2.2e-16
BARD1 p.Leu239Gln	rs200359745	0.00531915	0.00010450	Uncertain significance	6.19e-9
BARD1 p.Lys423Arg	rs749383704	0.00531915	0.00000604	Uncertain significance	2.2e-16
BRCA1 p.Asn550His	rs56012641	0.00531915	0.00082100	Benign	0.1277
BRCA1 p.Asp693Asn	rs4986850	0.03723404	0.05665000	Benign	0.2732
BRCA1 p.Cys197Cys	rs1799965	0.00531915	0.00142045	Benign	0.1574
BRCA1 p.Gln356Arg	rs1799950	0.03191489	0.04926100	Benign	0.2953
BRCA1 p.Glu1038Gly	rs16941	0.27127660	0.28243000	Benign	0.7515
BRCA1 p.Met1673Ile	rs1799967	0.00531915	0.00821000	Benign	0.6754
BRCA1 p.Ser1040Asn	rs4986852	0.04787234	0.03530400	Benign	0.3954
BRCA1 p.Ser153Ile	rs1800744	0.00531915	0.00328400	Benign	0.6626
BRCA1 p.Ser1613Gly	rs1799966	0.28191489	0.30098700	Benign	0.5948
BRCA2 p.Ala2717Ser	rs28897747	0.00531915	0.00150636	Benign	0.1793
BRCA2 p.Ala2951Thr	rs11571769	0.00531915	0.01313600	Benign	0.3614
BRCA2 p.Arg2034Cys	rs1799954	0.01595745	0.00574700	Benign	0.121
BRCA2 p.Asn289His	rs766173	0.03191489	0.04515600	Benign	0.4068
BRCA2 p.Asp1420Tyr	rs28897727	0.00531915	0.00164200	Benign	0.3092
BRCA2 p.Cys1290Tyr	rs41293485	0.00531915	0.00246300	Benign	0.4938
BRCA2 p.Gly271Arg	rs786204274	0.00531915	0.00164200	Conflicting interpretations of pathogenicity (LB, US)	0.3092
BRCA2 p.Lys2950Asn	rs28897754	0.00531915	0.00082100	Conflicting interpretations of pathogenicity (B, LB, US)	0.1277
BRCA2 p.Met2775Arg	rs80359073	0.00531915	0.00018763	Conflicting interpretations of pathogenicity (LB, US)	2.75e-13
BRIP1 p.Arg173Cys	rs4988345	0.01006383	0.00246300	Benign/Likely benign	0.07966
BRIP1 p.Gln740His	rs45589637	0.00531915	0.00410500	Conflicting interpretations of pathogenicity (LB, US)	0.8121
BRIP1 p.Ile1231Met	rs876659290	0.01006383	NA	Uncertain significance	NA
CDH1 p.Ala617Thr	rs33935154	0.01595745	0.00656800	Benign	0.008427
CHEK2 p.Arg137Gln	rs368570187	0.00531915	0.00020600	Likely benign	0.00004967
CHEK2 p.Ile160Met	rs575910805	0.00531915	0.00003911	Conflicting interpretations of pathogenicity (LB, US)	2.7e-16
FAM175A p.Ala348Thr	rs12642536	0.39893617	0.34154400	-	0.5934
FAM175A p.Arg163Ser	rs535462791	0.00531915	0.00000000	-	2.2e-16
MLH1 p.Glu102Glu	rs63751665	0.00531915	0.00001119	Conflicting interpretations of pathogenicity (LB, US)	2.2e-16
MLH1 p.Ser406Asn	rs41294980	0.00531915	0.00164200	Benign	0.3092
MLH1 p.Val213Met	rs2308317	0.00531915	0.00492600	Benign	0.9432
MLH1 p.Val716Met	rs35831931	0.00531915	0.00144070	Benign	0.1625
MRE11A p.Met701Val	rs1805362	0.00531915	0.00246300	Benign/Likely benign	0.4938
MRE11A p.Ser337Arg	rs115244417	0.00531915	0.00821000	Benign/Likely benign	0.6754
MSH2 p.Asn127Ser	rs17217772	0.02127660	0.00985200	Benign	0.1693
MSH2 p.Ala834Thr	rs63750757	0.00531915	0.00082100	Likely benign	0.1277
MSH2 p.Gly322Asp	rs4987188	0.02127660	0.00903100	Benign	0.1282
MSH6 p.Leu396Val	rs2020908	0.01006383	0.00328400	Benign	0.1499
MSH6 p.Val878Ala	rs2020912	0.00531915	0.01067300	Benign	0.4913
NBN p.Leu68Val	rs1200599843	0.00531915	NA	Uncertain significance	NA
NBN p.Pro266Leu	rs769420	0.00531915	0.00328400	Benign	0.6626
PALB2 p.Lys18Arg	rs138789658	0.01006383	0.00164200	Conflicting interpretations of pathogenicity (B, LB, US)	0.03112
PALB2 p.Thr317Pro	rs587780223	0.00531915	0.00000603	Uncertain significance	2.2e-16
PALB2 p.Val932Met	rs45624036	0.00531915	0.00328400	Benign/Likely benign	0.6626
PMS2 p.Arg20Gln	rs10254120	0.05851064	0.05917200	Benign	0.9097
PMS2 p.Arg563Leu	rs63750668	0.01595745	0.00903100	Likely benign	0.3733
PMS2 p.Asp784Asn	rs143340522	0.00531915	0.00106835	Conflicting interpretations of pathogenicity (LB, US)	0.05498
PMS2 p.Asp792Asn	rs587781265	0.00531915	0.00001219	Uncertain significance	2.2e-16
PMS2 p.Met622Ile	rs1805324	0.00531915	0.01313600	Benign	0.3614
PMS2 p.Val717Met	rs201671325	0.01006383	0.00090540	Conflicting interpretations of pathogenicity (B, LB, US)	0.000421
RAD51 p.Ala55Val	rs145617142	0.00531915	0.00164200	-	0.3092
UIMC1 p.Arg15Trp	rs13167812	0.01006383	0.00738900	-	0.6379
UIMC1 p.Pro435Leu	rs3733876	0.10638298	0.11822700	-	0.6377

A

BARD1

B

CHK2

C

PALB2

D

PMS2

Carvalho et al. Fig. 3

Table 1.

Variable	Mutational status						p-value&
	BRCA pathogenic*		BRCA Benign and VUS ¥		non-BRCA		
	n=17	%	n=65	%	n=12	%	
Gender							
Man			1	1.5			
Woman	17	18.1	64	98.5	12	100	
Age at diagnosis (median)							
	24 - 57 (34)		22 - 72 (37)		31 - 47 (36.5)		
Deaths	1	5.9	11	16.9	2	16.6	0.0927
Survival in years (median)	8		3		8		
Familial history							
Present	14	82.3	52	80	10	83.3	0.294
Absent	3	17.7	11	16.9	2	16.7	
NI			2	3.1			
Tumor site							
Breast	17	100	57	87.7	12	100	0.6034
Ovary			6	9.3			
Edometrium			1	1.5			
Stomach			1	1.5			
Tumor distribution							
Unilateral or located	12	70.6	48	73.8	10	83.3	0.2376
Bilateral (breast)	5	29.4	6	9.3	1	8.3	
Multiple tumors			5	7.7			
NI			6	9.3	1	8.3	
Breast molecular subtype							
Luminal	4	23.5	20	30.8	5	41.7	0.4425
Luminal HER	2	11.8	11	16.9	3	25	
HER2	2	11.8	7	10.8	1	8.3	
TN	9	52.9	13	20	1	8.3	
PR			1	1.5			
NI			13	20	2	16.7	
Tumor grade							
I	1	5.9	7	10.8	1	8.3	0.03686
II	3	17.6	29	44.6	5	41.7	
III	11	64.7	11	16.9	4	33.3	
NI	2	11.8	18	27.7	2	16.7	

Lymphonode						
Present	7	41.2	31	47.7	7	58.3
Absent	8	47.1	16	24.6	3	25
NI	2	11.8	18	27.7	2	16.7
Distant Metastasis						
M0	1	5.9	38	58.5	7	58.3
M1	15	88.2	14	21.5	3	25
NI	1	5.9	13	20	2	16.7

Table 2.

Gene	Variant name		dbSNP ID	Mutation Type	ClinVar	In silico Predictions						Samples
	Coding DNA	Protein				CADD	Align GVG D	UMD PREDICTOR	SIFT	PolyPhen	Mutation Taster	
HR genes												
ATM	c.1541G>A	p.Gly514Asp	rs2235000	missense	Benign/Likely Benign	25.7	Class C65	Polymorphism	Tolerated	Probably damaging	Disease Causing	3664; 4146
	c.1636C>G	p.Leu546Val	rs2227924	missense	Benign/Likely Benign	11.58	Class C25	Polymorphism	Damaging	Possibly damaging	Polymorphism	3617; 3634
	c.1810C>T	p.Pro604Ser	rs2227922	missense	Benign/Likely Benign	23.3	Class C65	Probably polymorphism	Tolerated	Possibly damaging	Disease Causing	2775
	c.2442C>A	p.Asp814Glu	rs3218695	missense	Benign	15.88	Class C35	Polymorphism	Tolerated	Benign	Polymorphism	2753; 2784
	c.2572T>C	p.Phe858Leu	rs1800056	missense	Conflicting interpretations of pathogenicity	13.50	Class C15	Polymorphism	Damaging	Possibly damaging	Polymorphism	4268
	c.4258C>T	p.Leu1420Phe	rs1800058	missense	Conflicting interpretations of pathogenicity	15.47	Class C15	Polymorphism	Tolerated	Benign	Disease Causing	3650
	c.5557G>A	p.Asp1853Asn	rs1801516	missense	Benign/Likely Benign	23.2	Class C15	Polymorphism	Tolerated	Benign	Polymorphism	2699; 2724; 2775; 3002; 3132 (homoz); 3141; 3166; 3187; 3728 (homoz); 4063; 4133; 4135; 4137; 4138 (homoz); 4147; 4226 (homoz)
	c.6995T>C	p.Leu2332Pro	rs4988111	missense	Benign/Likely Benign	15.87	Class C65	Polymorphism	Tolerated	Benign	Polymorphism	3617; 3634
	c.7740A>C	p.Arg2580Ser	rs199915459	missense	Uncertain significance conflicting	15.65	Class C65	Pathogenic	Tolerated	Benign	Disease Causing	3671
c.5558A>T	p.Asp1853Val	rs1801673	missense	interpretations of pathogenicity	24.2	Class C65	Pathogenic	Damaging	Possibly damaging	Disease Causing	4186; 4264	
ATR	c.2794C>A	p.Pro932Thr	-	missense	-	27.0	Class C35	Pathogenic	Damaging	Probably damaging	Disease Causing	4020
	c.7300C>G	p.Pro2434Ala	rs33972295	missense	Benign/Likely Benign	23.5	Class C25	Polymorphism	Damaging	Probably damaging	Disease Causing	4228
	c.946G>A	p.Val316Ile	rs28897764	missense	Benign/Likely Benign	18.61	Class C25	Probably polymorphism	Tolerated	Benign	Disease Causing	2726; 3116; 3671; 3703; 4228
BARD1	c.-83C>T	-	rs71579840	5'UTR premature start codon gain	-	8.485	-	-	-	-	-	3002

	c.1972C>T	p.Arg658Cys	rs3738888	missense	Benign/Likely Benign	26.5	Class C65	Probably pathogenic	Damaging	Probably damaging	Disease Causing	3671
	c.1268A>G	p.Lys423Arg	rs749383704	missense	Uncertain significance	21.8	Class C25	Probably pathogenic	Tolerated	Benign	Disease Causing	2995
	c.764A>G	p.Asn255Ser	rs138904906	missense	Uncertain significance	16.75	Class C45	Polymorphism	Tolerated	Probably damaging	Polymorphism	3716
	c.716T>A	p.Leu239Gln	rs200359745	missense	Uncertain significance	1.061	Class C65	Probably pathogenic	Tolerated	Benign	Polymorphism	3051

	c.*421G>T	-	rs8176318	3'UTR	Benign	4.78	-	-	-	-	-	2697 (homoz); 2699; 2742; 2750; 2753; 2779; 2801; 2815; 2972; 2977; 3002; 3056; 3078; 3097; 3114; 3115; 3116; 3132; 3141; 3166; 3227; 3462; 3617; 3671; 3703; 3728; 3806; 3842; 3897; 4016; 4020; 4135; 4138; 4146; 4147; 4161; 4166; 4177; 4220; 4226; 4228; 4279 2699; 2785; 2995; 3002; 3114; 3876; 4020; 4037; 4132
	c.3119G>A	p.Ser1040Asn	rs4986852	missense	Benign	14.86	Class C45	Polymorphism	Tolerated	Probably damaging	Polymorphism	
	c.5019G>A	p.Met1673Ile	rs1799967	missense	Benign	22.0	Class C0	Probably polymorphism	Tolerated	Benign	Disease Causing	3897
	c.4598G>T	p.Ser1533Ile	rs1800744	missense	Benign	16.14	Class C65	Polymorphism	Damaging	Possibly damaging	Polymorphism	3113
	c.1648A>C	p.Asn550His	rs56012641	missense	Benign	17.67	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	4132
	c.1067A>G	p.Gln356Arg	rs1799950	missense	Benign	17.80	Class C35	Polymorphism	Damaging	Probably damaging	Polymorphism	2724; 2775; 3187; 3703; 4133; 4139
BRCA1	c.2077G>A	p.Asp693Asn	rs4986850	missense	Benign	15.84	Class C15	Polymorphism	Damaging	Benign	Polymorphism	2815; 2977; 3097; 3116; 3671; 4122; 4220
	c.5507G>A	p.Trp1836Ter	rs80356962	stop gained	Pathogenic	44	-	Pathogenic	-	-	Disease Causing	3051
	c.5329dupC	p.Gln1756Profs*74	rs397507247	frameshift	Pathogenic	35	-	-	-	-	-	2812; 3132; 3141; 3155; 3639; 3722; 3728; 4093; 4135; 4137; 4186
	c.3331_3334delC AAG	p.Gln1111Asnfs*5	rs80357701	frameshift	Pathogenic	23.7	-	-	-	-	-	2723
	c.2612C>T	p.Pro871Leu	rs799917	missense	Benign	17.97	Class C65	-	Tolerated	Benign	Polymorphism	2697 (homoz); 2699; 2724; 2726; 2742 (homoz); 2750; 2753; 2779 (homoz); 2801; 2812; 2815; 2972; 2977; 3002; 3056 (homoz); 3078 (homoz); 3083; 3097; 3114; 3115 (homoz); 3116 (homoz); 3132; 3141; 3166; 3227; 3441(homoz); 3462; 3617(homoz); 3650(homoz); 3651; 3664(homoz); 3671; 3703; 3728; 3782; 3802; 3806; 3842(homoz); 3897; 3920; 4016; 4020; 4037; 4063; 4093; 4122;

	c.3548A>G	p.Lys1183Arg	rs16942	missense	Benign		Class C25	-	Tolerated	Benign	Polymorphism	4135; 4138; 4139; 4144; 4146(homoz); 4147; 4161; 4166 (homoz); 4177; 4186 (homoz); 4214; 4220; 4226 (homoz); 4228; 4250 (homoz); 4262; 4268 (homoz); 4279 2697 (homoz); 2699; 2742; 2750; 2753; 2801; 2815; 2972; 2977; 3002; 3056; 3078 (homoz); 3097; 3114; 3116; 3141; 3166; 3462; 3617 (homoz); 3651; 3703; 3806; 3842 (homoz); 3897; 4016; 4020; 4122; 4135; 4138; 4146; 4161; 4166; 4262; 4268 (homoz); 4279 2697; 2699; 2742; 2750; 2753; 2779; 2784; 2801; 2815; 2972; 2977; 3002; 3056; 3078 (homoz); 3097; 3114; 3115; 3116; 3132; 3141; 3166; 3227; 3462; 3617 (homoz); 3651; 3671; 3703; 3728; 3782; 3806; 3842 (homoz); 3897; 4016; 4020; 4122; 4135; 4138; 4146; 4147; 4161; 4166; 4177; 4220; 4226; 4228; 4262; 4268 (homoz); 4279
	c.4900A>G	p.Ser1613Gly	rs1799966	missense	Benign		Class C55	Polymorphism	Damaging	Possibly damaging	Polymorphism	3651; 3671; 3703; 3728; 3782; 3806; 3842 (homoz); 3897; 4016; 4020; 4122; 4135; 4138; 4146; 4147; 4161; 4166; 4177; 4220; 4226; 4228; 4262; 4268 (homoz); 4279
	c.536A>G	p.Tyr179Cys	rs56187033	missense	Benign	24.7	Class C65	Pathogenic	Damaging	Probably damaging	Disease Causing	4132
	c.591C>T	p.Cys197Cys	rs1799965	splice region&synonymous	Uncertain significance	14.63	-	Probably pathogenic	-	-	-	4063
	c.3113A>G	p.Glu1038Gly	rs16941	missense	Benign	22.2	Class C65	Polymorphism	Damaging	Possibly damaging	Polymorphism	2697; 2699; 2742; 2750; 2753; 2779; 2784; 2801; 2815; 2972; 2977; 3002; 3056; 3078 (homoz); 3097; 3114; 3115; 3116; 3132; 3141; 3166; 3227; 3462; 3617 (homoz); 3651; 3671; 3703; 3728; 3782; 3806; 3842; 3897; 4016; 4020; 4122; 4135; 4138; 4166; 4220; 4226; 4228; 4262; 4268 (homoz); 4279
BRCA2	c.156_157insAlu	-	-	insertion	Pathogenic	-	-	-	-	-	-	-
	c.811G>A	p.Gly271Arg	rs786204274	missense	Conflicting interpretations of pathogenicity	20.6	Class C65	Probably pathogenic	Damaging	Benign	Polymorphism	2977
	c.3869G>A	p.Cys1290Tyr	rs41293485	missense	Benign	14.39	Class C65	Probably pathogenic	Tolerated	Benign	Polymorphism	3662

	c.4258G>T	p.Asp1420Tyr	rs28897727	missense	Benign	15.81	Class C65	Polymorphism	Damaging	Benign	Polymorphism	3649
	c.6100C>T	p.Arg2034Cys	rs1799954	missense	Benign	20.4	Class C65	Polymorphism	Damaging	Possibly damaging	Polymorphism	3441; 3617; 4279
	c.8149G>T	p.Ala2717Ser	rs28897747	missense	Benign	15.38	Class C65	Polymorphism	Tolerated	Possibly damaging	Polymorphism	3002
	c.8850G>T	p.Lys2950Asn	rs28897754	missense	Conflicting interpretations of pathogenicity	22.4	Class C65	Probably polymorphism	Damaging	Probably damaging	Disease Causing	2781
	c.865A>C	p.Asn289His	rs766173	missense	Benign	17.23	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	2995; 3051; 3056; 3651; 3722; 4166
	c.2808_2811delA CAA	p.Ala938Profs*2 1	rs80359351	frameshift	Pathogenic	24.3	-	-	-	-	-	4147
	c.8851G>A	p.Ala2951Thr	rs11571769	missense	Benign	26.2	Class C55	Polymorphism	Damaging	Probably damaging	Disease Causing	3051
	c.9026_9030delA TCAT	p.Tyr3009Serfs* 7	rs80359741	frameshift	Pathogenic	-	-	-	-	-	-	2785
	c.9382C>T	p.Arg3128Ter	rs80359212	stop gained	Pathogenic	48	-	Pathogenic	-	-	Disease Causing	4262
	c.9976A>T	p.Lys3326Ter	rs11571833	stop gained	Benign	36	-	Pathogenic	-	-	Disease Causing	3650
BRIP1	c.8324T>G	p.Met2775Arg	rs80359073	missense	Conflicting interpretations of pathogenicity	24.1	Class C65	Pathogenic	Tolerated	Benign	Disease Causing	3116
	c.3693A>G	p.Ile1231Met	rs87665929 0	missense	Uncertain significance	13.29	Class C0	Polymorphism	Damaging	Benign	Polymorphism	3132; 3728
	c.517C>T	p.Arg173Cys	rs4988345	missense	Benign/Likely Benign	24.9	Class C65	Pathogenic	Damaging	Probably damaging	Disease Causing	4122; 4173
CHEK2	c.2220G>T	p.Gln740His	rs45589637	missense	Uncertain significance	12.53	Class C15	Probably pathogenic	Damaging	Probably damaging	Disease Causing	3078
	c.410G>A	p.Arg137Gln	rs36857018 7	missense	Likely Benign	16.54	Class C35	Probably polymorphism	Tolerated	Benign	Disease Causing	3116
	c.480A>G	p.Ile160Met	rs57591080 5	missense	Conflicting interpretations of pathogenicity	22.6	Class C0	Probably polymorphism	Damaging	Probably damaging	Disease Causing	3097
FAM175A	c.489G>T	p.Arg163Ser	rs53546279 1	missense	-	22.3	Class C65	Pathogenic	Tolerated	Possibly damaging	Disease Causing	3187
	c.1042G>A	p.Ala348Thr	rs12642536	missense	-	13.44	Class C55	Polymorphism	Damaging	Possibly damaging	Polymorphism	2697; 2699; 2723; 2724; 2726; 2742; 2750; 2754 (homoz); 2775; 2779; 2781 (homoz); 2784 (homoz); 2812; 2972; 2977; 3002; 3078; 3097; 3113 (homoz); 3114

(homozygous); 3115; 3132 (homozygous); 3141 (homozygous), 3155 (homozygous); 3166 (homozygous); 3441; 3462; 3617; 3634; 3639; 3649 (homozygous); 3651; 3662; 3706; 3716; 3722; 3728 (homozygous); 3761; 3772; 3806; 3876; 3897; 3920; 4016; 4037; 4063; 4093; 4132 (homozygous); 4137; 4144; 4145; 4147; 4161; 4166; 4177; 4226 (homozygous); 4228; 4250; 4259; 4264; 4279

MRE11A	c.1011C>G	p.Ser337Arg	rs11524441 7	missense	Benign	21.9	Class C65	Probably polymorphism	Tolerated	Benign	Disease Causing	4135
	c.2101A>G	p.Met701Val	rs1805362	missense	Benign/Likely Benign	16.49	Class C15	Polymorphism	Damaging	Benign	Disease Causing	3650
NBN	c.202T>G	p.Leu68Val	rs12005998 43	missense	Uncertain significance	15.95	Class C25	Polymorphism	Tolerated	Benign	Disease Causing	2785
	c.797C>T	p.Pro266Leu	rs769420	missense	Benign	24.6	Class C65	Polymorphism	Damaging	Probably damaging	Disease Causing	3078
PALB2	c.2794G>A	p.Val932Met	rs45624036	missense	Benign/Likely Benign	25.3	Class C15	Polymorphism	Tolerated	Probably damaging	Disease Causing	3842
	c.53A>G	p.Lys18Arg	rs13878965 8	missense	Conflicting interpretations of pathogenicity	24.3	Class C25	Polymorphism	Damaging	Probably damaging	Polymorphism	3897; 4037
	c.949A>C	p.Thr317Pro	rs58778022 3	missense	Uncertain significance	4.012	Class C35	Polymorphism	Tolerated	Possibly damaging	Polymorphism	4139
RAD51	c.164C>T	p.Ala55Val	rs14561714 2	missense	Uncertain significance	23.6	Class C55	Probably polymorphism	Tolerated	Possibly damaging	Disease Causing	3116
UIMC1	c.43C>T	p.Arg15Trp	rs13167812	missense	-	22.5	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	2750; 4250
	c.1304C>T	p.Pro435Leu	rs3733876	missense	-	23.8	Class C65	Polymorphism	Damaging	Probably damaging	Polymorphism	2699; 2724; 2754; 2815; 3056; 3083; 3132 (homozygous); 3634; 3649; 3671; 3761; 3842; 3920; 4020; 4063 (homozygous); 4138; 4173; 4214
MMR genes												
MLH1	c.306G>A	p.Glu102Glu	rs63751665	splice region&synonymous	uncertain significance	22.4	-	Pathogenic	-	-	Disease Causing	4020

	c.637G>A	p.Val213Met	rs2308317	missense	Benign	23.9	Class C15	Polymorphism	Tolerated	Possibly damaging	Disease Causing	4020
	c.1217G>A	p.Ser406Asn	rs41294980	missense	Benign	13.09	Class C45	Polymorphism	Tolerated	Possibly damaging	Polymorphism	2963
	c.2146G>A	p.Val716Met	rs35831931	missense	Benign	24.4	Class C15	Probably polymorphism	Damaging	Probably damaging	Disease Causing	2754
MSH2	c.2500G>A	p.Ala834Thr	rs63750757	missense	Likely Benign	33.00	Class C55	Pathogenic	Damaging	Probably damaging	Disease Causing	4214
	c.380A>G	p.Asn127Ser	rs17217772	missense	Benign	22.7	Class C45	Probably polymorphism	Damaging	Possibly damaging	Disease Causing	3650; 3920; 4146; 4228
	c.965G>A	p.Gly322Asp	rs4987188	missense	Benign	23.0	Class C65	Probably polymorphism	Tolerated	Possibly damaging	Disease Causing	2815; 3113; 3441; 4264
MSH6	c.1186C>G	p.Leu396Val	rs2020908	missense	Benign	16.97	Class C25	Polymorphism	Tolerated	Possibly damaging	Disease Causing	2699; 2754
	c.2633T>C	p.Val878Ala	rs2020912	missense	Benign	10.23	Class C55	Polymorphism	Tolerated	Benign	Disease Causing	4147
PMS2	c.59G>A	p.Arg20Gln	rs10254120	missense	Benign	16.65	Class C35	Polymorphism	Tolerated	Possibly damaging	Polymorphism	4146; 2963; 3097; 3116; 3722; 3782; 3806; 4137; 4145; 4220; 4262
	c.2374G>A	p.Asp792Asn	rs587781265	missense	Uncertain significance	29.8	Class C15	Probably pathogenic	Damaging	Probably damaging	Disease Causing	3802
	c.2350G>A	p.Asp784Asn	rs143340522	missense	Uncertain significance	27.8	Class C15	Polymorphism	Damaging	Probably damaging	Disease Causing	4264
	c.2149G>A	p.Val717Met	rs201671325	missense	Conflicting interpretations of pathogenicity	20.7	Class C15	Probably polymorphism	Damaging	Probably damaging	Disease Causing	3116; 3462
	c.1866G>A	p.Met622Ile	rs1805324	missense	Benign	18.96	Class C0	Polymorphism	Tolerated	Possibly damaging	Disease Causing	3772
	c.1688G>T	p.Arg563Leu	rs63750668	missense	Benign/Likely Benign	10.65	Class C65	Polymorphism	Tolerated	Possibly damaging	Polymorphism	2972; 3227; 4016
Other genes												
CDH1	c.1849G>A	p.Ala617Thr	rs33935154	missense	Conflicting interpretations of pathogenicity	15.13	Class C55	Polymorphism	Tolerated	Benign	Disease Causing	3664; 4145; 4166
TP53	c.1010G>A	p.Arg337His	rs121912664	missense	Pathogenic	22.6	Class C25	-	Damaging	Probably damaging	Disease Causing	2699; 3056; 3227; 3662; 4264
	c.818G>A	p.Arg273His	rs28934576	missense	Pathogenic/Likely pathogenic	24.0	Class C25	Pathogenic	Damaging	Probably damaging	Disease Causing	3227

Table S1.

Patient ID	Year of birth	Gender	Deceased	Year of Death	Group of age at diagnosis	Age at diagnosis	Tumor site	Cancer type	Breast molecular subtype	Bilateral breast	Age of contralateral diagnosis	Synchronous	Multifocal	Family history	Tumor grade	Tumor size	Lymph node staging	Distant metastasis	Distant metastasis site	Clinical staging	Genotype
2697	1963	Female	NO	–	NI	NI	Breast	Invasive Ductal Carcinoma	NI	NI	NI	NI	NI	YES	NI	NI	NI	NI	–	NI	BRCA VUS
2699	1957	Female	NO	–	31-50	36	Breast	Invasive Ductal Carcinoma	NI	NO	–	–	NO	YES	NI	NI	NI	NI	–	NI	BRCA VUS
2723	1967	Female	NO	–	31-50	31	Breast	Invasive Ductal Carcinoma	Triple-negative	YES	43	NO	NO	YES	III	T4D	N0	M0	–	NI	BRCA1 pathogenic
2724	1979	Female	NO	–	20-30	29	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	NI	T3	N0	M0	–	IIIA	BRCA VUS
2726	1982	Female	NO	–	20-30	26	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	NO	II	T4b	N2	M0	–	IIIB	BRCA VUS
2742	1968	Female	NO	–	31-50	41	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	NO	I	NI	N1	M0	–	NI	BRCA VUS
2750	1974	Female	YES	2010	31-50	35	Breast	Invasive Ductal Carcinoma	Triple-negative	YES	36	NO	NO	YES	II	T4D	N2	M1	Skin and Bone Marrow	IIIB	BRCA VUS
2753	1967	Female	NO	–	31-50	40	Breast	Invasive Ductal Carcinoma	HER2	NO	–	–	NO	YES	II	NI	NI	NI	–	NI	BRCA VUS
2754	1961	Female	YES	2014	31-50	41	Breast	Invasive Ductal Carcinoma	HER2	YES	47	NO	NO	YES	III	T4	N2	M1	Lung, Bones and CNS	IV	non-BRCA
2775	1947	Female	NO	–	51-60	58	Breast	Invasive Ductal Carcinoma	HER2	NO	–	–	NO	YES	NI	T2	N1	M0	–	IIB	BRCA VUS
2779	1972	Female	NO	–	31-50	37	Breast	Invasive Ductal Carcinoma	HER2	NO	–	–	NO	NO	III	T2	N0	M0	–	IIA	BRCA VUS
2781	1957	Female	YES	2015	51-60	52	Ovary	Papillary Serous Ovarian Carcinoma	NI	–	–	–	NO	YES	NI	NI	NI	M1	Liver and Pleural	IV	BRCA VUS
2784	1972	Female	NO	–	31-50	37	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	NI	T3	N2	M1	Bones	NI	BRCA VUS
2785	1967	Female	NO	–	31-50	42	Breast	Invasive Ductal Carcinoma	LuminalHER	YES	49	NO	NO	YES	III	T1C	N0	M0	–	NI	BRCA1 pathogenic
2801	1979	Female	NO	–	20-30	30	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	NO	II	T2	N0	M0	–	IC	BRCA VUS
2812	1984	Female	NO	–	20-30	25	Breast	Invasive Ductal Carcinoma	HER2	NO	–	–	NO	YES	III	T2	N1	M0	–	IIB	BRCA1 pathogenic
2815	1975	Female	NO	–	20-30	25	Breast	Ductal Carcinoma	NI	NI	NI	NI	YES-Thyroid	YES	NI	NI	NI	NI	–	NI	BRCA VUS
2963	1946	Female	NO	–	31-50	47	Breast	Invasive Ductal Carcinoma	NI	NO	–	–	NO	YES	NI	NI	NI	NI	–	NI	non-BRCA
2972	1965	Female	NO	–	31-50	45	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	T4B	N1	M0	–	IIIB	BRCA VUS
2977	1973	Female	YES	2012	20-30	30	Breast and Ovary	Invasive Ductal Carcinoma and Mature Cystic Teratoma of the Ovary	Luminal	NI	NI	NI	YES-Ovary	YES	NI	T3	NI	M1	Bones and Liver	IV	BRCA VUS
2995	1974	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	HER2	NO	–	–	NO	NI	NI	T4D	N2	M0	–	NI	BRCA VUS
3002	1977	Female	NO	–	31-50	32	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	II	T1	N0	M0	–	I	BRCA VUS
3051	1976	Female	NO	–	20-30	30	Breast	Invasive Ductal Carcinoma	HER2	YES	33	NO	NO	YES	III	T1C	NI	M0	–	IIA	BRCA1 pathogenic

3056	1952	Female	NO	–	51-60	55	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	III	T1	N0	M0	–	I	BRCA VUS
3078	1970	Female	NO	–	31-50	41	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	T1C	N0	M0	–	I	BRCA VUS
3083	1985	Female	NO	–	20-30	22	Ovary	Serous Borderline Ovarian Tumor	NI	–	–	–	NO	NO	NI	NI	NI	NI	–	NI	BRCA VUS
3097	1975	Female	YES	2011	31-50	34	Breast	Invasive Ductal Carcinoma	Luminal	YES	34	YES	NO	YES	II	T3	N0	M1	Lung, Liver and Bones	IV	BRCA VUS
3113	1979	Female	NO	–	31-50	33	Breast	Invasive Ductal Carcinoma	Triple- negative	NO	–	–	NO	YES	II	T3	N1	M0	–	NI	BRCA VUS
3114	1946	Female	NO	–	61+	64	Breast	Invasive Ductal Carcinoma	Luminal	NI	NI	NI	YES- Endometrium	YES	II	T2	N2	M1	–	IV	BRCA VUS
3115	1976	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	NI	II	T2	N2	M1	Liver	NI	BRCA VUS
3116	1953	Female	NO	–	51-60	52	Breast	Invasive Ductal Carcinoma	LuminalHER	NI	NI	NI	YES-Kidney	YES	II	T2	N2	M0	–	NI	BRCA VUS
3132	1979	Female	NO	–	31-50	33	Breast	Invasive Ductal Carcinoma	Triple- negative	NO	–	–	NO	YES	III	T3	N0	M0	–	IIB	BRCA1 pathogenic
3141	1981	Female	NO	–	31-50	31	Breast	Invasive Ductal Carcinoma	Triple- negative	NO	–	–	NO	YES	III	T2	N1	M0	–	IIB	BRCA1 pathogenic
3155	1963	Female	NO	–	31-50	47	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	T2	N0	M0	–	I	BRCA1 pathogenic
3166	1950	Female	NO	–	NI	NI	Breast	Ductal Carcinoma in situ	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	–	NI	BRCA VUS
3187	1968	Female	NO	–	31-50	43	Ovary	Ovary Adenocarcinoma	NI	–	–	–	NO	YES	II	NI	NI	NI	–	III	BRCA VUS
3227	1987	Female	YES	2016	20-30	25	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	II	T4D	N2	M1	Brain, Liver, and Bones	IV	TP53 pathogenic
3441	1952	Female	NO	–	51-60	60	Endometrium	Endometrial Adenocarcinoma	NI	–	–	–	NO	YES	I	T1A	NI	M0	–	NI	BRCA VUS
3462	1985	Female	NO	–	20-30	26	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	I	T3	N3	M0	–	IIIC	BRCA VUS
3617	1976	Female	NO	–	20-30	29	Breast	Invasive Ductal Carcinoma	Luminal	NI	NI	NI	NI	NO	NI	NI	NI	NI	–	NI	BRCA VUS
3634	1946	Female	NO	–	31-50	45	Breast	Ductal Carcinoma in situ	NI	NI	NI	NI	YES-Lung	YES	NI	NI	NI	NI	–	NI	non-BRCA
3639	1987	Female	NO	–	20-30	27	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	III	T2	NI	M0	–	NI	BRCA1 pathogenic
3649	1970	Female	NO	–	31-50	41	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	II	T1	N1	M0	–	IIB	BRCA VUS
3650	1976	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	II	T1	N1	M0	–	IIA	BRCA VUS
3651	1972	Female	NO	–	31-50	32	Breast	Invasive Ductal Carcinoma	PR	NI	NI	NI	NI	YES	II	T1	N0	M0	–	I	BRCA VUS
3662	1974	Female	NO	–	31-50	39	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	III	T2	N1A	M0	–	IIB	TP53 pathogenic
3664	1970	Female	NO	–	20-30	28	Breast	Invasive Ductal Carcinoma	HER2	NI	NI	NI	NI	YES	NI	NI	NI	NI	–	NI	BRCA VUS
3671	1978	Female	NO	–	31-50	32	Breast	Invasive Ductal Carcinoma	HER2	NO	–	–	NO	YES	II	T4D	N1C	M0	–	IIIB	BRCA VUS
3703	1980	Female	NO	–	31-50	33	Ovary	Mucinous Adenocarcinoma of Ovary and Mucinous Borderline Ovarian Tumor	NI	–	–	–	NO	YES	II	NI	NI	NI	–	NI	BRCA VUS
3706	1979	Female	NO	–	31-50	34	Breast	Invasive Ductal Carcinoma	Luminal	YES	NI	NI	NO	YES	II	T4B	N2	M1	Liver	NI	non-BRCA

3716	1974	Female	NO	–	31-50	39	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	II	T2	N1	M0	–	IIIA	non-BRCA
3722	1973	Female	NO	–	31-50	39	Breast	Ductal Carcinoma in situ	Triple-negative	NO	–	–	NO	YES	NI	T0	N1	M0	–	NI	BRCA1 pathogenic
3728	1974	Female	YES	2016	31-50	34	Breast	Ductal Carcinoma in situ	Triple-negative	NO	–	–	NO	YES	NI	TX	NX	M1	Lung	IV	BRCA1 pathogenic
3761	1977	Female	NO	–	31-50	36	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	T2	N0	M0	–	IIA	non-BRCA
3766	1977	Female	NO	–	31-50	34	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	III	T1C	N3A	M0	–	IIIC	non-BRCA
3772	1977	Female	NO	–	31-50	37	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	NO	II	T2	N1A	M0	–	IIB	non-BRCA
3782	1967	Female	NO	–	31-50	47	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	I	T1C	N0	M0	–	I	BRCA VUS
3802	1967	Male	YES	2016	31-50	46	Stomach	Diffuse Gastric Cancer	NI	–	–	–	NO	YES	NI	T1B	N0	M0	–	IA	BRCA VUS
3806	1969	Female	YES	2015	31-50	42	Breast	Invasive Ductal Carcinoma	LuminalHER	YES	NI	NI	NO	YES	II	T4D	N1	MX	Lung and CNS	NI	BRCA VUS
3842	1976	Female	NO	–	31-50	37	Breast	Invasive Ductal Carcinoma	LuminalHER	NI	NI	NI	NI	YES	II	T3	N1	M0	–	IIIA	BRCA VUS
3876	1977	Female	NO	–	31-50	37	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	II	T3	N1	M0	–	IIIA	BRCA VUS
3897	1973	Female	NO	–	31-50	42	Breast	Intraductal Papilloma	NI	NO	–	–	NO	YES	NI	NI	NI	NI	–	NI	BRCA VUS
3920	1985	Female	NO	–	20-30	30	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	NO	I	T1	NI	M0	–	I	BRCA VUS
4016	1983	Female	NO	–	31-50	32	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	III	T1C	NI	NI	–	NI	BRCA VUS
4020	1975	Female	NO	–	31-50	40	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	T3	N1A	M0	–	IIIA	BRCA VUS
4037	1982	Female	NO	–	31-50	33	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	NO	NI	T2	N0	M0	–	NI	BRCA VUS
4063	1957	Female	NO	–	31-50	49	Breast	Invasive Ductal Carcinoma	Luminal	YES	57	NO	NO	YES	II	T1	N1	M0	–	NI	BRCA VUS
4094	1964	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T2	N0	M0	–	IIB	BRCA1 pathogenic
4122	1963	Female	NO	–	51-60	52	Breast	Invasive Lobular Carcinoma	NI	NO	–	–	NO	YES	I	T1A	N0	M0	–	NI	BRCA VUS
4132	1980	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	NI	I	T3	N0	M0	–	NI	BRCA1 pathogenic
4133	1974	Female	NO	–	31-50	40	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	T2	N0	M0	–	IIA	BRCA VUS
4135	1973	Female	NO	–	31-50	40	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T3	N1	M0	–	IIB	BRCA1 pathogenic
4137	1980	Female	NO	–	31-50	31	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	NI	I	T3	N0	M0	–	NI	BRCA2 pathogenic
4138	1967	Female	YES	2018	31-50	47	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T3	N0	MX	Lung	IIA	BRCA VUS
4139	1981	Female	NO	–	31-50	34	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T3	N1	M0	–	NI	BRCA VUS
4144	1988	Female	NO	–	20-30	27	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T1C	N2A	M0	–	III	BRCA VUS
4145	1983	Female	YES	2018	31-50	31	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	III	T3	N1	MX	Bone and Liver	IIIA	non-BRCA
4146	1976	Female	NO	–	31-50	37	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	II	NI	NI	NI	–	NI	BRCA VUS
4147	1986	Female	NO	–	20-30	24	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	NI	II	T1c	N2a	M0	–	IIIA	BRCA2 pathogenic

4161	1948	Female	YES	2017	61+	66	Ovary	High-Grade Serous Ovarian Cancer	NI	–	–	–	YES-Uterine tubes	YES	NI	NI	NI	MX	Pleural and Peritoneum	IV	BRCA VUS BRCA VUS
4166	1970	Female	NO	–	31-50	45	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T2	N1	M0	–	IIB	BRCA VUS
4173	1983	Female	NO	–	31-50	31	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T3	N1	M0	–	NI	non-BRCA
4177	1988	Female	NO	–	20-30	27	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	NO	II	T2	N0	M0	–	IIB	BRCA VUS
4186	1956	Female	NO	–	51-60	57	Breast	Invasive Ductal Carcinoma	Triple-negative	YES	NI	NI	YES-OVARY	YES	III	T4D	N2	M0	–	NI	BRCA2 pathogenic
4214	1981	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	NI	T2	N0	M0	–	NI	BRCA VUS
4220	1960	Female	NO	–	51-60	53	Breast	Invasive Ductal Carcinoma	HER2	YES	53	YES	NO	YES	III	T1c	N2a	M0	–	I/IIIA	BRCA VUS
4226	1981	Female	YES	2018	31-50	34	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	YES	I	T3	N1	M1	Spine, Bones and CNS	III	BRCA VUS
4228	1974	Female	NO	–	31-50	41	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	II	T4B	N1	M0	–	NI	BRCA VUS
4250	1979	Female	NO	–	31-50	35	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	III	T4B	N1	M1	Bones, Liver and Lung	IV	BRCA VUS
4259	1982	Female	NO	–	31-50	33	Breast	Invasive Ductal Carcinoma	Luminal	NO	–	–	NO	NO	II	T2	N0	M0	–	IIA	non-BRCA
4262	1962	Female	NO	–	31-50	36	Breast	Invasive Ductal Carcinoma	LuminalHER	YES	49	NO	NO	YES	II	NI	NI	NI	–	NI	BRCA2 pathogenic
4264	1974	Female	NO	–	31-50	41	Breast	Invasive Ductal Carcinoma	LuminalHER	NO	–	–	NO	YES	I	T1	N0	M0	–	NI	non-BRCA
4268	1942	Female	YES	2016	61+	72	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	III	T4D	N1	M1	CNS	IIIB	BRCA VUS
4279	1977	Female	NO	–	31-50	39	Breast	Invasive Ductal Carcinoma	Triple-negative	NO	–	–	NO	YES	II	T2	N1	M0	–	NI	BRCA VUS

Table S2.

Haplotypes						Brazilian Cancer Centers											
Hp	p.Pro871Leu (C <u>C</u> G→C <u>T</u> G)	p.Glu1038Gly (G <u>A</u> A→G <u>G</u> A)	p.Lys1183Arg (A <u>A</u> A→A <u>G</u> A)	p.Ser1613Gly (<u>A</u> GT→ <u>G</u> GT)	c.*421G>T (G→T)	HPOA				ACC				HCB			
						Hap. freq.	Control (n=94)	HBOC (n=94)	p- value	Hap. freq.	Control (n=185)	HBOC (n=171)	p-value	Hap. freq.	Control (n=72)	HBOC (n=72)	p-value
1	Pro	Glu	Lys	Ser	G	0.601	0.574	0.627	0.271	0.576	0.546	0.609	0.082	0.558	0.549	0.568	0.729
2	Leu	Gly	Arg	Gly	T	0.279	0.287	0.271	0.724	0.087	0.086	0.088	0.967	0.304	0.306	0.304	0.999
3	Leu	Glu	Lys	Ser	G	0.056	0.085	0.027	0.027	0.096	0.124	0.064	0.007	0.088	0.125	0.05	0.026
4	Leu	Gly	Arg	Gly	G	0.035	0.043	0.027	0.388	0.23	0.224	0.237	0.643	0.014	0.014	0.014	1
5	Pro	Gly	Arg	Gly	T	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	Leu	Gly	Arg	Gly	G	NA	NA	NA	NA	NA	NA	NA	NA	0.014	0.007	0.021	0.311