



UNIVERSIDADE ESTADUAL PAULISTA

JÚLIO DE MESQUITA FILHO

FACULDADE DE MEDICINA

ANISSE MARQUES CHAMI FERRAZ

IMPACTO DA ANÁLISE “IN SILICO” NA PREDIÇÃO DE PATOGENICIDADE E RECLASSIFICAÇÃO DE VARIANTES EM GENES DE SUSCETIBILIDADE AO CÂNCER GINECOLÓGICO

Tese apresentada à Faculdade de Medicina da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Botucatu, para obtenção do título de Doutor(a) pelo Programa de Graduação em Tocoginecologia: Área de Cancerologia

Orientador: Prof. Dr. Agnaldo Lopes da Silva Filho

Coorientadora: Prof. Dra. Letícia da Conceição Braga

Botucatu. 2022



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"



ANISSE MARQUES CHAMI FERRAZ

IMPACTO DA ANÁLISE "IN SILICO" NA PREDIÇÃO DE PATOGENICIDADE E RECLASSIFICAÇÃO DE VARIANTES EM GENES DE SUSCETIBILIDADE AO CÂNCER GINECOLÓGICO

Tese apresentada à Faculdade de Medicina da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de Botucatu, para obtenção do título de Doutor(a) pelo Programa de Graduação em Tocoginecologia: Área de Cancerologia

Orientados: Prof. Dr. Agnaldo Lopes da Silva Filho

Coorientadora: Prof. Dra. Letícia da Conceição Braga

Botucatu, 2022

F381i

Ferraz, Anisse Marques Chami

Impacto da análise "in silico" na predição de patogenicidade e reclassificação de variantes em genes de suscetibilidade ao câncer ginecológico / Anisse Marques Chami Ferraz. -- Botucatu, 2022
152 p. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp),
Faculdade de Medicina, Botucatu

Orientador: Agnaldo Lopes da Filho

Coorientadora: Leticia da Conceição Braga

1. Genética Humana. 2. Tocoginecologia. 3. Aconselhamento Genético. 4. Câncer. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Medicina, Botucatu. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

*Pos meus amados filhos
Stella e Matheus,
dedico.*

AGRADECIMENTOS

À Deus.

À minha família, especialmente ao Daniel pelo apoio e compreensão.

Ao Prof. Agnaldo pela oportunidade do aprendizado e evolução na minha formação acadêmica.

À Prof. Letícia pelo suporte, ensinamentos e confiança.

À Prof. Maria Raquel, Thalia, Carol e Thamires do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais (UFMG) pela parceria e empenho em manter as atividades mesmo nas situações adversas ao longo dessa trajetória. Um agradecimento especial à Thalia, pela parceria e dedicação.

Ao Grupo Pardini (MG) e Laboratório da Clínica Personal de Oncologia (Belo Horizonte, MG) pela colaboração e participação ativa no trabalho através das colaboradoras Maíra (Grupo Pardini) e Juliana (Personal).

Aos colegas médicos pela colaboração para produção científica.

Aos colegas geneticistas do Hospital das Clínicas da UFMG que fizeram e fazem parte da minha formação e convívio. Especialmente, à Dra. Letícia Leão pelo apoio e exemplo de pessoa e profissional.

Aos pacientes que são a fonte de inspiração para busca do aprimoramento do conhecimento e da prática clínica na minha área de atuação.

Aos professores e colaboradores da UNESP que nos conduzem durante a pós-graduação.

À todas as pessoas que me auxiliam carinhosamente nos cuidados com meus filhos e meu lar.

Muito obrigada!!

“É o espírito que conduz o mundo e não a inteligência.”

Antoine de Saint-Exupéry

SUMÁRIO

RESUMO	8
ABSTRACT	9
LISTA DE ABREVIATURAS	10
INTRODUÇÃO	12
I - O Câncer de mama e ginecológico hereditários	12
II) Aconselhamento genético e testes genéticos na abordagem do câncer ginecológico	18
JUSTIFICATIVA	25
OBJETIVO GERAL	26
OBJETIVOS ESPECÍFICOS	26
APROVAÇÃO DO COEP	27
DESENVOLVIMENTO E METODOLOGIA PARA ANÁLISES DAS VARIANTES	27
CAPÍTULO I	31
Evidence for pathogenicity of <i>MSH2</i> c.1894_1898; p.Ile633Lysfs*9 mutation in Lynch syndrome: genotype-phenotype correlation and bioinformatic approaches	31
CAPÍTULO II	61
Title: <i>In deep</i> Bioinformatic characterization of a novel fumarate hydratase mutation <i>FH</i> c.199T>G; (p.Tyr67Asp) in hereditary leiomyomatosis and renal cell carcinoma.....	61
CAPÍTULO III	84
Análise de variantes no gene <i>RAD51C</i> em pacientes testadas pelo painel por <i>Next Generation Sequencing</i> (NGS) para mutações germinativas associadas ao câncer hereditário	84
CONSIDERAÇÕES FINAIS	123
REFERÊNCIAS BIBLIOGRÁFICAS	124
ANEXO I - PARECER DE APROVAÇÃO NO COMITÊ DE ÉTICA EM PESQUISA EM SERES HUMANOS	133

IMPACTO DA ANÁLISE “IN SILICO” NA PREDIÇÃO DE PATOGENICIDADE E RECLASSIFICAÇÃO DE VARIANTES EM GENES DE SUSCETIBILIDADE AO CÂNCER GINECOLÓGICO

RESUMO

A medicina genômica tem um importante papel na determinação do manejo clínico de mulheres com diagnóstico ou risco para neoplasias mamárias e/ou ginecológicas. Estima-se que até 10% dos casos de câncer de mama, em torno de 24%, de câncer de ovário epitelial e até 3%, de câncer de endométrio tem um padrão reconhecidamente hereditário associado a uma síndrome ou gene de predisposição ao câncer. Os testes genéticos são ferramentas fundamentais para estudo das mutações germinativas. Resultados desses exames podem guiar não somente a estratificação do risco das pacientes, mas também, medidas de redução de risco, decisões terapêuticas e assistência aos familiares. A interpretação das variantes desses testes é muito variável e, muitas vezes, conflitantes. Mesmo diante da identificação de uma variante patogênica mecanismos moleculares podem amenizar o impacto dessas variantes. A partir de métodos de bioinformática é possível analisar o impacto das variantes em nível de RNA e de proteína. Os critérios estabelecidos pelo Colégio Americano de Genética Médica é o recomendado para classificação das variantes e levam em conta vários aspectos, dentre eles, estudos computacionais. Neste trabalho, avaliamos por métodos de bioinformática, variantes em genes de interesse do grupo de trabalho com intuito de definir patogenicidade ou benignidade e rever a classificação de variantes em genes associados ao risco para neoplasias ginecológicas. Foi realizado o levantamento de casos e coleta de variantes de serviços médicos e dados laboratoriais de pacientes que fizeram teste genético para avaliação de câncer hereditário. Para os relatos de caso, aprofundamos nos estudos “in silico” da variante patogênica *MSH2* c.1894_1898; (p.Ile633Lysfs*9) e sugerimos uma reclassificação para a variante *FH* c.199T>G; (p.Tyr67Asp). Analisamos ainda 19 variantes coletadas no gene *RDA51C*. Desta, dentre as variantes de significado incerto (VUS), apenas três seriam passíveis reclassificação. O trabalho retrata o cenário em nossa comunidade das análises de variantes em genes de predisposição ao câncer e a importância da colaboração clínico-laboratorial para refinamento das interpretações do real impacto em nível molecular.

Palavras chaves: câncer ginecológico hereditário; testes genéticos; interpretação de variantes; critérios ACMG, análise bioinformática, *RAD51C*

THE "IN SILICO" ANALYSIS IMPACT IN THE PREDICTION OF PATHOGENICITY AND RECLASSIFICATION OF VARIANTS IN GYNECOLOGICAL CANCER SUSCEPTIBILITY GENES

ABSTRACT

Genomic medicine plays an important role in determining the clinical management of women diagnosed with or at risk for breast and/or gynecological malignancies. It is estimated that up to 10% of breast cancer cases, around 24% of epithelial ovarian cancer and up to 3% of endometrial cancer have a recognized hereditary pattern associated with a cancer predisposition syndrome or gene. Genetic tests are fundamental tools for studying germline mutations. Results of these tests can guide not only the patient risk stratification, but also risk reduction management, therapeutic decisions and family members genetic counseling. The variant interpretation is highly variable and often conflicting. Even if the identified variant is assumed to be pathogenic, molecular mechanisms can mitigate these variants functional impacts. Using bioinformatics tools, it is possible to analyze the impact of variants at the RNA and protein levels. The criteria established by the American College of Medical Genetics are recommended for their classification and take into account several aspects, including computational studies. In this work, it is evaluated, using bioinformatics, variants in genes of interest to the working group in order to define pathogenicity or benignity and review the variant classification in genes associated with risk for gynecological tumors. A case survey was carried out and variants were collected from medical services and laboratory data from patients who underwent genetic testing to assess hereditary cancer. For case reports, we delved deeper into the "in silico" studies of the pathogenic variant *MSH2* c.1894_1898; (p.Ile633Lysfs*9) and we suggest a reclassification to the *FH* variant c.199T>G; (p.Tyr67Asp). Also, we analyzed 19 variants collected in the *RDA51C* gene. Among the variants of uncertain meaning (VUS), only three tends to be reclassified. The work portrays our community scenario of the analysis of variants in cancer predisposition genes and the importance of clinical-laboratory collaboration to refine the interpretations of the real impact at the molecular level.

Key words: hereditary gynecological cancer; genetic testing; variant interpretation; ACMG criteria, bioinformatics analysis, *RAD51C* gene.

LISTA DE ABREVIATURAS

ACMG/AMP – The American College of Medical Genetics and Genomics and the Association for Molecular Pathology

ANS – Agência Nacional de Saúde Suplementar

BCAC – Breast Cancer Association Consortium

BLAST - Basic Local Alignment Search Tool

CE – Câncer de endométrio

CM – Câncer de mama

CM – Câncer de mama

CO – Câncer de ovários

db SNP - Single Nucleotide Polymorphism database

DGV – Database of Genomic Variants

DNA - ácido desoxirribonucleico

EVA - European Variation Archive

GTE_x - Genotype-Tissue Expression

HSF – Human Splice Finder

iPARP – inibidor de PARP

M-CAP - Mendelian Clinically Applicable Pathogenicity Score

mRNA – RNA mensageiro

MRR – mastectomia redutora de risco

NCCN – National Comprehensive Cancer Network

OMIM – Online Mendelian Inheritance in Men

OR - Odds Ratio

P – Prevalência

Pfam – Protein domain database

PolyPhen-2 - Polymorphism Phenotyping versão 2

PROMPT – Prospective Registry of multiplex testing

PROVEAN - protein Variation Effect Analyzer

RNA – ácido ribonucleico

RR – risco relativo

SC – Síndrome de Cowden

SIFT – Sorting intolerant from tolerant

SL – Síndrome de Lynch

SORR – Salpingooforectomia redutora de risco

SUS – Sistema único de saúde

UniProt – Universal Protein Resource

VB – variante benigna

VEP – Variant Effect Predictor

VP – variante patogênica

VB – variante benigna

VPB – variante provavelmente benigna

VPP – variante provavelmente patogênica

VUS – Variant of uncertain significance /variante de significado incerto

INTRODUÇÃO

I - O Câncer de mama e ginecológico hereditários

O conhecimento sobre a predisposição hereditária ao câncer evoluiu nas últimas décadas e muito se deve ao progresso da biologia molecular e do estudo de mutações somáticas e germinativas. A partir de 1980, direcionou-se o foco para questões de predisposição genética ao câncer a partir de estudos populacionais que revelaram fatores de risco para câncer, além dos fatores ambientais, em famílias com alto risco para câncer de mama (Neuman *et al.*;1988; Claus *et al.*; 1991).

O número estimado de novos casos em 2020 para todos os tipos de câncer em mulheres em todas as idades é de 9.227.484, com 750.007 (8,1%) concentrados na América Latina e Caribe e uma mortalidade global de 4.429.323 (*Global Cancer Observatory,2020*). O câncer de mama é o tipo com maior incidência (2.261.419). Considerando os cânceres ginecológicos, o câncer de colo de útero, útero e ovários representam respectivamente o 4º, 6º e 8º lugares, com números de novos casos em 2020, de 604.27; 417.367; 313.959 (*Global Cancer Observatory,2020*). De acordo com dados epidemiológicos publicados quanto ao panorama do câncer ginecológico no Brasil, o câncer de ovário (CO) e o câncer de endométrio (CE) foram o sétimo e oitavo cânceres mais comuns, com 6.650 e 6.540 novos casos, respectivamente, em 2020 (Paulino *et al.*, 2020).

Embora o câncer de ovário fosse o menos comum dos citados, é o mais letal. Representa a quinta causa de morte por câncer nos Estados Unidos com uma taxa de mortalidade de 48,6% em cinco anos. Não há rastreamento efetivo para detecção precoce da doença e, a maioria das mulheres diagnosticadas com CO tem uma doença já avançada (Siegel *et al.*, 2020; Henderson *et al.*, 2018). Por essa razão, a salpingooforectomia redutora de risco (SORR) é um procedimento profilático que pode fazer parte do manejo de redução de risco de mulheres com mutações em genes de predisposição para CO (Eleje *et al.*, 2018).

Estima-se que 5 a 10% dos casos de câncer de mama tem associação com componente hereditário e mutações patogênicas como em genes *BRCA1* e *BRCA2*, que representam a etiologia genética hereditária mais comum (Balmain et al.; 2003). Tanto o CO quanto o CE também representam fenótipos importantes de síndromes de predisposição ao câncer hereditário. Três a 5% dos casos de CE tem um padrão hereditário e mais de um quinto dos CO têm suscetibilidade hereditária (Randal & Ponturi, 2016; Toss et al., 2015).

Além das síndromes hereditárias de câncer de mama e ovário, associadas aos genes *BRCA1* e *BRCA2*, outras síndromes genéticas predisõem os indivíduos ao risco maior para câncer de mama e ginecológico. Genes supressores tumorais com mutações patogênicas são herdados de forma autossômica dominante e essas mutações aumentam significativamente o risco de câncer ao longo da vida em diferentes órgãos, dependendo da síndrome ou gene associado. Tais síndromes incluem a síndrome de Lynch, a causa hereditária mais comum de câncer endometrial e a síndrome de Peutz-Jeghers, que aumenta os riscos de câncer de mama e câncer de ovário. Já, a síndrome de Li-Fraumeni e de síndrome de Cowden aumentam o risco de câncer de mama; a síndrome de Gorlin aumenta o risco para neoplasias, como fibromas ovarianos e a Leiomiomatose hereditária relacionada ao exuberante quadro de leiomiomas uterinos e risco para câncer renal (Chung et al., 2020; Furuya et al., 2020).

Com os avanços nos testes genéticos, principalmente da tecnologia para os painéis genéticos pela metodologia de *Next Generation Sequencing* (NGS), numerosos genes adicionais associados aos cânceres de mama e ovário, além de *BRCA1* e *BRCA2*, tornaram-se reconhecidos como o *ATM*, *CHEK2*, *PALB2*, *RAD51 C*, *RAD51 D* entre outros. Conseqüentemente, o conhecimento sobre as manifestações clínicas ou fenótipos das mutações nesses genes estão se expandindo rapidamente. SAUNDRA e colaboradores (2017) avaliaram o genótipo do perfil de mutações germinativas em 35.409 mulheres com câncer de mama por um painel NGS de 25 genes. Mutações patogênicas em *BRCA1* e *BRCA2* representaram 24% e 24,4% dos casos,

respectivamente. Em outros genes, as mutações totalizaram 51,5%, com maior representatividade exatamente para *ATM*, *CHEK2* e *PALB2* (Buys et al., 2017). Protocolos e consensos das sociedades médicas publicam informações sobre a importância de reconhecer pacientes portadores dessas mutações e como manejá-los para redução do risco de câncer (*National Comprehensive Cancer Network, 2021*). Entretanto, para muitos genes, a informação sobre a proposta de redução de risco de câncer ainda carece de evidências robustas para elaboração de condutas médicas específicas. Dentre as recomendações sugeridas há procedimentos que são invasivos, cirúrgicos e que têm impacto relevante na qualidade de vida das pacientes. Isso dificulta a tomada de decisões frente à mastectomia redutora de risco (MRR), à SORR, definição de prole e até sobre a realização do teste genético, mesmo para os genes de alta penetrância para câncer de mama e/ou ovários (*BRCA1*, *BRCA2*, *CDH1*, *PALB2* e *TP53*) (*National Comprehensive Cancer Network, 2021*).

Além disso, deparamo-nos com situações conflitantes como, por exemplo, o risco absoluto para o desenvolvimento de câncer ao longo da vida em pacientes portadoras de mutação no gene *ATM*, é muito variável, de 15 a 40% para câncer de mama e para ovários, um risco menor que 3% (*National Comprehensive Cancer Network, 2021*). Via de regra, a literatura ainda não apresenta dados com evidências suficientes para as cirurgias redutoras de risco em pacientes portadores de mutações patogênicas no gene *ATM*. Mesmo assim, dados apresentados em análises do registro nacional norte americano de indivíduos com variantes germinativas detectadas em painéis multigenes, *PROSPECTIVE REGISTRY OF MULTIPLEX TESTING (PROMPT)*, revelam mulheres submetidas à SORR pelo fato de serem portadoras de mutações no gene *ATM*, por esse gene ter uma associação de risco também com CO, mas com evidência insuficiente quanto à redução de risco por SORR. Essa circunstância abrange outros genes, principalmente de penetrância moderada para câncer de mama, cujo risco relativo (RR) para o câncer ao longo da vida é de duas a cinco vezes o risco basal/populacional (Liu et al., 2021; Balmaña et al., 2016). Quanto ao risco para CO, estima-se que o risco populacional é de 1,3%. Apenas o histórico de CO, como doença na família, é considerado um importante fator de risco que eleva o RR para cerca de 3, embora não represente sozinho um fator de risco que justifique uma intervenção cirúrgica redutora de risco (Liu et al, 2021; Stratton et al., 1998). Se o risco

cumulativo ao longo da vida ultrapassa 4%, a avaliação de custo-efetividade aponta em direção à cirurgia redutora de risco. No entanto, é preciso considerar cuidadosamente o momento da RRSO, tendo em vista eventos pós cirúrgicos, como a menopausa precoce dessas mulheres e impactos na qualidade de vida e saúde a longo prazo (Manchanda *et al.*, 2016).

A maioria dos dados atuais sobre o manejo de redução de risco se baseia em estudos voltados para mulheres de alto risco e estudos familiares (*Breast Cancer Association Consortium/BCAC*, 2021). Estudos populacionais permitem dados mais robustos para guiar melhor a seleção de pacientes, rastreamento de neoplasias e outras estratégias para esse manejo necessário em mulheres, na população geral, com variantes patogênicas em genes de predisposição. Recentemente, dois estudos multicêntricos populacionais caso-controle sobre genes de risco para câncer de mama foram publicados com objetivo de avaliar a estimativa de risco absoluto (Naros, 2021; Hu *et al.*, 2021; *Breast Cancer Association Consortium/BCAC*, 2021). Outros dados reportados incluem o risco de câncer associado a subtipo tumoral e análise do risco de acordo com o tipo de variantes patogênicas e a consequência da mutação na proteína. Ambos os estudos revelaram informações importantes quanto aos genes de penetrância alta a moderada para câncer de mama como *BRCA1*, *BRCA2*, *BARD1*, *RAD51C*, *RAD51D*, *ATM*, *CHEK2*, *PALB2*, *NBN*, *TP53*. No estudo europeu para variantes patogênicas que codificam proteínas truncadas nos genes *BRCA1*, *BRCA2* e *PALB2*, os efeitos correspondem ao fenótipo de alto risco para câncer de mama. Já para os genes *ATM*, *CHEK2*, *RAD51C* e *RAD51D*, como genes de penetrância moderada (BCAC, 2021).

O CO, de origem epitelial, tem uma forte etiologia genética hereditária, representada em 25% dos casos por mutações em genes *BRCA1/2*, 10% por mutações em outros genes e ainda 6%, por fatores poligênicos, considerá-lo um câncer possivelmente hereditário é o primeiro passo para identificação de mulheres e famílias de alto risco (Jevris *et al.*, 2015). LIU *et alli.*, 2021 publicaram recentemente uma revisão sobre os genes de predisposição ao CO considerando as evidências para a salpingooforectomia bilateral redutora de risco.

Para genes de alta penetrância em que o percentual de risco cumulativo é alto, como *BRCA1* e *BRCA2*, os quais chegam a 40% e 20%, respectivamente, os *guidelines* mais reconhecidos recomendam a SORR pois reduz a incidência do CO e diminui a mortalidade. Para outros genes em que a estimativa é maior que 10%, há evidências menos robustas, mas é considerável a abordagem cirúrgica em momento não antes dos 35-40 anos, como para os genes *MLH1* e *MSH2*. Para os demais dentro dessa faixa, deve-se considerar o momento cirúrgico mais próximo da menopausa (entre 45-50 anos) (Liu *et al.*, 2021). A tabela abaixo resume as informações sobre os genes de acordo com a associação de risco para CO e o grau do risco (Liu *et al.*, 2021).

Tabela 1: Associação e estimativa de risco para genes relacionados ao câncer de ovários de origem epitelial.

Gene	Associação de risco	grau do risco (%)
<i>BRCA1</i>	sim	>20
<i>BRCA2</i>	sim	>10
<i>MLH1</i>	sim	>10
<i>MSH2</i>	sim	>10
<i>MSH6</i>	sim	3-10
<i>PMS2</i>	incerto	3
<i>EPCAM</i>	incerto	sem dados
<i>BRIP1</i>	sim	>10
<i>RAD51C</i>	sim	>10
<i>RAD51D</i>	sim	>10
<i>PALB2</i>	sim	3-5
<i>ATM</i>	sim	3-4
<i>CHEK2</i>	Não	-

Fonte: Adaptado de Liu *et al.*,2021

Já em relação à predisposição hereditária a câncer de endométrio, os principais genes em que variante patogênicas estão associadas ao risco são os genes *MLH1*, *MSH2*, *MSH6*, relacionados à Síndrome de Lynch (SL) e o gene *PTEN*, à Síndrome de Cowden (SC).

Estima-se que cerca de 3% dos diagnósticos de câncer de endométrio representem pacientes com SL. Vale ressaltar que SL é a síndrome de predisposição ao câncer hereditário com maior prevalência (P) populacional (P=0,442%), para a qual o risco relativo para o CE é maior que 10 e, portanto, um alto risco (Dork *et al.*, 2020; Haraldsdottir *et al.*,2020). Embora a SC fosse uma síndrome rara em que a prevalência estimada é de 1 para 200.00 indivíduos, o risco para o CE também está dentro de uma faixa de alto risco. (Dork, *et al.*,2020; Yenias & Eng,2021). Existem outros gene para os quais há uma estimativa de risco para CE dentro de uma faixa intermediária, como o próprio gene *BRCA1*, entretanto, os estudos ainda mostram dados conflitantes para definição de medidas de redução de risco específicas voltada para o câncer de endométrio em mulheres portadoras de mutações em gene *BRCA1*(Dork *et al.*,2020).

Perspectivas de dados nacionais para o câncer de mama e ginecológico hereditário

No cenário nacional, os estudos populacionais brasileiros publicados ainda concentram dados advindos predominantemente da região sul e sudeste do país e abrangem principalmente mutações nos genes *BRCA1* e *BRCA2* (Palmero *et al.*, 2018; de Oliveira *et al.*, 2018; Cipriano *et al.*, 2018). O maior estudo publicado até o momento reportou 649 variantes patogênicas ou, provavelmente, patogênicas em *BRCA1* e *BRCA2*. Destas, 16 foi o número de casos advindos do estado de Minas Gerais. Embora fosse o maior estudo, há diferenças significativas de dados dos pacientes quanto ao tipo de teste genético realizado e limitações quanto aos dados clínicos e demográficos dos pacientes (Palmero *et al.*, 2018). Estudos nacionais

dentro do contexto de painéis multigenes ainda são pouco expressivos em relação ao “n” amostral para definição de padrões populacionais, mas fazem parte do esforço da comunidade científica nesta área do conhecimento (Cipriano et al., 2019). A Agência Nacional de Saúde Suplementar (ANS) incluiu, na lista de exames com cobertura pelos planos de saúde da rede privada, a pesquisa de mutações em genes *BRCA1* e *BRCA2* em casos específicos de câncer de mama e Painel NGS para os genes associados ao risco para câncer de mama, ovários e outros tipos de câncer em determinadas situações em que há dados clínicos para se avaliar diagnósticos diferenciais de mais de uma síndrome de predisposição ao câncer hereditário. Com relação à saúde pública, há possibilidade do teste de mutações germinativas nos genes *BRCA1* e *BRCA2* no Sistema Único de Saúde (SUS), de acordo com Projeto de Lei 6262/2013, de autoria da Deputada Carmem Zanotto (PPS/SC), que delega ao SUS a cobertura dos testes para pesquisa de mutações nos genes *BRCA 1* e *BRCA2* em mulheres com histórico familiar de diagnóstico de câncer de mama e/ou ovário.

Sendo assim, a perspectiva para nossa população é de que um número cada vez maior de pacientes, tanto da rede privada quanto pública, realizarão testes genéticos para guiar a escolha de tratamento do câncer, manejo de redução de risco e aconselhamento familiar. Portanto, é necessário compreender melhor o genótipo e fenótipo das variantes nesses genes, mais comuns ou mais raros, dentro de um contexto populacional específico.

II) Aconselhamento genético e testes genéticos na abordagem do câncer ginecológico

O Aconselhamento Genético

A medicina genômica é uma prioridade estratégica e crescente para muitos sistemas de saúde e o aconselhamento genético (AG) faz parte disso. O AG é um processo de comunicação que capacita pacientes e famílias a tomar decisões autônomas e usar essas informações diante de uma condição genética hereditária. Esse processo é conduzido por um profissional treinado para realização do AG (Patch & Middleton, 2018). Uma das principais ferramentas utilizadas no AG são os testes genéticos que podem ter várias finalidades, que vão desde exames simples como análise de uma

única mutação, análise de inúmeros genes por NGS ou mesmo estudo do exoma ou genoma completos. Mundialmente, já se reconhece que os avanços na tecnologia genômica fornecerão dados em escala acessível e tem o potencial para fornecer diagnóstico direcionado de baixo custo através de testes genéticos de alta complexidade (Davies, 2017). Assim, é importante atentarmos para medicina genômica dentro de um contexto ético e multidisciplinar na abordagem dos pacientes e com profissionais capacitados para lidar com essas informações.

Quanto aos princípios para avaliação de risco de câncer hereditário, a decisão para realização de teste genético e abordagem do manejo clínico do paciente envolve três etapas. A primeira se trata do aconselhamento pré-teste; a segunda, é voltada para a escolha-teste mais apropriada e quais as implicações e limitações do mesmo e a terceira etapa se refere ao aconselhamento pós-teste e, conseqüentemente, condutas diante do resultado do teste (*National Comprehensive Cancer Network, 2021*).

O aconselhamento genético é um processo dinâmico, que evolui não só com o aprimoramento do conhecimento, mas também com a evolução e mudanças que ocorrem na vida do indivíduo. Questões importantes são abordadas no AG voltadas para fatores sociais, econômicos e psicológicos. Muitas vezes a percepção de risco do paciente depende muito da forma como é transmitida pelo profissional assistente. Assim, uma mesma informação pode ter interpretações diferentes para cada pessoa. Isso vale para reações emocionais ou psíquicas. Geralmente, a avaliação de risco para câncer se baseia em estimativa de risco ao longo da vida (*lifetime risk*). Isso não representa, necessariamente, o risco de ter câncer a qualquer momento ou nos próximos cinco a dez anos. O risco ao longo da vida, por exemplo, para uma mulher sadia de 25 anos e portadora de mutação em gene BRCA1, é de 70% ao longo dos próximos 60 anos, mas é de menos de 1% de ter um câncer de mama naquele momento (Henson & Resta, 2021).

Os critérios para indicação de teste genético que contemplam, pelo menos, os genes de alta penetrância para risco de câncer de mama e ovários, incluem: idade precoce (menor ou igual a 45 anos), câncer de mama associada ao histórico pessoal e/ou familiar de câncer de mama, pâncreas e/ou próstata, dados sobre tipo e grau

histológico do câncer de próstata, ascendência familiar judaica Ashkenazi, história clínica compatível com alguma síndrome de predisposição associada aos genes (Síndromes de Cowden, Li-Fraumeni, Peutz-Jeghers ou Síndrome do câncer gástrico difuso associada ao câncer de mama lobular), dentre outros. Alguns critérios, são independentes do contexto da idade de diagnóstico de câncer e da história familiar. Para estes, o teste genético deve ser indicado e podem ter impacto, inclusive na decisão terapêutica quimioterápica (*National Comprehensive Cancer Network, 2021*).

São eles:

- Terapia com inibidor de PARP (iPARP) em câncer de mama metastático;
- uso de Olaparib (iPARP) em câncer adjuvante em câncer de mama de alto risco Her2 negativo;
- câncer de mama triplo negativo;
- câncer de ovários;
- câncer de mama em homens.

Possibilidade de resultados e complexidade para interpretação de variantes

A classificação das variantes de acordo com os critérios estabelecidos pelo Colégio Americano de Genética Médica e Genômica e a Associação de Patologia Molecular - *The American College of Medical Genetics and Genomics and the Association for Molecular Pathology - ACMMG/AMP*, descritos no quadro 1 (Richards *et al.*, 2015) se resumem em:

- Benignas (VB) e provavelmente benignas (VPB), cuja interpretação não está associada a um evento causal para uma doença genética, geralmente não reportada nos laudos dos testes genéticos;
- Variante de significado incerto (VUS), em que não haja evidências para classificar a variante e nem associar à doença
- Provavelmente patogênica (VPP) e patogênica (VP), cuja variante representa um achado genotípico de associação causal.

O passo a passo para análise de uma variante depende de informações de base clínica e laboratorial que permitem guiar a classificação também de acordo com o

ACMMG/AMP. Essa análise se baseia nos seguintes quesitos, como (Henson & Resta, 2021):

(i) Correlação genótipo-fenótipo: existe alguma associação entre o achado laboratorial e a clínica do paciente?

(ii) Patogenicidade:

- O efeito da mutação em nível molecular é relevante?
- Trata-se de uma mutação com frequência alélica baixa de acordo com banco de dados? Mutações com baixa frequência alélica populacional tendem a ser patogênicas;
- A mutação já foi descrita nos bancos de dados ?;
- Haploinsuficiência - fenômeno comum em genes de suscetibilidade ao câncer hereditário, em que uma cópia mutante é suficiente para predispor o fenótipo variante/mutado;
- Variação do aminoácido (aa): o tipo de substituição do aa pode ser uma evidência para patogenicidade da mutação;
- Conservação dos aa através de espécies - análise conservacional;
- Análise de modelo de predição "*in silico*" - programas computacionais podem prever o efeito da variante em questão;

(iii) Análise constitucional;

(iv) Avaliação da fração alélica - dependendo da fração alélica encontrada, sugere-se que as mutações são de linhagem germinativa, somática ou pode se tratar de um evento pós-zigótico associado ao mosaicismo;

(v) Estudo de segregação familiar - análise da segregação da variante em familiares afetados e assintomáticos pode ser uma pista para interpretação e correlação genótipo-fenótipo das variantes;

(vi) Análise da variante e outros tecidos do paciente - auxílio para esclarecer se a mutação é somática ou pode ser um mosaicismo.

Considerando essas análises, classifica-se cada variante de acordo com os critérios que são de patogenicidade e de benignidade. A partir desta interpretação, usa-se uma regra para unir esses critérios e, então, classificar cada variante. Uma variante é classificada como VUS quando não é possível obter informações suficientes para classificá-la como VP, VPP, VB ou VPB (Quadro 2).

Quadro 1: Descrição da classificação das variantes e nível de patogenicidade e benignidade de acordo com ACMG/AMP

DESCRIÇÃO	INTERPRETAÇÃO
CRITÉRIO DE PATOGENICIDADE	
PVS1 - Variante nula (<i>nonsense</i> , <i>frameshift</i> , canônico \pm 1 ou 2 locais de <i>splice</i> , códon de iniciação, deleção única ou multiexon) em um gene e <i>LOF</i> é um mecanismo conhecido de doença.	EVIDÊNCIA MUITO FORTE DE PATOGENICIDADE
PS1 - Mesma mudança de aminoácido que uma variante patogênica previamente estabelecida, independentemente da mudança de nucleotídeo.	EVIDÊNCIA FORTE DE PATOGENICIDADE
PS2 - De Novo confirmado	EVIDÊNCIA FORTE DE PATOGENICIDADE
PS3 - Estudos funcionais <i>in vitro</i> ou <i>in vivo</i> bem estabelecidos que apoiam um efeito prejudicial no gene ou produto gênico.	EVIDÊNCIA FORTE DE PATOGENICIDADE
PS4 - Prevalência em pessoas afetadas maior do que nos controles	EVIDÊNCIA FORTE DE PATOGENICIDADE
PM1 - Localizado em uma região <i>hot spot</i> mutacional ou domínio funcional crítico e bem estabelecido (por exemplo, sítio ativo de uma enzima) sem variantes benignas frequentes	EVIDÊNCIA MODERADA DE PATOGENICIDADE
PM2 - Ausente ou com baixa frequência em bancos de dados populacionais	EVIDÊNCIA MODERADA DE PATOGENICIDADE
PM3 - Para recessivos - detectável com uma mutação patogênica em trans	EVIDÊNCIA MODERADA DE PATOGENICIDADE
PM4 - O comprimento da proteína muda como resultado de deleções/inserções <i>in-frame</i> em uma região não repetida ou variantes de <i>stop-loss</i> .	EVIDÊNCIA MODERADA DE PATOGENICIDADE
PM5 - <i>Missense</i> que altera um resíduo de aminoácido, no qual uma <i>missense</i> diferente é previamente estabelecida como patogênica	EVIDÊNCIA MODERADA DE PATOGENICIDADE
PM6 - De Novo não confirmado	EVIDÊNCIA MODERADA DE PATOGENICIDADE
PP1 - Variante segregada com doença em múltiplos membros da família afetados	A FAVOR DE PATOGENICIDADE
PP2 - Variante <i>missense</i> em um gene que tem uma baixa taxa de variantes <i>missense</i> benigna e em que as variantes <i>missense</i> são um mecanismo comum de doença.	A FAVOR DE PATOGENICIDADE
PP3 - Múltiplas linhas de evidência computacional com um efeito deletério sobre o gene ou produto do gene (conservação, evolução, impacto de <i>splicing</i> , etc.)	A FAVOR DE PATOGENICIDADE
PP4 - O fenótipo do paciente ou história familiar é altamente específico para uma doença com um único gene etiologia	A FAVOR DE PATOGENICIDADE

PP5 - Fonte confiável relatou recentemente a variante como patogênica, mas a evidência não está disponível para o laboratório realizar uma avaliação independente.	A FAVOR DE PATOGENICIDADE
CRITÉRIOS DE BENIGNIDADE	
BP1 - Variante <i>missense</i> em um gene, sendo que, principalmente, as variantes truncadas são conhecidas por causar doenças.	A FAVOR DE BENIGNIDADE
BP2 - Observado em trans com uma variante patogênica para um gene/distúrbio dominante totalmente penetrante ou observado em cis com uma variante patogênica em qualquer padrão de herança	A FAVOR DE BENIGNIDADE
BP3 - Deleções/inserções in-frame em uma região repetitiva sem uma função conhecida.	A FAVOR DE BENIGNIDADE
BP4 - Múltiplas linhas de evidência computacional não sugerem nenhum impacto no gene ou produto do gene (conservação, evolução, impacto de <i>splicing</i> , dentre outras)	A FAVOR DE BENIGNIDADE
BP5 - Causa alternativa (variante encontrada em um caso com uma base molecular alternativa para doença)	A FAVOR DE BENIGNIDADE
BP6 - Fonte confiável relatou recentemente a variante como benigna, mas a evidência não está disponível para o laboratório realizar uma avaliação independente.	A FAVOR DE BENIGNIDADE
BP7 - Uma variante sinônima (silenciosa) para a qual algoritmos de previsão de <i>splicing</i> não preveem nenhum impacto na sequência de consenso de <i>splice</i> nem a criação de um novo local de <i>splice</i> . E o nucleotídeo não é altamente conservado.	A FAVOR DE BENIGNIDADE
BS1 - Alta frequência (comum)	FORTEMENTE BENIGNA
BS2 - Presente no controle saudável	FORTEMENTE BENIGNA
BS3 - Sem consequência funcional	FORTEMENTE BENIGNA
BS4 - Falta de segregação nos membros afetados de uma família.	FORTEMENTE BENIGNA
BA1 - Muito comum (frequência do alelo é > 5%)	critério independente para BENIGNIDADE

Nota LOF: *Loss of function*. Termo em inglês para representar perda de função

Quadro 2: Critérios do ACMG/AMP para definição da interpretação das variantes como Patogênica e provavelmente patogênica.

Variante Patogênica	Variante Provavelmente Patogênica
<p>1- 1 Muito fortemente (PVS1) e</p> <p>a) ≥ 1 fortemente (PS1-PS4) ou</p> <p>b) ≥ 2 moderada (PM1-PM6)</p> <p>c) ≥ 1 moderada (PM1-PM6) e 1 a favor de patogenicidade (PP1-PP5)</p> <p>2- ≥ 2 fortemente (PS1-PS4) ou</p> <p>3- 1 Fortemente (PS1-PS4 e</p> <p>a) ≥ 3 moderada (PM1-PM6)</p> <p>b) 2 moderadamente (PM1-PM6) e ≥ 2 a favor de patogenicidade (PP1-PP5) ou</p> <p>c) 1 moderada (PM1-PM6) e ≥ 4 de patog.(PP1-PP4)</p>	<p>1- 1 muito fortemente (PVS1) e 1 moderada (PM1-PM6)</p> <p>2- Forte (PS1-PS4) e 1-2 moderadas (PM1-PM6)</p> <p>3- 1 forte (PS1-PS4) e $2 \geq$ a favor de patog.(PP1-PP5) ou</p> <p>4- ≥ 3 moderada (PM1-PM6) ou</p> <p>5- 2 moderadas (PM1-PM6) e ≥ 2 a favor de patog. (PP1-PP5) ou</p> <p>6- 1 moderada (PM1-PM6) e ≥ 4 a favor de (PP1-PP5)</p>

Fonte: adaptado de (Henson & Resta,2021)

Quadro 3: Critérios do ACMG/AMP para definição da interpretação das variantes como Benigna e provavelmente benigna.

Variante benigna	Variante provavelmente benigna
<p>1- 1 critério independente (BA1) ou</p> <p>2- \geq fortemente (BS1-BS5)</p>	<p>1- 1 fortemente (BS1-BS4) e 1 a favor de benign. (BP1-BP7) ou</p> <p>2- ≥ 2 a favor de (BP1-BP7)</p>

Fonte: adaptado de (Henson & Resta,2021)

Os resultados sobre a classificação das variantes em genes incorporados nos painéis NGS podem variar entre informativos quando positivos para variantes patogênicas ou não informativos quando negativos para variantes patogênicas ou presença de variantes de significado incerto. Entretanto, mesmo que existam critérios para padronizar a classificação dessas variantes, há discordância de resultados. Então, a colaboração entre laboratórios para tornar essas mutações evidentes na literatura e nos bancos de dados é fundamental (Pepin *et al.*, 2016; Harrison *et al.*, 2017). Para muitos genes de suscetibilidade ao câncer, incorporados nos painéis NGS, a magnitude de risco é pouco conhecida e a baixa prevalência de mutações torna difícil evoluir com essa estimativa de risco (Easton *et al.*, 2015). Muitas interpretações conflitantes são de baixa significância clínica porque a discrepância da classificação varia entre uma interpretação benigna/provavelmente benigna ou VUS. Portanto, a conduta clínica acaba por se basear mais na história clínica pessoal e familiar do paciente do que no resultado do teste. No entanto, a identificação de uma VUS pode causar uma grande incerteza para os pacientes e profissionais de saúde, gerar uma “ansiedade” frente ao resultado e aumentar o risco de uma conduta médica voltada para um alto risco de câncer e, conseqüentemente, inadequada para o paciente (Balmaña *et al.*, 2016).

A interpretação das variantes depende, principalmente, de bons testes funcionais e banco de dados de variantes bem estabelecidos. Além disso, compartilhar informações com o laboratório, ou seja, apresentação de dados fenotípicos passíveis de serem analisados com o genótipo e a colaboração por parte dos laboratórios clínicos com especialistas reconhecidos, provavelmente, melhoraria a consistência na interpretação de variantes entre laboratórios e possibilitaria melhor assistência ao paciente (Pepin *et al.*, 2016).

JUSTIFICATIVA

Avaliar e manejar a condição de risco de câncer de mama e ginecológico hereditários de maneira mais específica depende de testes genéticos que revelam resultados, muitas vezes, pouco elucidativos, conflitantes e desafiadores para guiar condutas médicas. Elevar a análise e reanálise dessas variantes para um nível mais

aprofundado pode favorecer a compreensão sobre o real impacto das variantes quanto a um efeito causal ou não.

OBJETIVO GERAL

Avaliar o impacto da análise “in silico” na predição de patogenicidade e reclassificação de variantes em genes de suscetibilidade a neoplasias ginecológicas.

OBJETIVOS ESPECÍFICOS

- i) Documentar variantes específicas no gene *FH* e *MSH2* não reportadas ou passíveis de reclassificação por estudo de casos.
- ii) Realizar levantamento de casos associados a variantes patogênicas ou VUS em gene *RAD51C* e propor reclassificação de variantes

Esse estudo foi realizado a partir do levantamento e análise do estudo apresentado no formato pôster no Congresso Internacional da Sociedade Brasileira de Oncologia em parceria com a *American Association of Cancer Research*, em outubro de 2020 e submetido na Revista Brasileira de Ginecologia e Obstetrícia: “GERMINATIVE MUTATIONS LANDSCAPE IN MINAS GERAIS, BRAZIL PATIENTS UNDERGONE TO GENETIC COUNSELING FOR GYNECOLOGICAL AND BREAST CANCER” (Carvalho *et al.*,2020). Neste trabalho 382 pacientes foram submetidos ao aconselhamento genético foi identificado variantes que despertaram o interesse dos colaboradores, principalmente voltado para variantes não descritas ou raras. Duas delas geraram os trabalhos apresentados nos relatos de casos. Foi montado um grupo de trabalho de análise de variantes composto por médicos, incluindo médicos geneticistas e biólogos com experiência na área de bioinformática para auxiliar nas análises e interpretação das variantes. Com a evolução do trabalho e parceria de laboratórios privados foi possível evoluir com coleta de variantes em genes *RAD51C*, gene de interesse de estudo do grupo de trabalho.

APROVAÇÃO DO COEP

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais sob CAAE.01758418.0.0000.5149 e CAAE.48770621.6.0000.5149 (Anexo I).

DESENVOLVIMENTO E METODOLOGIA PARA ANÁLISES DAS VARIANTES

Os resultados deste trabalho serão apresentados na forma de capítulos, em que os dois primeiros capítulos correspondem aos artigos científicos produzidos durante o doutoramento. O terceiro capítulo corresponde ao estudo de análise direcionada para o gene *RAD51C*, cujo artigo ainda está em processo de elaboração.

O primeiro capítulo corresponde ao trabalho submetido à revista "*Genetics and Molecular Biology 2022*", cujo título é "*Evidence for pathogenicity of MSH2 c.1894_1898; p.Ile633Lysfs*9 mutation in Lynch syndrome: genotype-phenotype correlation and bioinformatic approaches*". Os objetivos principais deste trabalho foram:

- (i) descrever passo a passo a análise bioinformática da mutação germinativa no gene *MSH2*, previamente não descrita, em uma família com Síndrome de Lynch e
- (ii) investigar as evidências de patogenicidade em nível proteico e de RNA,
- (iii) estabelecer uma correlação genótipo-fenótipo para essa mutação.

O segundo capítulo corresponde ao trabalho intitulado "*In deep Bioinformatic characterization of a novel fumarate hydratase mutation (FH c.199T>G; p.Tyr67Asp) in hereditary leiomyomatosis and renal cell carcinoma*" submetido à revista *Familial Cancer 2022*. Teve como objetivos:

- (i) promover a reclassificação da VUS para "provavelmente patogênica" encontrada em uma família com fenótipo clínico da síndrome;

- (ii) descrever a análise detalhada para evidências de patogenicidade para a variante.

O terceiro capítulo descreve o trabalho realizado para reanálise de possíveis reclassificações de variantes em gene *RAD51C*, associado ao risco para câncer ginecológico e de mama, coletadas em diferentes serviços médicos e laboratoriais.

A análise das variantes é apresentada em cada capítulo de maneira aprofundada e descrevendo o passo a passo da metodologia utilizada (Figura 1). Para cada etapa das análises, as seguintes ferramentas foram disponibilizadas:

Predição dos impactos funcionais das mutações na estrutura da proteína

- Para predição do impacto funcional, foram usados: Poly-Phen-2 (Adzhubei *et al.*, 2010), SIFT (Vaser *et al.*, 2016), MCAP (Jagadeesh *et al.*, 2016), PROVEAN (Choi *et al.*, 2012), Mutation Taster (Steinhaus *et al.*, 2021). Nesta etapa, foram usadas outras ferramentas como: GeneCards (Stelzer *et al.*, 2016), 1000 Genomes (Danecek *et al.*, 2011), *Universal Protein Resource* (UniProt) (UniProt Consortium, 2015), Pfam (Finn *et al.*, 2014); DGV (MacDonald *et al.*, 2013); *European Variation Archive* (EVA, EMBL-EBI, 2021), entre outras.

Avaliação da conservação evolutiva

- Uma evidência indireta de funcionalidade pode ser obtida com a análise da conservação evolutiva da posição nucleotídica e do contexto genômico no qual a variante está inserida - Clustal Omega (Madeira *et al.*, 2019).

Análise de Splicing alternativo

- *Human Splicing Finder* (HSF) (Desmet *et al.*, 2009).

Análise de Isoformas

- *GeneCards*, *BLAST* (Altschul *et al.*, 1990).

Avaliar o padrão de expressão tecido específico

- *GTex* (*GTEx Consortium*, 2013), *BioGPS* (Wu et al., 2016), *Proteomics* (Patroklos et al., 2020).

Avaliar os efeitos fenotípicos

- OMIM – *Online Mendelian Inheritance in Men* (Amberger et al., 2015).

Verificar quanto ao estado de homocigoto ou heterocigoto

- Para cada indivíduo, verificar se a variante está presente em homocigose ou heterocigose.

Avaliar a presença em regiões conservadas evolutivamente

- Avaliação de conservação evolutiva de regiões do genoma: *BLASTP*, *The Consurf Server* (Ashkenazy et al., 2016).

Modelagem por homologia de estruturas tridimensionais de proteínas

- Modeller (Webb et al., 2016), PyMol (Zhu et al., 2014).

Reclassificação final das variantes e definição de nomenclatura

- Diretrizes do Colégio Americano de Genética Médica e Genômica (*American College of Medical Genetics and Genomics*) e atualizações (Richards et al., 2015; Nykamp et al., 2017; Abou Tayon et al., 2018).

Os capítulos I e II estão redigidos no formato das revistas para as quais foram submetidas.

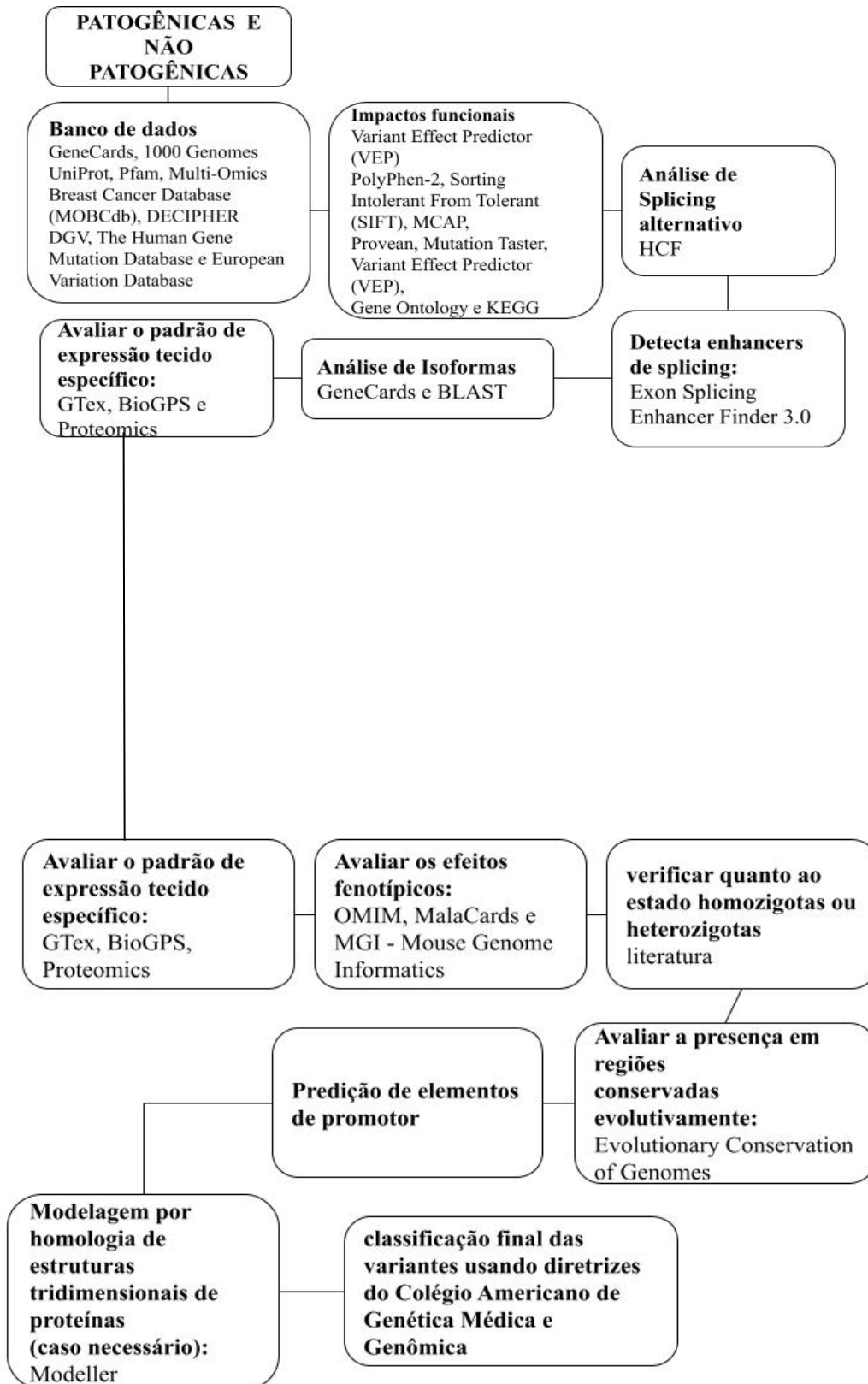


Figura 1: Fluxograma com descrição das etapas para análises das variantes e ferramentas utilizadas

CAPÍTULO I

Evidence for pathogenicity of *MSH2* c.1894_1898; p.Ile633Lysfs*9 mutation in Lynch syndrome: genotype-phenotype correlation and bioinformatic approaches

Running title - *MSH2* c.1894_1898; p.Ile633Lysfs*9

Anisse Chami^{1,2}

¹Programa de Pós-Graduação em Tocoginecologia da Universidade Estadual de São Paulo - UNESP, Botucatu, SP, Brazil

²Serviço Especial de Genética Médica do Hospital das Clínicas da Universidade Federal de Minas Gerais - UFMG, Belo Horizonte, MG, Brazil

ORCID: 0000-0002-2651-0995

Thalía Rodrigues de Souza³

Belo Horizonte, MG, Brazil

ORCID: 0000-0001-8473-2548

Carolina Guimarães Ramos Matozinho³

³Programa de Pós-Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

ORCID: 0000-0001-8070-1031

Agnaldo Lopes da Silva Filho^{4,5}

⁴ Departamento de Ginecologia e Obstetrícia, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

⁵ Programa de Pós-Graduação em Tocoginecologia da Universidade Estadual de São Paulo - UNESP, Botucatu, SP, Brazil

ORCID: 0000-0002-8486-7861

Maria Raquel Santos Carvalho⁶

⁶ Programa de Pós-Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

mraquel-carvalho@ufmg.br

ORCID: 0000-0002-1744-448-X

Letícia da Conceição Braga^{7,8}

⁷ Núcleo de Pesquisa Básica e Translacional, Instituto Mário Penna, Belo Horizonte, MG, Brazil

⁸ OncoTag Desenvolvimento de Produto e Serviços para Saúde Humana, Belo Horizonte, MG, Brazil

ORCID: 0000-0002-6181-9410

Anisse Chami and Thalía de Souza contributed equally to this work.

Corresponding author - Maria Raquel Santos Carvalho

Departamento de Genética, Ecologia e Evolução

Instituto de Ciências Biológicas

Universidade Federal de Minas Gerais

Av. Antônio Carlos, 6627 - Campus Pampulha

31.270-901

Fone: +55 31 3409-2598

Fax: +55 31 3409-2570

mraquel-carvalho@ufmg.br

ma.raquel.carvalho@gmail.com

ORCID: 0000-0002-1744-448-X

ABSTRACT

Mutation screening in familial cancer syndromes supports clinical decisions concerning prophylaxis, treatment, and genetic counseling. Typically, genetic variants are classified according to evolutive conservation, populational frequency, and substitution effects on amino acid class and polarity. Variants of uncertain significance (VUS) are common. Frame-shift mutations (fsm) are presumably pathogenic, but their effects may be counteracted by cellular mechanisms functioning at the RNA or protein levels. Damage reducing mechanisms include nonsense-mediated decay (NMD), alternative splicing, and exon skipping. Consequently, even fsm should be carefully analysed. Here, we describe a multi-step Bioinformatic evaluation of the germline MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation

found in a Lynch syndrome family. Evidence for pathogenicity includes, at protein level, the loss of the 294 C-terminal residues, including the ATPase domain. Domains involved in dimerization and DNA binding are present, therefore, negative dominance is possible. At the RNA level, this fsm causes a premature termination code highly suggestive of NMD, a better explanation for the loss of MSH2 expression, detected in the proband's tumor immunohistochemistry. No evidence of rescue mechanisms (alternative splicing, exon skipping) were found. These findings provide evidence for the classification of MSH2 c.1894_1898del; p.Ile633Lysfs*9 as a pathogenic mutation.

Keywords - Lynch syndrome; hereditary cancer; colorectal cancer, mismatch repair; MSH2

INTRODUCTION

Lynch syndrome (LS; OMIM 120435), autosomal dominant, is the most common hereditary cancer predisposition syndrome (Lynch, Snyder et al. 2015; Gupta, Gellert et al. 2012). It is characterized by incomplete penetrance and heterogeneous clinical manifestations due to germline deleterious mutations in the post-replicative Mismatch Repair (MMR) genes *MLH1* (OMIM 120436), *MSH2* (OMIM 609309), *MSH6* (OMIM 600678), and *PMS2* (OMIM 600259). In addition, germline mutations in the *EPCAM* gene can influence *MSH2* expression and cause LS. Mutations in *MLH1* and *MSH2* present the highest penetrance, with a risk for colorectal cancer exceeding 50%, while *PMS2* and *MSH6* mutations present lower penetrance. The MMR proteins correct DNA replication errors, reducing mutation rates; control the fidelity of meiotic recombination, blocking recombination between non-homologous sequences; help dealing with environmental mutagens, mediating DNA damage-induced apoptosis, among other functions (Crouse 2016; Toft, Winton et al. 1999; Li, Chao et al. 2020). Different cancer types including gastrointestinal, gynecological, skin, urinary tract, and

central nervous system have been associated with LS and the risk of developing them depends on the gene being affected (Lynch, Snyder et al. 2015). Colorectal and endometrial cancers are the most frequent in LS. It has been estimated that 3% of all colorectal cancer and 1.8% of endometrial cancers are associated with LS (Biller, Syngal et al. 2019).

Over the last decades, knowledge on genotype-phenotype correlations for each LS gene has allowed the development of risk reduction management protocols for patients with pathogenic mutations in the MMR genes (Dominguez-Valentin, Sampson et al. 2020). Complete anamnesis with family history, tumor pathology, microsatellite stability analysis, and immunohistochemistry of MMR proteins are crucial to guide LS diagnosis and management. Thus, the Revised Bethesda Guidelines and Amsterdam (I and II) clinical criteria enable the recognition of highly suspected patients or families (Umar, Boland et al. 2004; Vasen, Mecklin et al. 1991; Vasen, Watson et al. 1999). Prediction models such as PREMM5 also help the screening of individuals that could be mutation carriers. Furthermore, the immunohistochemistry and/or microsatellite stability analysis of all colorectal and, even, endometrial cancers as part of the Universal Testing is now proposed to be considered in order to identify more individuals at risk (Thompson, Spurdle et al. 2014). However, even with all these tools, the definitive diagnosis of LS is established through the identification of a pathogenic mutation (Lynch, Snyder et al. 2015). Consequently, the genetic-molecular studies are fundamental, not only for the identification of germline and/or tumor mutations associated with LS, but also for establishing their phenotype impacts.

Several mutations in the MMR genes have already been described in cancer syndromes and shared in the main international variants databases, such as InSiGHT, Varsome, HGVD, among others). The identification and registration of mutations and associated phenotypes, also assist in the pursuit of better care and in the identification of specific population or families risk groups (Thompson, Spurdle et al. 2014). However, mutations are frequent in the

general population and may not be associated with cancer causation. General criteria for the classification of mutations in MMR genes have been established (Plazzer, Sijmons et al. 2013). Nonetheless, the main limitation in the clinical use of sequencing methodologies for molecular diagnosis is the high frequency of variants of uncertain significance (VUS). Consequently, it is fundamental to better characterize the pathogenic potential of newly identified genetic variants and the strategies needed will vary according to each variant. Here, we describe the phenotype associated and the results of the Bioinformatic analysis conducted in order to evaluate the pathogenic potential of a *MSH2* mutation. We also outlined the mutation impact in the protein domain in order to guide a better understanding of the molecular interaction mechanisms of MMR genes/proteins.

PATIENTS AND METHODS

The proband was referred to genetic counseling considering LS as a diagnostic hypothesis. She was 55 years old and was diagnosed with metachronous nonpolypoid colorectal cancer in the previous year. The first tumor was localized in the sigmoid colon and the second one was diagnosed 12 months later in the cecum, measuring 4.0 x 3.6 x 2.0 cm. This second tumor was not visualized at the time of the first colonoscopy and left colectomy for treating the first tumor. According to her colonoscopy records, the cecum was not totally visualized at that time. Consequently, she underwent a second colectomie for the cecum tumor treatment. Both tumors were pathologically described as moderately differentiated invasive adenocarcinomas. The pathological records did not describe mucinous, signet-ring cell histologic features, a medullary growth pattern, not even infiltrating lymphocytes in the tumor. Immunohistochemical analysis of MMR proteins showed loss of MSH2 and MSH6 expression in the first tumor and loss of only MSH2 expression, in the second tumor. At cancer staging, both tumors were classified as initial, without signals of invasion or spreading.

Consequently, both tumors were resected and further oncological treatments were not required.

The family history was relevant for colorectal cancer and it was reported that her grandmother would have had gynecological/uterine cancer. Unfortunately, the proband has had little contact with members of her extended family and was unable to provide details about the family history and to add confirmation reports or documents regarding the type of tumor for distant family members. The family comes from Teofilo Otoni, a city in the North of Minas Gerais state, Southeastern Brazil, where medical care is quite limited. Even so, the family history meets Amsterdam II criteria as illustrated in the pedigree below (Fig. 1).

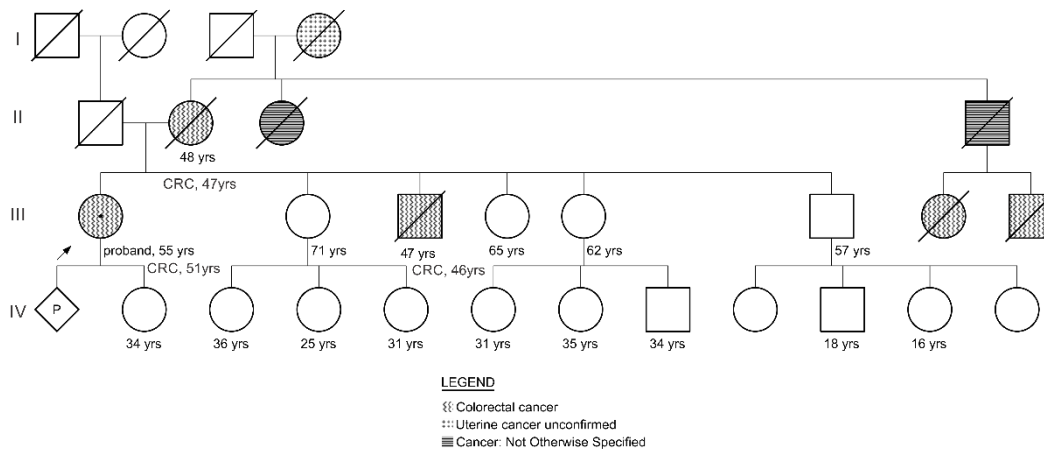


Figure 1. Proband pedigree showing familial cancer, age of onset, and cancer types. Age in years refers to the current age of the person. A well defined age at diagnosis was obtained for III.1, (51 years old), II.2 (47 years old), III.3 (46 years old).

After the pre-test genetic counseling, a Next-Generation Sequencing (NGS) based, multi-gene panel for hereditary cancer genes was performed at a CLIA-approved laboratory using DNA extracted from the patient's blood. The panel included all MMR genes associated with LS and *EPCAM*. Detection of copy number variations was not conducted. A novel, germline variant was identified, *MSH2* c.1894_1898del; p.Ile633Lysfs*9, present in heterozygosis. A multi-step Bioinformatic investigation of this variant was conducted in the search of evidence for

pathogenicity and molecular mechanisms involved. These Bioinformatic analyses included: 1. The search for previous descriptions of this variant in ClinVar (Landrum, Chitipiralla et al. 2020), Database of Genomic Variants (MacDonald, Ziman et al. 2014), Ensembl (Hunt, McLaren et al. 2018), OMIM (McKusick, 1998), PubMed (Agarwala, Barrett et al. 2018), European Variation Archive (EMBL-EBI 2020), GeneCard (Stelzer, Rosen et al. 2016), InSIGHT DNA Variant Database (Thompson, Spurdle et al. 2014) (International Society for Gastrointestinal Hereditary Tumours; Thompson et al., 2014), Varsome (Kopanos, Tsiolkas et al. 2019), and ABraOM (Naslavsky, Scliar et al. 2020) ; 2. Mutation impacts were predicted using the Variant Effect Predictor (VEP) (Hunt, McLaren et al. 2018); 3. The *MSH2* c. 1894_1898del CDS sequence was produced by editing the FASTA format sequence of the *MSH2* CDS (NM_000251.3) downloaded from the National Center for Biotechnology Information (NCBI) (Agarwala, Barrett et al. 2018); 4. Mutation effects at protein level were evaluated by translating the *MSH2* c. 1894_1898del CDS, created as described above, using ExPASy Translate (Artimo, Jonnalagedda et al. 2012); 5. The search for MSH2 protein isoforms was conducted in the UNIPROT database (Consortium 2019). Initially, two MSH2 protein isoforms were found (P43246-1 and P43246-2) and compared using Basic Local Alignment Search Tool-Proteins (BLASTP) (Altschul, Gish et al. 1990). Isoforms were also predicted using the Variant Effect Predictor (VEP) (Hunt, McLaren et al. 2018) and GTEx (GTEx, 2020); 6. The hypothesis of nonsense-mediated decay (NMD) and exon skipping were evaluated by CDS and protein database search and in the literature; 7. Information on MSH2 protein domains was obtained using Pfam (El-Gebali, Mistry et al. 2019); 8. The impact of this mutation in MSH2 tridimensional (3D) structure was evaluated using modeling by homology using Modeller 9.24 (Šali and Blundell 1993), with eight molds (2O8B, 2O8C, 2O8D, 2O8F, 3THW, 3THX, 3THY, and 3THZ) obtained from PDB (Heinz, Baase et al. 1993). Ten models of the *MSH2* c.1894_1898del; p.Ile633Lysfs*9 were produced. Among

them, those five presenting the lowest DOPE (discrete optimizer protein energy) values were selected for the production of the Ramachandran plots in PDBsum (Laskowski, Jabłońska et al. 2018) and for the estimation of Z-scores using PROSA-Web (Protein Structure Analyses) software (Wiederstein and Sippl 2007). The better models were selected based on the higher PROCHECK/Z-score. Once validated the models, the one with the highest PROCHECK/Z-score value was selected for the comparison of the 3D structures (MSH2 wild-type and *MSH2* c.1894_1898del; p.Ile633Lysfs*9) using the PyMOL software (Schrödinger, 2017).

RESULTS

In this study, a patient presenting LS phenotype was investigated clinical-, immunohistochemical-, and molecularly. The clinical and familial presentation showing metachronous colorectal cancer and family members also presenting colorectal cancer at an early age or possible gynecological cancer pointed towards a clinical diagnostic for LS. The immunohistochemical analysis of the tumors, evidencing loss of MSH2 expression, reinforces the LS syndrome hypothesis. A germline mutation identified (*MSH2* c.1894_1898del; p.Ile633Lysfs*9) (Fig. 2) was not found in all the mutation databases searched, except for Varsome, where it is listed as pathogenic; however, molecular mechanisms or associated phenotypes are not provided. This mutation was also not described in studies which included Brazilian samples published so far (Vaccaro, López-Kostner et al. 2019; Schneider, Pastor et al. 2018).

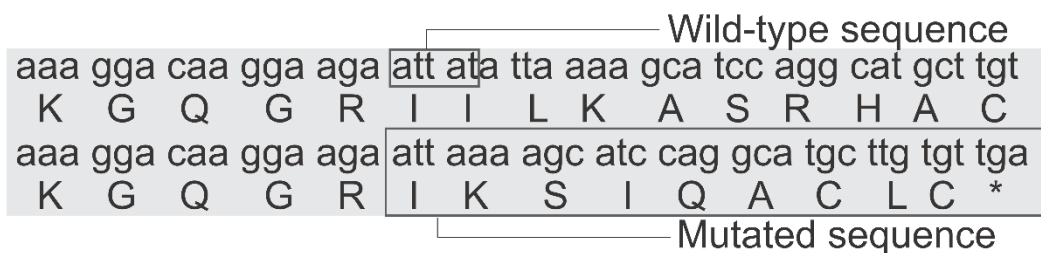


Figure 2. Comparative alignment of MSH2 mRNA and protein, wild-type vs. MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation. The 5-nucleotides deletion is shown within a box in the wild-type mRNA sequence. Observe that the first codon created after the 5-nucleotide deletion is an ATT and, therefore, the first amino acid in the mutated segment is still an isoleucine (position 632). The sequence changes from the codon 633 on, to include a stop codon at position 640.

Frame-shift mutations are usually assumed as potentially pathogenic or pathogenic variants. It is mostly true. However, there are compensatory cellular mechanisms functioning at both the RNA and protein levels. Some are damage-reducing mechanisms, such as NMD, alternative splicing, and exon skipping (Lindeboom, Supek et al. 2016; Lambert, Ashi et al. 2020). Some others are damaging, such as dominant negative effects, a phenomenon that occurs in dimeric or multimeric proteins, in which the entry of a mutated subunit in the complex induces its precipitation (Hanna, Guerra-Moreno et al. 2019). Consequently, the mutated allele reduces the function of the normal allele. However, dominant negative effects are not observed when the mutation affects the protein surfaces which interact, preventing it from forming the complex (Vieler and Sanyal 2018). Therefore, even fsm impact shall be carefully evaluated.

To investigate the pathogenicity of *MSH2* c.1894_1898del; p.Ile633Lysfs*9 mutation, we adopted a multi-step procedure, investigating the possible effects at the RNA and protein levels. At the mRNA level, we evaluated the hypotheses of NMD, alternative splicing, and exon skipping. At protein level, we investigated the impact of *MSH2* c.1894_1898del; p.Ile633Lysfs*9 mutation on the protein sequence and 3D structure of MSH2, and 3D quaternary structure of the MSH2/MSH6 heterodimer (Fig. 3).

After confirming that this mutation has not yet been analysed, we generated the *MSH2* CDS containing the mutation, with the corresponding five nucleotide deletion (Supplementary Fig. 1). Using ExPASy Translate (Artimo, Jonnalagedda et al. 2012), the effects at the protein level were predicted. The wild-type MSH2 protein contains 934 amino acids (isoform P43246, in UniProt). The deletion causes a reduction in protein length to 640 amino acids; therefore, 294 amino acids at the C-terminal were lost. The five-nucleotide deletion affects

codons 632 and 633. After the frameshift, amino acid 632, an isoleucine, does not change. Therefore, the first change observed in the protein sequence is in amino acid residue 633, a substitution of an isoleucine with a lysine. A premature stop codon follows in the 9th position (Fig. 2). Using BLASTP (Altschul, Gish et al. 1990), we searched for the protein containing *MSH2* c.1894_1898del; p.Ile633Lysfs*9 mutation. No match for the sequence after the frameshift mutation was identified, supporting the idea that this mutation has not been investigated before.

```
>CDS sequence mutated MSH2 (MSH2 c.1894 1898del; p.Ile633Lysfs*9)
ATGGCGGTGCAGCCGAAGGAGACGCTGCAGTTGGAGAGCGCGCCGAGGTCGGCTTCGTGCGCTTCTTTTC
AGGGCATGCCGGAGAAGCCGACCACCACAGTGCAGCCTTTTCGACCGGGGCGACTTCTATACGGCGCACGG
CGAGGACGCGCTGCTGGCCGCCCGGAGGTGTTCAAGACCAGGGGGTGATCAAGTACATGGGGCCGGCA
GGAGCAAAGAATCTGCAGAGTGTGTGCTTAGTAAAATGAATTTTGAATCTTTGTAAAAGATCTTCTTC
TGGTTCGTCAGTATAGAGTTGAAGTTTATAAGAATAGAGCTGGAATAAGGCATCCAAGGAGAATGATTG
GTATTTGGCATATAAGGCTTCTCCTGGCAATCTCTCTCAGTTTGAAGACATCTCTTTGGTAACAATGAT
ATGTCAGCTTCCATTGGTGTGTGGGTGTTAAAATGTCCGAGTTGATGGCCAGAGACAGGTTGGAGTTG
GGTATGTGGATTCCATACAGAGGAACTAGGACTGTGTGAATTCCTGATAATGATCAGTTCTCCAATCT
TGAGGCTCTCCTCATCCAGATTGGACCAAAGGAATGTGTTTTACCCGGAGGAGAGACTGCTGGAGACATG
GGGAAACTGAGACAGATAATTCAAAGAGGAGGAATTCGTATCACAGAAAAGAAAAAAGCTGACTTTTCCA
CAAAAGACATTTATCAGGACCTCAACCGGTTGTTGAAAGGCAAAAAGGGAGAGCAGATGAATAGTGTGT
ATTGCCAGAAATGGAGAATCAGGTTGCAGTTTCACTACTGCTGCGGTAATCAAGTTTTTGAAGCTTTA
TCAGATGATTTCAACTTTGGACAGTTTGAAGTACTACTTTTACTTCAAGCAGTATATGAATTTGGATA
TTGCAGCAGTCAGAGCCCTTAACTTTTTCAGGGTCTGTGTGAAGATACCCTGGCTCTCAGTCTCTGGC
TGCCTTGCTGAATAAGTGTAAAACCCCTCAAGGACAAAGACTTGTAAACAGTGGATTAAGCAGCCTCTC
ATGGATAAAGAACAGAAATAGAGGAGAGATTGAATTTAGTGAAGCTTTTGTAGAAGATGCAGAATTGAGGC
AGACTTTACAAGAAGATTTACTTCGTGATTTCCAGATCTTAAACCGACTTGCACAAAGATTTCAAAGACA
AGCAAAACTTACAAGATTTTACCAGCTCTATCAGGGTATAAAATCAACTCACTAATGTTATACAGGCT
CTGGAAAAACATGAAGGAAAACACCAGAAATTTGTTGGCAGTTTTTGTGACTCCTTACTGATCTTC
GTTCTGACTTCTCCAAGTTTCAGGAAATGATAGAAACAACTTTAGATATGGATCAGGTGGAAAACCATGA
ATTCCTTGTAACCTTCAATTTGATCCATACTCAGTGAATTAAGAGAAATAATGAATGACTTGGAAAAG
AAGATGCAGTCAACATTAATAAGTGCAGCCAGAGATCTTTGGCTTGGACCCCTGGCAACAGATTAACCTGG
ATTCAGTGCACAGTTTGGATATTACTTTCTGTGTAACCTGTAAGGAAGAAAAAGTCTTTCGTAACAATAA
AACTTTAGTACTGTAGATATCCAGAAGAATGGTGTAAATTTACCAACAGCAAAATGACTTCTTTAAAT
GAAGAGTATACAAAAATAAAACAGAAATGAAGAAGCCAGGATGCCATTGTTAAAGAAATTTGCAATA
TTTTCTCAGGCTATGTAGAACCAATGCAGACACTCAATGATGTGTGACTCAGCTAGATGCTGTTGTCAG
CTTTGCTCAGTGTCAAATGGAGCACCTGTCCATATGTACGACCAGCCATTTTGGAGAAAGGACAAGGA
AGA-----ATTAAGAGCATCCAGGCATGCTGTGTGAAGTTCAAGATGAATTCATTTATTCCTAATG
ACGTATACTTTGAAAAGATAAACAGATGTTCCACATCATTACTGGCCCCAATATGGGAGGTAATCAAC
ATATATTCGACAAACTGGGGTGATAGTACTCATGGCCCAATTTGGGTGTTTTGTGCCATGTGAGTCAGCA
GAAGTGTCCATTGTGGACTGCATCTTAGCCGAGTAGGGGCTGGTGACAGTCAATGAAAGGAGTCTCCA
CGTTCATGGCTGAAATGTTGAAAACCTGCTTCTATCCTCAGGTCTGCAACCAAGATTCATTAATAATCAT
AGATGAATGGGAAGAGGAACTTACCTACGATGGATTTGGGTAGCATGGGCTATATCAGAATACATT
GCAACAAAGATTTGGTCTTTTTCATGTTTTGCAACCCATTTTCATGAACCTTACTGCCTTGGCCAATCAGA
TACCAACTGTTAATAATCTACATGTACAGCACTCACCCTGAAGAGACCTTAACTATGCTTTATCAGGT
GAAGAAAGGTGCTGTGATCAAAGTTTGGGATTCATGTTGCAGAGCTTGCTAATTTCCCTAAGCATGTA
ATAGAGTGTGCTAAACAGAAAGCCCTGGAAGTGGAGTTTTCAGTATATTGGAGAATCGCAAGGATATG
ATATCATGGAACCCAGCAGCAAAAGAGTGTATCTGGAAAAGAGCAAGGTGAAAAAATTTATTCAGGAGTT
CCTGTCCAAGGTGAAACAAATGCCCTTTACTGAAATGTGCAAGAAAACATCACAATAAAGTTAAAAACAG
CTAAAAGCTGAAGTAAATAGCAAAGAATAATAGCTTTGTAATGAAATCATTTACGAATAAAAGTTACTA
CGTGA
```

```
>Protein sequence mutated of the DNA mismatch repair protein MSH2 (MSH2 c.1894_1898del;
p.Ile633Lysfs*9)
MAVQPKETLQLESAAEVGFVRFVQGMPEKPTTTVRLFDRGDFYTAHGEDA
LLAAREVFKTQGVIKYMGPAKAKNLQSVVLSKMNFEFVKDLLLVQRV
EVYKNRAGNKASKENDWYLAYKASPGNLSQFEDILFGNNDMSASIGVVG
KMSAVDGRQVGVGYVDSIQRKLGLCEFPDNDQFSNLEALLIQIGKECV
LPGGETAGDMKQLRQIIQRGGILITERKKADFSTKDIYQDLNRLKGGK
EQMNSAVLPEMENVAVSSLSAVIKFLELLSDDSNFGQFELTTFFDFSQYM
KLDIAAVRALNLFQGSVEDTTGSQSLAALLNKCKTPQGRQVLVQWIKQPL
```

MDKNRIEERLNLVEAFVEDAELRQTLQEDLLRRFPDLNRLAKKFQRQAAN
LQDCYRLYQGINQLPNVIQALEKHEGKHQKLLAVFVTPLTDLRSDFSKE
QEMIETTLMDQVENHEFLVKPSFDPNLSELREIMNDLEKKMQSTLISAA
RDLGLDPGKQIKLDSSAQFGYYFRVTCKEEKVLRNKNFSTVDIQKNGVK
FTNSKLTSLNEEYTKNKTEYEEAQDAIVKEIVNISSGYVEPMQTLNDVLA
QLDAVVSFAHVSNGAPVPYVRPAILEKGQGRIKSIQACL

Supplementary Figure 1 - *MSH2* CDS and protein sequences predicted in the presence of the *MSH2*

c.1894_1898del; p.Ile633Lysfs*9.

Note: The first sequence denotes a mutated CDS from the *MSH2* gene. The five dashes (-----) denote the 5-nucleotide deletion in position 1894-1898. The second sequence refers to the 640 amino acids, mutated *MSH2* protein. The 5-nucleotide deletion produces a frame-shift mutation. Consequently, the protein sequence changes from amino acid residue 633 on and a stop codon is inserted after residue 640.

Next, we evaluated the impact of the frameshift mutation in *MSH2* protein domains.

According to Pfam (El-Gebali, Mistry et al. 2019), *MSH2* presents 33 domains and features.

This mutation disrupts exon 12 and leads to the deletion of exons 13 to 16 (in isoforms P43246-1 and P43246-2, see below). *MSH2* c.1894_1898del; p.Ile633Lysfs*9 is located in the loop between MutS domains III/IV and V, causing the loss of the V domain, a highly conserved ATPase domain, which provides energy for the first steps of MMR (Lakhani, Thayer et al. 2017).

To evaluate the impact of this frameshift mutation on the ability of *MSH2* to form dimers with *MSH3* and *MSH6*, we modeled the 3D structure of the mutated protein using the Modeller software. The wild-type *MSH2* protein tridimensional model selected by Modeller (Šali and Blundell 1993) was 2O8B. The comparative modeling of the 3D structures of the wildtype and the putative mutated protein suggests that the mutation probably affects the interaction of the *MSH2* protein with *MSH6* (Fig. 3). This finding suggests that, if expressed, *MSH2* protein would eventually be able to enter MutS complexes, but the lack of ADP binding site would be harmful.

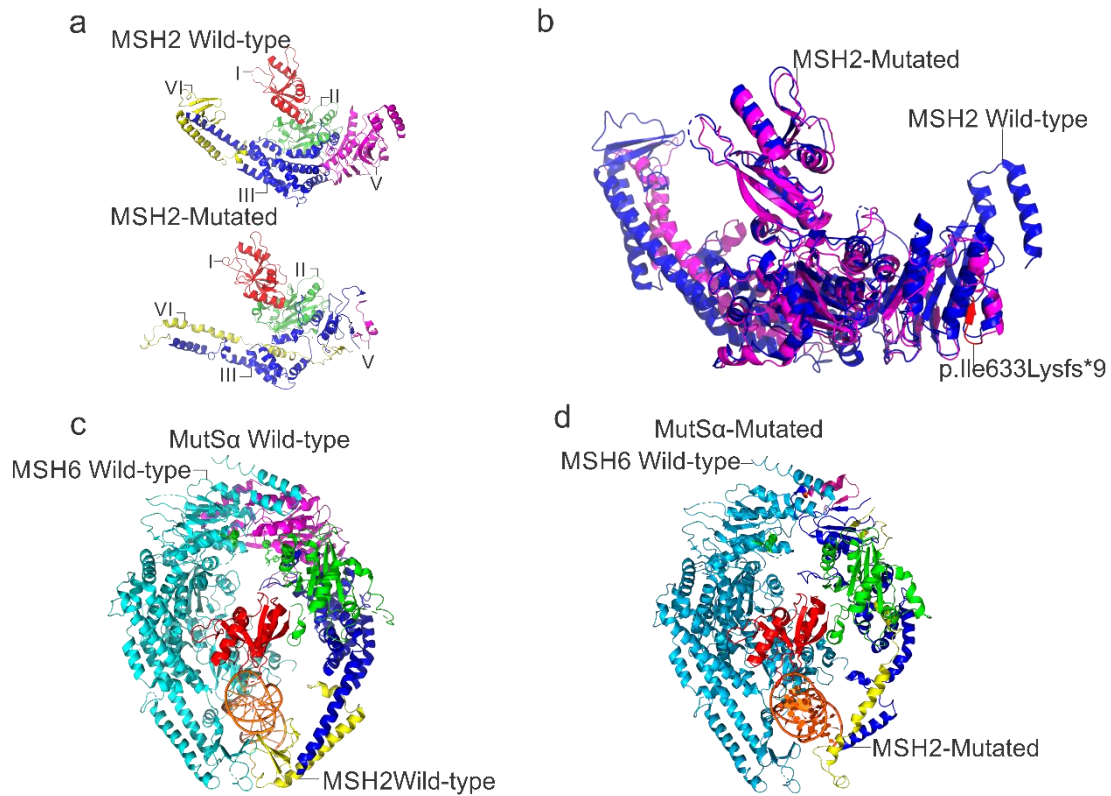


Figure 3. a) Comparison between MSH2 wild-type and *MSH2* c.1894_1898del; p.Ile633Lysfs*9 proteins. MutS I domain is represented in red; MutS II domains in green; MutS III domain in blue; MutS IV domain in yellow; and, in magenta the MutS V domain; b) Overlapping blue stained wild-type MSH2 protein and magenta stained mutated MSH2 protein. In red, the portion of the protein, whose sequence was changed as a consequence of the frame-shift mutation; c) Wild-type MutS α complex. Blue contains wild-type MSH2 protein and orange contains wild-type MSH6 protein; d) Mutated MutS α complex. In blue, the mutated MSH2 protein and in orange the wild MSH6 protein.

In addition, we evaluated the impact of this mutation on the different *MSH2* isoforms and overlapping genes. We also investigated the hypothesis of mechanisms that could reduce the impact of the mutation, such as alternative splicing and exon skipping. *MSH2* is located on chromosome 2p21-p16.3 region, CHR2:47403067-47634501 (GRCh38 p.13). Typically, *MSH2* is described as composed of 16 exons, but the number of exons in the region varies from 16 to 28, depending on the consulted database. *MSH2* presents alternative splicing. One difficulty faced in predicting the impact of *MSH2* mutations is that the number of CCDS and protein isoforms, as well as their identification numbers vary according to the source consulted. For example, nine isoforms were identified in Ensembl Genome Browser (Kent,

Sugnet et al. 2002), VEP (Hunt, McLaren et al. 2018) and UNIPROT (Consortium 2019) (ENST00000233146.7, ENST00000543555.6, ENST00000645506.1, ENST00000406134.5, ENST00000644092.1, ENST00000646415.1, ENST00000645339.1, ENST00000644900.1, and ENST00000467323.1); 10 isoforms are informed in GTEx (GTEx 2020), in the UCSC Genome Browser, and in the NCBI Genome Browser (Agarwala, Barrett et al. 2018), which include ENST00000454849, an isoform excluded from the Ensembl database. The information on the CCDS and MSH2 protein isoforms, linking different nomenclatures, is organized in Supplementary Table 1. Some of these isoforms are predicted to undergo NMD. The two main MSH2 isoforms (P43246-1 and P43246-2) are affected by MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation. P43246-1 is 66 amino acids longer than P43246-2. This difference is due to alternative splicing of the first exon, which encodes a flexible loop in the N-terminal of the protein and does not change the MutS domains. Therefore, sequence differences induced by MSH2 c.1894_1898del; p.Ile633Lysfs*9 start at residue 633 in P43246-1 isoform and in residue 597 in P43246-2 isoform. However, the sequences of both isoforms are the same in this region, as well as the mutation impact. Additional information on the impact of this mutation on the different isoforms is provided in Supplementary Table 1.

Supplementary Table 1 MSH2 isoforms, according to different databases, respective annotation status, and impact of MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation.

Ensembl Isoform Name	Transcript ID	bp	Protein	Biotype	Uniprot	Description
MSH2-201	ENST00000233146.7	3115	934 aa	Protein coding	P43246-1	Reviewed (Swiss-Prot): Manually annotated - Peptide RefSeq
MSH2-204	ENST00000543555.6	3018	868 aa	Protein coding	P43246-2	The sequence of this isoform differs from the peptide RefSeq by present alternative splicing, with 66 the loss
MSH2-208	ENST00000645506.1	7893	924 aa	Protein coding	A0A2R8Y6P0	Unreviewed (TrEMBL):Computationally analyzed -

						Records that await full manual annotation
MSH2-202	ENST00000406134.5	3628	921 aa	Protein coding	E9PHA6	Unreviewed (TrEMBL): Computationally analyzed - Records that await full manual annotation
MSH2-205	ENST00000644092.1	5343	558aa	Nonsense mediated decay	A0A2R8Y7S8	Unreviewed (TrEMBL): Computationally analyzed - Records that await full manual annotation
MSH2-209	ENST00000646415.1	4523	938aa	Nonsense mediated decay	A0A2R8YG02	Unreviewed (TrEMBL): Computationally analyzed - Records that await full manual annotation.
MSH2-207	ENST00000645339.1	3113	918 aa	Nonsense mediated decay	A0A2R8YFH0	Unreviewed (TrEMBL): Computationally analyzed - Records that await full manual annotation
MSH2-206	ENST00000644900.1	1663	168 aa	Nonsense mediated decay	A0A2R8Y713	Unreviewed (TrEMBL): Computationally analyzed - Records that await full manual annotation
MSH2-203	ENST00000467323.1	313	No protein	Processed transcript	-	-
MSH2-003	ENST00000454849.1	560	140 aa	Protein coding	C9J809	This entry is obsolete
MSH2-005	ENST00000461394.1	416	No protein	Processed transcript	-	-

Note - Edited from VEP isoform prediction. [Hunt S, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, Parton A, Armean I, 2018. Ensembl variation resources. Database (Oxford). bay119

There are two additional genes in this region, MSH2-OT1 and KCNK2 (NCBI, November 2020). MSH2-OT1 and KCNK12 overlap with the long MSH2 isoforms (ENST00000648301.1, ENST00000435331.3, and ENST00000439182.2). MSH2-OT1 is a processed transcript expressed in the testis and presenting three isoforms. KCNK12 (ENST00000327876.5 and ENST00000493527.1) is a protein coding gene expressed from the DNA complementary strand. These genes do not share exons with MSH2. In addition, two other genes associated with LS map to the same region, EPCAM maps approximately 16 kb upstream and MSH6 maps approximately 150 kb downstream to MSH2 (NCBI, 2020). None of these genes/isoforms would be affected by the MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation.

We also investigated the evidence for alternative splicing, which could rescue the mutated allele, by generating a shorter but functional protein (Koenig, Beggs et al. 1989; Monaco, Bertelson et al. 1988). In Fig. 4, we present tissue specific expression patterns of MSH2 CCDS isoforms according to GTex. No MSH2 protein isoform without exon 12, but preserving NH₂- and COOH-end is supported. None has been deposited so far in both the RNA and protein databases. The predicted impact of MSH2 c.1894_1898del; p.Ile633Lysfs*9 on the different MSH2 RNA and protein isoforms is presented in Supplementary Table 1.

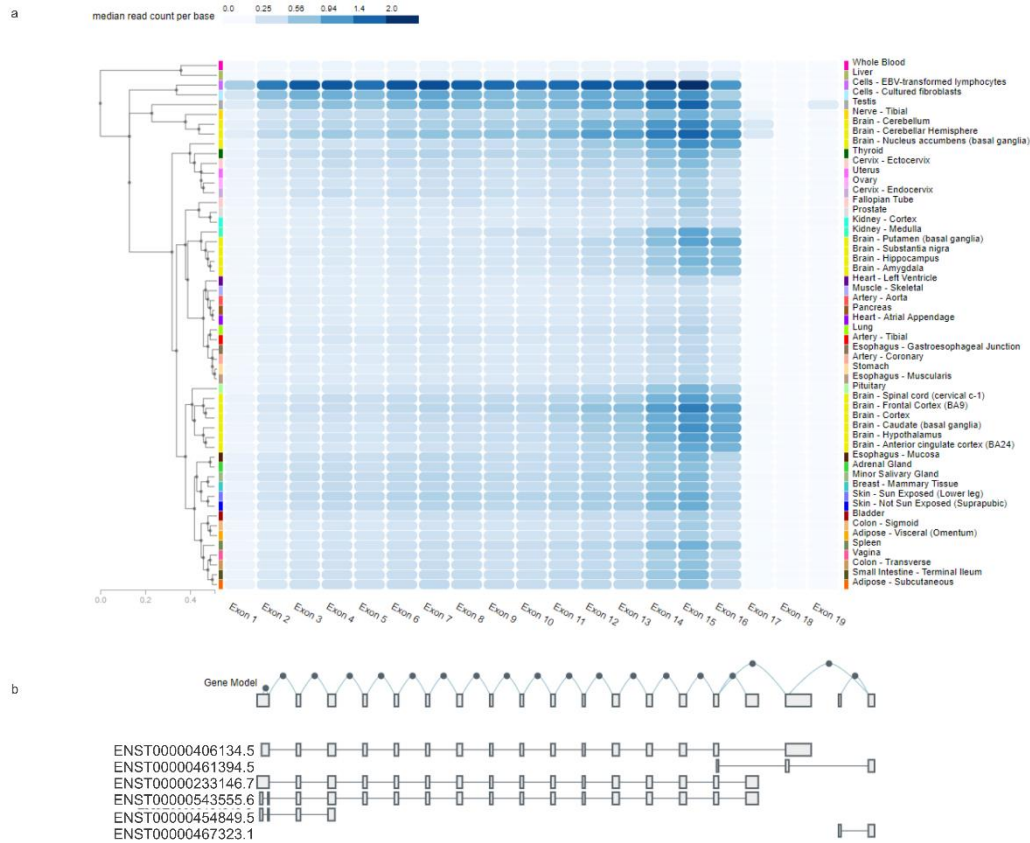


Figure 4. Data on tissue specific MSH2 isoforms, according to GTex (<https://gtexportal.org/home>). In a), observe that there are tissue specific differences in expression pattern. Exons 13-16 expression is higher in most tissues. In b), the splicing map of the MSH2 exons. Alternative splicing is depicted within exon 1, differentiating isoforms ENST00000233146.7 and ENST00000543555.6. From top to bottom, isoforms ENST00000461394.5 is a processed transcript expressed in the testis. ENST00000454849.1 has been excluded from UNIPROT and ENSEMBL databases, and is most probably a databasis artifact.

Considering the findings of the tumor immunohistochemistry, showing loss of expression of the MSH2 protein in one tumor and MSH2 and MSH6 proteins in the other one, we investigated the hypothesis of NMD. The PTC inserted in consequence of this frame-shift mutation occurs at 780 nucleotides after the translation start codon and 709 nucleotides before the last exon-exon boundary (NM_000251.3). According to the literature, the main theoretical evidence for NMD comes from the position of the newly created premature termination codon (PTC) in the transcript (Lambert, Ashi et al. 2020; Schweingruber, Rufener et al. 2013; Miller and Pearce 2014;). Therefore, the PTC introduced by MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation fulfills the main requirements of NMD, meaning: 1. The gene contains introns; 2. The distance between the stop codon created by MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation and the last exon-exon boundary of the transcript is larger than 50-54 nucleotides; and, 3. The distance between the new stop codon and the translation start codon is larger than 200 nucleotides.

DISCUSSION

Variants of Uncertain Significance (VUS) are a frequent problem in clinical practice, affecting decisions related to both patient follow-up and genetic counseling. Considering the large number of individual genetic differences in humans, any genetic variant has a larger probability of being non-pathogenic. This scenario changes when we consider genetic variants affecting the protein sequence; however, each variant requires careful analysis (Thompson, Spurdle et al. 2014). Frame-shift mutations are usually assumed as pathogenic, but the underlying mechanism may vary from a complete loss of RNA expression to the expression of a completely different protein product. In addition, effects on isoforms may vary according to gene, tissue, developmental period, and health conditions. However, there are some protective mechanisms that may reduce the fsm effects. Damaging prone mechanisms include negative dominance, in which the entry of a mutated subunit in a protein complex induces its

precipitation, stealing function of the normal allele. Therefore, mutations preventing a subunit from entering the complex are less damaging than those rendering a subunit still able to enter protein complexes. Out of prudence, fsm are usually classified as pathogenic, but additional studies are required. Here, we describe the clinical, immunohistochemical, and molecular findings of a patient presenting LS, and the Bioinformatic evidence for pathogenicity of the MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation.

One of the dilemmas of evaluating persons with mutations that predispose to some inherited disease, but not necessarily determine the disease, is how to better guide them through the decision-making process. Main questions concern medical support and balancing risks and benefits of the available procedures, because some of them are highly invasive. Since LS is a heterogeneous disease, an optimized management might be limited, depending on well-documented and complete family data, so that each case could fit guidelines available (Network 2020). Complete family data is sometimes difficult to obtain, particularly the actual number of affected relatives, and, for which one of them, information on type of neoplasia, age of onset, tumor documentation. In addition, patients may have restrictions on contacting some family members and improving the quality of data may be difficult.

In the present study, we came across similar situations. Although the clinical findings in the proband were highly suggestive, the LS genetic diagnosis was made only after the detection of the second tumor. The second tumor was considered metachronous, although the cecon has not been completely evaluated in the first colonoscopy. Even considering that the tumorigenesis in LS can follow a faster path (Dominguez-Valentin, Sampson et al. 2020; Rondagh, Gulikers et al. 2013), the tumors present in this patient could be too large to have developed in one year. Another problem observed in this study was the poor contact among the family members. Consequently, the type of tumor, age at diagnosis, and age at death were not accessible for some family members. To date, only a few family members have accessed

genetic counseling and manifested any interest in mutation screening. Two factors are crucial for this, family detachment and the difficulty of access to the genetic services in remote areas.

Evidence pointing to the germline MSH2 c.1894_1898del; p.Ile633Lysfs*9 mutation in the causation of LS in this family is provided by its low population frequency, but emerged both in immunohistochemistry and in the Bioinformatic analyses of the mutation identified in the patient. In immunohistochemistry, the two tumors analyzed presented different results: loss of MSH2 expression was detected in one tumor and loss of MSH2 and MSH6 expression in the other one. These findings suggest several different hypotheses. First, although at the familial level the condition is autosomal dominant, there was a loss of expression of the two *MSH2* alleles at the tumor level. This loss of function may reflect a second mutation, affecting the other *MSH2* allele, in the tumor. This finding was described in the context of Rb mutations in retinoblastoma, but is typically seen in tumor suppressor genes (Dryja, Rapaport et al. 1986)(e.g., RB1). Alternatively, the protein may have been transiently expressed, but may have been precipitated due to dominant negative effects, for example. Considering the immunohistochemistry results, where the MSH2 protein was not detected, the hypothesis of dominant negative effects seems less probable.

The loss of both *MSH2* and *MSH6* expression in one of the tumors may have at least four additional explanations. First, the loss of MSH2 protein expression would cause MSH6 precipitation, once MSH6 does not dimerize with other proteins. Second, mutations in MSH2 dimerization regions could induce MSH6 precipitation, a dominant negative effect. Third, the loss of MSH6 expression in tumors from individuals having germline MSH2 mutations may also reflect the gain of MSH6 somatic mutations accumulated in the course of the clonal evolution of the tumor. Fourth, MSH6 silencing in the presence of MSH2 mutations has already been described and corresponded to somatic epimutations in the tumor tissue

(Olkinuora, Gylling et al. 2020). At this point, our investigation has a limitation, due to the lack of molecular analyses at the RNA level. *MSH2* c.1894_1898del; p.Ile633Lysfs*9 mutation may induce the loss of MSH2 protein expression by mechanisms at the RNA level (NMD) or at the protein level. This analysis would also help to understand the mechanisms underlying the loss of expression of the second *MSH2* allele and of *MSH6*. However, the results of the Bioinformatic analysis reported here suggest that *MSH2* c.1894_1898del; p.Ile633Lysfs*9 is a pathogenic mutation. This mutation fulfills the typical requirements for NMD and no evidence of alternative splicing, producing an isoform excluding exon 12 was found in any database searched or in the literature. Isoforms described in *MSH2* in both healthy as well as colon cancer cell lines include the deletion of some exons, e.g., -5, -13, -2/3/4/5/6/7, and -2/3/4/5/6/7/8 (Genuardi, Viel et al. 1998). No isoform excluding exon 12 has been described in the literature or deposited in RNA or protein databases. Consistently, assuming NMD as highly probable in this case, this mutation would be a loss of function, germline allele.

Nonetheless, we investigated *MSH2* c.1894_1898del; p.Ile633Lysfs*9 effects at protein level. With the results of the in silico investigation, we have demonstrated that the *MSH2* protein loses the 294 C-terminal amino acids, causing the loss of the MutS domain V and, consequently, the loss of the interaction with ADP. *MSH2* interacts with *MSH6* and *MSH3* proteins to form MutSalpha and MutSbeta, respectively, via MutS domains I and IV; these domains are also responsible for DNA binding. MutS domain V presents ATPase activity, while domains MutS domain II and MutS domain III act as connectors between MutS domain V and the MutS domains I and IV (Warren, Pohlhaus et al. 2007; Schneider, Pastor et al. 2018; (Gupta, Gellert et al. 2012). At first sight, the loss of MutS domain V would not prevent the mutated *MSH2* monomer from entering the MutSalpha or MutSbeta complexes, once these interactions depend on the integrity of MutS domains I and IV. However, it would affect

ADP binding and replacement by ATP, needed for mismatch recognition and DNA bending in the very beginning of the MMR (Warren, Pohlhaus et al. 2007; Schneider, Pastor et al. 2018; Gupta, Gellert et al. 2012). The interaction with DNA mismatches or insertion/deletion loops occurs initially with the MutS domains I of both subunits (MSH2/MSH6 or MSH3/MSH6) accessing the DNA minor groove. The MutSbeta complex induces the distortion of the sugar-phosphate structure in the DNA segment presenting the mutation (Gupta, Gellert et al. 2012). The MutSbeta and MutSalpha complexes exchange ADP for ATP to move the DNA strand and to interact with downstream effectors, such as PCNA, MLH1/PMS2, and EXO1. Once recruited, PCNA is responsible for starting the excision of the newly replicated DNA strand. The MLH1/PMS2 complex has endonucleolytic activity, thus being able to create a cut in the nucleotide segment that needs to be repaired, allowing EXO1 to remove hundreds of nucleotides (Fukui 2010; Martín-López and Fishel 2013; Edelbrock, Kaliyaperumal et al. 2009). Consequently, the downstream MMR cascade would be affected by the loss of MutSbeta function. After evaluating the consequences of *MSH2* c.1894_1898del; p.Ile633Lysfs*9 mutation, we consider that it fulfills the criteria for pathogenicity accepted for fsm in *MSH2* (Plazzer, Sijmons et al. 2013).

CONCLUSIONS

Earlier diagnosis of LS can help to optimize treatments and reduce patient risk for new cancers by prophylactic surgery, for example. Genetic counseling as part of health care is essential and still challenging in some cases. Updates in non- or poorly-described mutations must be done as a contribution to better understand clinical and molecular phenotypes. Currently, Bioinformatics provides a plethora of tools to apply in the search for the necessary evidence for classifying genetic variants. Identifying germline mutations help define the management of patients and respective families. Mutations classified as VUS immobilize clinicians and deserve a particularly careful evaluation.

In the present study, we used several different Bioinformatic tools to evaluate *MSH2* c.1894_1898del; p.Ile633Lysfs*9 germline mutation, identified in a patient belonging to a LS family and presenting colorectal cancers, probably metachronous. Evidence was provided by immunohistochemistry, showing the lack of expression of MSH2 in one of the patient's tumors and of MSH2 and MSH6 in the other one.

In synthesis, the results of the Bioinformatic analyses point to: 1. This mutation was not found in any populational database. It has been included in Varsome already, but without additional information; 2. It causes the loss of MutS domain V, therefore, the ADP binding ability of MSH2; 3. There is no evidence that the effects of this mutation could be ameliorated by physiological alternative splicing or exon skipping induced by the mutation. Actually, the large number of exon 12 pathogenic mutations in InSIGHT, suggest that such a mechanism really does not exist; 4. There is evidence for NMD; 5. If this mRNA escapes NMD, the protein produced would be possibly able to enter MutSalpha complexes, but these complexes would be less compact than those established by the normal allele and probably not functional, due to the loss of the MSH2 ADP binding site.

In conclusion, all the evidence summed up in the current study indicate that *MSH2* c.1894_1898del; p.Ile633Lysfs*9 is a pathogenic mutation. The phenotype associated is LS, with probably metachronous colorectal cancer.

Acknowledgments

We thank the patient and her family for their engagement in this investigation. MRSC was awarded a fellowship from the Brazilian National Research Council (Conselho Nacional de Pesquisa, CNPq, grant number 312068/2015-8, 312405/2018-9). TRS and CGRM were awarded a fellowship from the Coordenação de Pessoal de Nível Superior (CAPES); MRSC was awarded a fellowship from the Brazilian National Research Council (Conselho Nacional

de Pesquisa, CNPq, grant numbers 312068/2015-8 and 312405/2018-9); Pró-Reitoria de Pesquisa de Universidade Federal de Minas Gerais (PRPq/UFMG). ALSF was awarded with a Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) APQ-02373-17.

Conflicts of interest/Competing interests

The authors have no conflicts of interest or competing interests to declare.

Ethics approval

This project was approved by the Ethics in Research Committee of the Universidade Federal de Minas Gerais under CAAE.01758418.0.0000.5149 and CAAE.48770621.6.0000.5149.

Availability of data and material

Not applicable.

Code availability

Not applicable

Authors' contributions

Anisse Marques Chami conducted the clinical evaluation; Thalia Rodrigues de Souza, Carolina Guimarães Ramos Matosinho, Letícia da Conceição Braga, Agnaldo Lopes da Silva Filho, and Maria Raquel Santos Carvalho conducted the analyses, which results were evaluated by all authors. Manuscript was written and reviewed by all the authors.

REFERENCES

- Agarwala R, Barrett T, Beck J, Benson DA, Bollin C, Bolton E, Bourexis D, Brister J. R, Bryant SH and Canese K (2018) Database resources of the national center for biotechnology information. *Nucleic Acids Res* 46 (D1):D8-D13.
- Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410.
- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, Castro E De, Duvaud S, Flegel V, Fortier A and Gasteiger E (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic acids Res* 40(W1): W597-W603.
- Biller LH, Syngal S and Yurgelun MB (2019) Recent advances in Lynch syndrome. *Fam Cancer* 18:211-219.
- Consortium U (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic acids Res* 47(D1):D506-D515.
- Crouse, GF (2016) Non-canonical actions of mismatch repair. *DNA repair* 38: 102-109.
- Dominguez-Valentin M, Sampson JR, Seppälä TT, Ten Broeke SW, Plazzer J-P, Nakken S, Engel C, Aretz S, Jenkins MA and Sunde L (2020) Cancer risks by gene, age, and gender in 6350 carriers of pathogenic mismatch repair variants: findings from the Prospective Lynch Syndrome Database. *Genet Med* 22:1142-1148.
- Dryja TP, Rapaport JM, Joyce JM and Petersen RA (1986) Molecular detection of deletions involving band q14 of chromosome 13 in retinoblastomas. *PNAS* 83:7391-7394.
- Edelbrock MA, Kaliyaperumal S and Williams KJ (2009) DNA mismatch repair efficiency and fidelity are elevated during DNA synthesis in human cells. *Mutat Res* 662:59-66.

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA and Smart A (2019) The Pfam protein families database in. *Nucleic acids Res* 47(D1):D427-D432.

Fukui K. (2010) DNA mismatch repair in eukaryotes and bacteria. *J Nucleic Acids*.

Genuardi M, Viel A, Bonora D, Capozzi E, Bellacosa A, Leonardi F, Valle R, Ventura A, Pedroni M and Boiocchi M (1998) Characterization of MLH1 and MSH2 alternative splicing and its relevance to molecular testing of colorectal cancer susceptibility. *Hum Genet* 102:15-20.

Gupta S, Gellert M and Yang W (2012) Mechanism of mismatch recognition revealed by human MutS β bound to unpaired DNA loops. *Nat Struct Mol Biol* 19: 72.

Hanna J, Guerra-Moreno A, Ang J and Micoogullari Y (2019). Protein degradation and the pathologic basis of disease. *Am J Clin Pathol* 189:94-103.

Heinz DW, Baase WA, Dahlquist FW and Matthews BW (1993) How amino-acid insertions are allowed in an α -helix of T4 lysozyme. *Nature* 361:561-564.

Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, Parton A and Armean I (2018) Ensembl variation resources. Database (Oxford). bay119.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM and Haussler D (2002). The human genome browser at UCSC. *Genome Res* 12:996-1006.

Koenig M, Beggs A, Moyer M, Scherpf S, Heindrich K, Bettecken T, Meng G, Müller C, Lindlöf M and Kaariainen H (1989) The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am J Hum Genet* 45:498.

Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Aguilera MA, Meyer R and Massouras A (2019) VarSome: the human genomic variant search engine. *Bioinformatics* 35: 1978.

Lakhani B, Thayer KM, Hingorani MM and Beveridge DL (2017) Evolutionary covariance combined with molecular dynamics predicts a framework for allostery in the MutS DNA mismatch repair protein. *J Phys Chem B* 121:2049-2061.

Lambert J-M, Ashi MO, Srour N, Delpy L and Saulière J (2020) Mechanisms and Regulation of Nonsense-Mediated mRNA Decay and Nonsense-Associated Altered Splicing in Lymphocytes. *Int J Mol Sci* 21:1335.

Landrum, M. J., S. Chitipiralla, G. R. Brown, C. Chen, B. Gu, J. Hart, D. Hoffman, W. Jang K, Kaur and Liu C (2020) ClinVar: improvements to accessing data. *Nucleic Acids Res* 48(D1): D835-D844.

Laskowski RA, Jabłońska J, Pravda L, Vařeková RS and Thornton JM (2018) PDBsum: Structural summaries of PDB entries. *Protein Sci* 27:129-134.

Li MM, Chao E, Esplin ED, Miller DT, Nathanson KL, Plon SE, Scheuner MT and Stewart DR (2020) Points to consider for reporting of germline variation in patients undergoing tumor testing: a statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 22:1142-1148.

Lindeboom RG, Supek F and Lehner B (2016) The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat Genet* 48:1112-1118.

Lynch HT, Snyder CL, Shaw TG, Heinen CD and Hitchins M P (2015) Milestones of Lynch syndrome: 1895–2015. *Nat Rev Cancer* 15: 181-194.

MacDonald JR, Ziman R, Yuen RK, Feuk L and Scherer SW (2014) The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic acids Res* 42(D1): D986-D992.

Martín-López JV and Fishel R (2013) The mechanism of mismatch repair and the functional analysis of mismatch repair defects in Lynch syndrome. *Fam Cancer* 12: 159-168.

McKusick VA (1998) Mendelian inheritance in man: a catalog of human genes and genetic disorders, JHU Press.

Miller JN and Pearce DA (2014) Nonsense-mediated decay in genetic disease: friend or foe? *Mutat Res Rev Mutat Res* 762:52-64.

Monaco AP, Bertelson CJ, Liechti-Gallati S, Moser H and Kunkel LM (1988) An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics* 2: 90-95.

Naslavsky M, Scliar M, Yamamoto G, et al (2020) Whole-genome sequencing of 1,171 elderly admixed individuals from the largest Latin American metropolis. *Research Square*. doi: <https://doi.org/10.1101/2020.09.15.298026>

Olkinuora A, Gylling A, Almusa H, Eldfors S, Lepistö A, Mecklin J-P, Nieminen TT and Peltomäki P (2020) Molecular Basis of Mismatch Repair Protein Deficiency in Tumors from Lynch Suspected Cases with Negative Germline Test Results. *Cancers* 12:1853.

Plazzer J-P, Sijmons RH, Woods MO, Peltomäki P, Thompson B, Den Dunnen JT and Macrae F (2013) The InSiGHT database: utilizing 100 years of insights into Lynch syndrome. *Fam Cancer* 12:175-180.

Rondagh E, Gulikers S, Gómez-García E, Vanlingen Y, Detisch Y, Winkens B, Vasen H, Masclee A and Sanduleanu S (2013) Nonpolypoid colorectal neoplasms: a challenge in endoscopic surveillance of patients with Lynch syndrome. *Endoscopy* 45:257-264.

Šali A, and Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779-815.

Schneider NB, Pastor T, Paula AEd, Achatz MI, Santos ÂRd, Vianna FSL, Rosset C, Pinheiro M, Ashton-Prolla P and Moreira MÂM (2018) Germline MLH 1, MSH 2 and MSH 6 variants in Brazilian patients with colorectal cancer and clinical features suggestive of Lynch Syndrome. *Cancer Med* 7:2078-2088.

Schrödinger LLC (2017) The PyMOL Molecular Graphics System, Version 2.0,

<https://www.pymol.org>

Schweiggruber C, Rufener SC, Zünd D, Yamashita A and Mühlemann O (2013) Nonsense-mediated mRNA decay—mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim Biophys Acta Gene Regul Mech.* 1829:612-623.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I and Mazor Y (2016) The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinform* 54:1.30. 31-31.30. 33.

Thompson BA, Spurdle AB, Plazzer J-P, Greenblatt MS, Akagi K, Al-Mulla F, Bapat B, Bernstein I, Capellá G and Den Dunnen JT (2014) Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet* 46:107-115.

Toft NJ, Winton DJ, Kelly J, Howard LA, Dekker M, Riele H te, Arends MJ, Wyllie A H, Margison GP and Clarke AR (1999) Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. *PNAS* 96:3911-3915.

Umar A, Boland CR, Terdiman JP, Syngal S, Chapelle ADL, Rüschoff J, Fishel R, Lindor NM, Burgart LJ and Hamelin R (2004) Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 96:261-268.

Vaccaro CA, López-Kostner F, Adriana DV, Palmero EI, Rossi BM, Antelo M, Solano A, Carraro DM, Forones NM and Bohorquez M (2019) From colorectal cancer pattern to the characterization of individuals at risk: Picture for genetic research in Latin America. *Int J Cancer* 145:318-326.

Vasen H, Mecklin J-P, Khan PM and Lynch HT (1991) The international collaborative group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). *Dis Colon Rectum* 34:424-425.

Vasen HF, Watson P, Mecklin JP and Lynch HT (1999) New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* 116:1453-1456.

Vieler M and Sanyal S (2018) p53 isoforms and their implications in cancer. *Cancers* 10:288.

Warren JJ, Pohlhaus TJ, Changela A, Iyer RR, Modrich PL and Beese LS (2007) Structure of the human MutS α DNA lesion recognition complex. *Mol Cell*. 26:579-592.

Wiederstein M and Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35(suppl_2):W407-W410.

Internet Resources

Brazilian genomic variants (ABraOM), <https://abraom.ib.usp.br/search.php> (April, 2021)

Basic Local Alignment Search Tool-Proteins (Blastp), <https://blast.ncbi.nlm.nih.gov/Blast.cgi> (April, 2021)

ClinVar, <https://www.ncbi.nlm.nih.gov/clinvar/> (April, 2021)

Database of Genomic Variants, <http://dgv.tcag.ca/dgv/app/faq> (April, 2021)

EMBL-EBI. European-variation Archives, <https://www.ebi.ac.uk/eva/> (April, 2021)

Ensembl, <https://www.ensembl.org/index.html> (April, 2021)

ExPASy Translate, <https://web.expasy.org/translate/> (April, 2021)

GeneCard, <https://www.genecards.org/> (April, 2021)

GTEEx, <https://gtexportal.org/home/transcriptPage> (April, 2021)

InSIGHT, <https://www.insight-group.org/variants/databases/> (April, 2021)

Modeller, <https://salilab.org/modeller/> (April, 2021)

National Center for Biotechnology Information (NCBI), <https://www.ncbi.nlm.nih.gov/> (April, 2021)

Network, NCC Clinical Practice Guidelines in Oncology: genetic/familial high risk assessment: colorectal, https://www.nccn.org/guidelines/category_1 (April, 2021)

Online Mendelian Inheritance in Man (OMIM), <https://www.omim.org/> (April, 2021)

Protein Data Bank (PDB), <https://www.rcsb.org/> (April, 2021)

PDBsum, <http://www.ebi.ac.uk/pdbsum> (April, 2021)

Pfam, <https://pfam.xfam.org/> (April, 2021)

PREMM5, <https://premm.dfc.harvard.edu/> (April, 2021)

Protein Structure Analysis PROSA-Web, <https://prosa.services.came.sbg.ac.at/prosa.php> (April, 2021)

PubMed, <https://pubmed.ncbi.nlm.nih.gov> (April, 2021)

PyMOL, <https://www.pymol.org> (April, 2021)

UNIPROT database, <https://www.uniprot.org/> (April, 2021)

VarSome, <https://varsome.com> (April, 2021)

Variant Effect Predictor (VEP), <https://grch37.ensembl.org/info/docs/tools/vep> (April, 2021)

CAPÍTULO II

Title: *In deep* Bioinformatic characterization of a novel fumarate hydratase mutation *FH* c.199T>G; (p.Tyr67Asp) in hereditary leiomyomatosis and renal cell carcinoma

Anisse Chami^{1,2}

¹Programa de Pós-Graduação em Tocoginecologia da Universidade Estadual de São Paulo - UNESP, Botucatu, SP, Brazil

²Rede Mater Dei de Saúde, Belo Horizonte, MG, Brazil

ORCID: 0000-0002-2651-0995

Thália Rodrigues de Souza Zózimo³

³Programa de Pós-Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

ORCID: 0000-0001-8473-2548

Thamiris Matias Alves⁴

⁴Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

ORCID: 0000-0003-4681-4723

Carolina Guimarães Ramos Matosinho³

³Programa de Pós-Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

ORCID: 0000-0001-8070-1031

Cleydson Santos²

²Rede Mater Dei de Saúde, Belo Horizonte, MG, Brazil

Marcela Mattos Simões⁵

⁵ Av. Brasil, 1438/704, Santa Efigênia. CEP: 30140-003, Belo Horizonte, MG, Brazil

ORCID:0000-0001-7142-9312

Walter Luiz Ribeiro Cabral²

²Rede Mater Dei de Saúde, Belo Horizonte, MG, Brazil

Bernardo Ferreira de Paula Ricardo⁶

⁶Anatomia Patologia Diagnóstica, Belo Horizonte, MG, Brazil

ORCID: 0000-0002-1891-1457

Agnaldo Lopes da Silva Filho^{1,7}

¹Programa de Pós-Graduação em Tocoginecologia da Universidade Estadual de São Paulo - UNESP, Botucatu, SP, Brazil

⁷Departamento de Ginecologia e Obstetrícia, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

ORCID: 0000-0002-8486-7861

Maria Raquel Santos Carvalho^{3,4}

³Programa de Pós-Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

⁴ Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais - UFMG, Campus Pampulha, Belo Horizonte, MG, Brazil

mraquel-carvalho@ufmg.br

ORCID: 0000-0002-1744-448-X

Letícia da Conceição Braga^{8,9}

⁸ Núcleo de Pesquisa Básica e Translacional, Instituto Mário Penna, Belo Horizonte, MG, Brazil

⁹ OncoTag Desenvolvimento de Produto e Serviços para Saúde Humana, Belo Horizonte, MG, Brazil

ORCID: 0000-0002-6181-9410

Corresponding author
Prof. Maria Raquel Santos Carvalho
Departamento de Genética, Ecologia e Evolução
Instituto de Ciências Biológicas - Bloco E3 - Sala 175
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627 - Campus Pampulha
31270-901 - Belo Horizonte - Brazil
Fone +55 31 34092598
Fax +55 3134092570
ma.raquel.carvalho@gmail.com
mraquel-carvalho@ufmg.br

AC and TRSZ contributed equally to this study.

Running title:

FH c.199T>G; p.Tyr67Asp in HLRCC

ABSTRACT

Hereditary leiomyomatosis and renal cell carcinoma (HLRCC) is a rare autosomal dominant tumor predisposition syndrome characterized by the variable development of multiple skin and uterus leiomyomas and aggressive forms of kidney cell carcinoma (RCC). Mutations in *fumarate hydratase* (FH), one of the proteins in homologous recombination repair, precede the development of HLRCC with high penetrance. Considering the risk of early metastasization of RCC, FH has been included in mutation screening panels. The identification of a pathogenic FH mutation would guide the screening for tumors in the carriers. However, variants of uncertain significance (VUS) are a frequent finding, limiting the clinical value of the mutation screening. Here, we describe the phenotype associated and an in-depth, multi-step Bioinformatic evaluation of the germline FH c.199T>G p.Tyr67>Asp variant segregating in an HLRCC family. Evidence for FH c.199T>G; p.Tyr67Asp pathogenicity include the variant segregation with the disease in three affected family members; its absence in population database; and the deep evolutionary conservation of the Tyr67 residue. At protein level, this residue substitution causes the loss of molecular bonds and ionic interactions, affecting molecular dynamics and protein stability. Considering ACMG/AMP criteria, we propose the reclassification of the FH c.199T>G; p.Tyr67Asp variant as “likely pathogenic”. In addition, the in-depth, *in silico* approach used here allowed the understanding of how and why FH c.199T>G; p.Tyr67Asp would cause HLRCC, helping in clinical management decisions concerning the monitoring of unaffected family members having this variant.

Key-words: Hereditary leiomyomatosis and renal cell carcinoma; cancer; fumarate hydratase; Bioinformatics; mutation; recombination repair

Acknowledgments

We thank the patient and her family for their engagement in this investigation. MRSC was awarded a fellowship from the Brazilian National Research Council (Conselho Nacional de Pesquisa, CNPq, grant number 312068/2015-8, 312405/2018-9). TRS and CGRM were awarded fellowships from the Brazilian Coordenação de Pessoal de Nível Superior (CAPES); Pró-Reitoria de Pesquisa da Universidade Federal de Minas Gerais (PRPq/UFMG). ALSF was awarded with a grant from the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) APQ-02373-17.

Conflicts of interest/Competing interests

The authors have no conflicts of interest nor competing interests to declare.

Ethics approval

This project was approved by the Ethics in Research Committee of the Universidade Federal de Minas Gerais under CAAE.01758418.0.0000.5149 and CAAE.48770621.6.0000.5149.

Availability of data and material

Not applicable.

Code availability

Not applicable

Authors' contributions

Anisse Marques Chami, Marcela Mattos Simões, Walter Luiz Ribeiro Cabral, Bernardo Ferreira de Paula Ricardo, and Cleydson Santos conducted the genetic counseling and cancer risk management, dermatological, cirurgical, pathological, and oncological evaluation, respectively. Thalia Rodrigues de

Souza Zózimo, Carolina Guimarães Ramos Matosinho, Letícia da Conceição Braga, Aginaldo Lopes da Silva Filho, and Maria Raquel Santos Carvalho conducted the Bioinformatic analyses, which results were evaluated by all authors. Manuscript was written by Anisse Marques Chami, Thalia Rodrigues de Souza Zózimo, Thamiris Matias Alves, Letícia da Conceição Braga, and Maria Raquel Santos Carvalho, and reviewed by all the authors.

INTRODUCTION

Hereditary leiomyomatosis and renal cell carcinoma (HLRCC; OMIM #150800) is a rare autosomal dominant hereditary cancer syndrome, predisposing to the development of multiple skin and uterus leiomyomas and aggressive forms of kidney cancer [1,2]. Four main characteristics are observed in HLRCC syndrome: (i) multiple cutaneous piloleiomyomas, which present as firm reddish skin papules or nodules; (ii) affected females usually develop early-onset leiomyomata of the uterus (fibroids), diagnosed by their 20-35 years, leading to myomectomies or hysterectomy; (iii) early-onset form of type 2 papillary renal cell cancer, which has a propensity to metastasize early; and, (iv) cumulative, age-dependent risk of renal cell carcinoma (RCC); the risk for the carriers of mutations in fumarate hydratase (*FH*; EC 4.2.1.2) is 1-2% before 20 years and 15% between 35 and 45 years [3].

FH is a protein with different cytoplasmic, mitochondria, and nuclear functions. In the cytoplasm, *FH* catalyzes the reversible hydration of malate to fumarate, as part of the urea cycle and arginine metabolism. In the mitochondria, *FH* converts fumarate to malate in the tricarboxylic acid cycle. In addition, *FH* plays a role in response to DNA double-strand breaks (DSB). Following DSB induction, the cellular levels of *FH* increase and the cytoplasmic enzyme is transposed to the nucleus [4-6]. In the nucleus, *FH* is part of the human DNA-dependent protein kinase catalytic subunit (DNA-PKcs) that plays a key role in mediating non-homologous end joining and human homologous recombination signaling [4]. The loss of tumor suppressor genes, such as *FH*, impacts DNA repair and can both induce and potentiate tumor progression. In addition, *FH* loss of function renders tumor cells highly susceptible to synthetic lethal targeting with PARP inhibitors and provides a novel therapeutic strategy that could potentially be translated into the clinic for these otherwise difficult to treat malignancies [5].

FH mutations have been associated with both recessive (fumaric aciduria) and dominant (HLRCC) conditions [3]. Although defined by the presence of cutaneous and uterine leiomyomas, and kidney cancer, *FH* mutations have been associated also with pheochromocytomas, paragangliomas, adrenocortical carcinoma, neuroblastomas, glioma, ependymoma, osteosarcoma, and Ewing's sarcoma [7]. Consequently, for those individuals having a *FH* mutation, surveillance includes dermatological evaluation of skin lesions every two years after the age of 30; screening for renal cancer using abdominal magnetic resonance imaging (MRI) annually from age 8 years; annual gynecological exam from 20 years old [3,5,8]. HLRCC phenotypes can be subtle and the penetrance is incomplete [9]. Pathogenic germline *FH* mutations are found in 76-100% of patients with HLRCC [3]. Several *FH* mutations have been included in international variant databases, such as Varsome, HGVD, among others [10,11]. However, the use of high-throughput next-generation sequencing methods in cancer diagnosis and genetic counseling lead to a large number of variants of uncertain significance (VUS) [12,13]. Consequently, it is fundamental to better characterize the pathogenic potential of newly identified genetic variants. Here, we describe the phenotype associated and the results of the Bioinformatic evaluation of the pathogenic potential of an *FH* variant, *FH* c.199T>G p.Tyr67>Asp.

Patient and methods

In November 2018, a 56-years-old female presented with progressive and limiting back pain. An abdominal computed tomography (CT) scan showed a large mass in the left kidney. As previous pathological history, she was diagnosed with systemic arterial hypertension, bronchial asthma, hypothyroidism, and visible and/or palpable typical skin leiomyomas. She presented a surgical history of a subtotal hysterectomy at 28 years old and total hysterectomy with cervical preservation at 38 years old, both due to leiomyomatosis.

The PET-CT for staging showed an 8.5 cm (QSUV 27.61) mass occupying the upper and middle thirds of the left kidney and invading structures of the renal hilum and with extracapsular extension, without a clear cleavage plane with the adrenal and the Psoas muscle (Figure S1). There were also hypermetabolic ipsilateral retroperitoneal lymphadenomegaly (QSUV 18.12) and a 31 mm hypercaptant nodule in the cervix (QSUV 10.54). Skull MRI imaging also without evidence of secondary lesions. No distant lesions were detected.

The patient underwent a left nephrectomy with a retroperitoneal lymphadenectomy. Anatomopathological analysis of the tumor revealed high-grade pT4N1M0 papillary renal carcinoma (IV). At immunohistochemistry, the tumor was classified as a fumarate-hydratase-deficient renal cell carcinoma, ISUP (International Society of Urological Pathology) Grade III, characterized by nucleolar prominence. Subsequently, she also performed partial hysterectomy (cervix) due to a hypermetabolic nodule identified at PET-CT, but the anatomopathology was compatible with leiomyoma without signs of malignancy. Management and follow-up of this patient are reported in Supplementary Box 1.

During her oncological follow-up, she was referred to genetic counseling due to suspected hereditary leiomyomatosis. The *FH* gene sequencing was performed in a Clinical Laboratory Improvement Amendments (CLIA) approved laboratory and only a VUS was detected (FH c.199T>G p.Tyr67Asp). Familial history and the variant segregation in the family are illustrated in Figure 1. According to family history, her mother (II.8) also had a hysterectomy due to leiomyomatosis (at 33 years old) and has similar skin lesions. Her daughter (IV.1) has skin and uterine leiomyomas. II.8, III.1, and IV.1 tested positive for the FH c.199T>G p.Tyr67Asp variant. On the contrary, her maternal cousin (III.7) tested negative for the FH c.199T>G p.Tyr67Asp variant. She has also been submitted to a partial hysterectomy due to leiomyomas. This patient reports “two or three myomas”, but she has kept no anatomopathological reports from this surgery.

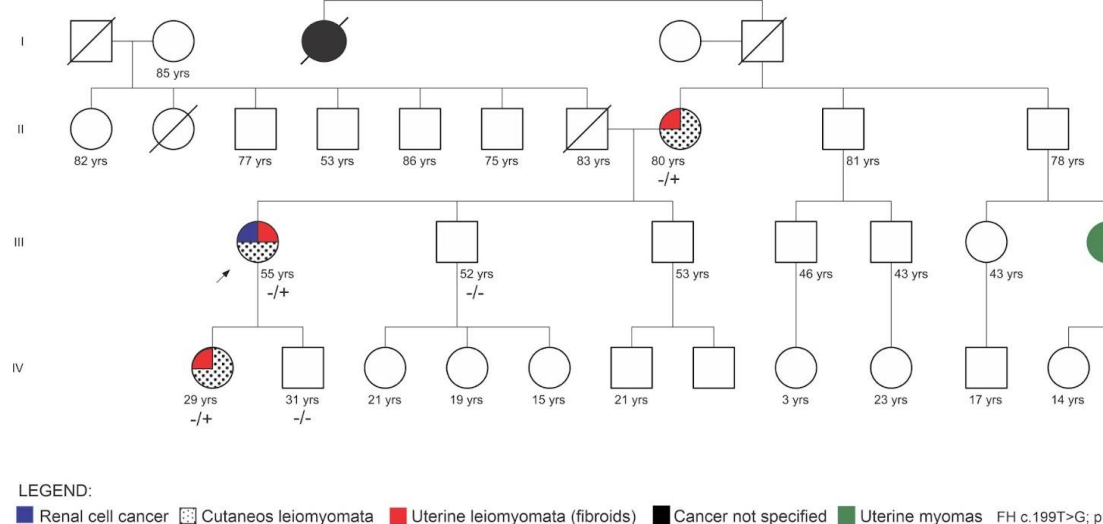


Fig. 1 Family history. FH c.199T>G p.Tyr67Asp segregation in affected individuals II-8, III-1, and IV-1. III-7 does not have FH c.199T>G p.Tyr67Asp variant; however, she reported a few uterus myomas.

Bioinformatic analyses

The following questions were investigated: 1. What is the population frequency of FH c.199T>G; p.Tyr67>Asp? Database consulted were ClinVar [15], COSMIC [16], dbSNP [17], Database of Genomic Variants [18], Ensembl [19], European Variation Archive [20], GeneCards [21], GnomAD [22], OMIM[23], PubMed [24], VarSome [10] UniProt [25], and 1000 Genomes [26]; 2. Which characteristics of the FH gene, RNA, and protein may affect the interpretation of the bioinformatic analysis of FH c.199T>G; p.Tyr67>Asp impacts? The nucleotide (NM_000143.4) and protein (P07954-1 and P07954-2) sequences were uploaded from the NCBI [27] and UniProt databases. Comparative alignments with these sequences were carried out using BlastN [28] and Clustal Omega [29], to verify the existence of differences between the nucleotide assemblies and to search for differences between the two FH protein isoforms. The transcripts encoded by the FH, the exon-intron structure, and the tissue expression pattern were identified using Ensembl [19] and GTEEx [30]. Possible impacts on splicing sites and

regulatory elements of splicing were assessed using Human Splicing Finder [31] and Mutation Taster [32]; 3. *How do softwares for pathogenicity analysis classify FH c.199T>G p.Tyr67>Asp?* Pathogenic impact of this amino acids substitution was evaluated using Polymorphism Phenotyping v2 (PolyPhen) [33], SIFT [34], Provean [35], and Mendelian Clinically Applicable Pathogenicity (M-CAP) [36]. The impact on amino acid and protein interactions was evaluated using I-Mutant[37], mCSM[38], DynaMut [39], and DUET [40]. The evolutionary conservation of the mutated amino acid was evaluated using Clustal Omega [29] and BlastP [28]; 4. *How does FH c.199T>G p.Tyr67>Asp impact FH secondary, tertiary, and quaternary structure and function?* Protein domains, post-synthesis modifications, and motifs were investigated using PDBSum [41], Pfam [42], InterPro [43], neXtProt [44], and PhosphoSitePlus [45]. The FH mutated isoform was predicted using ExPasy Translate [46]. The 3D structure was predicted using the Modeller [47] program. In addition, visualization of the wild-type and mutated 3D protein structure were performed using PYMOL [48]; 5. *Considering all the analysis developed here, what is the FH c.199T>G; p.Tyr67Asp pathogenicity classification according to American College of Genetics and Medical Genomics and the Association of Molecular Pathology (ACMG/AMP) criteria [49-51]?* The complete list of the software used is provided in Table S1.

Table S1 Bioinformatic analyses and resources

Step	Analyze	Softwares/programs used	Ref
New variant?	FHc.199T>G; p.Tyr67Asp database search	ClinVar	Landrum et al., 2020
		COSMIC	Tate et al., 2018
		dbSNP	Sherry et al., 2001
		Ensembl	Howe et al., 2021
		EVA	EVA, EMBL-EBI, 2021
		GeneCards	Stelzer et al., 2016
		gnomAD	Karczewski et al., 2020
		OMIM	McKusick, 1998
		PubMed	O'Leary et al., 2016
		VarSome	Kopanos et al., 2018
		UniProt Consortium, 2015	
		1000 Genomes	Fairley et al., 2020
Impact DNA/RNA	Canonical and alternative splicing	Human Finder Splicing	Desmet et al, 2009
	Nonsense-mediated decay (NMD)	Mutation Taster	Steinhaus et al., 2021

Impact on FH protein	Amino acid substitution	PolyPhen	Adzhubei et al., 2010
		SIFT	Vaser et al., 2016
		Provean	Choi et al., 2012
		Mutation Taster	Steinhaus et al., 2021
	Evolutionary conservation of amino acids	M-CAP	Jagadeesh et al., 2016
		Blastp	Altschul et al, 1990
	Domains	Clustal omega	Madeira et al, 2019
		Pfam	Mistry et al., 2020
		InterPro	Blum et al., 2020
	Motifs, post-synthesis modifications and protein features	PDBSum	Laskowski et al., 2018
		neXtProt	Zahn-Zabal et al., 2020
		PhosphoSitePlus	Hornbeck et al., 2015
		I-Mutant Suite	Capriotti et al., 2005
		Stability of the protein and ability to interact with other molecules	mCSM
DynaMut	Rodrigues et al, 2018		
DUET	Pires et al, 2014		
Impact on 3D structure	RCSB PDB	Burley et al., 2021	
	Modeller	Webb et al., 2016	
	PyMOL	Zhu et al., 2021	
Classification	ACMG criteria	---	Sue et al., 2015; Biesecker et al., 2018; Tavtigian et al., 2020

Note: Software accessed in October 2021

Results

Results will be presented in the same order and linked to the questions formulated in Materials and Methods.

1. *What is the population frequency of FH c.199T>G; p.Tyr67>Asp?* The FH c.199T>G; p.Tyr67Asp variant has not been identified in any database for allele frequency in the general population or in any cancer variant database searched (Table 1). It was described only once in ClinVar, dbSNP, and VarSome, classified as VUS in both.

Table 1 Frequency and pathogenicity classification of FH c.199T>G; p.Tyr67Asp according to different database

Frequency database	
COSMIC	No entry for this variant. Other variants in FH p.Tyr67: FH c.199T>C; p.Tyr67His and c.200A>C;p.TyrY67Sser
dbSNP	FH c.199T>G; p.Tyr67Asp (rs1660243497)
Ensembl	No entry for this variant Other variants in FH p.Tyr67: FH c.201T>A; p.Tyr67Ter and FH c.201T>G; p.Tyr67Ter
EVA	No entry for this variant
gnomAD	No entry for this variant
1000 Genomes	No entry for this variant Other variants in FH p.Tyr67: FH c.201T>A; p.Tyr67Ter and FH c.201T>G; p.Tyr67Ter
Database for pathogenicity classification*	
ClinVar	VUS
VarSome	FH c.199T>G; p.Tyr67Asp, VUS: PM2, PP2, and PP3 Other variants in Tyr67 found: FH c.200A>C; p.Tyr67Ser, c.201T>A; p.Tyr67Ter, c.199T>C; p.Tyr67His
Miscellaneous	
Ensembl	No entry for this variant Other variants in FH p.Tyr67: FH c.201T>A; p.Tyr67Ter and FH c.201T>G; p.Tyr67Ter
GeneCards	The variant GeneCards due to ClinVar registration Other variants in FH p.Tyr67: FH c.201T>G; p.Tyr67Ter (rs1558402241)
OMIM	No entry for this variant
PubMed	No publication
UniProt	No entry for this variant

Note: *according to the guidelines of the American College of Genetics and Medical Genomics and the Association of Molecular Pathology (ACMG/AMP) [49-51]. Programs accessed in October 2021.

2. Which characteristics of the *FH* gene, RNA, and protein may affect the interpretation of the bioinformatic analysis of *FH* c.199T>G; p.Tyr67>Asp impacts?

FH c.199T>G; p.Tyr67Asp occurs in the middle of *FH* exon 3 (Figure S2). At RNA level, impacts were evaluated by considering splicing, splicing regulation, and alternative splicing. According to GTEx, *FH* has 11 exons and, except for exon 2, all exons are highly expressed in the 54 tissues evaluated (Figure S3). *FH* presents an alternative translation initiation, starting at exon 1 or exon 2, which would not be affected by this variant. Three *FH* transcripts were found in Ensembl and GTEx. The canonical transcript is ENST00000366560.4/NM 000143.4. It has 1791 bp and 10 exons (absence of exon two) and encodes a 510 amino acid protein. The second and third transcripts, ENST00000493477.1 and ENST00000497042.1, are non-coding RNAs. Using Human Finder Splicing and Mutation Taster, no exonic splicing enhancers or silencers were detected overlapping the nucleotide position affected by this variant and no significant impacts of the *FH* c.199T>G; p.Tyr67Asp variant on splicing, splicing regulation, alternative splicing were predicted .

Figure S2 *FH* c.199T>G; p.Tyr67Asp location in the exon

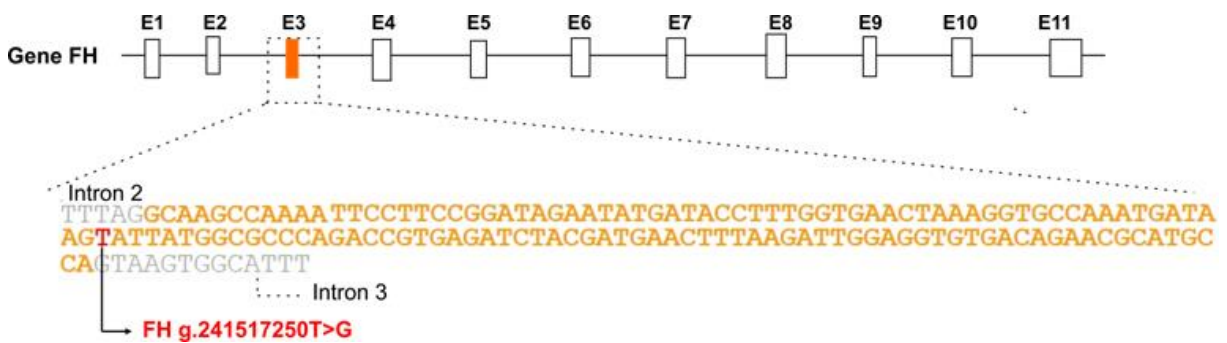


Fig. S2 In the top, a model of the *FH* exon-intron structure according to GRCh38.p13; below, *FH* c.199T>G; p.Tyr67Asp nucleotide substitution in exon 3. No effects on exon borders, exonic splicing enhancers or silencers were detected using Human Finder Splicing and Mutation Taster.

Figure S3 Tissue-specific expression of *FH* exons

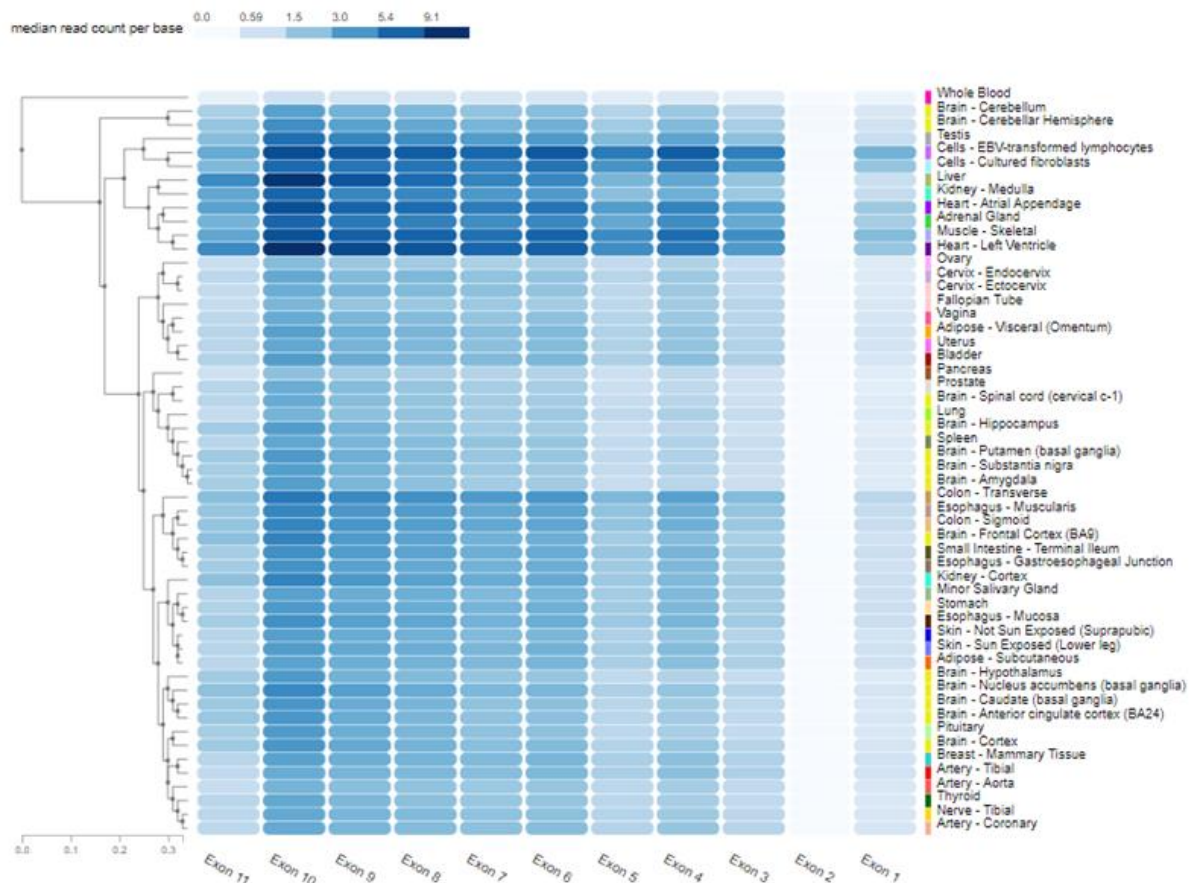


Fig. S3 The *FH* has 11 exons and, except for exon 2, all exons show moderate to high expression in the 54 tissues evaluated. The measurement is given in different shades of blue, in which the darker the blue, the higher the expression of the exon in the tissue. Created using GTEX, accessed in October 2021 (<https://www.gtexportal.org/home/gene/FH>).

There are two *FH* protein isoforms: the mitochondrial fumarate hydratase (P07954-1), with 510 amino acids, and the cytoplasmic fumarate hydratase (P07954-2), with 467 amino acids. The mitochondrial *FH* presents a 44 amino acids mitochondrial targeting sequence and the mature isoform starts in the amino acid residue 45. The cytoplasmic *FH*, produced by alternative initiation, starts in methionine at position 44. However, this methionine is removed by a methionine aminopeptidase [52]. Consequently, active cytoplasmic and mitochondrial *FH* have the same sequence. The variant *FH* c.199T>G; p.Tyr67Asp affects the residue 67 in the long, mitochondrial isoform (P07954-1) (before the exclusion of the mitochondrial targeting signal) and residue 23 in the short, cytoplasmic protein isoform (P07954-2).

FH active enzyme is homotetrameric [53]. Each monomer has three domains: N-terminal Lyase 1 domain (D1), Central domain (D2), and C-terminal Fumarase C domain (D3). *FH* c.199T>G; p.Tyr67Asp is located in the beginning of the Lyase 1 N-terminal domain (Figure S4). *FH* quaternary structure has been characterized [54 - 57, 53]. Two kinds of functional sites have been identified, one catalytic and one allosteric site, which are discussed below.

Figure S4 Illustration of the *FH* complex, the monomers and the location of the *FH* p.Tyr67Asp variant

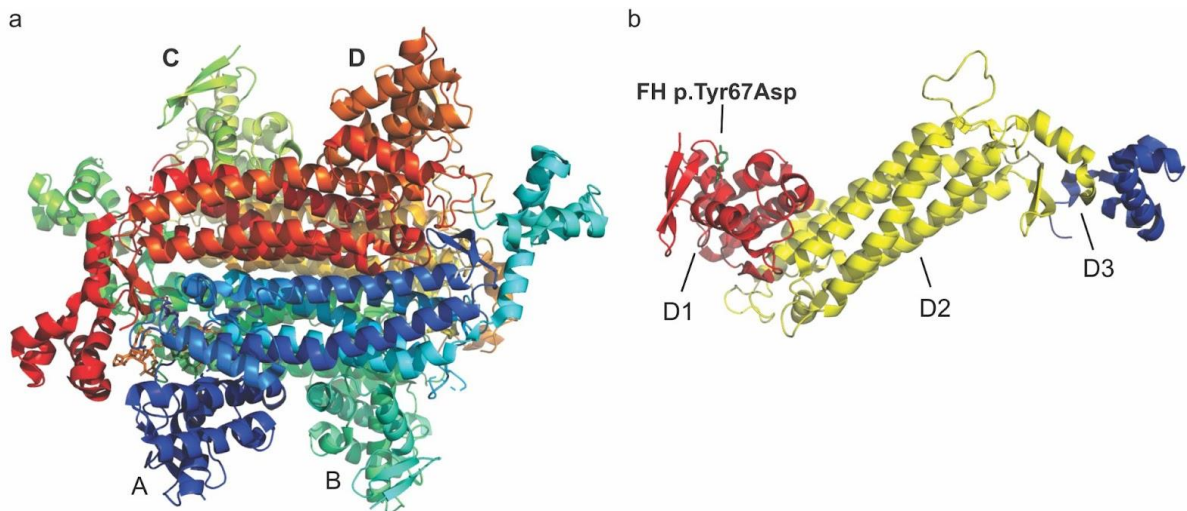


Fig. S4 In a, the wild-type FH complex formed by the four monomers, A, B, C, and D. In b, the FH monomer domains. D1, in red: Lyase1 N-terminal; D2 in yellow: Central; and D3, in violet: Fumarase C C-terminal. The position of FH p.Tyr67Asp in the monomer is shown in b. This variant is located in a loop between an alpha-helix and a beta-sheet. Adapted from Picaud et al., 2011; Bulku et al., 2018).

FH c.199T>G; p.Tyr67Asp causes the substitution of an uncharged polar by an acid amino acid. Impacts predicted according to PolyPhen, SIFT, Provean, M-CAP, Mutation Taster, I-Mutant Suite, mCSM, DUET, DynaMut, ClustalOmega, and Blastp are presented in Table 2. I-Mutant, mCSM, and DUET software indicate that FH c.199T>G; p.Tyr67Asp variant affects protein stability and protein-protein affinity. DynaMut predicted changes in FH vibrational entropy. FH c.199T>G; p.Tyr67Asp would increase flexibility between amino acids 49-88 and 109-186 (Figure 2a). In addition, the results suggested FH c.199T>G; p.Tyr67Asp affects interatomic interactions in the protein, causing the loss of halogen bonds, hydrogen bonds, water mediated hydrogen bonds, and ionic interactions (Figure 2b and 2c).

3. How do software for pathogenicity analysis classify FH c.199T>G p.Tyr67>Asp?

Over the last few years, many software have been developed aiming to classify genetic variants in terms of pathogenicity. The information and criteria vary widely among different software. Most included population frequency of the variant, evolutionary conservation of the amino acid, electric charge of the amino acids involved, and effects on protein stability and protein-protein complex affinity change upon mutation. Results of the prediction of the functional impact and/or pathogenicity classification are shown in Table 2.

Table 2 FH c.199T>G; p.Tyr67Asp predicted functional impact

Software	Score	Prediction
PolyPhen	0.99	Probably pathogenic
SIFT	1.00	Intolerant (polar non-charged to acid)
Provean	-8.07	Deleterious
M-CAP	95%	Likely pathogenic
Mutation Taster	0.99	Deleterious

	(accuracy)	
I-Mutant Suite	-1.64 kcal/mol	Decrease stability
mCSM (Protein stability change upon mutation)	-2.813 Kcal/mol	Highly destabilizing
mCSM (Protein-protein complex affinity change upon mutation)	-1.20 Kcal/mol	Affinity destabilizing
DUET Predicted Stability Change	-2.238 kcal/mol	Destabilizing
DUET SDM (Predicted Stability Change)	-0.300 kcal/mol	Destabilizing
DynaMut	-2.238 kcal/mol	Destabilizing
DynaMut ENCoM	-1.073 kcal/mol	Destabilizing
ClustalOmega and Blastp	---	Conserved amino acid

Note: Scores have different meanings according to the software. In PolyPhen, Mutation Taster, M-Cap, scores are presented in probability of the prediction. SIFT $p < 0.05$; in Provean, variants with a score equal to or below -2.5 are considered deleterious and variants with a score above -2.5 are considered neutral; I-Mutant Suite $\Delta\Delta G$ (delta-delta Gibbs, difference in free energy), value prediction $\Delta\Delta G < 0$ reflects decreased stability and $\Delta\Delta G > 0$ reflects increased stability; mCSM, DUET, and DynaMut $p < 0.001$ = destabilizing.

Figure 2 FH c.199T>G; p.Tyr67Asp effects on FH vibrational entropy energy and interatomic interactions
a Vibrational Entropy Energy

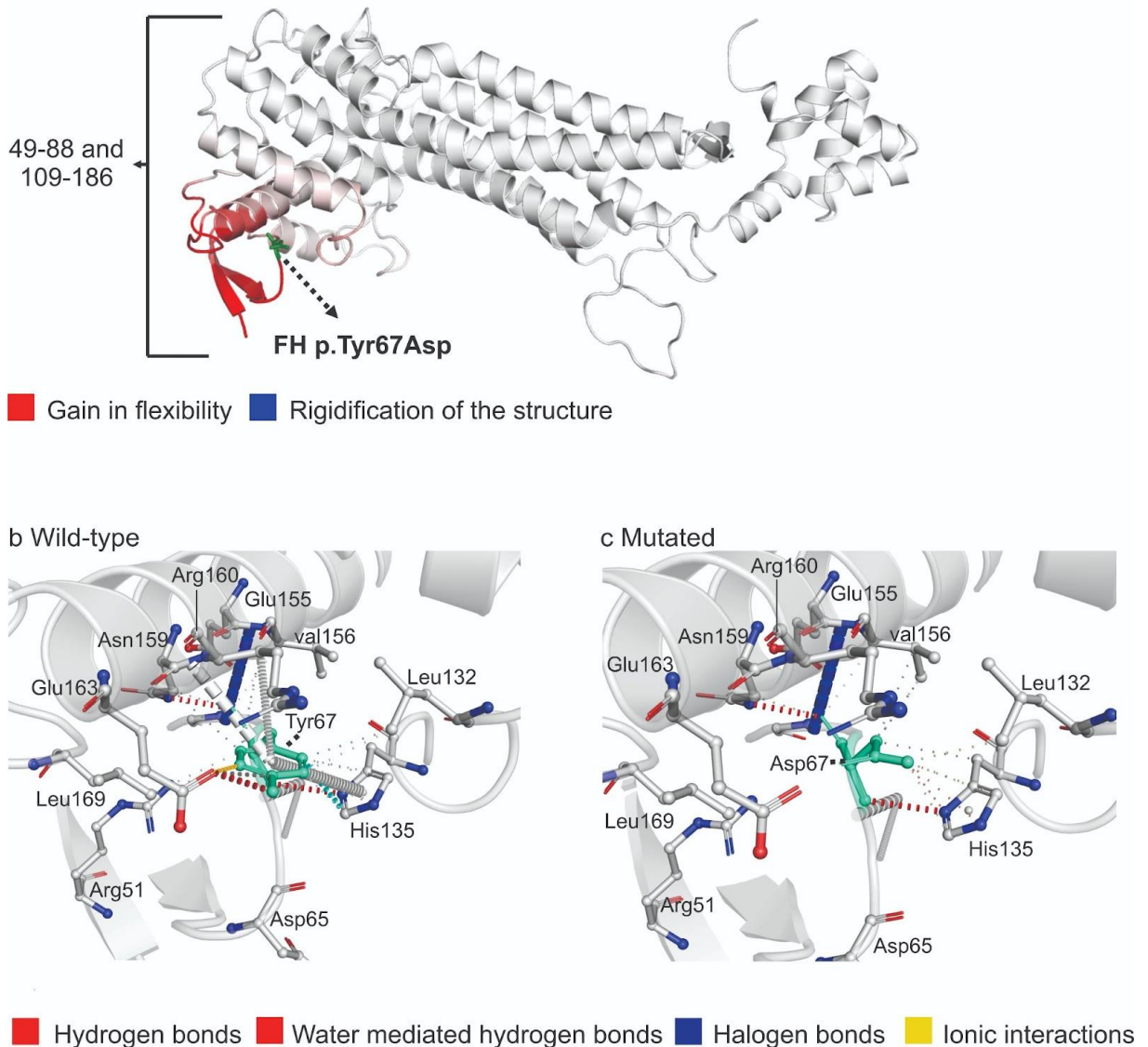


Fig. 2 - a) FH vibrational entropy energy in the presence of FH c.199T>G; p.Tyr67Asp, using DynaMut. Amino acids that presented enhanced vibrational entropy energy due to FH c.199T>G; p.Tyr67Asp are shown varying from rose to red, representing the gradual gain in flexibility of the loop; b and c) Interatomic interactions of wild-type (b) and FH c.199T>G; p.Tyr67Asp mutated protein (c), predicted using DynaMut. Residues Tyr67 (b) and Asp67 (c) are shown in light green. The loss of specific interatomic interactions, including loss of hydrogen bonds, water-mediated hydrogen bonds, halogen bonds, and ionic interactions were observed.

Another aspect to consider is evolutionary conservation. FH c.199T>G; p.Tyr67Asp is in the Lyase 1 N-terminal domain. This domain is highly conserved with approximately 90% homology in mammals, birds, reptiles, and fish, and 70% homology in fungi, bacteria, and plants. Within this domain, we identified a highly conserved, 12-amino acid segment, DKYYGAQTVRST, which contains this two tyrosine, Y67 and Y68, residues (Blastp results for DKYYGAQTVRST are shown in Figure S5). Mammals, birds, reptiles, and fish have 100% homology in this segment of the protein, while bacteria, fungi, and plants present some amino acid substitutions and/or absence of one of the two tyrosine residues, having approximately 90% of identity. In some species, only one tyrosine is present in this segment of the protein, which can be both Y67 (which can be replaced by a leucine) or Y68 (which can be replaced by a tryptophan) (e.g., *Mycobacterium tuberculosis*, see Fig S14 in Kasbekar et al., 2016). In plants, both these tyrosines can be replaced by leucine and tryptophan residues. In no species, considering all the FH sequences deposited in NCBI, an aspartic acid replacing Y67 or Y68 has been identified (accessed in October, 2021). In an effort to better characterize this DKYYGAQTVRST, we searched motif databases; this sequence has

not been included in any of the databases searched using BlastP, PhosPho.ELM [58], and ELM-the eukaryotic linear motif resource [59].

Fig. S5 Comparative alignment using the Clustal Omega tool. The blue color is the amino acids acidic; magenta color is the basic amino acids; red color is the polar amino acid; and green are non-polar amino acids; asterisks (*) indicate amino acids that are identical; and the colon (:) shows the amino acids that are not identical but have similar physicochemical character.

Homo DKYYGAQTVRST
Pan_troglodytes DKYYGAQTVRST
Gorilla DKYYGAQTVRST
Pan DKYYGAQTVRST
Pongo DKYYGAQTVRST
Nomascus DKYYGAQTVRST
Trachypithecus DKYYGAQTVRST
Theropithecus DKYYGAQTVRST
Cercocebus DKYYGAQTVRST
Ptilocolobus DKYYGAQTVRST
Mandrillus DKYYGAQTVRST
Macaca_fascicularis DKYYGAQTVRST
Rhinopithecus DKYYGAQTVRST
Chlorocebus DKYYGAQTVRST
Mulatta DKYYGAQTVRST
Papio DKYYGAQTVRST
Callithrix DKYYGAQTVRST
Colobus DKYYGAQTVRST
Saimiri DKYYGAQTVRST
Aotus DKYYGAQTVRST
Sapajus DKYYGAQTVRST
cebus DKYYGAQTVRST
Macaca DKYYGAQTVRST
Carlito DKYYGAQTVRST
Microcebus DKYYGAQTVRST
Chinchilla DKYYGAQTVRST
Pteropus DKYYGAQTVRST
Ceratotherium DKYYGAQTVRST
Erinaceus DKYYGAQTVRST
Castor DKYYGAQTVRST
Nannospalax DKYYGAQTVRST
Rousettus DKYYGAQTVRST
Odobenus DKYYGAQTVRST
Zalophus DKYYGAQTVRST
Fukomys DKYYGAQTVRST
Callorhinus DKYYGAQTVRST
Molossus DKYYGAQTVRST
Canis DKYYGAQTVRST
Vicugna DKYYGAQTVRST
Urocitellus DKYYGAQTVRST
Ictidomys DKYYGAQTVRST
Orycteropus DKYYGAQTVRST
Oryctolagus DKYYGAQTVRST
Marmota DKYYGAQTVRST
Hipposideros DKYYGAQTVRST
Balaenoptera_musculus DKYYGAQTVRST
Equus DKYYGAQTVRST
Panthera DKYYGAQTVRST
Ursus DKYYGAQTVRST
Physeter DKYYGAQTVRST
Loxodonta DKYYGAQTVRST
Phoca DKYYGAQTVRST
Acinonyx DKYYGAQTVRST
Trichechus DKYYGAQTVRST
Lynx DKYYGAQTVRST
Lipotes DKYYGAQTVRST
Balaenoptera DKYYGAQTVRST
Halichoerus DKYYGAQTVRST
Phyllostomus DKYYGAQTVRST
Neophocaena DKYYGAQTVRST
Phocoena DKYYGAQTVRST
Puma DKYYGAQTVRST
Dasybus DKYYGAQTVRST
Ailuropoda DKYYGAQTVRST
Enhydra DKYYGAQTVRST
Monodon DKYYGAQTVRST
Bos DKYYGAQTVRST
Mus DKYYGAQTVRST
Anas DKYYGAQTVRST
Meleagris DKYYGAQTVRST
Egretta DKYYGAQTVRST
Columba DKYYGAQTVRST
Calypte DKYYGAQTVRST
Dermochelys DKYYGAQTVRST
Mauremys DKYYGAQTVRST
Terrapene DKYYGAQTVRST
Chelonoidis DKYYGAQTVRST
Chelonia DKYYGAQTVRST
Pelodiscus DKYYGAQTVRST
Danio DKYYGAQTVRST
Acipenser DKYYGAQTVRST
Erpetoichthys DKYYGAQTVRST
Chanos DKYYGAQTVRST
Paramormyrops DKYYGAQTVRST
Scleropages DKYYGAQTVRST
Rhizopus DKYYGAQTVRST
Glomus DKYYGAQTVRST
Choanephora DKYYGAQTVRST
Actinomortierella DKYYGAQTVRST
Gigaspora DKYYGAQTVRST
Proteobacteria DKYYGAQTVRST
Desulfobacteriales DKYYGAQTVRST
Candidatus DKYYGAQTVRST
Chlamydiae DKYYGAQTVRST
Ignavibacteriales DKYYGAQTVRST
Rhodamnia DKYYGAQTVRST
Quercus DKYYGAQTVRST
Morus DKYYGAQTVRST
:::****,*

Mammalia
Aves
Reptiles
Fish
Fungi
Bacteria
Plants

Homo DKYYGAQTVRST
Pan_troglodytes DKYYGAQTVRST
Gorilla DKYYGAQTVRST
Pan DKYYGAQTVRST
Pongo DKYYGAQTVRST
Nomascus DKYYGAQTVRST
Trachypithecus DKYYGAQTVRST
Theropithecus DKYYGAQTVRST
Cercocebus DKYYGAQTVRST
Ptilocolobus DKYYGAQTVRST
Mandrillus DKYYGAQTVRST
Macaca_fascicularis DKYYGAQTVRST
Rhinopithecus DKYYGAQTVRST
Chlorocebus DKYYGAQTVRST
Mulatta DKYYGAQTVRST
Papio DKYYGAQTVRST
Callithrix DKYYGAQTVRST
Colobus DKYYGAQTVRST
Saimiri DKYYGAQTVRST
Aotus DKYYGAQTVRST
Sapajus DKYYGAQTVRST
cebus DKYYGAQTVRST
Macaca DKYYGAQTVRST
Carlito DKYYGAQTVRST
Microcebus DKYYGAQTVRST
Chinchilla DKYYGAQTVRST
Pteropus DKYYGAQTVRST
Ceratotherium DKYYGAQTVRST
Erinaceus DKYYGAQTVRST
Castor DKYYGAQTVRST
Nannospalax DKYYGAQTVRST
Rousettus DKYYGAQTVRST
Odobenus DKYYGAQTVRST
Zalophus DKYYGAQTVRST
Fukomys DKYYGAQTVRST
Callorhinus DKYYGAQTVRST
Molossus DKYYGAQTVRST
Canis DKYYGAQTVRST
Vicugna DKYYGAQTVRST
Urocitellus DKYYGAQTVRST
Ictidomys DKYYGAQTVRST
Orycteropus DKYYGAQTVRST
Oryctolagus DKYYGAQTVRST
Marmota DKYYGAQTVRST
Hipposideros DKYYGAQTVRST
Balaenoptera_musculus DKYYGAQTVRST
Equus DKYYGAQTVRST
Panthera DKYYGAQTVRST
Ursus DKYYGAQTVRST
Physeter DKYYGAQTVRST
Loxodonta DKYYGAQTVRST
Phoca DKYYGAQTVRST
Acinonyx DKYYGAQTVRST
Trichechus DKYYGAQTVRST
Lynx DKYYGAQTVRST
Lipotes DKYYGAQTVRST
Balaenoptera DKYYGAQTVRST
Halichoerus DKYYGAQTVRST
Phyllostomus DKYYGAQTVRST
Neophocaena DKYYGAQTVRST
Phocoena DKYYGAQTVRST
Puma DKYYGAQTVRST
Dasybus DKYYGAQTVRST
Ailuropoda DKYYGAQTVRST
Enhydra DKYYGAQTVRST
Monodon DKYYGAQTVRST
Bos DKYYGAQTVRST
Mus DKYYGAQTVRST
Anas DKYYGAQTVRST
Meleagris DKYYGAQTVRST
Egretta DKYYGAQTVRST
Columba DKYYGAQTVRST
Calypte DKYYGAQTVRST
Dermochelys DKYYGAQTVRST
Mauremys DKYYGAQTVRST
Terrapene DKYYGAQTVRST
Chelonoidis DKYYGAQTVRST
Chelonia DKYYGAQTVRST
Pelodiscus DKYYGAQTVRST
Danio DKYYGAQTVRST
Acipenser DKYYGAQTVRST
Erpetoichthys DKYYGAQTVRST
Chanos DKYYGAQTVRST
Paramormyrops DKYYGAQTVRST
Scleropages DKYYGAQTVRST
Rhizopus DKYYGAQTVRST
Glomus DKYYGAQTVRST
Choanephora DKYYGAQTVRST
Actinomortierella DKYYGAQTVRST
Gigaspora DKYYGAQTVRST
Proteobacteria DKYYGAQTVRST
Desulfobacteriales DKYYGAQTVRST
Candidatus DKYYGAQTVRST
Chlamydiae DKYYGAQTVRST
Ignavibacteriales DKYYGAQTVRST
Rhodamnia DKYYGAQTVRST
Quercus DKYYGAQTVRST
Morus DKYYGAQTVRST
:::****,*

Mammalia
Aves
Reptiles
Fish
Fungi
Bacteria
Plants

In addition, we investigated the impacts of FH c.199T>G; p.Tyr67Asp on FH the motifs and post-synthesis modification sites. Based on the results of PhosphoSitePlus, 45 amino acids in FH are involved in different post-synthesis modification processes, including ubiquitination, phosphorylation, acetylation, mono-methylation, and succinylation (Figure 3). The two residues, Y67 and Y68, are predicted to be phosphorylated. Other software (PhosPho.ELM and ELM-the eukaryotic linear motif resource) did not predict the phosphorylation of these two tyrosines.

Using neXtProt, three highly conserved regions were identified in the monomeric FH protein, consisting of the residues 145-147 (SerGlyThr), 186-188 (SerSerAsn), and 371-373 (LysValAsn), that enhance the connection of substrates and ligands (Figure 3). Another region highlighted is the B site, which is composed of residues 173-176 (HisProAsn) (Figure 3). According to Picaud and coworkers (2011), the FH tetrameric protein has two sites, A and B, that allow the binding of substrates and allosteric effectors [53]. Site A has catalytic activity and is composed of residues S145, T147, and N188.

Figure 3 FH post-synthesis modifications

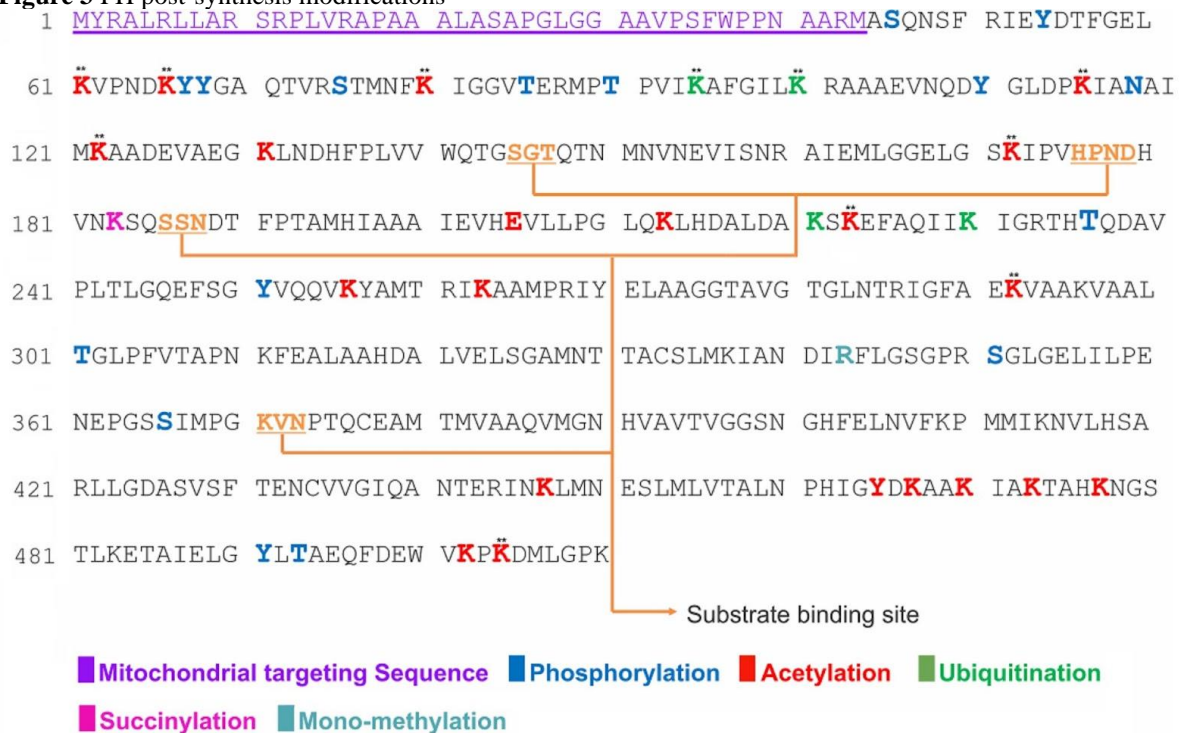


Fig. 3 FH domains and post-synthesis modification sites, according to PhosphoSitePlus. The mitochondrial targeting sequence has 44 amino acids and is shown in gold. The double asterisks (**) indicate those amino acids that may have more than one post-synthesis modification: K61: acetylation and succinylation; K66: acetylation and ubiquitination; K80: acetylation, ubiquitination, and succinylation; K94: ubiquitination, acetylation, and succinylation; K100: ubiquitination and mono-methylation; K115: acetylation and succinylation; K122: acetylation, ubiquitination and succinylation; K172: acetylation and ubiquitination; K223: acetylation and succinylation; K292: acetylation, ubiquitination and succinylation; K502: acetylation and ubiquitination.

Using the Modeller and the human FH 3D protein model (5UPP) available in the PDB [60] database (<https://www.rcsb.org/structure/5UPP>), no conformational changes in the FH 3D protein structure, nor changes in the active site were observed in spite of the loss of a highly conserved phosphorylation site, the Tyr67. These analyses were also conducted with PDB ID: 3E04 (<https://www.rcsb.org/structure/3E04>) human FH model, with similar results (results not shown).

Considering all the results described here, we reevaluated the pathogenicity of FH c.199T>G; p.Tyr67Asp. This variant fulfills the following ACMG/AMP criteria: PM1, PM2, PP1, PP2, and PP3 (Table 3). Consequently, we propose the reclassification of FH p.Tyr67Asp from VUS to likely pathogenic variant.

Table 3 ACMG/AMP guidelines applied to FH c.199T>G; p.Tyr67Asp variant

Criterion	Analysis description
PM1	<p><i>Criterion - Located in a mutational hot spot and/or critical and well-established functional domain without benign variation</i></p> <p>Answer - FH Tyr67 is not a mutation hotspot. FH p.Tyr67Asp is located in a D1 (Lyase1 N-terminal) functional domain; Y67 is a predicted phosphorylation site. No benign/likely benign has been identified in this residue. Two nonsense variants have been identified in this residue and classified as pathogenic.</p>
PM2	<p><i>Criterion - Absent from controls (or extremely low frequency if recessive) in Exome Sequencing Project [61], 1000 Genomes Project, or Exome Aggregation Consortium</i></p> <p>Answer - This variant has not been included in Exome Sequencing Project, 1000 Genomes Project, or gnomAD (which includes <i>Exome Aggregation Consortium</i>). FH p.Tyr67Asp has been recently included in ClinVar, GeneCard, dbSNP, and VarSome, classified as VUS.</p>
PP1	<p><i>Criterion - Co-segregation with disease in multiple affected family members of a single family in a gene definitively known to cause the disease</i></p> <p>Answer - FH c.199T>G; p.Tyr67Asp co-segregates with the disease in three affected family members</p>
PP2	<p><i>Criterion - Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease</i></p> <p>Answer - According to VarSome, the gene has 991 mutations, of which 279 are classified as pathogenic and likely pathogenic, 206 are benign/likely benign, and 506 are VUS. Considering the missense variants only, three were classified as benign/likely benign, 89 are classified as pathogenic/likely pathogenic</p>
PP3	<p><i>Criterion - Multiple lines computational evidence support a deleterious effect on the gene or gene product (evolutionary conservation, splicing impact, etc.)</i></p> <p>Answer - <i>In silico</i> functional impact prediction programs (PolyPhen, SIFT, Provean, M-CAP, Mutation Taster, I-Mutant Suite, mCSM, DynaMut, DUET, neXtProt, PhosphoSitePlus) characterize the FH c.199T>G; p.Tyr67Asp as pathogenic, deleterious, or destabilizing. No impact is predicted on splicing, splicing regulation or alternative splicing. Tyr67 is located in a functional domain, lyase 1 N-terminal). Within this domain, we identified a highly conserved protein segment (DKYYGAQTVRST), present in mammals, reptiles, birds, bacteria, fungi, and even plants</p>

Note: ACMG/AMP: American College of Genetics and Medical Genomics and the Association of Molecular Pathology [49-51]. (accessed in November, 2021).

Discussion

VUS are frequent in genetic tests, especially when it comes to NGS panels. In clinical settings, VUS brings uncertainty about the patient's management and can produce as much anxiety as clearly pathogenic variants [62, 63]. Even though it is not recommended to change the clinical management of patients based on VUS results, it is encouraged to engage the patient in follow up studies and to assure long time effort to reclassify each and every VUS [64]. Currently, efforts aiming at VUS reclassification are based on bioinformatic methods or functional studies of each variant [65]. The methodology chosen depends on how well established is the association of a

specific gene and the disease, but also on how feasible functional studies for every single candidate variant are. The analysis of the co-segregation of the variant and the phenotype in affected families is particularly informative.

In the present study, HLRCC and FH c.199T>G; p.Tyr67Asp co-segregate in three members of the family reported here (the proband, her mother, and her daughter). Interestingly, the maternal cousin (III.7), daughter of a brother of her mother, reported having had several leiomyomas (“two or three”) and does not have the FH c.199T>G; p.Tyr67Asp variant. It is noteworthy, that the manifestation of uterine fibroids in this cousin was not as exuberant as in the proband. Uterine fibroids, known as leiomyomata, represent the most common benign uterine tumor and occur in 50 to 60% of the women [66]. Although the specific cause is not well understood, the risk factors are multiple, including race, age, early menarche, obesity, lifestyle, and genetic factors [67]. Given the negative test, the cousin's clinical manifestations would represent a phenocopy, given the high frequency of this condition in the population. Considering the incomplete penetrance and clinical variability, the presence of a phenocopy in a family where HLRCC is segregating, may represent the genetic background (meaning polygenic component) associated with this condition.

Currently, most genetic variations are classified using Bioinformatic tools. Most studies are based on a small number of software, evaluating: 1. variant class (nonsense, missense, or silent); 2 low population frequency; 3. high evolutionary conservation of the nucleotide/amino acid position [49-51]. The usefulness of such criteria depends on the variant itself. Rare, nonsense variants, affecting highly conserved nucleotide/amino acid positions are generally accepted as pathogenic/likely pathogenic. However, most of the variants that do not fulfill one of these criteria are classified as VUS. Over the last decade, a plethora of Bioinformatic methods have been developed for variation analysis, investigating all and every molecular genetic mechanism from DNA to protein function. FH c.199T>G; p.Tyr67Asp is a rare variant, affecting a highly conserved amino acid residue. The impact of the residue substitution suggests functional repercussion, but additional Bioinformatic analysis is necessary.

First, we evaluate the impact at RNA level. No impacts on splicing, splicing regulation and alternative splicing were detected. Once FH c.199T>G; p.Tyr67Asp does not insert a stop codon, RNA-based rescue mechanisms, such nonsense mediated decay would not be expected.

Second, at protein level, we evaluated the impacts on protein secondary, tertiary, and quaternary structures. FH c.199T>G; p.Tyr67Asp maps to a five residue loops, between an alpha-helix and a beta-sheet. Considering the 3D structure, FH c.199T>G; p.Tyr67Asp is in the D1, lyase 1 N-terminal, which contains part of the catalytic site (residues 143–147 and 186–188). The vibrational entropic energy analysis identified increased flexibility affecting residues 49-88 e 109-186 (Fig. 2). This increased flexibility fades gradually in the proximity of residue 186, the first of the catalytic site, and therefore, its impact can not be straightforwardly assumed. To better understand the physico-chemical basis of this increased flexibility, we compared the atomic interactions established in the region in the presence of the wild-type and the mutated protein (Fig. 2a and 2b). The replacement of Tyr67 by an Asp67 causes the loss of hydrogen or water mediated hydrogen bonds, halogen bonds, and ionic interactions. The increase in free energy allows the prediction of loss of protein stability and change in protein-protein or protein-other partners affinity. The interpretation of this result is complex, because there are two models of quaternary structure for FH. On one hand, the model predicted by Kasbekar and coworkers (2016) suggests that the FH tetramer would present four catalytic sites, one in each monomer and two allosteric sites, composed of amino acids from two different subunits (or monomers). On the other hand, the model proposed by Subasri and coworkers (2017) predicted that the catalytic site would be composed of residues from three subunits and the allosteric site would be composed of residues from one subunit. Kasbekar and coworkers' model is based on a *Mycobacterium tuberculosis* protein model (PDB ID: 5F92, 5F91, 3NO9, and 4ADL) and Subasri and coworkers model is based on (PDB ID: 3E04), a human FH model also tested in the present study, with results similar to those obtained using 5UPP (results not shown). The sequence and tertiary structure observed in the *Mycobacterium tuberculosis* FH are quite similar to the human FH and we are not perfectly sure that the differences in the models are due to species-specific differences. Therefore, although the software (DynaMut, mCSM, and DUET) used predicted that the higher free energy levels would induce protein destabilization, the extension of this impact on the functions of catalytic and allosteric sites, as well as the ability of the monomer to enter tetramerization, are not specifically predicted.

PolyPhen, SIFT, Provean, M-CAP, and Mutation Taster software were used to assess the effects of the FH c.199T>G; p.Tyr67Asp amino acid substitution. These softwares take into account: a) the physicochemical characteristics of the amino acids; b) the effects on protein and regulatory characteristics; and c) the population frequency and the evolutionary conservation of the residue. Using the *in-silico* predictors mentioned above, FH c.199T>G; p.Tyr67Asp was classified as likely pathogenic and with deleterious effects, due to physicochemical characteristics of the exchanged residues (an uncharged polar to an acidic amino acid) and the absence of FH c.199T>G;p.Tyr67Asp in populational variant databases.

Additionally, Tyr67 is part of a highly conserved, 12 residues FH segment (DKYYGAQTVRST), which has not been recognized by the motif detection software used (PhosPho.ELM and ELM-the eukaryotic linear motif). No aspartic acid (or glutamic acid) at Tyr67 position, even considering all the species analyzed by Blastp, suggesting that an acidic amino acid is not tolerated in this position. The deep evolutionary conservation of Tyr67 suggests that it may be a target for post-synthesis modifications. Using PhosphoSitePlus, several post-synthesis modifications were predicted throughout FH, including the phosphorylation of the Tyr67 and Tyr68. In fact, considering Blastp results, we observed that even when an amino acid substitution affected Tyr67 or Tyr68, at least one phosphorylatable residue remains. Either the Tyr67 was replaced by a leucine in plants *or* a Tyr68 is replaced by a tryptophan in plants and fungi. However, no study investigating post-translational FH modifications in cellular or animal models described specifically the phosphorylation of FH Tyr67 or Tyr68 to date [68, 69]. Contrarily, carboxylation of human FH active sites has been described in the literature [54], but not detected by PhosPhoSitePlus.

Conclusion

In conclusion, considering the results and criteria of the ACMG/AMP [49-51], we propose the reclassification of the FH c.199T> G; p.Tyr67Asp variant from VUS to likely pathogenic. The criteria used for this reclassification were: 1) PM1, because the variant is located in a functional domain; 2) PM2, because the variant is absent in population frequency and in cancer variation database. It has been included in ClinVar, VarSome and dbSNP; 3) PP1, because the variant segregates with the disease in three family members; 4) PP2, because it is a missense variant in a gene with a low rate of benign missense variants; 5) PP3, because the variant shows evidence of deleterious effects *in silico* analyses. The in-depth *in silico* approach used here allowed not only the reclassification but also to understand how and why, FH c.199T> G; p.Tyr67Asp is a likely pathogenic variant. FH c.199T> G; p.Tyr67Asp reclassification is relevant for clinical management and genetic counseling. Considering the clinical diagnosis, the affected individuals in this family were already being managed as HLRCC as proposed for the main phenotype with a dermatological evaluation of skin lesions, screening for renal cancer using abdominal MRI and gynecological exam [3,5,8]. The reclassification of FH c.199T>G; p.Tyr67Asp will help in decisions related to clinical management of the unaffected members of this family.

References

1. Furuya M, Iribe Y, Nagashima Y, Kambe N, Ohe C, Kinoshita H, Sato C, Kishida T, Okubo Y, Numakura K, Nanjo H, Nakaigawa N, Makiyama K, Hasumi H, Iwashita H, Ohta J, Kitamura H, Nakajima T, Yoshida T, Nakagawa M, Tanaka R, Yao M (2020) Clinicopathological and molecular features of hereditary leiomyomatosis and renal cell cancer-associated renal cell carcinomas. *J Clin Pathol* 73:819-825. <https://doi.org/10.1136/jclinpath-2020-206548>.
2. Sanz-Ortega J, Vocke C, Stratton P, Linehan WM, Merino MJ (2013) Morphologic and molecular characteristics of uterine leiomyomas in hereditary leiomyomatosis and renal cancer (HLRCC) syndrome. *Am J Surg Pathol* 37:74-80. <https://doi.org/10.1097/PAS.0b013e31825ec16f>.
3. Menko FH, Maher ER, Schmidt LS, Middleton LA, Aittomäki K, Tomlinson I, Richard S, Linehan WM (2014) Hereditary leiomyomatosis and renal cell cancer (HLRCC): renal cancer risk, surveillance and treatment. *Fam Cancer* 13:637-644. <https://doi.org/10.1007/s10689-014-9735-2>.
4. Leshets M, Silas YBH, Lehming N, Pines O (2018) Fumarase: From the TCA cycle to DNA damage response and tumor suppression. *Front Mol Biosci* 5:68. <https://doi.org/10.3389/fmolb.2018.00068>.
5. Sulkowski PL, Sundaram RK, Oeck S, Corso CD, Liu Y, Noorbakhsh S, Niger M, Boeke M, Ueno D, Kalathil AN, Bao X, Li J, Shuch B, Bindra RS, Glazer PM (2018) Krebs-cycle-deficient hereditary cancer syndromes are defined by defects in homologous-recombination DNA repair. *Nat Genet* 50:1086-1092. <https://doi.org/10.1038/s41588-018-0170-4>.
6. Hol JA, Jongmans MCJ, Littooij AS, de Krijger RR, Kuiper RP, van Harssel JJT, Mensenkamp A, Simons M, Tytgat GAM, van den Heuvel-Eibrink MM, van Grotel M (2020) Renal cell carcinoma in young FH mutation carriers: case series and review of the literature. *Fam Cancer* 19: 55-63. <https://doi.org/10.1007/s10689-019-00155-3>.

7. Schmidt C, Sciacovelli M, Frezza C (2019) Seminars in cell and developmental biology-fumarate hydratase in cancer: A multifaceted tumour suppressor. Elsevier pp. 15-25. <https://doi.org/10.1016/j.semcdb.2019.05.002>.
8. Henson JW, Resta RG (2021) Hereditary leiomyoma and renal cell carcinoma. In: Diagnosis and management of hereditary cancer, 1st edn. Elsevier, London pp 277-278.
9. Tomlinson IP, Alam NA, Rowan AJ, Barclay E, Jaeger EE, Kelsell D, Leigh I, Gorman P, Lamlum H, Rahman S, Roylance RR, Olpin S, Bevan S, Barker K, Hearle N, Houlston RS, Kiuru M, Lehtonen R, Karhu A, Vilkkki S, Laiho P, Eklund C, Vierimaa O, Aittomäki K, Hietala M, Sistonen P, Paetau A, Salovaara R, Herva R, Launonen V, Aaltonen LA (2002) Multiple Leiomyoma Consortium. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat Genet* 30:406-410. <https://doi.org/10.1038/ng849>.
10. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Aguilera MA, Meyer R, Massouras A (2019) VarSome: the human genomic variant search engine. *Bioinformatics* 35:1978. <https://doi.org/10.1093/bioinformatics/bty897>.
11. Higasa K et al (2016) Human genetic variation database, a reference database of genetic variations in the Japanese population. *J Hum Genet* 61:547-553. <https://doi.org/10.1038/jhg.2016.12>.
12. Forde C, Lim DHK, Alwan Y, Burghel G, Butland L, Cleaver R, Dixit A, Evans DG, Hanson H, Lalloo F, Oliveira P, Vilard L, Wallis Y, Maher E, Woodward ER (2020) Hereditary leiomyomatosis and renal Cell cancer: Clinical and molecular, and screening features in a cohort of 185 affected individuals. *Eur Urol Oncol* 3:764-772. <https://doi.org/10.1016/j.euo.2019.11.002>.
13. Federici G, Soddu S (2020) Variants of uncertain significance in the era of high-throughput genome sequencing: a lesson from breast and ovary cancers. *J Exp Clin Cancer Res* 39:1-12. <https://doi.org/10.1186/s13046-020-01554-6>.
14. Ko JJ, Xie W, Kroeger N, Lee JL, Rini BI, Knox JJ, Bjarnason GA, Srinivas S, Pal SK, Yuasa T, Smoragiewicz M, Donskov F, Kanesvaran R, Wood L, Ernst DS, Agarwal N, Vaishampayan UN, Rha SY, Choueiri TK, Heng DY (2015) The International Metastatic Renal Cell Carcinoma Database Consortium model as a prognostic tool in patients with metastatic renal cell carcinoma previously treated with first-line targeted therapy: a population-based study. *Lancet Oncol* 16:293-300. [https://doi.org/10.1016/S1470-2045\(14\)71222-7](https://doi.org/10.1016/S1470-2045(14)71222-7).
15. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, Hoffman D, Jang KW, Kaur and Liu C (2020) ClinVar: improvements to accessing data. *Nucleic Acids Res* 48(D1):D835-D844. <https://doi.org/10.1093/nar/gkx1153>.
16. Tate John G et al (2019) COSMIC: The Catalogue of Somatic Mutations In Cancer. *Nucleic Acids Res* 47(D1):D941-D947. <https://doi.org/10.1093/nar/gky1015>.
17. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29:308-11. <https://doi.org/10.1093/nar/29.1.308>.
18. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW (2014) The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* 42(D1): D986-D992. <https://doi.org/10.1093/nar/gkt958>.
19. Howe KL et al (2021) Ensembl. *Nucleic Acids Res* 49(D1):D884-D891. <https://doi.org/10.1093/nar/gkaa942>.
20. The European Bioinformatics Institute - EMBL-EBI (2020) European Variation Archive. <https://www.ebi.ac.uk/eva/>. Last edited: Feb. 25, 2020. Accessed: Oct 25, 2021.
21. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I and Mazor Y (2016) The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinform* 54:1.30. 31-31.30. 33. <https://doi.org/10.1002/cpbi.5>.

22. Karczewski KJ, Francioli LC, Tiao G et al (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581:434–443. <https://doi.org/10.1038/s41586-020-2308-7>.
23. McKusick VA (1998) Mendelian Inheritance in Man: a catalog of human genes and genetic disorders. JHU Press. <https://doi.org/10.1086/514346>.
24. O'Leary NA et al (2015) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44(D1):D733–45. <https://doi.org/10.1093/nar/gkv1189>.
25. Consortium U (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47(D1):D506–D515. <https://doi.org/10.1093/nar/gky1049>.
26. Fairley S, Lowy-Gallego E, Perry E, Flicek P (2020) The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. *Nucleic Acids Res* 48D1:D941–D947. <https://doi.org/10.1093/nar/gkz836>.
27. NCBI Resource Coordinators (2018) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 46(D1):D8–D13. <https://doi.org/10.1093/nar/gkx1095>.
28. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
29. Madeira F, Park YM, Lee J et al (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47(W1): W636–W641. <https://doi.org/10.1093/nar/gkz268>.
30. GTEx Consortium (2015) The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* 348:648–660. <https://doi.org/10.1126/science.1262110>.
31. Desmet FO, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 37: e67. <https://doi.org/10.1093/nar/gkp215>.
32. Steinhaus R, Proft S, Schuelke M, Cooper DN, Schwarz JM, Seelow D (2021) MutationTaster. *Nucleic Acids Res* 49(W1): W446–W451. <https://doi.org/10.1093/nar/gkab266>.
33. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249. <https://doi.org/10.1038/nmeth0410-248>.
34. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC (2016) SIFT missense predictions for genomes. *Nat Protocols* 11:1–9. <https://doi.org/10.1038/nprot.2015.123>.
35. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE* 7: e46688. <https://doi.org/10.1371/journal.pone.0046688>.
36. Jagadeesh KA, Wenger AM, Berger MMJ, Guturu H, Stenson PD, Cooper DN, Bernstein J, Bejerano G (2016) M-CAP eliminates a majority of variants with uncertain significance in clinical exomes at high sensitivity. *Nat Genet* 48:1581–1586. <https://doi.org/10.1038/ng.3703>.
37. Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res* 33: W306–W310. <https://doi.org/10.1093/nar/gki375>.
38. Pires DEV, Ascher DB, Blundell TL (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* 30:335–342. <https://doi.org/10.1093/bioinformatics/btt691>.

39. Rodrigues CHM, Pires DEV, Ascher DB (2018) DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res* 46W1:W350–W355. <https://doi.org/10.1093/nar/gky300>.
40. Pires DEV, Ascher DB, Blundell TL (2014) DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res* 42W1:W314–W319. <https://doi.org/10.1093/nar/gku411>.
41. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM (2018) PDBsum: Structural summaries of PDB entries. *Protein Sci* 27:129-134. <https://doi.org/10.1002/pro.3289>.
42. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A (2021) Pfam The protein families database in 2021. *Nucleic Acids Res* 49(D1):D412-D419. <https://doi.org/10.1093/nar/gkaa913>.
43. Blum M, Chang H-I, Schugurensky S, Grego T, Kandasamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, Richardson L, Salazar GA, Williams L, Bork P, Bridge A, Gough J, Haft DH, Letunic I, Marchler-Bauer A, Mi H, Natale DA, Necci M, Orengo CA, Pandurangan AP, Rivoire C, Sigrist CJA, Sillitoe I, Thanki N, Thomas PD, Tosatto SCE, Wu CH, Bateman A, Finn RD (2020) The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* 49(D1):D344-D354. <https://doi.org/10.1093/nar/gkaa977>.
44. Zahn-Zabal M, Michel PA, Gateau A, Nikitin F, Schaeffer M, Audot E, Gaudet P, Duek Roggli P, Teixeira D, Rech de Laval V, Samarasinghe K, Bairoch A, Lane L (2020) The neXtProt knowledgebase in 2020: data, tools and usability improvements. *Nucleic Acids Res* 48(D1): D328-D334. <https://doi.org/10.1093/nar/gkz995>.
45. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E (2015) PhosphoSitePlus mutations, 2014: PTMs and recalibrations. *Nucleic Acids Res* 43D1:D512-520. <https://doi.org/10.1093/nar/gkr1122>.
46. Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C (2021) ExPASy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Res* 49(W1): W216-W227. <https://doi.org/10.1093/nar/gks225>.
47. Webb B, Sali A (2016) Comparative Protein Structure Modeling Using Modeller. *Curr Protoc Bioinform* 54: 5.6.1-5.6.37. <https://doi.org/10.1002/cpbi.3>.
48. Zhu K, Day T, Warshaviak D, Murrett C, Friesner R, Pearlman D (2014) Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins* 82:1646-1655. <https://doi.org/10.1002/prot.24551>.
49. Richards S, Aziz N, Bale S et al (2015) 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* 17:405423. <https://doi.org/10.1038/gim.2015.30>.
50. Biesecker LG, Harrison SM & the ClinGen Sequence Variant Interpretation Working Group (2018) The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genet Med* 20:1687-1688. <https://doi.org/10.1038/gim.2018.42>.
51. Tavtigian SV, Harrison SM, Boucher KM, Biesecker LG (2020) Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. *Human Mutation* 41:1734-1737. <https://doi.org/10.1002/humu.24088>.
52. Dik E, Naamati A, Asraf H, Lehming N, Pines O (2016) Human fumarate hydratase is dual localized by an alternative transcription initiation mechanism. *Traffic* 17:720-32. <https://doi.org/10.1111/tra.12397>.
53. Picaud S, Kavanagh KL, Yue WW, Lee WH, Muller-Knapp S, Gileadi O, Sacchettini J, Oppermann U (2011) Structural basis of fumarate hydratase deficiency. *J Inher Metab Dis* 34:671-6. doi: <https://doi.org/10.1007/s10545-011-9294-8>.

54. Ajalla Aleixo MA, Rangel VL, Rustiguel JK, de Pádua RAP, Nonato MC (2019) Structural, biochemical, and biophysical characterization of recombinant human fumarate hydratase. *FEBS J* 286:1925-1940. <https://doi.org/10.1111/febs.14782>.
55. Bulku A, Weaver TM, Berkmen MB (2018) Biochemical characterization of two clinically relevant human fumarase variants defective for oligomerization. *Open Biochem J* 12:1-15. <https://doi.org/10.2174/1874091X01812010001>.
56. Kasbekar M, Fischer G, Mott BT, Yasgar A, Hyvönen M, Boshoff HI, Abell C, Barry CE 3rd, Thomas CJ (2016) Selective small molecule inhibitor of the Mycobacterium tuberculosis fumarate hydratase reveals an allosteric regulatory site. *Proc Natl Acad Sci U S A* 113:7503-8. <https://doi.org/10.1073/pnas>.
57. Subasri S, Chaudhary SK, Sekar K, Kesharwani M, Velmurugan D (2017) Molecular docking and molecular dynamics simulations of fumarate hydratase and its mutant H235N complexed with pyromellitic acid and citrate. *J Bioinform Comput Biol* 15:1750026. <https://doi.org/10.1142/S0219720017500263>.
58. Diella F, Gould CM, Chica C, Via A, Gibson TJ (2008) Phospho.ELM: a database of phosphorylation sites--update 2008. *Nucleic Acids Res* 36:D240-4. <https://doi.org/10.1093/nar/gkm772>.
59. Kumar M, Gouw M, Michael S, Sámano-Sánchez H, Pancsa R, Glavina J, Diakogianni A, Valverde JA, Bukirova D, Čalyševa J, Palopoli N, Davey NE, Chemes LB, Gibson TJ (2020) ELM-the eukaryotic linear motif resource in 2020. *Nucleic Acids Res* 48(D1):D296-D306. <https://doi.org/10.1093/nar/gkz1030>.
60. Burley SK et al (2021) RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res* 49D1:D437-D451. <https://doi.org/10.1093/nar/gkaa1038>.
61. Exome Variant Server (2021) NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (<http://evs.gs.washington.edu/EVS/>) [Accessed: Oct 25, 2021].
62. O'Neill SC, Rini C, Goldsmith RE et al (2009) Distress among women receiving uninformative BRCA1/2 results: 12-month outcomes. *Psychooncology* 18:1088-1096. <https://doi.org/10.1002/pon.1467>.
63. Hamilton JG, Lobel M, Moyer A et al (2009) Emotional distress following genetic testing for hereditary breast and ovarian cancer: a meta-analytic review. *Health Psychol* 28:510-518. <https://doi.org/10.1037/a0014778>.
64. Yadav S, Couch FJ (2019) Germline genetic testing for breast cancer risk: The past, present, and future. *Am Soc Clin Oncol Educ Book* 39:61-74. https://doi.org/10.1200/EDBK_238987.
65. Wai HA, Lord J, Lyon M, Gunning A, Kelly H, Cibin P, Seaby EG, Spiers-Fitzgerald K, Lye J, Ellard S, Thomas NS, Bunyan DJ, Douglas AGL, Baralle D, Splicing and disease working group (2020) Blood RNA analysis can increase clinical diagnostic rate and resolve variants of uncertain significance. *Genet Med* 22:1005-1014. <https://doi.org/10.1038/s41436-020-0766-9>.
66. Donnez J, Dolmans MM (2016) Uterine fibroid management: from the present to the future. *Hum Reprod Update* 22:665-686. <https://doi.org/10.1093/humupd/dmw023>.
67. Segars JH, Al-Hendy A (2017) Uterine leiomyoma: new perspectives on an old disease. *Semin Reprod Med* 35:471-472. <https://doi.org/10.1055/s-0037-1606569>.
68. Beeckmans S, Kanarek L (1983) The modification with tetranitromethane of an essential tyrosine in the active site of pig fumarase. *Biochim Biophys Acta* 743:370-8. [https://doi.org/10.1016/0167-4838\(83\)90395-3](https://doi.org/10.1016/0167-4838(83)90395-3).
69. Wang S, Ramamurthy D, Tan J, Liu J, Yip J, Chua A, Yu Z, Lim TK, Lin Q, Pines O, Lehming N (2020) Post-translational modifications of fumarase regulate its enzyme activity and function in respiration and the DNA damage response. *J Mol Biol* 432:6108-6126. <https://doi.org/10.1016/j.jmb.2020.09.021>.

CAPÍTULO III

Análise de variantes no gene *RAD51C* em pacientes testadas pelo painel por *Next Generation Sequencing* (NGS) para mutações germinativas associadas ao câncer hereditário

INTRODUÇÃO

O gene *RAD51C* é um gene supressor tumoral e foi descrito em 1998 como um gene da família *RAD51*, representada por outros cinco: *RAD51B*, *RAD51C*, *RAD51D*, *XRCC2* e *XRCC3*. Está localizado distalmente no braço longo do cromossomo 17, na posição 17q21-q24, e codifica uma proteína que atua no processo de reparo do *DNA* por recombinação homóloga (Dosanjh et al, 1998; Masson et al.,2001). A recombinação homóloga desempenha um papel importante no reparo de quebras da fita dupla de *DNA* causadas por efeitos mutagênicos. O reparo do *DNA* preciso por esse mecanismo, usando a cromátide “irmã” como modelo, é necessário para a manutenção da estabilidade do genoma. Defeitos nesse processo podem levar à introdução de mutações, translocações cromossômicas, apoptose e eventos carcinogênicos (Masson et al,2001).

Segundo o Ensembl (Howe KL et al, 2021) e o GTEx (GTEx Consortium, 2015), o gene *RAD51C* possui 12 exons. Os de maior expressão tecidual são os exons 2, 3, 5, 8, 9 e 10, portanto, permanecem conservados em todos os transcritos que os apresentam. Os outros exons 1, 4, 6, 7, 11 e 12, têm um padrão de baixa expressão (Figura 2A). Inclusive, os exons 4, 6 e 7 raramente são vistos em transcritos e proteínas, estando presentes apenas nos transcritos ENST00000622327.4, ENST00000487525.5 e ENST00000425173 (Figura 2B).

O gene tem um papel importante na avaliação de risco de câncer hereditário, pois está associado ao risco de câncer de ovário (CO) e de mama (CM). Estima-se que 1% dos CO estejam associados a mutações em *RAD51C* ou *RADA51D*. Quanto ao CM, há forte evidência de associação de mutações em *RAD51C* com o padrão tumoral imunoistoquímico triplo-negativo, ou seja, que não expressam os receptores de estrógeno, progesterona e Her2 (Couch et al.,2015; Breast Cancer Association Consortium, 2022).

O risco estimado para CO em portadoras de variantes patogênicas (VP) em *RAD51C* é muito variável (*Odds ratio* de 3.4 a 15.8) e a idade de diagnóstico da doença tende a ser após 50 anos (Yang *et al.*, 2020; Cummings *et al.*, 2021). Yang *et al.*, 2020 de forma mais precisa estimou um risco cumulativo de 11% para CO e de 20% para CM em mulheres portadoras de VP. O estudo envolveu 6178 famílias, das quais 125 tinham VP em *RAD51C*. Uma consideração importante que o próprio estudo aponta é uma variação do risco dependente do histórico familiar. Para o CO o risco poderia ser tão alto quanto 32% a 36% em caso de dois familiares de primeiro grau com CO. Para o CM, considerando a mesma situação de familiares de primeiro grau afetados, o risco pode elevar-se entre 44 e 46%. O mesmo estudo ainda estima a frequência populacional de VP no gene *RAD51C*, equivalente a 0,00022 (22:100.000) (Yang *et al.*, 2020).

Em se tratando de medidas de redução de risco, a literatura reforça sobre as limitações dos níveis de evidências até o momento, mas considera-se discutível abordagem da salpingooforectomia redutora de risco entre 45 e 50 anos, ou antes, dependendo do contexto familiar. Já em relação às medidas para redução de risco de CM, não há consenso específico para guiar o manejo baseado na presença de VP, mas sim, na presença de uma história familiar positiva para câncer de mama ou ovários (NCCN, 2021).

Variantes patogênicas em *RAD51C* podem estar associadas à Anemia de Fanconi, uma doença congênita rara, de etiologia genética heterogênea. Assim, a identificação de variantes em *RAD51C*, tem impacto na abordagem clínica, além da avaliação de risco para câncer hereditário (Mehta & Ebens, 2021).

De acordo com os dados do VARSOME sobre o gene *RAD51C*, das 1211 variantes já descritas, 216 são patogênicas, 250, benignas e a grande maioria, 745, são VUS. Considerando ainda estudos sobre a análise conflitante da interpretação de variantes, publicada no estudo *PROMPT* (citado na introdução), o *RAD51C* foi um dos genes apontados com maior concentração de resultados conflitantes, assim como *CHEK2* e *ATM* (Balmaña *et al.*, 2016). Guiar condutas clínicas baseando-se em resultados de testes genéticos, principalmente para genes de penetrância moderada com variante patogênica, já é um desafio devido à carência de evidências científicas robustas. Quando se trata de resultados conflitantes, a situação fica ainda mais crítica. O

RAD51C é um gene determinante para avaliação de risco de câncer ginecológico e, por isso, foi tema deste estudo para análise aprofundada de variantes encontradas em nossa população. Portanto, o objetivo do estudo é, analisar de modo aprofundado por modelos de bioinformática (*in silico*) a classificação de variantes no gene *RAD51C* com intuito de verificar se alguma seria passível de reclassificação e, se possível, correlacionar o genótipo fenótipo, comparando com dados da literatura.

Aprovação COEP

O projeto foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais sob CAAE.01758418.0.0000.5149 e CAAE.48770621.6.0000.5149.

Materiais e Métodos

A pesquisa foi realizada em parceria com dois grupos de serviços laboratoriais da rede privada vinculados à clínicas de genética médica em Belo Horizonte (Grupo Pardini e Personal - Oncologia Personalizada) em Belo Horizonte, Minas Gerais. Dois tipos de dados foram coletados: i) variantes compartilhadas pelos médicos geneticistas assistentes dos próprios pacientes que realizaram painel *NGS* para investigação de câncer hereditário ou ii) os dados foram compartilhados de laboratórios que realizam teste genéticos para avaliação de perfil de mutações em genes associados ao câncer hereditário por painel *NGS*. Para o item i) dispunhamos de informações clínicas de dos pacientes e dados familiares de alguns. Para o item ii) os dados clínicos não foram disponibilizados. Um total de 19 variantes do gene *RAD51C* foram compartilhadas entre mutações patogênicas, provavelmente patogênicas e VUS.

Os dados das variantes com relação ao tipo, número de vezes encontrada, dados clínicos e familiares de cada indivíduo estão descritos no quadro 1 abaixo.

Quadro 1: Dados da casuística: tipo de variante, número de vezes reportada por cada serviço, classificação da variante e dados clínicos disponibilizados pelos serviços

Serviço	N	Variante	Nº de vezes encontrada	Classificação reportada	Fenótipo disponibilizado de pacientes	História familiar
Serviço 1	1	<i>RAD51C</i> c.323A>G (p.Asp108Gly)	1	VUS	mama luminal 34a	tia-avó ovário >50a
	2	<i>RAD51C</i> c.444delT (p.Phe148Leufs*23)	1	Patogênica	mama triplo neg 45a	mama, pâncreas, próstata, CCR, gástrico >70a
Serviço 2	3	<i>RAD51C</i> c.32A>G; (p.Gln11Arg)	1	VUS	CDI de mama esquerda T3N0M0 – Triplo negativo	x
	4	<i>RAD51C</i> c.571+4A>G	1	VUS	Tumor desmoide e câncer de tireoide.	
	5	<i>RAD51C</i> c.656T>C; (p.Leu219Ser)	3	VUS	- assintomática - ca de mama bilateral -ca de mama bilateral	
	6	<i>RAD51C</i> c.784T>G; (p.Leu262Val)	2	VUS	- x - câncer de mama aos 39 anos - câncer de mama idade precoce	x
Serviço 3	7	<i>RAD51C</i> c.3G>A; (p.Met1Ile)	2	VUS	x	x
	8	<i>RAD51C</i> c.170C>T; (p.Ala57Val)	1	VUS	x	x
	9	<i>RAD51C</i> c.200A>G; (p.Glu67Gly)	1	VUS	x	x
	10	<i>RAD51C</i> c.431T>C; (p.Ile144Thr)	1	VUS	x	x
	11	<i>RAD51C</i> c.440G>A; (p.Cys147Tyr)	1	VUS	x	x
	12	<i>RAD51C</i> c.492T>G. (p.Phe164Leu)	1	VUS	x	x
	13	<i>RAD51C</i> c.656T>C; (p.Leu219Ser)	1	VUS	x	x
	14	<i>RAD51C</i> c.705+3A>G	1	VUS	x	x
	15	<i>RAD51C</i> c.709C>T; (p.Arg237Ter)	1	Patogênica	x	x
	16	<i>RAD51C</i> c.779G>A; (p.Arg260Gln)	1	VUS	x	x
	17	<i>RAD51C</i> c.790G>A; (p.Gly264Ser)	1	VUS	x	x
	18	<i>RAD51C</i> c.890_899del; (p.Leu297Hisfs*2)		Patogênica		
	19	<i>RAD51C</i> c.904G>A; (p.Gly302Arg)	1	VUS	x	x
20	<i>RAD51C</i> c.1102C>T; (p.Arg368Trp)	1	VUS	x	x	

Obs: as variantes de 3 e 20 são as mesmas, por isso o total é de 19 variantes

(X) dados clínicos não disponíveis

Conforme descrito na introdução, foram utilizadas ferramentas para conduzir a reanálise das variantes em *RAD51C*. Abaixo segue a metodologia específica utilizada para esta etapa do estudo.

As variantes foram submetidas a uma série de análises subsequentes *in silico* para detalhar o mecanismo de mutação: 1) inicialmente, foi feita uma busca pelas variantes

nas bases de dados para identificar se já haviam sido descritas anteriormente; 2) para as análises seguintes foram utilizadas as sequências de codificação (CDS) (NM_058216.3 e NM_002876.4) e proteína (O43502-1 e O43502-2) no formato FASTA, obtidas nas bases de dados NCBI (National Center for Biotechnology Information - 2021) e Uniprot (Universal Protein Resource - 2021), respectivamente. Essas sequências foram comparadas por meio do Clustal Omega, com o objetivo de identificar as diferenças entre as montagens; 3) Isoformas de *RAD51C* foram identificadas no banco de dados Ensembl, sendo 18 transcritos e 13 proteínas. Essas sequências foram comparadas usando Clustal Omega; 4) A expressão tecidual foi analisada usando o programa GTEx (V8) para todos os transcritos codificados por *RAD51C*; 5) Os impactos das variantes foram avaliados com dois critérios. O primeiro em nível do mRNA, no qual as mudanças nos locais de *splicing* foram avaliadas por meio da ferramenta Human Splicing Finder (V3.1) e, pelo Mutation Taster (GRCh37, Ensembl 102 - 2021), avaliou-se a classificação quanto ao Nonsense Mediated Decay (NMD - trata-se de um mecanismo compensatório utilizado pelas células eucariotas, o qual reduz erros na expressão de genes a partir da eliminação de transcritos de RNAm que contém códons de parada prematura (Baker & Parker, 2014). O segundo critério considera a proteína, usando ferramentas que exploram o efeito das substituições de aminoácidos, principalmente na estabilidade da proteína e os impactos gerados nos domínios, *motifs* e modificações pós-síntese. *Motifs* são pequenas regiões da proteína definidas tridimensionalmente ou sequência de aminoácidos compartilhada entre diferentes proteínas, reconhecíveis que podem ter, ou não, uma função química ou biológica na proteína (Johansson *et al.*, 2012). Nesta etapa, foram utilizados os programas PolyPhen-2 (Polymorphism Phenotyping v2), SIFT (sorting intolerant from tolerant - 2017), PROVEAN (Protein Variation Effect Analyzer- V1.0), M-CAP (Mendelian Clinically Applicable Pathogenicity - v1.0), DynaMut (V1.0), DUET (), mCSM (), InterPro (EMBL-EBI 2021), neXtProt (v2.41.0), PhosphoSitePlus (v6.6.0.3) e Pfam (v35.0); 6) Ao final, foram utilizados os critérios American College of Medical Genetics (ACMG) para classificação das variantes.

Resultados

1) Busca pelas variantes nas bases de dados para identificar se já haviam sido descritas anteriormente.

Com a identificação de variantes pelo rs/ID, foi verificado sobre o *status* das 19 variantes no Clinvar, VarSome, dbSNP. Apenas uma variante, *RAD51C* c.170C>T; (p.Ala57Val), não possui rs/ID. Para estas variantes, tanto a frequência populacional quanto a localização em relação aos exons é apresentada na tabela 1. A descrição das variantes coletadas, classificação, informações fenotípicas e levantamento de dados já descritos na literatura estão apresentadas na tabela 1 abaixo.

Tabela 1. Descrição das variantes no gene *RAD51C* e levantamento de dados em relação a frequência populacional, dados clínicos e referência da literatura de cada variante em estudo

Anotação	Rs/ID - Classificação	Tipo de variante	Localização DNA	Efeito da troca de aa	Frequência*	Dados fenotípicos nos bancos de dados	Referências literatura para cada variante
<i>RAD51C</i> c.3G>A; p.Met1Ile	rs76905388 6	Variante em códon iniciador	Exon 1	Apolar/ Apolar	A=0.00003 /3:100.000	ClinVar: Fanconi anemia, complementation group O, Hereditary breast and ovarian cancer syndrome.	Não há estudo funcional disponível. Duas publicações citam esta variante, e uma a classifica como possivelmente patogênica através de algoritmos <i>in silico</i> (Cybulski <i>et al</i> , 2019). Outra publicação demonstra que a proteína iniciada na metionina seguinte (Met10) é funcional (French <i>et al</i> , 2003). Esta variante recebe o critério PVS1 devido a perda do códon de início (VarSome)
<i>RAD51C</i> c.32A>G; p.Gln11Arg	rs73088193 7 VUS	Missense	Exon 1	Polar não carregado/ Básico	G=0.000004/4:1.000.000	ClinVar: Fanconi anemia, complementation group O, Hereditary breast and ovarian cancer syndrome. Esta variante foi relatada em um indivíduo afetado com câncer de	Há uma publicação associada a esta variante que a classifica como provavelmente patogênica (Song <i>et al</i> , 2015)

ovário (PMID:
26261251).

<i>RAD51C</i> c.170C>T; p.Ala57Val	-- VUS	Missense	Exon 2	Apolar/ Apolar	Variant not found in gnomAD	Esta variante não foi relatada na literatura em indivíduos com doença relacionada ao RAD51C.	Ainda não há publicação associada a esta variante.
<i>RAD51C</i> c.200A>G; p.Glu67Gly	rs37545195 5 VUS	Missense	Exon 2	Acid/ Apolar	G=0.000032/32:1.00 0.000	Esta variante não foi relatada na literatura em indivíduos com doença relacionada ao RAD51C.	Ainda não há publicação associada a esta variante.
<i>RAD51C</i> c.323A>G; p.Asp108Gly	rs15555937 45 VUS	Missense	Exon 2	Acido/ Apolar	Variant not found in gnomAD	ClinVar: Esta variante não foi relatada na literatura em indivíduos com doença relacionada ao RAD51C.	Ainda não há publicação associada a esta variante.
<i>RAD51C</i> c.431T>C; p.Ile144Thr	rs28363307 VUS	Missense	Exon 3	Apolar/Unchar ged Polar	C=0.000048/ 48:1.000.000	Esta alteração foi relatada em pacientes com história pessoal e/ou familiar de câncer de mama e/ou ovário (PMID: 21537932, 22538716, 23117857, 25470109, 26261251, 26740214)	Há uma publicação que classificou a variante como possivelmente patogênica através de análises in silico (Jønson <i>et al</i> , 2016) Outra publicação que cita a variante como provavelmente benigna (Song <i>et al</i> , 2015); Outras publicações: (Sopik V <i>et al</i> , 2015; Yorczyk A <i>et al</i> , 2015; Kushnir A <i>et al</i> , 2012-; Loveday C <i>et al</i> , 2012.; Romero A <i>et al</i> , 2011).
<i>RAD51C</i> c.440G>A; p.Cys147Tyr	rs76417783 8 VUS	Missense	Exon 3	Apolar/ Apolar	A=0.000004/ 4:1.000.000	Esta variante não foi relatada na literatura em indivíduos com doença relacionada ao RAD51C.	Ainda não há publicação associada a esta variante.

<i>RAD51C</i> c.444delT; p.Phe148Leufs *23	rs15555946 17	Frameshi ft	Exon 3	---	Variant not found in gnomAD	HF de Câncer de mama, pâncreas, próstata, CCR, gástrico >70a.	Ainda não há publicação associada a esta variante.
<i>RAD51C</i> c.492T>G; p.Phe164Leu	rs57399210 1	Missense VUS	Exon 3	Apolar/ Apolar	G=0.000040/4:100.0 00	Esta variante foi relatada em indivíduos afetados com câncer de mama e / ou ovário, e um deles carregava uma variante BRCA1 patogênica (PMID: 28528518).	Há uma publicação associada a esta variante (Cock-Rada <i>et al</i> , 2018))
<i>RAD51C</i> c.656T>C; p.Leu219Ser	rs20152979 1	Missense VUS	Exon 5	Apolar/ Uncharged Polar	C=0.000008/ 8:1.000.000	ClinVar: Esta variante foi observada em indivíduos afetados com câncer de ovário e de mama (PMID: 22451500, 25086635).	Há publicações que descrevem a variante como possivelmente patogênica (Blanco <i>et al</i> , 2014- doi:10.1007/s10549- 014-3078-4; Sopik V <i>et al</i> , 2015- doi:10.1111/cge.1254 8; Osorio <i>et al</i> , 2012- doi:10.1093/hmg/dds1 15. Estudo funcional que não caracteriza se é patogênica, relata a redução na atividade HR (Somyajit <i>et al</i> , 2015- doi:10.1093/carcin/bg u211)
<i>RAD51C</i> c.709C>T; p.Arg237Ter	rs77063762 4	Missense Patogênica	Exon 8	Básico	T=0.000016/ 16:1.000.000	ClinVar: Esta variante foi relatada em indivíduos afetados com câncer de ovário (PMID: 25086635, 30322717, 33008098) e câncer de mama (PMID: 31300551).	Uma publicação descreve a variante como deletéria (Blanco <i>et al</i> , 2014; Cybulski <i>et al</i> , 2019)
<i>RAD51C</i> c.779G>A; p.Arg260Gln	rs73088192 6	Missense VUS	Exon 8	Básico/ Polar não carregado	A=0.000016/16:1.00 0.000		Ainda não há publicação associada a esta variante.

<i>RAD51C</i> c.784T>G; p.Leu262Val	rs14933153 7 VUS	Missense	Exon 8	Apolar/ Apolar	G=0.000056/ 56:1.000.000	ClinVar: Na literatura, esta variante também foi relatada em vários pacientes com câncer de mama e / ou ovário, câncer colorretal e câncer pancreático (Thompson_2011, Cunningham_2014, Song_2015, Jonson_2016, Yurgelun_2017, Shindo_2017, etc)	Publicações que descreve a variante como patogênica (Thompson <i>et al</i> , 2011; Jonson <i>et al</i> ., 2016), e uma publicação que cita como variante rara provavelmente benigna (Song <i>et al</i> , 2015) Outras publicações (Kushnir <i>et al</i> ., 2012; Sopik V <i>et al</i> , 2015)
<i>RAD51C</i> c.790G>A; p.Gly264Ser	rs14724170 4 VUS	Missense	Exon 8	Apolar/ Polar não carregado	A=0.001746/ 1746:1.000.000	ClinVar: Na literatura, esta variante também foi relatada em vários pacientes com câncer de mama e/ou ovário	Publicação que cita a variante como benigna (n Song <i>et al</i> , 2015; Jonhson <i>et al</i> ., 2016. Outras publicações: Loveday <i>et al</i> .,2012; Coulet <i>et al</i> ., 2012; Vuorela <i>et al</i> ., 2011; Sopik <i>et al</i> ., 2015; Meindl <i>et al</i> ,2010; Thompson <i>et al</i> , 2011; Leeneer <i>et al</i> , 2012)
<i>RAD51C</i> c.890_899del10; p.Leu297Hisfs*2	rs15556021 41 Patogênica	Frameshift	Exon 9	---	Variant not found in gnomAD	ClinVar: Na literatura, esta variante também foi relatada em vários pacientes com câncer de mama e/ou ovário	Há uma publicação que descreva a mutação como deletéria: Cybulski <i>et al</i> , 2019; Blanco <i>et al</i> ,2014)
<i>RAD51C</i> c.904G>A; p.Gly302Arg	rs15556021 58 VUS Patogênica (VarSome)	Missense	Junção do exon 9 e 10	Apolar/ Básico	Variant not found in gnomAD	Esta variante não foi relatada na literatura em indivíduos afetados com câncer hereditário.	Ainda não há publicação associada a esta variante.
<i>RAD51C</i> c.1102C>T; p.Arg368Trp	rs58778025 3 VUS	Missense	Exon 12	Básico/ Apolar	T=0.000048/ 48/1.000.000	Esta variante não foi relatada na literatura em indivíduos afetados com câncer hereditário.	Ainda não há publicação associada a esta variante.
<i>RAD51C</i> c.571+4A>G	rs58778025 7 VUS	Intron Variant	Intron 3	---	16/1.000.000	ClinVar: Essa variante foi relatada em um indivíduo afetado com câncer de mama	Estudo funcional disponível que demonstra um efeito prejudicial devido ao salto do exon 3 (Dawson <i>et al</i> , 2019)

<i>RAD51C</i> c.705+3A>G	rs87885517 9 VUS	Variante intrônica	Intro n 4	---	Variant not found in gnomAD	ClinVar: Fanconi anemia, complementati on group O, Hereditary breast and ovarian cancer syndrome. Essa variante não foi relatada na literatura em indivíduos afetados com câncer hereditário.	Ainda não há publicação associada a esta variante.
-----------------------------	------------------------	-----------------------	--------------	-----	--------------------------------	--	--

Nota: (aa) aminoácidos.

2) Identificação das diferenças entre as montagens da sequência cromossômica.

Para identificar as diferenças entre as montagens da sequência cromossômica, duas montagens foram usadas, uma anterior, GRCh37.p13 (GCF_000001405.25) e uma atual, GRCh38.p13 (GCF_000001405.39). Pelo fato de não haver diferença entre elas, foi usada a referência mais recente.

3) Identificação das isoformas de *RAD51C*

Para identificar as isoformas de *RAD51C*, foi empregado o banco de dados Ensembl. O banco de dados Ensembl identificou 18 transcritos codificados por *RAD51C* (Tabela 2), em que apenas dois são transcritos canônicos: ENST00000337432.9, que possui 9 exons e codifica a proteína RAD51C (O43502-1), em comparação aos demais transcritos, este é o mais fiel ao gene; ENST00000421782.3, que tem 2 exons e codifica a proteína RAD51C (O43502-2).

Tabela 2. Identificação dos transcritos e proteínas codificados pelo *RAD51C*

Transcritos (Ensembl)	Biotipos	ID proteína UniProt	MaxQB*
ENST00000337432.9	Codificante proteína	de O43502-1	Proteína canônica/ Proteína verdadeira
ENST00000421782.3	Codificante proteína	de O43502-2	Proteína canônica/ Proteína verdadeira
ENST00000583539.5	Codificante proteína	de J3QKK3	Proteína falsa
ENST00000413590.5	Codificante proteína	de H7C1R0	Proteína falsa
ENST00000461271.5	Codificante proteína	de J3QR58	Proteína falsa
ENST00000622327.4	Codificante proteína	de A0A087WZ35	Proteína falsa
ENST00000425173.5	Codificante proteína	de H7C2Q5	Proteína falsa
ENST00000584804.1	Codificante proteína	de J3QLQ2	Proteína falsa
ENST00000475762.5	<i>Nonsense mediated decay</i>	E9PI66	Proteína falsa
ENST00000486827.1	<i>Nonsense mediated decay</i>	E9PI66	Proteína falsa
ENST00000487525.5	<i>Nonsense mediated decay</i>	Q7KZJ0	Proteína falsa
ENST00000482007.5	<i>Nonsense mediated decay</i>	Q7KZJ0	Proteína falsa

ENST00000584617.5	<i>Nonsense mediated decay</i>	J3QLB5	Proteína falsa
ENST00000476741.2	Transcrito processado	--	--
ENST00000487921.5	Transcrito processado	--	--
ENST00000581221.5	Intron retido	--	--
ENST00000578151.1	Intron retido	--	--
ENST00000461706.1	Intron retido	--	--

Tabela adaptada do banco de dados

Ensembl (https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000108384;r=17:58692573-58735611).

*MaxQB é uma base de dados de proteômica com o objetivo de avaliar quais isoformas descritas para o *RAD51C* são realmente funcionais

<http://maxqb.biochem.mpg.de/mxdb/searchProteins/simpleSearch?proteinName=RAD51C&showResults=true>

4) Expressão tecidual dos exons do *RAD51C*

Nesta etapa, foi utilizado o programa GTEX para todos os transcritos codificados por *RAD51C* (Figura 1). Conforme evidenciado, o *RAD51C* apresenta 12 exons, sendo os exons 2, 3, 5, 8, 9 e 10 os mais expressos nos tecidos avaliados, como está ilustrado na figura 1A. Desta forma, é evidente que as variantes que situam-se entre esses exons poderiam desencadear uma supressão da expressão, por se tratarem de exons conservados. A representação das isoformas de *RAD51C* e expressão tecidual estão mostrados na figura 2.

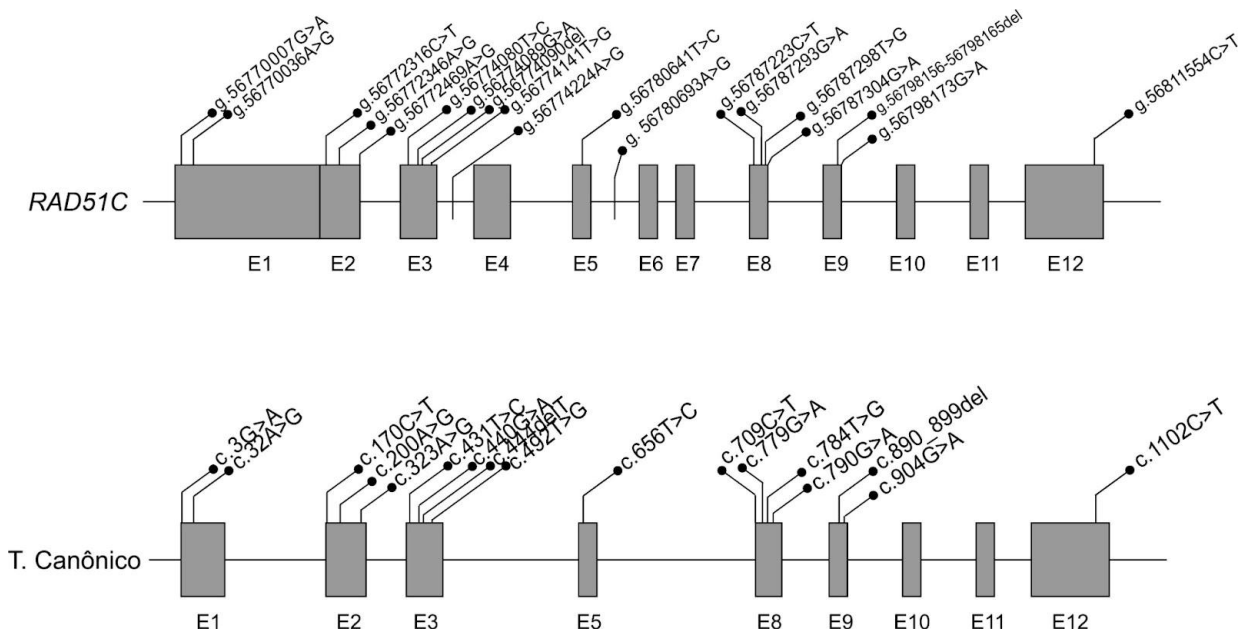
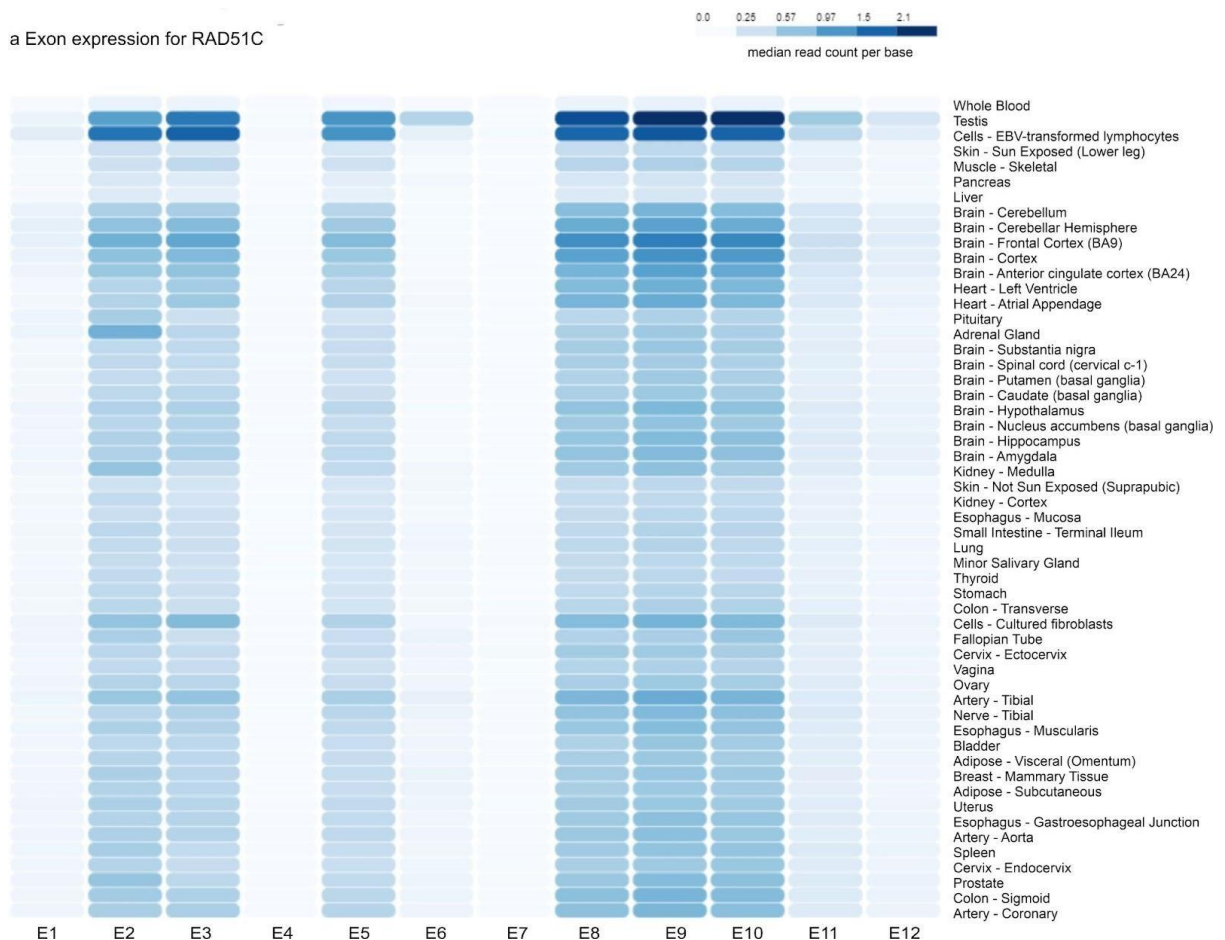


Figura 1. Representação esquemática do gene *RAD51C*, indicando a localização das variantes em relação a posição cromossômica e os exons. No esquema acima mostra a posição das variantes no transcrito canônico do *RAD51C*



b Expression of exons in RAD51C transcripts

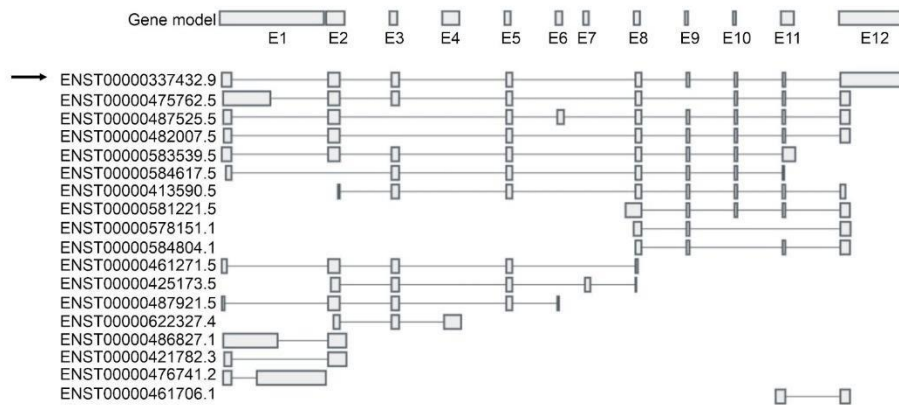


Figura2.A- Disposição da expressão dos exons do *RAD51C* em diferentes tecidos. A expressão é dada em tons gradativos de azul, sendo o mais claro baixa expressão e o mais escuro alta expressão tecidual. Em **B** - Disposição dos exons ao longo dos 18 transcritos disponíveis. A seta indica o transcrito canônico principal do *RAD51C*.

5) Avaliação do impacto das variantes

Para avaliar o nível de patogenicidade de cada variante, as análises *in silico* foram divididas em dois tipos, um em de nível do RNA mensageiro (mRNA) e outro, de proteína.

- *Avaliação do impacto das variantes no mRNA:* Foi avaliada a possibilidade de alterações nos sítios de *splicing* já ativos (*acceptor* ou *donor*), ativação de novos sítios de *splicing* e alterações nos sítios reguladores de *splicing*. Com auxílio do *Human splicing finder*, foi observado que, das 19 mutações avaliadas, apenas 12 têm impactos significativos nos sítios de *splicing*, criando e/ou abolindo sítios de *acceptors*, *donor* e *branching point* e/ou alterando os elementos regulatórios de *splicing* (tabela 3).

Tabela 3. Descrição do impacto das variantes nos sítios de *splicing* observados através *software Human splicing finder*

Mutação	Região*	Criação/abolição de sítios donor, acceptor or branch sites*,§	Criação/abolição de elementos regulatórios de splicing *,§
<i>RAD51C</i> c.3G>A	Exon 1	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.32A>G	Exon 1	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.170C>T	Exon 2	Sem impacto significativo nos sítios de splicing	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)
<i>RAD51C</i> c.200A>G	Exon 2	Sem impacto significativo nos sítios de splicing	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)
<i>RAD51C</i> c.323A>G	Exon 2	New Donor splice site/ Activation of a cryptic Donor site/ Potential alteration of splicing	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)
<i>RAD51C</i> c.431T>C	Exon 3	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.440G>A	Exon 3	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing

<i>RAD51C</i> c.444delT; p.Phe148Leufs*23	Exon 3	Cria um novo sítio de splicing <i>Acceptor</i> Cria um novo sítio de splicing <i>Donor</i>	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.492T>G	Exon 3	Cria um novo sítio de splicing <i>Donor</i>	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.656T>C	Exon 5	Sem impacto significativo nos sítios de splicing	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)
<i>RAD51C</i> c.709C>T	Exon 8	Sem impacto significativo nos sítios de splicing	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)
<i>RAD51C</i> c.779G>A	Exon 8	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.784T>G	Exon 8	Cria um novo sítio de splicing <i>Donor</i>	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.790G>A	Exon 8	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.890_899del	Exon 9	Cria um novo sítio de splicing <i>Acceptor</i> (Ativação de um exon críptico)	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)

<i>RAD51C</i> c.904G>A	Junção do exon 9 e 10	Cria um novo sítio de splicing <i>Acceptor</i>	Alteração de sequências auxiliares (elementos regulatórios de splicing - <i>enhancers</i> e <i>silencers</i>)
<i>RAD51C</i> c.1102C>T	Exon 12	Sem impacto significativo nos sítios de splicing	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.571+4A>G	Intron 4	Quebra um sítio de splicing <i>Donor</i> selvagem Cria um novo sítio de splicing <i>Donor</i>	Sem impacto significativo nos elementos regulatórios de splicing
<i>RAD51C</i> c.705+3A>G	Intron 5	Quebra um sítio de splicing <i>Donor</i> selvagem Cria um novo sítio de splicing <i>Donor</i>	Sem impacto significativo nos elementos regulatórios de splicing

Note: * considering isoform O43502-1; according to MutationTaser - Access date: December, 2021; § detected using Human Splice Finder. Access date: December, 2021.

Três variantes criaram um novo sítio *acceptor*, *RAD51C* c.444delT, *RAD51C* c.890_899del, e *RAD51C* c.904G>A

- A variante *RAD51C* c.444delT criou um novo sítio *acceptor* e um novo *donor*. Para o novo *acceptor*, há um *branching site point* localizado em menos de dez aminoácidos, tornando-o *inativo*, já o novo sítio *donor*, gera um novo exon 3 com 52 nucleotídeos. A tradução deste transcrito predito gerou uma isoforma com 199 aminoácidos, formado pelos os exons 1, 2, 3, 5 e 8, sendo o exon 3 mais curto e o exon 8 com um stop códon prematuro;

- A variante *RAD51C* c.890_899del criou um novo sítio *acceptor* no exon nove, no qual há um *branching site point* a cerca de 20 nucleotídeos, permitindo a ativação deste novo sítio, o que resulta em um exon 9 com 176 nucleotídeos, sendo 157 nucleotídeos são derivados do intron 9. Ao incluir o exon 9 alterado ao transcrito canônico, gera-se uma isoforma com 287 aminoácidos, formado pelos exons 1, 2, 3, 5, 8 e 9, em que somente oito aminoácidos são provenientes do novo exon 9, seguido de um stop códon;
- A variante *RAD51C* c.904G>A criou um *acceptor* no exon 9, uma vez que para o novo *acceptor* há um *branching site* com cerca de 53 nucleotídeos. Devido a distância em que se encontra o *branching site*, o novo *acceptor* não será ativo, e, portanto, não apresenta efeitos sobre o exon 9. Outro fenômeno que inativa o novo sítio *acceptor*, são os silenciadores de *splicing* gerados próximo à região.

Outras três variantes apresentam alterações no sítio *donor*, *RAD51C* c.323A>G, *RAD51C* c.492T>G e *RAD51C* c.784T>G, (fig 3 a, b and c).

- A variante *RAD51C* c.323A>G criou um novo *donor* no exon 2, enquanto o exon 2 selvagem tem 411 nucleotídeos, após o novo *donor*, o exon 2 passa a ter 177 nucleotídeos. A perda de 234 nucleotídeos altera a fase de leitura do transcrito e gera uma isoforma curta, com 112 aa, por criar um stop códon prematuro no final do exon 2. A nova isoforma não está descrita nas bases de dados;
- A variante *RAD51C* c.492T>G criou um novo *donor* no exon 3, conduzindo a perda de 86 nucleotídeos. Na nova isoforma, observa-se a inserção de três aminoácidos (Lys163, His164 e Arg165) e a deleção de 29 aminoácidos (SFMVDRVVDLATAACIQHLQLIAEKHKGEE). Esta nova isoforma não está descrita nas bases de dados;
- A variante *RAD51C* c.784T>G criou um novo *donor* no exon 8. O novo exon 8 selvagem tem 132 nucleotídeos, e, após o *donor* 78 nucleotídeos são perdidos,

o que também ocasiona a perda de 17 aminoácidos do exon 3. A nova isoforma gerada não está descrita nas bases de dados (Fig. 3c)

As variantes intrônicas, *RAD51C* c.705+3A>G e *RAD51C* c.571+4A, conduzem a perda do sítio *donor* ativo e gera um novo sítio *donor* (Fig. 4a e b). Das quatro variantes restantes que apresentaram impactos ao sítio de *splicing*, *RAD51C* c.170C>T, *RAD51C* c.200A>G, *RAD51C* c.656T>C e *RAD51C* c.709C>T, houve alterações quanto aos elementos regulatórios de *splicing*.

Usando o Mutation Taster, foram avaliadas as variantes que conduzem ao *nonsense mediated decay* (NMD). Ambas as variantes de *frameshifts* (*RAD51C* c.890_899del10; (p.Leu297Hisfs*2) e *RAD51C* c.444delT; (p.Phe148Leufs*23) e a variante de terminação prematura (*RAD51C* c.709C> T; (p.Arg237Ter), por gerar proteínas truncadas, podem levar ao NMD do mRNA (Tabela 4).

Tabela 4 - Avaliação das variantes quanto ao NMD

Mutação	Região*	NMD**
<i>RAD51C</i> c.444delT; (p.Phe148Leufs*23)	Exon 3	Afeta
<i>RAD51C</i> c.709C>T; (p.Arg237Ter)	Exon 8	Afeta
<i>RAD51C</i> c.890_899del; (p.Leu297Hisfs*2)	Exon 9	Afeta

Nota: De acordo com Mutation Taster. Data de acesso: Dezembro, 2021.

Para avaliar o impacto das variantes na proteína, foram identificadas duas isoformas canônicas codificadas por *RAD51C* no UniProt. A isoforma identificada por O43502-1 é a isoforma longa e a O43502-2, é a isoforma curta. A primeira isoforma, devido ao *splicing* alternativo, inicia-se no exon 1, possui 376 aminoácidos e é comumente usada como sequência de referência na análise *in silico*. Já a segunda isoforma inicia-se no

exon 1 e possui 135 aminoácidos. Na isoforma curta, os resíduos de aminoácidos 136-376 estão ausentes, resultando em 241 aminoácidos a menos do que a isoforma longa canônica. Com isso, a isoforma curta não tem o *motif* de sinal de localização nuclear (do aminoácido 366 a 370). Ambas as isoformas são semelhantes até o aminoácido 134, porém, no aminoácido 135 da isoforma canônica longa há uma cisteína, enquanto na isoforma canônica curta, há um triptofano. Todas as mutações neste estudo acontecem ao longo da isoforma longa, O43502-1.

Os preditores *in silico* usados para investigar o nível de patogenicidade das variantes, tanto no nível de substituição de aminoácidos quanto no nível de deleção e/ou ganho de códon de parada, são descritos na tabela 5. A conservação evolutiva foi avaliada utilizando o ConSurf que prediz um score de conservação evolutiva de cada resíduo de aminoácido presente na proteína, através de um múltiplo-alinhamento (tabela 5). Utilizando o, mCSM (*Protein Stability Change Upon Mutation*, *Protein-Protein Affinity Change Upon Mutation* e *Protein-DNA Affinity Change Upon Mutation*), DynaMut e DUET, previu-se os efeitos de cada variante na estrutura da proteína. Para essa análise é considerado a possível variação quanto a estabilidade da proteína e quanto a capacidade de interação da RAD51C com outras proteínas e com a molécula de DNA, após a mutação. Os programas que classificam as variantes com um valor menor que 0.001 ($p < 0.001$), consideram-na como desestabilizantes, pois diminuem a estabilidade da proteína. Outro valor de predição utilizado é o delta-delta-G ($\Delta\Delta G$), em que um valor menor que 0 (< 0) refere-se a diminuição da estabilidade e, quando maior que 0 (> 0) significa estabilidade da proteína, ou seja, sem efeitos da variante em relação a estrutura da proteína. Esses resultados são apresentados na tabela 6.

Tabela 5 - Impacto das variantes *in silico* em nível da proteína - Avaliação da conservação evolutiva, da substituição de aa e ganho de stop códon,

Mutation	Evolutionary conservation	PolyPhen	Sift	Provean	M-CAP	Mutation Taster
<i>RAD51C p. Met1Ile</i>	Aminoácido variável	Benigna	Não tolerada	Neutra	Benigna	Benigna
<i>RAD51C p.Gln11Arg</i>	Conservado	Prejudicial	Tolerada	Neutra	Possivelmente benigna	Benigna
<i>RAD51C p.Ala57Val</i>	Altamente conservado	Prejudicial	Não tolerada	Neutra	Possivelmente benigna	Deletéria
<i>RAD51C p.Glu67Gly</i>	Aminoácido variável	Benigna	Tolerada	Neutra	Benigna	Deletéria

<i>RAD51C</i> <i>p.Asp108Gly</i>	Altamente conservado	Prejudicial	Não tolerada	Deletéria	Possivelmente Patogênica	Deletéria
<i>RAD51C</i> <i>p.Ile144Thr</i>	Altamente conservado	Prejudicial	Não tolerada	Deletéria	Possivelmente benigna	Deletéria
<i>RAD51C</i> <i>p.Cys147Tyr</i>	Pouco conservado	Prejudicial Benigna	Não tolerada	Deletéria	Possivelmente benigna	Deletéria
<i>RAD51C</i> <i>p.Phe148Leufs*23</i>	---	---	---	---	---	Deletéria
<i>RAD51C</i> <i>p.Phe164Leu</i>	Altamente conservado	Prejudicial	Não tolerada	Deletéria	Possivelmente benigna	Deletéria
<i>RAD51C</i> <i>p.Leu219Ser</i>	Altamente conservado	Prejudicial	Não tolerada	Deletéria	Patogênica	Deletéria
<i>RAD51C</i> <i>p.Arg237Ter</i>	---	---	---	---	---	Deletéria
<i>RAD51C</i> <i>p.Arg260Gln</i>	Altamente conservado	Prejudicial	Não tolerada	Deletéria	Patogênica	Benigna
<i>RAD51C</i> <i>p.Leu262Val</i>	Altamente conservado	Prejudicial	Não tolerada	Neutra	Patogênica	Benigna
<i>RAD51C</i> <i>p.Gly264Ser</i>	Conservado	Benigna	Não tolerada	Deletéria	Possivelmente Benigna	Benigna
<i>RAD51C</i> <i>p.Leu297Hisfs*2</i>	---	---	---	---	---	Deletéria
<i>RAD51C</i> <i>p.Gly302Arg</i>	Altamente conservado	Prejudicial	Não tolerada	Deletéria	Patogênica	Deletéria
<i>RAD51C</i> <i>p.Arg368Trp</i>	Altamente conservado	Benigna	Não tolerada	Neutra	Patogênica	Benigna

Nota: No PolyPhen, Mutation Taster, M-Cap, os *scores* são apresentados em probabilidade da previsão; SIFT $p < 0.05$; No Provean as variantes com *scores* igual ou inferior a -2,5 são consideradas deletérias e variantes com *scores* acima de -2,5 são consideradas neutras. (---) Variantes frameshifts, para as quais os scores não se aplicam.

Tabela 6 - Impacto das variantes em nível da proteína quanto a estabilidade estrutural e sua atividade de interação com outras proteínas e molécula de DNA

Mutação	mCSM PSCUM* ($\Delta\Delta G$)	mCSM PPACUM* ($\Delta\Delta G$)	mCSM PDACUM*** ($\Delta\Delta G$)	DynaMut NMA ($\Delta\Delta G$)	DynaMut ENCoM	DynaMut Vibrationa l Entropy Energy ($\Delta\Delta S$)	DUET Predicted Stability Change ($\Delta\Delta G$)	DUET SDM- Predicted Stability Change ($\Delta\Delta G$)
<i>RAD51C</i> p.Met11e	-0.469 kcal/mol Desestabiliza	0.068 kcal/mol Estabiliza	-1.004 kcal/mol Desestabiliza	-0.540 kcal/mol Desestabiliza	0.012 kcal/mol Desestabiliza	-0.015 kcal.mol ⁻¹ Diminuiçã o da flexibilidade da molécula	-0.018 kcal/mol Desestabiliza	0.0 kcal/mol Desestabiliza

<i>RAD51C</i> p. Gln11Arg	0.152 kcal/mol Estabiliza	-0.241 kcal/mol Desestabiliza	-0.231 kcal/mol Desestabiliza	0.930 kcal/mol Estabiliza	0.593 kcal/mol Estabiliza	-0.741 kcal.mol-1.K-1 Diminuiçã o da flexibilidade da molécula	0.502 kcal/mol Estabiliza	0.63 kcal/mol Estabiliza
<i>RAD51C</i> p. Ala57Val	-0.634 Desestabiliza	-1.628 Desestabiliza	0.64 kcal/mo Estabiliza	-0.215 kcal/mol Desestabiliza	0.438 kcal/mol Desestabiliza	-0.547 kcal.mol-1.K-1 Diminuiçã o da flexibilidade da molécula	-0.509 kcal/mol Desestabiliza	-1.03 kcal/mol Desestabiliza
<i>RAD51C</i> p. Glu67Gly	-1.238 Desestabiliza	-0.761 Desestabiliza	0.693 kcal/mo Estabiliza	-1.197 kcal/mol Desestabiliza	-0.484 kcal/mol Desestabiliza	0.606 kcal.mol-1.K-1 aumento da flexibilidade da molécula	-1.47 kcal/mol Desestabiliza	-1.64 kcal/mol Desestabiliza
<i>RAD51C</i> p. Asp108Gly	-0.631 Desestabiliza	-0.79 Desestabiliza	1.711 kcal/mo Estabiliza	-0.224 kcal/mol Desestabiliza	-0.427 kcal/mol Desestabiliza	0.534 kcal.mol-1.K-1 aumento da flexibilidade da molécula	-0.573 kcal/mol Desestabiliza	-0.82 kcal/mol Desestabiliza
<i>RAD51C</i> p. Ile144Thr	-2.569 kcal/mol Altamente desestabiliza nte	-0.52 kcal/mol Desestabiliza	-0.015 kcal/mol Desestabiliza	-1.412 kcal/mol Desestabiliza	-0.132 kcal/mol Desestabiliza	0.165 kcal.mol-1.K-1 aumento da flexibilidade da molécula	-2.785 kcal/mol Desestabiliza	-2.41 kcal/mol Desestabiliza
<i>RAD51C</i> p. Cys147Tyr	-1.189 kcal/mol Desestabiliza	-1.403 Desestabiliza	-2.897 kcal/mol Altamente desestabilizante	0.767 kcal/mol Estabiliza	0.399 kcal/mol Desestabiliza	-0.499 kcal.mol-1.K-1 Diminuiçã o da flexibilidade da molécula	-0.836 kcal/mol Desestabiliza	0.15 kcal/mol Estabiliza
<i>RAD51C</i> p. Phe148Leufs*23	---	---	---	---	---	---	---	---
<i>RAD51C</i> p. Phe164Leu	-1.294 kcal/mol Desestabiliza	-0.963 kcal/mol Desestabiliza	-0.202 kcal/mol Desestabiliza	-0.870 kcal/mol Desestabiliza	-0.065 kcal/mol Desestabiliza	0.081 kcal.mol-1.K-1 aumento da flexibilidade da molécula	-1.45 kcal/mol Desestabiliza	-1.95 kcal/mol Desestabiliza
<i>RAD51C</i> p. Leu219Ser	-3.95 kcal/mol Altamente desestabiliza nte	-0.362 kcal/mol Desestabiliza	2.008 kcal/mol Altamente desestabilizante	-3.184 kcal/mol Desestabiliza	-0.747 kcal/mol Desestabiliza	0.934 kcal.mol-1.K-1 aumento da flexibilidade da molécula	-4.196 kcal/mol Desestabiliza	-4.31 kcal/mol Desestabiliza

<i>RAD51C</i> p.Arg237Ter	---	---	---	---	---	---	---	---
<i>RAD51C</i> p.Arg260Gln	-1.135 kcal/mol Desestabiliza	-0.21 kcal/mol Desestabiliza	-0.538 kcal/mol Desestabiliza	-1.406 kcal/mol Desestabiliza	-1.053 kcal/mol Desestabiliza	1.316 kcal.mol- 1.K-1 aumento da flexibilida de da molécula	-1.12 kcal/mol Desestabiliza	-1.0 kcal/mol Desestabiliza
<i>RAD51C</i> p.Leu262Val	-1.044 kcal/mol Desestabiliza	-1.38 kcal/mol Desestabiliza	-0.447 kcal/mol Desestabiliza	-1.060 kcal/mol Desestabiliza	-0.353 kcal/mol Desestabiliza	0.441 kcal.mol- 1.K-1 aumento da flexibilida de da molécula	-1.49 kcal/mol Desestabiliza	-3.21 kcal/mol Desestabiliza
<i>RAD51C</i> p.Gly264Ser	-1.722 kcal/mol Desestabiliza	-0.438 kcal/mol Desestabiliza	-0.124 kcal/mol Desestabiliza	0.993 kcal/mol Estabiliza	0.498 kcal/mol Desestabiliza	-0.622 kcal.mol- 1.K-1 Diminuiçã o da flexibilida de da molécula	-1.667 kcal/mol Desestabiliza	-1.39 kcal/mol Desestabiliza
<i>RAD51C</i> p.Leu297Hisfs*2	---	---	---	---	---	---	---	---
<i>RAD51C</i> p.Gly302Arg	-0.503 kcal/mol Desestabiliza	-0.494 kcal/mol Desestabiliza	0.152 kcal/mol Estabiliza	0.704 kcal/mol Estabiliza	0.412 kcal/mol Desestabiliza	-0.516 kcal.mol- 1.K-1 Diminuiçã o da flexibilida de da molécula	-0.572 kcal/mol Desestabiliza	-2.59 kcal/mol Desestabiliza
<i>RAD51C</i> p.Arg368Trp	-0.551 kcal/mol Desestabiliza	-0.145 kcal/mol Desestabiliza	0.116 kcal/mol Estabiliza	0.154 kcal/mol Estabiliza	0.107 kcal/mol Desestabiliza	-0.134 kcal.mol- 1.K-1 Diminuiçã o da flexibilida de da molécula	-0.524 kcal/mol Desestabiliza	0.17 kcal/mol Estabiliza

Nota: mSCSM, DUET, and DynaMut p < 0.001 = Desestabilizando. Valor de predição $\Delta\Delta G < 0$ refere-se a diminuição da estabilidade e $\Delta\Delta G > 0$ refere-se o aumento da estabilidade. mSCSM avalia a estabilidade da RAD51C e a afinidade da mesma com outras proteínas e com a molécula de DNA: *Mudança na estabilidade da proteína após a mutação, ** mudança na afinidade proteína-proteína após a mutação; ***Mudança na afinidade da proteína-DNA após a mutação; (---) Variantes frameshifts, para as quais os scores não se aplicam.

Para avaliar as variantes nos domínios, com base em InterPro e Pfam, a proteína RAD51C tem três domínios estruturais sobrepostos, sendo eles *RecA-Like C-terminal*, *RecA-Like ATP-binding* e o *RAD51/DMC1/RadA* (Figura 6). Foi identificado também um *motif* de sinal de localização nuclear na porção C-terminal da proteína (366-370 aa) (Figura 6). As variantes que são localizadas e apresentam efeitos nos domínios e no sinal de localização nuclear são mostradas na tabela 7.

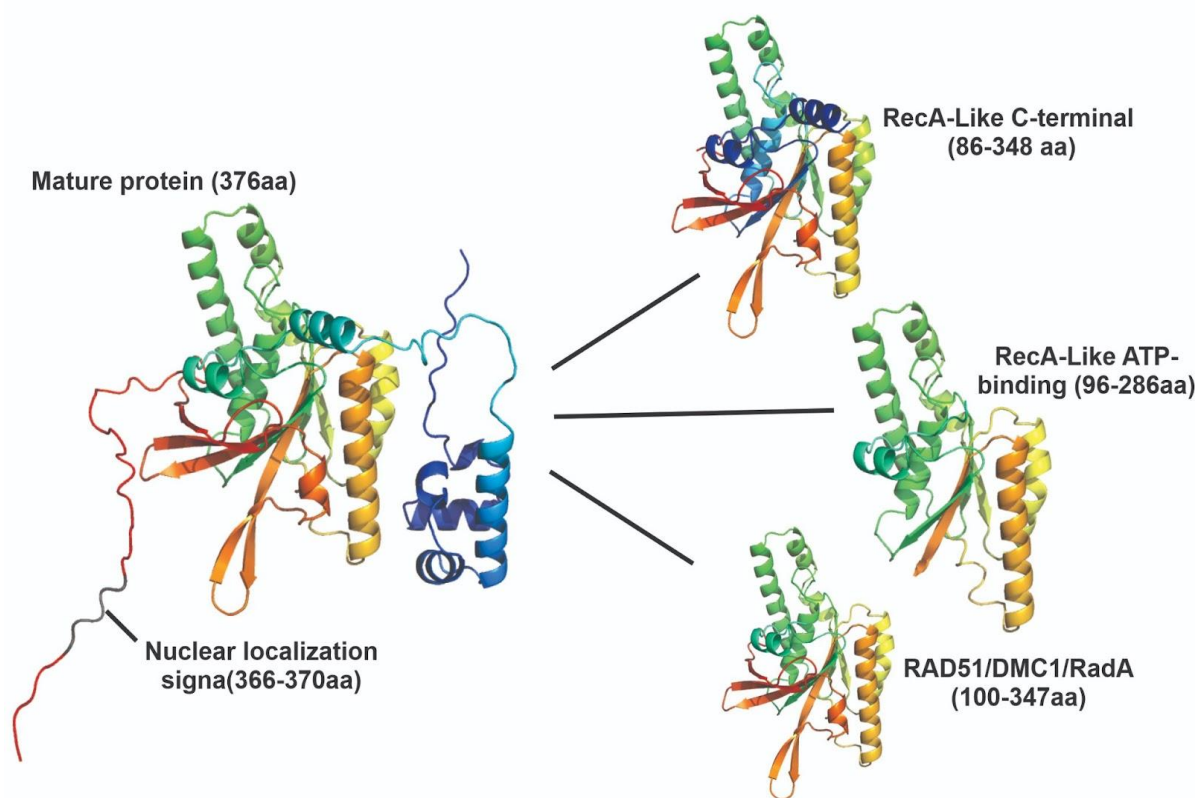


Figura 6. Demonstra a proteína madura com 376 aminoácidos, evidenciando o sinal de localização nuclear que é composto por cinco aminoácidos (366-370aa - cinza escuro). São vistos na figura os três domínios funcionais RecA-Like C-terminal (86-348aa), RecA-Like ATP binding (96-286aa) e RAD51/DMC1/RadA (100-347aa). Utilizou o PyMol para gerar a figura, através do modelo predito da proteína (AF-O43502-F1) no formato PDB.

Tabela 7. Efeitos observados em relação aos domínios funcionais, sítios de modificações pós-síntese, *motifs* e características de proteínas de cada variante em estudo

Mutação	Domínios	Modificações pós-síntese	Características da proteína e <i>motifs</i>
<i>RAD51C</i> c.3G>A (p.Met1Ile)	Não afeta	Não afeta	Está localizado na região de exigência para Junções Holliday (1-126)
<i>RAD51C</i> c.32A>G; (p.Gln11Arg)	Não afeta	Não afeta	Está localizado na região de exigência para Junções Holliday (1-126)
<i>RAD51C</i> c.170C>T; (p.Ala57Val)	Não afeta	Não afeta	Está localizado na região de exigência para Junções Holliday (1-126)

<i>RAD51C</i> c.200A>G; (p.Glu67Gly)	Não afeta	Não afeta	Está localizado na região de exigência para Junções Holliday (1-126)
<i>RAD51C</i> c.323A>G; (p.Asp108Gly)	RecA-Like C-terminal, ATP-binding, e Rad51_DMC1_RadA	Não afeta	Está localizado na região de exigência para Junções Holliday (1-126) e na região de interação da proteína RAD51C (79-136)
<i>RAD51C</i> c.431T>C; (p.Ile144Thr)	RecA-Like C-terminal, ATP-binding, e Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.440G>A; (p.Cys147Tyr)	RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.444delT (p.Phe148Leufs*23)	Perda dos domínios: RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	NMD	Devido a presença de um stop códon prematuro, os aminoácidos 170 a 376 são perdidos. Portanto, perde-se a região <i>multimer interface</i> (Fig. 3), <i>NP_Bindig motif</i> (Walker B 238-242 aa) e o sinal de localização nuclear
<i>RAD51C</i> c.492T>G (p.Phe164Leu)	RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.656T>C; (p.Leu219Ser)	RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.709C>T; (p.Arg237Ter)	Perda dos domínios: RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	NMD	O ganho de <i>stop</i> códon gera uma proteína truncada com 139 aminoácidos a menos. Além disso, perde-se a região de interação da proteína <i>multimer interface</i> (Fig. 3), <i>NP_Bindig motif</i> (Walker B 238-242aa), <i>ATP-binding site</i> (Fig. 2), e o sinal de localização nuclear
<i>RAD51C</i> c.779G>A; (p.Arg260Gln)	RecA-Like C-terminal, ATP-binding Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.784T>G; (p.Leu262Val)	RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	Não afeta	Não afeta

<i>RAD51C</i> c.790G>A; (p.Gly264Ser)	RecA-Like C-terminal, ATP-binding e Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.890_899del10; (p.Leu297Hisfs*2)	Perda dos domínios: RecA-Like C-terminal e Rad51_DMC1_RadA	NMD	Perde-se o sinal de localização nuclear
<i>RAD51C</i> c.904G>A; (p.Gly302Arg)	RecA-Like C-terminal e Rad51_DMC1_RadA	Não afeta	Não afeta
<i>RAD51C</i> c.1102C>T; (p.Arg368Trp)	Não afeta	Não afeta	Está localizada no sinal de localização nuclear (366-370 aa)

Como proposta de reclassificação das variantes pelos critérios do ACMG / AMP, após o estudo detalhado pelos métodos descritos e resultados obtidos, foi proposto uma possível reclassificação dessas variantes. Das 19 variantes analisadas, 3 VUS foram reclassificadas como possivelmente patogênicas; 13 VUS mantiveram-se como VUS e três variantes patogênicas mantiveram-se como patogênicas. Os dados das análises de reclassificações de acordo com critérios ACMG estão especificados no Quadro 3.

Quadro 3: Análise dos critérios ACMG e proposta de reclassificação das variantes. Proposta de reclassificação das variantes em análise, com base nos critérios de classificação de variantes ACMG e, na coleta de dados das variantes ao longo do estudo

Mutação	Critérios ACMG/AMP	Classificação recebida de colaboradores	Reclassificação o proposta	Observações quanto à análise
<i>RAD51C</i> c.3G>A; (p.Met11Ile)	PVS1 (moderado): Devido à perda da metionina inicial (códon iniciador), outros oito aminoácidos são perdidos, e a nova isoforma inicia-se na segunda metionina <i>in-frame</i> Met10.	VUS	Provavelmente patogênica	-PVS1 aplicável pelo Auto PVS1 como critério moderado - Duas outras variantes que abolem iniciação, reportadas como patogênica
	PM2: Frequência populacional rara (0.00003185 gnomAD); variante			

	<p>não encontrada no 1000 Genomes Browser e Exome Variant Server</p> <p>PM5: Segundo o VarSome as variantes <i>RAD51Cc.3G>T</i>; p.Met1Ile (rs769053886); <i>RAD51C c.3G>C</i>; p.Met1Ile são classificadas como patogênicas iniciação, com Fanconi ou predisposição à câncer.</p>			
<i>RAD51 C c.32A>G</i> ; (p.Gln11Arg)	<p>PM1: Além de ser uma região rica de mutações, a variante localiza-se em uma região requerida para Junções Holliday (1-126)</p> <p>PM2: Frequência populacional rara (0.000003977 gnomAD); variante não encontrada no 1000 Genomes Browser e Exome Variant Server</p> <p>BP4: “<i>Multifactorial likelihood analysis for Benign</i>” a maioria dos programas a classificam como benigna. Além disso, há uma publicação que a considera como possivelmente benigna. “<i>Missense change seen in multiple mammalian species</i>” A glutamina é conservada, porém, uma arginina nesta posição é vista em primata Sifaka-de-coquerel.</p>	VUS	VUS	Análises “in silico” realizadas apontam para características não impactantes ou deletérias do ponto de vista funcional.
<i>RAD51C c.170C>T</i> ; (p.Ala57Val)	PM2: A variante não é encontrada no gnomAD, 1000	VUS	VUS	Análises “in silico” sugerem impacto

	<p>Genomes Browser e Exome Variant Server</p> <p>PP3: Variante considerada patogênica na maioria dos preditores <i>in silico</i>. Além disso, é um aminoácido altamente conservado, somente uma espécie de planta apresentou uma valina (Alcachofra). Quando alterado, promove a desestabilização da proteína interferindo na afinidade de interação com outras proteínas.</p>			funcional, principalmente em nível de análise conservacional .
<i>RAD51C</i> c.200A>G; (p.Glu67Gly)	<p>PM2: A variante possui uma baixa frequência populacional (0.00003889 gnomAD; 0.00020 no 1000 Genomes Browser e 0.00015 Exome Variant Server)</p> <p>BP4: Multifactorial likelihood analysis for Benign” apenas o Mutation taster a considera como deletéria, por situar em uma região de requerimento para junções holliday. entretanto, os outros programas a classificam como benigna. Além disso, esse ácido glutâmico não é conservado</p>	VUS	VUS	Preditores “in silico” não sugerem impacto relevam para essa variante
<i>RAD51C</i> c.323A>G; (p.Asp108Gly)	<p>PM1: Variante localiza-se em um domínio funcional. E conforme observado, essa é uma região da proteína sem mutações benignas, nem patogênicas. Na região, tem patogênicas, mas</p>	VUS	VUS	Segundo o <i>Mutation Taster</i> ocorre a perda da região de exigência para Junções <i>Holliday</i> (1-126) e na região de

	<p>frequentemente truncantes</p> <p>PM2: A variante não é encontrada no gnomAD, 1000 Genomes Browser e Exome Variant Server</p> <p>PP3: Todos os preditores in silico utilizados neste estudo caracterizam a variante como prejudicial e com efeitos deletérios. Afeta sítios de <i>splicing</i></p>			<p>interação da proteína <i>RAD51C</i>. Embora a variante tenha sido considerada como patogênica nas análises in silico, não há evidências suficientes que comprove sua patogenicidade</p>
<p><i>RAD51C</i> c.431T>C; (p.Ile144Thr)</p>	<p>PM1: Variante localiza-se em um domínio funcional, uma região com poucas variantes benignas</p> <p>PM2: A variante possui uma baixa frequência populacional (0.00005656 gnomAD e 0.00015 Exome Variant Server); Não descrita no 1000 Genomes Browser</p> <p>PP3: Variante considerada patogênica na maioria dos preditores in silico. Além disso, é um aminoácido altamente conservado, e, quando alterado, promove a desestabilização da proteína interferindo na afinidade de interação com outras proteínas</p>	VUS	VUS	<p>Análises “in silico” sugerem impacto funcional, principalmente em nível de análise conservacional</p>
<p><i>RAD51C</i> c.440G>A; (p.Cys147Tyr)</p>	<p>PM2: A variante possui uma baixa frequência populacional (0.000003976 gnomAD); Não descrita no 1000</p>	VUS	VUS	<p>Análises “in silico” sugerem impacto funcional, principalmente em nível de análise</p>

	Genomes Browser e no Exome Variant Server			conservacional e de estabilidade de proteína
	PP3: Variante considerada patogênica na maioria dos preditores in silico. É um aminoácido pouco conservado, e, quando alterado, promove a desestabilização da proteína interferindo na afinidade de interação com outras proteínas			
RAD51C c.444del; (p.Phe148Leufs*23)	PVS1 (muito forte): É uma frameshift que conduz a deleção de multi-exons (5, 8, 9, 10, 11 e 12). Não tem 6 e 7. Gera NMD em gene com haploinsuficiência conhecida	Patogênica	Patogênica	PP4: O fenótipo da paciente é de câncer de mama triplo negativo. A história familiar inclui outros tipos de cânceres. Entretanto, não se trata de uma doença de etiologia genética única.
	PM2: A variante não é encontrada no gnomAD, 1000 Genomes Browser e Exome Variant Server			
	PM4: A variante tem efeitos significativos sobre o comprimento da proteína			
	PP3: A variante afeta os domínios, <i>motifs</i> , modificações pós-síntese e modifica (cria e abole) sítios de <i>splicing</i> e sítios regulatórios de <i>splicing</i> . Gera NMD em um gene com haploinsuficiência			
	PP5: ClinVar classifica esta variante como provavelmente patogênica, com 1 estrela. VarSome			

	classifica como patogênica.			
RAD51C c.492T>G; (p.Phe164Leu)	<p>PM2: A variante apresenta uma baixa frequência populacional (0.00003977 gnomAD e 0.00020 1000 Genomes Browser); Não está descrita no Exome Variant Server</p>	VUS	VUS	Análises <i>in silico</i> sugerem efeito deletério em diferentes níveis de eventos no processo de síntese proteica (<i>splicing</i> , interação e estabilidade proteica) e conservacional .
	<p>PP3: Variante considerada patogênica na maioria dos preditores <i>in silico</i>. A Phe164 é altamente conservada, e, quando alterada, promove a desestabilização da proteína, interferindo na afinidade de interação com outras proteínas. Além disso, situa-se entre os domínios funcionais. Em nível do DNA, há alterações quanto ao sítio de <i>splicing</i>, no qual cria-se um novo sítio <i>donor</i></p>			
RAD51C c.656T>C; (p.Leu219Ser)	<p>PM2: A variante apresenta uma frequência populacional extremamente baixa (0.000007959 gnomAD); Não descrita no 1000 Genomes Browser e Exome Variant Server</p>	VUS	Provavelmente patogênica	O fator impactante para possível reclassificação dessa variante envolve o critério PS3 associado ao estudo funcional reportado na literatura.
	<p>PS3: Estudo funcional reporta que a proteína mutada apresentou uma redução em suas atividades de recombinação homóloga. Sendo assim, é uma evidência boa de um efeito moderado em relação a perturbação de função da</p>			

	<p>proteína (Somyajit, K et al, 2014- Doi: 10.1093/carcin/bgu211</p> <p>PP3: Todos os preditores in silico utilizados neste estudo caracterizam a variante como prejudicial e com efeitos deletérios. A leu219 é altamente conservada e a mutação conduz à desestabilização da proteína interferindo na afinidade da <i>RAD51C</i> com outras proteínas. A variante localiza-se entre os domínios funcionais</p>			
<p><i>RAD51C</i> c.709C>T; (p.Arg237Ter)</p>	<p>PVS1 (Muito forte): Ganho de um códon de parada, no qual conduz a deleção de multi-exons (parte do 8, 9,10, 11 e 12)</p>	<p>Patogênica</p>	<p>Patogênica</p>	
	<p>PM2: A variante apresenta uma baixa frequência populacional (0.00001591 gnomAD); Não descrita no 1000 Genomes Browser e no Exome Variant Server</p>			
	<p>PP3: A variante conduz a perda de parte dos domínios, a perda da região de interação da proteína <i>multimer interface, motif</i> (Walker B 238-242aa), <i>ATP-binding site</i> e o sinal de localização nuclear</p>			
	<p>PP5: A variante é reportada no ClinVar e VarSome como patogênica</p>			

<p><i>RAD51C</i> c.779G>A; (p.Arg260Gln)</p>	<p>PM2: Variante com baixa frequência populacional, 0.000007955 0,000008 (8:1000000; semelhante a 1:100.000) gnomAD; Não encontrada no 1000 Genomes Browser e no Exome Variant Server</p>	<p>VUS</p>	<p>VUS</p>	<p>Análises <i>in silico</i> também sugerem efeitos deletérios</p>
<p>PP3: Aminoácido altamente conservado. Além disso, a maioria dos programas <i>in silico</i> a classificam como patogênica com efeitos deletérios. Além disso, após a mutação é evidente a perda estabilização da proteína, impactando na atividade de interação da <i>RAD51C</i> com outras moléculas</p>				
<p><i>RAD51C</i> c.784T>G; (p.Leu262Val)</p>	<p>PM2: Variante com baixa frequência populacional (0.00019 gnomAD e 0.00008 Exome Variant Server); Não descrita no 1000 Genomes Browser</p>	<p>VUS</p>	<p>VUS</p>	<p>Análises <i>in silico</i> também sugerem efeitos deletérios</p>
<p>PP3: A variante apresenta alteração quanto ao sítio de <i>splicing</i>, criando-se um novo sítio <i>donor</i>. Apesar de a Leu262 ser altamente conservada, a maioria dos preditores <i>in silico</i> classificam-a como patogênica. No entanto, a variante afeta os domínios funcionais e promove a desestabilização da proteína, interferindo na afinidade de interação com outras proteínas.</p>				

<p><i>RAD51C</i> c.790G>A; (p.Gly264Ser)</p>	<p>PM2: A variante apresenta baixa frequência populacional (0.001729 gnomAD e 0.00308 Exome Variant Server); Não descrita no 1000 Genomes Browser</p>	<p>VUS</p>	<p>VUS</p>	
<p><i>RAD51C</i> c.890_899del; (p.Leu297Hisfs*2)</p>	<p>PVS1 (Muito forte): Frameshift que conduz a deleção de multi-exons (9, 10, 11, e 12)</p>	<p>Patogênica</p>	<p>Patogênica</p>	
<p>PM2: A variante não é encontrada no gnomAD, 1000 Genomes Browser e Exome Variant Server</p>				
<p>PP3: Perde-se os domínios funcionais e o sinal de localização nuclear. Altera sítio de <i>splicing</i></p>				
<p>PP5: A variante é reportada no ClinVar e no VarSome como patogênica</p>				
<p><i>RAD51C</i> c.904G>A; (p.Gly302Arg)</p>	<p>PS1: Outra variante que altera o nucleotídeo, porém mantém a mesma troca de aminoácidos (<i>RAD51C</i> c.904G>C; p.Gly302Arg), descrita como patogênica no VarSome</p>	<p>VUS</p>	<p>Possivelmente Patogênica</p>	<p>O impacto da mutação em nível de <i>splicing</i> (PVS1) sugere afetar a guanina, que é conservada nas principais espécies de mamíferos, répteis e algumas plantas. Em nível de proteína, nenhuma arginina é vista nesta posição entre as espécies avaliadas</p>
<p>PM2: Variante não descrita no gnomAD, 1000 Genomes Browser e Exome Variant Server</p>				
<p>PP3: A variante altera o sítio de <i>splicing</i>, criando um novo sítio <i>acceptor</i>. Além disso, a Gly302 é altamente conservada, e todos preditores in silico</p>				

	<p>utilizados nesse estudo caracterizaram a variante como prejudicial e com efeitos deletérios</p> <p>PP5: Embora a variante esteja classificada como VUS no ClinVar, no VarSome ela está classificada como patogênica</p>			
<p><i>RAD51C</i> c.1102C>T; (p.Arg368Trp)</p>	<p>PM1: Variante localiza em um <i>transit peptide</i> composto por 5 aminoácidos (366-370), que sinaliza para o núcleo</p> <p>PM2: Variante com baixa frequência populacional (0.00004814 gnomAD); não descrita no 1000 Genomes Browser e Exome Variant Server</p>	VUS	VUS	
<p><i>RAD51C</i> c.571+4A>G</p>	<p>PS3: Estudo funcional disponível que demonstra um efeito prejudicial devido ao salto do exon 3 (Dawson <i>et al.</i>, 2019-PubMed: 31782267). Com o skipping do exon 3, é criada uma isoforma e criando uma fsm.</p> <p>PM2: Variante com baixa frequência populacional (0.00003183 gnomAD); não descrita no 1000 Genomes Browser e Exome Variant Server</p> <p>PP3: Durante a predição de alterações do sítio de <i>splicing</i>, a variante apresentou uma quebra de um sítio</p>	VUS	Provavelmente patogênica	<p>O estudo funcional sugere perda de função.</p> <p>O estudo classifica como patogênica considerando PP1 e PP4 devido aos dados clínicos das famílias e estudo de segregação.</p>

	<i>donor</i> e a criação de um novo sítio <i>donor</i> .			
<i>RAD51C</i> c.705+3A>G;	PM2: Variante não descrita No gnomAD, 1000 Genomes Browser e Exome Variant Server	VUS	VUS	A perda deste <i>donor</i> leva a um <i>splicing</i> alternativo, com fsm no exon 8 (previsto na análise <i>in silico</i>). Seria indicação de estudo <i>invitro</i>
	PP3: Durante a predição de alterações ao sítio de <i>splicing</i> , a variante abole um sítio <i>donor</i> e cria um novo <i>donor</i> , que se usado, provoca uma fsm			
	PP5: A variante é reportada como possivelmente patogênica no VarSome			

Discussão

A observações sobre cada variantes estão descritas no quadro acima no item observações. Ressalta-se que para algumas variantes, dependendo do critério de pontuação de acordo com o ACMG e atualizações, haveria possibilidade de reclassificação para algumas variantes e para outras, a análise proposta não teria impacto direto.

Vale ressaltar que o critério de análise de benignidade conhecido como BP1, o qual caracteriza-se por “BP1 - Variante *missense* em um gene, em que, principalmente, as variantes truncadas são conhecidas por causar doenças”, não foi adotado para esse gene. Assim como relatado acima, embora a maioria descritas no VarSome com patogênicas sejam de efeito truncado, a maioria das mutações descritas em *RAD51C* são VUS. Se tal critério fosse adotado, duas variantes seriam classificadas como provavelmente benignas: *RAD51C* c.32A>G; (p.Gln11Arg) e *RAD51C* c.200A>G; (p.Glu67Gly).

Outra questão relevante é que embora a maior parte das variantes mantivessem a classificação como VUS, foi discutido no grupo de estudo que algumas teriam critérios “*in silico*” predizendo efeitos deletérios pelos métodos utilizados para estudos neste

trabalho (Quadro 3). Entretanto, isso não quer dizer que há um real efeito de patogenicidade e mais estudos de análise molecular seriam necessários para validar os achados.

A variante *RAD51C* c.3G>A; (p.Met1Ile), recebe um critério maior de patogenicidade PVS1 conforme os critérios ACMG/AMP e atualizações quanto a esse critério (Nykamp *et al*,2017; Tayon *et al.*, 2018). O critério seria adotado não como forte, mas sim como moderado, pois não há códon de início alternativo em outros transcritos e mais de uma variante patogênicas foi encontrada na sequência à montante de acordo com descrição nos bancos de dados (Xiang *et al*,2020). Em relação a Met1 (*RAD51C* c.3G>C (p.Met1Ile), *RAD51C* c.3G>A; (p.Met1Ile), *RAD51C* c.3G>T (p.Met1Ile), *RAD51C* c.1A>T; (p.Met1Leu), *RAD51C* c.1A>G; (p.Met1Val)); e em relação a Met10 (*RAD51C* c.28del; (p.Met10Cys*6))

Conseqüentemente, essa variante pode ganhar uma classificação como “provavelmente patogênica”, mas sabe-se das possibilidades de interpretações divergentes. Assim, outros estudos seriam importantes para analisar com mais critério o efeito dessa variante como, por exemplo, estudos funcionais por sequenciamento de *RNA*. Uma limitação presente no estudo e observada para algumas variantes foi a não disponibilização de dados clínicos. Por isso,não temos como relacionar o genótipo-fenótipo especificamente.

Já na variante *RAD51C* c.170C>T; (p.Ala57Val) a classificação do critério PP3 foi baseada na avaliação em que a maioria dos modelos de predição “*in silico*” acusam-na como uma variante deletéria, mas ainda seria “conflitante” e poderia não pontuar esse critério (Nykamp *et al.*, 2017). Entretanto, não há outros critérios para reclassificar a variante e, conseqüentemente, permanece como VUS. Isso chama a atenção para mais uma questão diante da variabilidade de interpretação de examinadores. O mesmo raciocínio pode-se aplicar para a variante *RAD51C* c.200A>G; (p.Glu67Gly), na qual poderia pontuar PP3, apenas por um modelo de predição, o Mutation taster, devido ao impacto da mutação na junção *Holliday*. Entretanto, a maioria dos modelos de predição de impacto *in silico*, acusam-na como não deletéria ou tolerável. Assim, não a pontuamos como PP3 e pontuamos o BP4.

A variante *RAD51C* c.656T>C; (p.Leu219Ser) foi outra variante analisada neste estudo que é passível de reclassificação de VUS para PP. Baseando-se em reporte de estudo funcional nos genes *RAD51C* e dessa mutação especificamente como comprometedora do evento de recombinação homóloga observado “*in vitro*”, a variante ganha o critério PS3 (Somyajit *et al.*, 2015). Como é uma variante com frequência populacional extremamente baixa e todos os modelos de predição “*in silico*” a classificaram como deletéria e prejudicial, a classificação seria provavelmente patogênica (PM2+PS3+PP3). Essa variante foi observada quatro vezes nesse estudo. Três pacientes com dados fenotípicos, duas com câncer de mama bilateral e a outra paciente assintomática. Para estes casos, não temos maiores detalhes clínicos e nem se são indivíduos da mesma família. Para outro indivíduo da casuística, não dispúnhamos de dados clínicos.

A variante em região intrônica *RAD51C* c.571+4A>G também poderia ser reclassificada como PP considerando os critérios descritos no quadro. Um dado relevante para essa mutação é que ela está descrita em uma frequência populacional maior por causa do efeito fundador reportado na Ilha de Terra Nova, uma província canadense localizada ao noroeste do oceano Atlântico (Dawson *et al.*, 2020). O estudo, sugere a reclassificação dessa variante para patogênica a partir de dados de 5 pacientes afetadas com câncer de mama e ovários que apresentaram teste negativo para pesquisa de mutações em genes BRCA1, BRCA2 e outros genes de penetrância alta e moderada para câncer de mama e ovários por painel NGS. O detalhamento de dados clínicos e segregação familiar reportados no estudo, permitiram a classificação da variante como patogênica. O dado fenotípico reportado para a paciente da nossa casuística, foi de um câncer de tireóide e tumor desmóide e sem dados familiares. O *RAD51C* não está associado ao fenótipo descrito e, portanto, trata-se de um achado incidental e se a variante for reclassificada, teria impacto para uma nova abordagem em relação à redução de risco de câncer ginecológico.

A variante *RAD51C* c.904G>A; (p.Gly302Arg) é rara nas bases de dados e localiza-se próximo a um sítio *donor* (na última base do exon 9). Embora a variante conduza um enfraquecimento do sítio *donor*, não o abole, cria um *acceptor*, que não será ativado pela distância do *branching site*. Além do impacto nos sítios de *splicing*, a variante apresenta efeitos deletérios quanto a função proteica, justificando a aplicação

do critério PP3. Observou em base de dados (VarSome) outra variante (RAD51C c.904G>C; p.Gly302Arg) descrita como patogênica. A mesma, apresenta diferenças quanto a troca de nucleotídeos, mas possui a mesma substituição de aminoácidos e, portanto, aplica-se o critério PS1. Por fim, emprega-se o critério PP5 devido a variante já estar classificada como patogênica em base de dados (VarSome).

A falta de dados clínicos neste estudo é uma limitação que impossibilita a investigação mais criteriosa das variantes, pois vários critérios de benignidade e patogenicidade propostos pelo ACMG, dependem de dados fenotípicos. Quando se trata de uma doença de etiologia genética multifatorial (doença complexa) como câncer, mesmo nos casos de câncer hereditário, para o qual há uma etiologia genética monogênica associada, mas a mutação confere risco e não determina a doença, dado clínico pode ser fundamental para guiar a escolha dos critérios de classificação das variantes.

Conclusão

A partir deste estudo percebemos que embora a investigação de patogenicidade ou benignidade, através de análises computacionais sejam necessárias para a análise das variantes, para efetivar a classificação pelos critérios ACMG, é essencial considerar outros dados, como contexto clínico e populacional de cada variante. Entretanto, uma avaliação aprofundada pelos modelos *in silico* contribuiu para reclassificação de algumas variantes. Ainda assim, a análise “in silico” chama a atenção para a investigação de quais variantes merecem estudo de impacto mutacional em nível molecular.

CONSIDERAÇÕES FINAIS

A avaliação de risco para câncer hereditário é uma área de atuação crescente e aliada à medicina genômica. É voltada para pacientes oncológicos e/ou pacientes assintomáticos, que tem risco para desenvolvimento de câncer. Para lidar com os pacientes, os profissionais de saúde dessa área dependem de um conhecimento que ultrapassa a *expertise* clínica e os coloca diante de resultados com interpretações complexas. Por outro lado, os profissionais do setor laboratorial que lidam com testes genéticos dependem de dados clínicos para interpretar testes que impactam diretamente nas decisões dos pacientes e condutas médicas. Essa interseção de conhecimento e atuação depende ainda de um esforço de ambas as partes para tornar mais clara a aplicabilidade dos testes e guiar melhor as definições diagnósticas, escolha de tratamentos, estimativas de risco e as medidas de redução de risco de câncer.

Inúmeras são as ferramentas disponíveis para estudo e análise dessas variantes. Métodos de bioinformática, banco de dados, acesso e compartilhamento de informações laboratoriais e da literatura são fundamentais para classificação de variantes em genes associados ao risco de câncer ou outras doenças hereditárias. Mesmo assim, em muitas situações, não se chega a um “denominador comum” para tomada de decisão diante dos resultados.

Em se tratando de saúde da mulher e, portanto, câncer de mama e ginecológico, é preciso dar atenção especial para essa área, uma vez que o câncer de mama é altamente frequente, o câncer de ovários tem alta mortalidade e ambos têm um componente hereditário de predisposição relevante. Já, o câncer de endométrio está associado à síndrome genética de predisposição mais comum, a síndrome de Lynch. Assim, é preciso compreender de forma mais específica possível o que os resultados de testes genéticos para mutações germinativas, de fato, representam para as pacientes e também para seus familiares.

Com relação ao nosso grupo de pesquisa, recentemente, foi elaborada uma proposta apresentada ao CNPq pelo projeto de pesquisa intitulado “Patogênica ou não? Avaliação do impacto de mutações descobertas em pessoas com câncer?”. O projeto estende o estudo das variantes, além de métodos de Bioinformática usando CRISPR/Cas(9) e siRNA, possibilitando avaliação do impacto sobre a função da proteína com possibilidade de avaliação, inclusive de terapia *in vitro*.

REFERÊNCIAS BIBLIOGRÁFICAS

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, *et al.* A global reference for human genetic variation. *Nature*. 2015 Oct 1;526(7571):68-74. DOI: <https://doi.org/10.1038/nature15393>.

Abou Tayoun AN, Pesaran T, DiStefano MT, *et al.* ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI). Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat*. 2018 Nov;39(11):1517-1524.

Adzhubei IA, Schmidt S, Peshkin L, *et al.* (2010) A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248-249. DOI: <https://doi.org/10.1038/nmeth0410-248>.

Altschul SF, Gish W, Miller W, Myers EW, Lipman. Basic local alignment search tool. *J Mol Biol* 1990; 215:403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).

Ashkenazy H., Abadi S., Martz E., *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucl. Acids Res*. 2016; DOI: <https://doi.org/10.1093/nar/gkw408>

Baker KE, Parker R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol*. 2004 Jun;16(3):293-9.

Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *Nat Genet*. 2003 Mar;33 Suppl:238-44. DOI: <https://doi.org/10.1038/ng1107> .

Balmaña J, Digiovanni L, Gaddam P, *et al.* Conflicting Interpretation of Genetic Variants and Cancer Risk by Commercial Laboratories as Assessed by the Prospective Registry of Multiplex Testing. *J Clin Oncol*. 2016 Dec;34(34):4071-4078. doi: <https://doi.or/10.1200/JCO.2016.68.4316> .

Blanco A, Gutiérrez-Enríquez S, Santamariña M, *et al.* A. RAD51C germline mutations found in Spanish site-specific breast cancer and breast-ovarian cancer families. *Breast Cancer Res Treat.* 2014 Aug;147(1):133-43.

Brasil. Agência Nacional de Saúde Suplementar - ANS. Rol de Procedimentos e Eventos em Saúde 2021 - Anexo II Diretrizes de utilização para cobertura de procedimentos na saúde suplementar. Resolução Normativa - RN nº 456/2021.

Breast Cancer Association Consortium, Dorling L, Carvalho S, Allen J, *et al.* Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med.* 2021 Feb 4;384(5):428-439. doi: <https://doi.org/10.1056/NEJMoa1913948> .

Breast Cancer Association Consortium, Mavaddat N, Dorling L, Carvalho S, *et al.* Pathology of Tumors Associated With Pathogenic Germline Variants in 9 Breast Cancer Susceptibility Genes. *JAMA Oncol.* 2022 Jan 27.

Buyss SS, Sandbach JF, Gammon A, *et al.* A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. *Cancer.* 2017 May 15;123(10):1721-1730. DOI: <https://doi.org/10.1002/cncr.30498> .

Carvalho CM; Chami AM, Lopes LMS; Braga LC; Filho ALS. Germinative mutations landscape em Minas Gerais, Brazil. Patients undergone to genetic counseling for gynecological and breast cancer. E-Pôster apresentado no congresso da Sociedade Brasileira de Oncologia juntamente com a American Society Cancer Research from oct 30-31, 2020. ONLINE

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE* 7:e46688. <https://doi.org/10.1371/journal.pone.0046688>

Chung SH, Woldenberg N, Roth AR, *et al.* *BRCA* and Beyond: Comprehensive Image-rich Review of Hereditary Breast and Gynecologic Cancer Syndromes. *Radiographics.* 2020 Mar-Apr;40(2):306-325. DOI: <https://doi.org/10.1148/rg.2020190084>.

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* [Internet]. 2019;26(3):397–405. DOI: <http://dx.doi.org/10.1007/s12282-018-00938-z>

Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet*. 1991 Feb;48(2):232-42.

Cock-Rada AM, Ossa CA, Garcia HI, Gomez LR. A multi-gene panel study in hereditary breast and ovarian cancer in Colombia. *Fam Cancer*. 2018 Jan;17(1):23-30.

Consortium U (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47(D1):D506-D515. <https://doi.org/10.1093/nar/gky1049>.

Couch FJ, Hart SN, Sharma P, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. *J Clin Oncol*. 2015 Feb 1;33(4):304-11.

Coulet F, Fajac A, Colas C, et al. Soubrier F. Germline *RAD51C* mutations in ovarian cancer susceptibility. *Clin Genet*. 2013 Apr;83(4):332-6.

Cummings S, Roman SS, Saam J, Bernhisel R, Brown K, Lancaster JM, Usha L. Age of ovarian cancer diagnosis among *BRIP1*, *RAD51C*, and *RAD51D* mutation carriers identified through multi-gene panel testing. *J Ovarian Res*. 2021 Apr 29;14(1):61.

Cybulski C, Kluźniak W, Huzarski T, et al. Polish Hereditary Breast Cancer Consortium. The spectrum of mutations predisposing to familial breast cancer in Poland. *Int J Cancer*. 2019 Dec 15;145(12):3311-3320.

Davies SC. Annual Report of the Chief Medical Officer 2016, Generation Genome. London: Department of Health, 2017.

Dawson LM, Smith KN, Werdyani S, *et al.* A dominant RAD51C pathogenic splicing variant predisposes to breast and ovarian cancer in the Newfoundland population due to founder effect. *Mol Genet Genomic Med.* 2020 Feb;8(2):e1070.

De Leeneer K, Van Bockstal M, De Brouwer S, *et al.* Evaluation of RAD51C as cancer susceptibility gene in a large breast-ovarian cancer patient population referred for genetic testing. *Breast Cancer Res Treat.* 2012 May;133(1):393-8.

Desmet FO, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2019; 37: e67. <https://doi.org/10.1093/nar/gkp215>.

Dörk T, Hillemanns P, Tempfer C, Breu J, Fleisch MC. Genetic Susceptibility to Endometrial Cancer: Risk Factors and Clinical Management. *Cancers (Basel).* 2020 Aug 25;12(9):2407. DOI: <https://doi.org/10.3390/cancers12092407>.

Easton DF, Pharoah PD, Antoniou AC, *et al.* Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med.* 2015 Jun 4;372(23):2243-57. DOI: <https://doi.org/10.1056/NEJMs1501341> .

Eleje GU, Eke AC, Ezebialu IU, Ikechebelu JI, Ugwu EO, Okonkwo OO. Risk-reducing bilateral salpingo-oophorectomy in women with BRCA1 or BRCA2 mutations. *Cochrane Database Syst Rev.* 2018 Aug 24;8(8):CD012464. DOI: <https://doi.org/10.1002/14651858> .

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

Haraldsdottir S, Rafnar T, Frankel WL, *et al.* Comprehensive population-wide analysis of Lynch syndrome in Iceland reveals founder mutations in MSH6 and PMS2. *Nat Commun.* 2017 May 3; 8:14755.

Harrison SM, Dolinsky JS, Knight Johnson AE, *et al.* Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med.* 2017 Oct;19(10):1096-1104.

Henderson JT, Webber EM, Sawaya GF. Screening for Ovarian Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA.* 2018 Feb 13;319(6):595-606.

Henson JW, Resta RG (2021) ACMG Evidence of pathogenicity. In: *Diagnosis and management of hereditary cancer*, 1st edn. Elsevier, London, 2021; pp 68-69.

Henson JW, Resta RG (2021) Genetic Counseling Considerations. In: *Diagnosis and management of hereditary cancer*, 1st edn. Elsevier, London pp 40-41.

Henson JW, Resta RG. Analysis of a detected alteration. In: *Diagnosis and management of hereditary cancer*, 1st edn. Elsevier, London, 2021; 64-65.

Hu C, Hart SN, Gnanaolivu R, Huang H, *et al.* A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med.* 2021 Feb 4;384(5):440-451.

Jagadeesh KA, Wenger AM, Berger MMJ, *et al.* M-CAP eliminates a majority of variants with uncertain significance in clinical exomes at high sensitivity. *Nat Genet.* 2016; 48:1581-1586. <https://doi.org/10.1038/ng.3703>.

Jervis S, Song H, Lee A, *et al.* A risk prediction algorithm for ovarian cancer incorporating BRCA1, BRCA2, common alleles and other familial effects. *J Med Genet.* 2015 Jul;52(7):465-75.

Liu YL, Breen K, Catchings A, *et al.* Risk-Reducing Bilateral Salpingo-Oophorectomy for Ovarian Cancer: A Review and Clinical Guide for Hereditary Predisposition Genes. *JCO Oncol Pract.* 2021 Sep 28: OP2100382.

MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* 2014; 42(D1): D986-D992. <https://doi.org/10.1093/nar/gkt958>

Madeira F, Park YM, Lee J et al (2019) The EMBL-EBI search and sequence analysis tools APIs. *Nucleic Acids.* 2019; Res 47(W1): W636-W641. <https://doi.org/10.1093/nar/gkz268> .

Manchanda R, Legood R, Antoniou AC, Gordeev VS, Menon U. Specifying the ovarian cancer risk threshold of 'premenopausal risk-reducing salpingo-oophorectomy' for ovarian cancer prevention: a cost-effectiveness analysis. *J Med Genet.* 2016 Sep;53(9):591-9.

Mistry J, Chuguransky S, Williams L, *et al.* Pfam The protein families database in 2021. *Nucleic Acids Res.* 2021; 49(D1):D412-D419. <https://doi.org/10.1093/nar/gkaa913> .

Narod SA. Which Genes for Hereditary Breast Cancer? *N Engl J Med.* 2021 Feb 4;384(5):471-473.

National Cancer Comprehensive Network Clinical Practice Guidelines in Oncology: genetic/familial high risk assessment: colorectal. Version 1.2022. NCCN Guidelines for Genetic/Familial High-Risk Assessment. Disponível em: [Guidelines Detail \(nccn.org\)](#) - Acesso em dezembro de 2021.

Newman B, Austin MA, Lee M, King MC. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A.* 1988 May;85(9):3044-8.

Nykamp K, Anderson M, Powers M, *et al.* Westbrook M; Invitae Clinical Genomics Group, Topper S. Sherlock: a comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med.* 2017 Oct;19(10):1105-1117. doi: <https://10.1038/gim.2017.37> . Erratum in: *Genet Med.* 2020 Jan;22(1):240-242.

[Observatório Global do Câncer \(iarc.fr\)](http://gco.iarc.fr). Disponível em ><http://gco.iarc.fr><. Acesso em 26/12/2021

Oliveira ES, Soares BL, Lemos S, *et al*. Screening of the BRCA1 gene in Brazilian patients with breast and/or ovarian cancer via high-resolution melting reaction analysis. *Fam Cancer*. 2016;15(2):173–81.

Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), {december,2021}. World Wide Web URL: <https://omim.org/>

Palmero EI, Carraro DM, Alemar B, *et al*. The germline mutational landscape of BRCA1 and BRCA2 in Brazil. *Sci Rep*. 2018;8(1):1–10.

Patch C, Middleton A. Genetic counseling in the era of genomic medicine. *Br Med Bull*. 2018 Jun 1;126(1):27-36.

Paulino E, de Melo AC, Silva-Filho AL, Maciel LF, Thuler LCS, Goss P, Nogueira-Rodrigues A. Panorama of Gynecologic Cancer in Brazil. *JCO Glob Oncol*. 2020 Oct; 6:1617-1630.

Pepin MG, Murray ML, Bailey S, *et al*. The challenge of comprehensive and consistent sequence variant interpretation between clinical laboratories. *Genet Med*. 2016 Jan;18(1):20-4. doi: 10.1038/gim.2015.31.

Randall LM, Pothuri B. The Genetic Prediction of Risk for Gynecologic Cancer. *Gynecol Oncol*. 2016 Abr;141(1):10-6

Richards S, Aziz N, Bale S *et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405423. <https://doi.org/10.1038/gim.2015.30>.

Siegel RL, Miller KD, Jemal A. Cancer statistics. CA Cancer J Clin. 2020 Jan;70(1):7-30.

Steinhaus R, Proft S, Schuelke M, Cooper DN, Schwarz JM, Seelow D MutationTaster. Nucleic Acids Res. 2021; 49(W1):W446-W451. DOI: <https://doi.org/10.1093/nar/gkab266>.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, *et al.* The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinform 2016; 54:1.30. 31-31.30. 33. <https://doi.org/10.1002/cpbi.5>.

Stratton JF, Pharoah P, Smith SK, Easton D, Ponder BA. A systematic review and meta-analysis of family history and risk of ovarian cancer. Br J Obstet Gynaecol. 1998 May;105(5):493-9.

The European Bioinformatics Institute - EMBL-EBI (2020) European Variation Archive. <https://www.ebi.ac.uk/eva/>. Last edited: Feb. 25 2020. Acesso em outubro de 2021.

Toss A, Tomasello C, Razzaboni E, *et al.* Hereditary ovarian cancer: not only BRCA 1 and 2 genes. Biomed Res Int. 2015;2015:341723.

Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC (2016) SIFT missense predictions for genomes. Nat Protocols 11:1-9. <https://doi.org/10.1038/nprot.2015.123>.

Webb B, Sali A (2016) Comparative Protein Structure Modeling Using Modeller. Curr Protoc Bioinform 54: 5.6.1-5.6.37. <https://doi.org/10.1002/cpbi.3>.

Yehia L and Charis Eng. PTEN Hamartoma Tumor Syndrome. G. [PTEN Hamartoma Tumor Syndrome - GeneReviews® - NCBI Bookshelf \(nih.gov\)](#). Acesso em 28/02/2022

Zhu K, Day T, Warshaviak D, Murrett C, Friesner R, Pearlman D. Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Protein* 2014; 82:1646-1655. <https://doi.org/10.1002/prot.24551>.

ANEXO I - PARECER DE APROVAÇÃO NO COMITÊ DE ÉTICA EM PESQUISA EM SERES HUMANOS



MINISTÉRIO DA SAÚDE - Conselho Nacional de Saúde - Comissão Nacional de Ética em Pesquisas - CONEP
PROJETO DE PESQUISA ENVOLVENDO SERES HUMANOS

Projeto de Pesquisa:
Perfil de mutações germinativas em pacientes submetidas ao aconselhamento genético para câncer hereditário de mama, ovário e endométrio, em Minas Gerais, Brasil

Informações Preliminares

Responsável Principal

CPF/Documento: 814.686.366-34	Nome: Agnaldo Lopes da Silva Filho
Telefone: 3132939373	E-mail: agnaldo.ufmg@gmail.com

Instituição Proponente

CNPJ:	Nome da Instituição: Faculdade de Medicina da UFMG
-------	--

É um estudo internacional? Não

Assistentes

CPF/Documento	Nome
056.411.156-22	CAMILA MARTINS DE CARVALHO

Equipe de Pesquisa

CPF/Documento	Nome
035.400.556-14	LAURENCE RODRIGUES DO AMARAL
977.664.456-20	Luciana Maria Silva
056.411.156-22	CAMILA MARTINS DE CARVALHO
864.990.846-20	Letícia da Conceição Rezza
014.083.496-63	Matheus de Souza Gomes
055.110.156-36	ANISSE MARQUES CHAMI

Área de Estudo

Grande Área do Conhecimento (GNPq)

● Grande Área 4. Ciências da Saúde

Propósito Principal do Estudo (OMS)

● Ciências Básicas

Título Público da Pesquisa: Perfil de mutações germinativas em pacientes submetidas ao aconselhamento genético para câncer hereditário de mama, ovário e endométrio, em Minas Gerais, Brasil

Contato Científico: ANISSE MARQUES CHAMI

Desenho de Estudo / Apoio Financeiro

Desenho:

Avaliar o perfil de mutações germinativas em pacientes com câncer de mama, ovário e endométrio submetidas a aconselhamento genético para câncer hereditário, bem como associar a presença dessas mutações com as características clínicas das pacientes e o tipo de tumor, caracterizar as variantes genéticas detectadas nos genes entre as pacientes com câncer ginecológico hereditário e comparar os dados obtidos com dados já reportados na literatura da população brasileira e "The Cancer Genome Atlas".

Apoio Financeiro

CNPJ	Nome	E-mail	Telefone	Tipo
				Financiamento Próprio

Palavra Chave

Palavra-chave
MMR
MLH1
MSH2
MSH6
EPCAM
PMS2
neoplasia de mama
neoplasia de endométrio
neoplasia de ovário
síndromes neoplásicas hereditárias
BRCA1
BRCA2
TP53
PTEN
CDH1
ATM
BARD1
CHEK2

ATR
PALB2
BRIP1
RAD51
FANC

Detalhamento do Estudo

Resumo:

A maioria dos cânceres ginecológicos são esporádicos, mas cerca de 5 a 10% dos casos tem um padrão hereditário. Destes, grande parte apresentam uma mutação deletéria em um dos genes BRCA. Mutações em BRCA1 conferem um risco de 39 a 46% de câncer de ovário e 65 a 85% de câncer de mama; em BRCA2, o risco é de 10 a 27% para ovário e 45 a 85% para mama. Com menor frequência, o risco para estas neoplasias pode estar associado a outras síndromes de predisposição ao câncer hereditário como Síndrome de Li-Fraumeni, Síndrome de Lynch, Síndrome de Cowden. O entendimento da predisposição genética ao câncer leva a uma melhor identificação dos pacientes em risco e, assim, ginecologistas, mastologistas, oncologistas e geneticistas, poderão coordenar em conjunto estratégias para detecção, manejo e medidas preventivas contra a doença. O objetivo deste estudo é avaliar o perfil das mutações germinativas presentes em pacientes que foram submetidas a aconselhamento genético para avaliação de risco para o câncer de mama e ginecológico com possível padrão hereditário. Serão analisados os prontuários de todos os pacientes que realizaram o aconselhamento genético, a fim de avaliar o desfecho diagnóstico diante do resultado do teste genético solicitado. Este estudo permitirá um aprofundamento sobre a compreensão das principais mutações identificadas nas famílias no Estado de Minas Gerais e contribuir com os esforços dos estudos populacionais para avaliar o perfil de mutações na população brasileira.

Introdução:

O câncer é uma das maiores causas de morte no mundo e sua causa básica é o dano a genes dos sistemas regulatórios da célula chamados oncogenes e genes supressores de tumor. A ativação dos oncogenes pode ocorrer através de translocações, ampliações ou mutações intragênicas, que levam à desregulação do controle do ciclo celular e, assim, promovem o crescimento e proliferação desordenada. Os genes supressores de tumor bloqueiam a proliferação celular descontrolada que pode levar ao câncer. A perda destes genes através de deleções, inativações ou silenciamento epigenético faz com que o produto gênico reduza sua atividade e, assim, predisponha à tumorigênese. Instabilidade genômica pode ocorrer em razão de defeitos nas proteínas necessárias para a divisão celular precisa ou em proteínas responsáveis pelo reparo do DNA (Vogelstein & Kinzler, 2004). Mutações germinativas nesses genes causam predisposição ao câncer e não a doença em si. Os indivíduos afetados, em geral, apresentam múltiplos tumores e doença em idade precoce. Em quase todas as síndromes hereditárias dominantes causadas por genes supressores de tumor e de estabilidade, a primeira mutação somática afeta os genes normais herdados do parente não afetado. As formas mais comuns de predisposição ao câncer hereditário, como de cólon e mama, ocorrem devido a mutações em genes de estabilidade, ao invés de supressores de tumor ou oncogenes (Vogelstein & Kinzler, 2004). O câncer de mama (CM) é a neoplasia com maior incidência nas mulheres, excluindo-se os casos de pele não melanoma, representando 25% de todos os casos de câncer em 2012, com 1,7 milhões de novos casos naquele ano. O CM é a causa mais frequente de morte pela doença em mulheres, nos países subdesenvolvidos. (WORLD HEALTH ORGANIZATION). O câncer de ovário (CO) é a neoplasia maligna ginecológica mais difícil de ser diagnosticada e com menor chance de cura. A detecção em fase tardia, falta de técnicas de rastreamento eficaz e resistência à quimioterapia contribuem para as altas taxas de mortalidade por CO (Ohman; Hasan; Dinulesco, 2014). Tal doença foi responsável por 239 mil casos e 152 mil mortes em 2012, sendo o sétimo tipo de câncer mais comum e citava causa de morte por câncer em mulheres (3,6% dos casos e 4,3% de mortes). O câncer de endométrio (CE) é o sexto câncer mais comum em mulheres, com uma estimativa mundial de 320 mil novos casos e 76 mil óbitos, em 2012 (Ferlay et al., 2014). O CE é a malignidade dos órgãos reprodutores mais comum nos Estados Unidos e o seu pico de incidência ocorre na sexta década de vida (American Cancer Society). A maioria das pacientes são diagnosticadas com a doença confinada ao útero e possuem mais de 90% de sobrevida em 5 anos (Kimura T et al., 2004). A estimativa de novos casos das neoplasias estudadas, no Brasil, para o biênio 2018-2019 é: 6.150 mulheres com CO, por ano, com um risco estimado de 5,79 casos a cada 100 mil mulheres; 59.700 novos casos de CM, por ano, com um risco estimado de 56,33 casos a cada 100 mil mulheres; 6.600 mulheres com CE, por ano, com um risco estimado de 6,22 casos a cada 100 mil mulheres, segundo o Instituto Nacional do Câncer José Alencar Gomes da Silva (INCA). Em Minas Gerais, são esperados 560 novos casos de CO, 5360 de CM e 510 de CE, em 2018, segundo este mesmo Instituto. Em 2015, 3.536 mulheres morreram por CO, 15.593 por CM e 3.604 por CE, de acordo com o Sistema de Informação sobre Mortalidade, do Ministério da Saúde. Apesar da maioria dos casos de neoplasia maligna de mama, ovário e endométrio serem esporádicos, grande parte das pacientes com câncer hereditário apresentam uma mutação deletéria em um dos genes BRCA. Mutações germinativas em BRCA1/2 correspondem a maioria dos casos de câncer de ovário hereditário, podendo corresponder a 18% de todos os CO. Essas mutações são herdadas de forma autossômica dominante e seus portadores possuem risco de 80% de CM ao longo da vida. (Ring & Modesitt, 2018) A perda da função do BRCA1 está associada ao câncer de mama e ovário em idade precoce e, em geral, apresenta tumores de alto grau histológico com padrão imunohistoquímico triplo negativo. Os tumores associados à BRCA2 usualmente são de alto grau, receptor de estrogênio positivo e HER2 negativo. Carreadores de mutação BRCA apresentaram sobrevida maior nos primeiros anos após o diagnóstico de câncer de mama e ovário, o que pode refletir a maior sensibilidade à quimioterapia, principalmente, às platinas (Copson, Maishman et al, 2018). Em 2014, a Sociedade de Ginecologia Oncológica divulgou uma recomendação que orienta oferecer teste genético para mutações germinativas em BRCA 1 e 2 para todas as pacientes diagnosticadas com câncer de ovário, tuba uterina e peritoneal. Essa recomendação foi endossada pelo National Comprehensive Cancer Network (NCCN). A síndrome de Lynch (SL) é herdada de forma autossômica dominante e está associada a mutações em 1 dos 4 genes MMR (MSH2, MLH1, MSH6 e PMS2) ou no gene EPCAM, que é o regulador do MSH2. Os principais cânceres relacionados à SL são cólon, endométrio, ovários e estômago. Cerca de 15% dos cânceres de ovário e 2 a 6% dos de endométrio são causados pela SL (Ring, Garcia, et al., 2017; Leenen, van Lier, et al., 2012). Alguns estudos sugerem risco aumentado de câncer de mama nos portadores (Walsh, Buchanan, et al, 2010; Buerki, Gautier, et al., 2012). A síndrome Li-Fraumeni (SLF) está associada a mutações germinativas no gene supressor de tumor TP53 e seus carreadores possuem alto risco de desenvolver múltiplos cânceres primários, como mama, ovário, sarcomas, cérebro, adrenocortical e leucemias. O risco de CM em pacientes mutadas é de 50% aos 60 anos (Masciarì, Dillon, et al., 2012). Mutações no gene STK11 estão associadas à síndrome Peutz-Jeghers (SPJ), uma rara doença que eleva o risco de câncer do sistema gastrointestinal, mama e ovário (do tipo não epitelial). O risco ao longo da vida de CM é de, aproximadamente, 55%, com diagnóstico em idade média aos 37 anos, e CO 21%, com idade média de 28 anos (Beggs, Latchford, et al., 2010; Giardiello, Brensinger, et al., 2000). A síndrome de Cowden (SC) é o tipo mais comum da Síndrome dos Hamartomas Múltiplos e ocorre devido à mutação no gene supressor de tumor PTEN. A SC confere um de risco de até 85,2% para câncer de mama e 30% de endométrio, segundo estudo prospectivo realizado com 3.399 pacientes, por Tan e colaboradores. Além desses, pacientes com SC apresentam risco elevado de câncer colorretal, rim, tireoide e melanoma. (Tan, Mester, et al., 2012). Mutações no gene CDH1 aumentam risco de câncer de mama lobular, com risco acumulado de 60%. A maioria dos portadores da mutação desenvolvem câncer antes dos 40 anos (Kaurah & Huntsman, 2002). Em um estudo caso-controle com mais de 65.000 pacientes com câncer de mama submetidas a teste genético com painel multigene, após exclusão de BRCA1/2 e genes relacionados à síndrome do câncer de mama (CDH1, PTEN, TP53), foram encontradas variantes genéticas que aumentam o risco desta

neoplasia: variantes PALB2 foram associadas ao maior risco; variantes em CHEK2, ATM, BARD1, RAD51, risco moderado (Couch, Shimelis, et al., 2017). Mutações em RAD51C, RAD51D aumentam o risco de câncer de ovário. Além deles, outros genes estão em investigação para risco aumentado de CO, como PALB2 e BARD1 (Ring, Garcia, et al., 2017). O conhecimento acerca da genética do câncer está em crescimento, uma vez que as informações sobre novos genes e os já conhecidos são associadas à cânceres hereditários. Há um aumento nos achados de variantes genéticas de significado indeterminado com os testes de painéis multigenes (Yurgelun et al., 2015), assim a realização de pesquisas de registro deve ser encorajada, para tentar entender o significado dessas variantes e correlaciona-las ao câncer. As predisposições herdadas são significativamente determinantes no risco de um indivíduo desenvolver câncer, assim, a pesquisa de mutações genéticas permite identificar pacientes e familiares nos quais o risco é elevado. A identificação da predisposição de um câncer hereditário, associado a mutação genética, é de suma importância na saúde pública, pois muda o manejo do risco da doença e permite que os pacientes e seus familiares em risco se beneficiem de medidas voltadas para o diagnóstico precoce e intervenção, condutas profiláticas voltadas para redução considerável do risco para o câncer e o prognóstico seja melhor para esses indivíduos, com diminuição da morbidade e mortalidade. Além disso, em casos específicos, pode possibilitar tratamentos oncológicos específicos (Garnis, Buys, et al., 2004). Neste contexto, o aconselhamento genético assume importância primordial no processo de ajudar os pacientes a entender e adaptar-se às implicações médicas, psicológicas e familiares das contribuições genéticas para as doenças. Esse processo consiste em interpretar a história médica e familiar, para avaliar a chance de ocorrência ou recorrência da doença, educar a respeito da sua herança, testes, manejo e prevenção, além de aconselhar, para promover escolhas e adaptação do indivíduo e de seus familiares à doença ou ao seu risco (Resta, Biesecker, et al., 2006).

Hipótese:

A identificação de mutações genéticas associadas ao câncer apresenta grande relevância clínica pela possibilidade do desenvolvimento de tratamento personalizado para as malignidades ginecológicas e, assim, melhorar a resposta terapêutica e prognóstico das mulheres afetadas com consequente aumento da sobrevida destas pacientes.

Objetivo Primário:

Avaliar o perfil de mutações germinativas em pacientes com câncer de mama, ovário e endométrio submetidas a aconselhamento genético para câncer hereditário.

Metodologia Proposta:

Serão analisados os prontuários de todos pacientes que realizaram o aconselhamento genético, devido à suspeita de câncer hereditário. Para a composição da amostra deste estudo serão selecionados 300 pacientes que apresentaram história pessoal ou familiar de câncer de mama, ovário, endométrio e outros tipos de câncer associado a síndromes hereditárias.

Critério de Inclusão:

- 1- Pacientes com diagnóstico de câncer de mama, ovário, endométrio ou outros tipos de câncer hereditário, acima de 18 anos, que realizaram aconselhamento genético e que podem ter realizado teste para pesquisa de mutação genética associada à predisposição ao câncer hereditário
- 2- Pacientes assintomáticos com história familiar de câncer de mama, ovário, endométrio ou outros tipos de câncer hereditário, acima de 18 anos, que realizaram aconselhamento genético e que podem ter realizado teste para pesquisa de mutação genética associada à predisposição ao câncer hereditário
- 3- Termo de consentimento livre e esclarecido assinado.

Critério de Exclusão:

- 1- Pacientes cujos prontuários apresentam dados incompletos.

Riscos:

Os riscos decorrentes da participação nesta pesquisa são mínimos, pois podem decorrer de possível constrangimento dos participantes com a divulgação de dados confidenciais, invasão de privacidade e segurança dos prontuários. Para mimizar estes riscos, o pesquisador irá tratar a identidade dos participantes com padrões profissionais de sigilo, divulgação do nome ou qualquer material que indique o participante sem sua permissão.

Benefícios:

Não haverá benefício direto para as pacientes, exceto nos casos em que forem identificadas mutações genéticas que já estão associadas, na literatura, a síndromes neoplásicas, pois muda o manejo do risco da doença e permite que os pacientes e seus familiares em risco se beneficiem de medidas voltadas para o diagnóstico precoce e intervenção, condutas profiláticas voltadas para redução considerável do risco para o câncer e o prognóstico seja melhor para esses indivíduos, com diminuição da morbidade e mortalidade.

Metodologia de Análise de Dados:

Os dados serão armazenados no programa Excel do Microsoft Office e analisados por meio de ferramentas de mineração de dados e estatística apropriadas, a fim de identificar os tipos de mutação presentes na população em estudo, determinar a prevalência de cada uma, avaliar os desfechos clínicos determinados e comparar os dados obtidos com dados da população em geral e a literatura internacional, como The Cancer Genome Atlas.

Desfecho Primário:

Os resultados obtidos permitirão avaliar o perfil das mutações germinativas presentes nos pacientes e um aprofundamento sobre a compreensão das principais mutações identificadas nas famílias em Minas Gerais, além de contribuir com os esforços dos estudos populacionais para avaliar o perfil de mutações na população brasileira.

Tamanho da Amostra no Brasil: 300

Países de Recrutamento

País de Origem do Estudo	País	Nº de participantes da pesquisa
Sim	BRASIL	300

Outras Informações

Haverá uso de fontes secundárias de dados (prontuários, dados demográficos, etc)?

Sim

Detalhamento:

Serão analisados os prontuários de todos pacientes que realizaram o aconselhamento genético, devido à suspeita de câncer hereditário, além da análise dos dados já obtidos na população em geral na literatura internacional, como The Cancer Genome Atlas.

Informe o número de indivíduos abordados pessoalmente, recrutados, ou que sofrerão algum tipo de intervenção neste centro de pesquisa:

300

Grupos em que serão divididos os participantes da pesquisa neste centro

ID Grupo	Nº de Indivíduos	Intervenções a serem realizadas
Pacientes com câncer de mama	100	nenhuma
Pacientes com câncer de ovário	10	nenhuma
Pacientes com câncer de endométrio	10	nenhuma
Pacientes com história familiar de CA hereditário	100	nenhuma
Pacientes com outros tipo de câncer	40	nenhuma
Pacientes com história familiar de câncer	40	nenhuma

O Estudo é Multicêntrico no Brasil?

Não

Centros Coparticipantes

CNPJ	Nome da Instituição Co-participante	Nome do Responsável	Nome do Comitê de Ética	Instituição Selecionada Via Plataforma Brasil
19.378.769/0001-76	INSTITUTO HERMES PARDINI	Anisse Marques Chami Ferraz		Sim

Propõe dispensa do TCLE?

Não

Haverá retenção de amostras para armazenamento em banco?

Não

Cronograma de Execução

Identificação da Etapa	Início (DD/MM/AAAA)	Término (DD/MM/AAAA)
Análise dos dados	01/03/2019	01/09/2019
Redação de manuscritos, artigos e/ou patentes	01/03/2019	01/09/2019
Revisão da literatura	01/09/2018	01/09/2019
Seleção dos prontuários	01/03/2019	01/09/2019
Coleta dos dados dos prontuários selecionados	01/03/2019	01/09/2019
Submissão do projeto ao COEP e aprovação	01/09/2018	01/03/2019

Orçamento Financeiro

Identificação de Orçamento	Tipo	Valor em Reais (R\$)
Transporte para coleta de dados	Custeio	R\$ 400,00
Material de escritório	Custeio	R\$ 1.428,00
Total em R\$		R\$ 1.828,00

Bibliografia:

American Cancer Society Beggs, A. D., Latchford, A. R., Vasen, H. F. A., Moslein, G., Alonso, A., Aretz, S., ... Hodgson, S. V. (2010). Peutz-Jeghers syndrome: a systematic review and recommendations for management. *Gut*, 59(7), 975–986. <https://doi.org/10.1136/gut.2009.198499> Buerki, N., Gautier, L., Kovac, M., Marra, G., Buser, M., Mueller, H., & Heinimann, K. (2012). Evidence for breast cancer as an integral part of lynch syndrome. *Genes, Chromosomes and Cancer*, 51(1), 83–91. <https://doi.org/10.1002/gcc.20935> Copson, E. R., Maishman, T. C., Tapper, W. J., Cutress, R. I., Greville-Heygate, S., Altman, D. G., ... Eccles, D. M. (2018). Germline BRCA mutation and outcome in young-onset breast cancer (POSH): a prospective cohort study. *The Lancet. Oncology*, 19(2), 169–180. [https://doi.org/10.1016/S1470-2045\(17\)30891-4](https://doi.org/10.1016/S1470-2045(17)30891-4) Couch, F. J., Shimelis, H., Hu, C., Hart, S. N., Polley, E. C., Na, J., ... Dolinsky, J. S. (2017).

Data de Submissão do Projeto: 06/01/2019

Nome do Arquivo: PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1190397.pdf

Versão do Projeto: 2

Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. *JAMA Oncology*, 3(9), 1190. <https://doi.org/10.1001/jamaoncol.2017.0424>

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., ... Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5), E359–E386. <https://doi.org/10.1002/ijc.29210>

Garnis, C., Buys, T. P., & Lam, W. L. (2004). Genetic alteration and gene expression modulation during cancer progression. *Molecular Cancer*, 3(1), 9. <https://doi.org/10.1186/1476-4598-3-9>

Giardiello, F. M., Brensinger, J. D., Tersmette, A. C., Goodman, S. N., Petersen, G. M., Booker, S. V., ... Offerhaus, J. A. (2000). Very high risk of cancer in familial Peutz–Jeghers syndrome. *Gastroenterology*, 119(6), 1447–1453. <https://doi.org/10.1053/gast.2000.20228>

Kaurah, P., & Huntsman, D. G. (1993). Hereditary Diffuse Gastric Cancer. *GeneReviews®*. University of Washington, Seattle. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20301318>

Kimura, T., Kamiura, S., Yamamoto, T., Seino-Noda, H., Ohira, H., & Saji, F. (2004). Abnormal uterine bleeding and prognosis of endometrial cancer. *International Journal of Gynecology & Obstetrics*, 85(2), 145–150. <https://doi.org/10.1016/j.ijgo.2003.12.001>

Leenen, C. H. M., van Lier, M. G. F., van Doorn, H. C., van Leerdam, M. E., Kooi, S. G., de Waard, J., ... Steyerberg, E. W. (2012). Prospective evaluation of molecular screening for Lynch syndrome in patients with endometrial cancer 70 years. *Gynecologic Oncology*, 125(2), 414–420. <https://doi.org/10.1016/j.ygyno.2012.01.049>

Masciani, S., Dillon, D. A., Rath, M., Robson, M., Weitzel, J. N., Balmana, J., ... Garber, J. E. (2012). Breast cancer phenotype in women with TP53 germline mutations: a Li-Fraumeni syndrome consortium effort. *Breast Cancer Research and Treatment*, 133(3), 1125–1130. <https://doi.org/10.1007/s10549-012-1993-9>

Ohman, A. W., Hasan, N., & Dinulescu, D. M. (2014). Advances in tumor screening, imaging, and avatar technologies for high-grade serous ovarian cancer. *Frontiers in Oncology*, 4, 322. <https://doi.org/10.3389/fonc.2014.00322>

Pilarski, R., Berry, M. P., Jude, S., Buys, S. S., Friedman, S., Garber, J. E., ... Dwyer, M. (2018). FORCE: Facing Our Risk of Cancer Empowered NCCN Guidelines Version 2.2019 Genetic/Familial High-Risk Assessment: Breast and Ovarian. Retrieved from https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf

Pilarski, R., Buys, S. S., Farmer, M., Khan, S., Klein, C., Kohlmann, W., ... Dwyer, M. (2018a). NCCN Guidelines Index Table of Contents Discussion NCCN Guidelines Version 1.2018 Panel Members Genetic/Familial High-Risk Assessment: Breast and Ovarian. Retrieved from https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf

Pilarski, R., Buys, S. S., Farmer, M., Khan, S., Klein, C., Kohlmann, W., ... Dwyer, M. (2018b). NCCN Guidelines Index Table of Contents Discussion NCCN Guidelines Version 1.2018 Panel Members Genetic/Familial High-Risk Assessment: Breast and Ovarian. Retrieved from https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf

Resta, R., Biesecker, B. B., Bennett, R. L., Blum, S., Estabrook Hahn, S., Strecker, M. N., & Williams, J. L. (2006). A New Definition of Genetic Counseling: National Society of Genetic Counselors' Task Force Report. *Journal of Genetic Counseling*, 15(2), 77–83. <https://doi.org/10.1007/s10897-005-9014-3>

Ring, K. L., Garcia, C., Thomas, M. H., & Modesitt, S. C. (2017). Current and future role of genetic screening in gynecologic malignancies. *American Journal of Obstetrics and Gynecology*, 217(5), 512–521. <https://doi.org/10.1016/j.ajog.2017.04.011>

Ring, K. L., & Modesitt, S. C. (2018). Hereditary Cancers in Gynecology: What Physicians Should Know About Genetic Testing, Screening, and Risk Reduction. *Obstetrics and Gynecology Clinics of North America*, 45(1), 155–173. <https://doi.org/10.1016/j.J.OG.2017.10.011>

MINISTERIO DA SAÚDE Agência Brasileira do ISBN. (n.d.). Retrieved from <http://www.inca.gov.br/estimativa/2018/estimativa-2018.pdf>

Tan, M.-H., Mester, J. L., Ngeow, J., Rybicki, L. A., Orloff, M. S., & Eng, C. (2012). Lifetime cancer risks in individuals with germline PTEN mutations. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 18(2), 400–407. <https://doi.org/10.1158/1078-0432.CCR-11-2283>

Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine*, 10(8), 789–799. <https://doi.org/10.1038/nm1087>

Walsh, M. D., Buchanan, D. D., Cummings, M. C., Pearson, S.-A., Arnold, S. T., Clendenning, M., ... Young, J. P. (2010). Lynch syndrome-associated breast cancers: clinicopathologic characteristics of a case series from the colon cancer family registry. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 16(7), 2214–2224. <https://doi.org/10.1158/1078-0432.CCR-09-3058>

Yurgelun, M. B., Hiller, E., & Garber, J. E. (2015). Population-Wide Screening for Germline BRCA1 and BRCA2 Mutations: Too Much of a Good Thing? *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 33(28), 3092–3095. <https://doi.org/10.1200/JCO.2015.60.8596>

Upload de Documentos

Arquivo Anexos:

Tipo	Arquivo
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx
Projeto Detalhado / Brochura Investigador	projeto_detalhado.docx
Projeto Detalhado / Brochura Investigador	projeto.pdf
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx
Comprovante de Recepção	PB_COMPROVANTE_RECEPCAO_1190397.pdf
Outros	parecer_departamento.pdf
Folha de Rosto	FOLHA_DE_ROSTO.pdf
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx
Outros	Carta_resposta_ao_parecer.docx
Outros	parecer_departamento.pdf
TCLE / Termos de Assentimento / Justificativa de Ausência	dispensa_novo_tcle.pdf
Outros	carta_anuencia_hermes_pardini.pdf
Projeto Detalhado / Brochura Investigador	projeto_detalhado.docx
Outros	carta_anuencia_hermes_pardini.pdf
Folha de Rosto	FOLHA_DE_ROSTO.pdf
Projeto Detalhado / Brochura Investigador	projeto.pdf
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_para_pesquisa.doc
TCLE / Termos de Assentimento / Justificativa de Ausência	dispensa_novo_tcle.pdf

Data de Submissão do Projeto: 06/01/2019

Nome do Arquivo: PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1190397.pdf

Versão do Projeto: 2

Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1190397.pdf
Projeto Detalhado / Brochura Investigador	projeto.pdf

Finalizar

Manter sigilo da íntegra do projeto de pesquisa: Não

Data de Submissão do Projeto: 06/01/2019

Nome do Arquivo: PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1190397.pdf

Versão do Projeto: 2

Página 7 de 7

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Perfil de mutações germinativas em pacientes submetidas ao aconselhamento genético para câncer hereditário de mama, ovário e endométrio, em Minas Gerais, Brasil

Pesquisador: Agnaldo Lopes da Silva Filho

Área Temática:

Versão: 2

CAAE: 01758418.0.0000.5149

Instituição Proponente: Faculdade de Medicina da UFMG

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.134.990

Apresentação do Projeto:

Trata-se de versão do projeto de pesquisa que responde diligências elencadas no parecer de número 3.078.927, emitido por este CEP.

Desenho inalterado em relação àquele parecer:

Conforme informado no formulário de informações básicas:

"Avaliar o perfil de mutações germinativas em pacientes com câncer de mama, ovário e endométrio submetidas a aconselhamento genético para câncer hereditário, bem como associar a presença dessas mutações com as características clínicas das pacientes e o tipo de tumor, caracterizar as variantes genéticas detectadas nos genes entre as pacientes com câncer ginecológico hereditário e comparar os dados obtidos com dados já

reportados na literatura da população brasileira e 'The Cancer Genome Atlas'.

[...]

Metodologia Proposta:

Serão analisados os prontuários de todos pacientes que realizaram o aconselhamento genético, devido à suspeita de câncer hereditário. Para a composição da amostra deste estudo serão selecionados 300 pacientes que apresentaram história pessoal ou familiar de câncer de mama, ovário, endométrio e outros tipos de câncer associado a síndromes hereditárias".

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 3.134.990

Objetivo da Pesquisa:

Conforme formulário de informações básicas atual:

Objetivo Primário:

Avaliar o perfil de mutações germinativas em pacientes com câncer de mama, ovário e endométrio submetidas a aconselhamento genético para câncer hereditário.

Avaliação dos Riscos e Benefícios:

Conforme formulário de informações básicas atual, descrito pelos pesquisadores:

"Riscos:

Os riscos decorrentes da participação nesta pesquisa são mínimos, pois podem decorrer de possível constrangimento dos participantes com a divulgação de dados confidenciais, invasão de privacidade e segurança dos prontuários. Para minimizar estes riscos, o pesquisador irá tratar a identidade dos participantes com padrões profissionais de sigilo, divulgação do nome ou qualquer material que indique o participante sem sua permissão.

Benefícios:

Não haverá benefício direto para as pacientes, exceto nos casos em que forem identificadas mutações genéticas que já estão associadas, na literatura, a síndromes neoplásicas, pois muda o manejo do risco da doença e permite que os pacientes e seus familiares em risco se beneficiem de medidas voltadas para o diagnóstico precoce e intervenção, condutas profiláticas voltadas para redução considerável do risco para o câncer e o prognóstico seja melhor para esses indivíduos, com diminuição da morbidade e mortalidade".

Comentários e Considerações sobre a Pesquisa:

Pendências emitidas no parecer de número 3.078.927 foram atendidas, a saber:

- 1) inclusão do Hermes Pardini como centro coparticipante;
- 2) inclusão de novo TCLE exclusivo para a pesquisa (independente do documento de TCLE de aconselhamento genético). Este novo TCLE esclarece que serão utilizados dados de prontuários e que tais dados serão tratados sob sigilo, com uso exclusivo para pesquisa.

Considerações sobre os Termos de apresentação obrigatória:

Termos adequados.

Recomendações:

Afirmar, no TCLE, que os dados serão utilizados de maneira agregada, sem possibilitar identificação do participante.

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad Sl 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 3.134.990

Conclusões ou Pendências e Lista de Inadequações:

Tendo sido atendidas as diligências do parecer de número 3.078.927, aprova-se o projeto de pesquisa.

Considerações Finais a critério do CEP:

Tendo em vista a legislação vigente (Resolução CNS 466/12), o CEP-UFMG recomenda aos Pesquisadores: comunicar toda e qualquer alteração do projeto e do termo de consentimento via emenda na Plataforma Brasil, informar imediatamente qualquer evento adverso ocorrido durante o desenvolvimento da pesquisa (via documental encaminhada em papel), apresentar na forma de notificação relatórios parciais do andamento do mesmo a cada 06 (seis) meses e ao término da pesquisa encaminhar a este Comitê um sumário dos resultados do projeto (relatório final).

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1190397.pdf	06/01/2019 19:40:32		Aceito
Projeto Detalhado / Brochura Investigador	projeto_detalhado.docx	06/01/2019 19:40:18	CAMILA MARTINS DE CARVALHO	Aceito
Outros	Carta_resposta_ao_parecer.docx	05/01/2019 21:04:23	CAMILA MARTINS DE CARVALHO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_para_pesquisa.doc	05/01/2019 21:03:11	CAMILA MARTINS DE CARVALHO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx	15/10/2018 15:12:47	CAMILA MARTINS DE CARVALHO	Aceito
Outros	parecer_departamento.pdf	13/10/2018 17:38:22	CAMILA MARTINS DE CARVALHO	Aceito
Outros	carta_anuencia_hermes_pardini.pdf	13/10/2018 17:37:15	CAMILA MARTINS DE CARVALHO	Aceito
Folha de Rosto	FOLHA_DE_ROSTO.pdf	13/10/2018 17:35:09	CAMILA MARTINS DE CARVALHO	Aceito

Situação do Parecer:

Aprovado

Necessita apreciação da CONEP:

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad Sl 2005
Bairro: Unidade Administrativa II **CEP:** 31.270-901
UF: MG **Município:** BELO HORIZONTE
Telefone: (31)3409-4592 **E-mail:** coep@prpq.ufmg.br

Continuação do Parecer: 3.134.990

Não

BELO HORIZONTE, 06 de Fevereiro de 2019

Assinado por:
Eliane Cristina de Freitas Rocha
(Coordenador(a))

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Avaliação da expressão, metilação e mutação germinativa e somática nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2) em mulheres com câncer epitelial de ovário como preditivo de resposta ao tratamento quimioterápico

Pesquisador: Agnaldo Lopes da Silva Filho

Área Temática:

Versão: 2

CAAE: 85375317.0.0000.5149

Instituição Proponente: Hospital das Clínicas - Universidade Federal de Minas Gerais

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.667.131

Apresentação do Projeto:

O estudo "Avaliação da expressão, metilação e mutação germinativa e somática nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2) em mulheres com câncer epitelial de ovário como preditivo de resposta ao tratamento quimioterápico" é um estudo de coorte de mulheres diagnosticadas com câncer epitelial de ovário (CEO) tipo seroso.

Na metodologia explicitada pelos pesquisadores, os mesmos informam que serão estudadas 150 mulheres (amostra de conveniência) com diagnóstico de CEO com proposta de tratamento cirúrgico e 10 mulheres submetidas à ooforectomia sem evidências de neoplasias malignas ovariana (grupo controle), atendidas no Serviço de Ginecologia e Obstetrícia do Hospital das Clínicas/UFMG (HC-UFMG). A participação na pesquisa não implica modificações no tratamento proposto às pacientes, uma vez que apenas serão coletadas amostras de tecido e amostra de sangue total. Quatro grupos serão formados para avaliação das amostras clínicas:

Mulheres com diagnóstico patológico de CEO inicial com proposta de tratamento cirúrgico e quimioterápico ou apenas tratamento cirúrgico (n=50).

Mulheres com diagnóstico de CEO metastático (n=50).

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

Mulheres com recidiva, independentemente do tempo (n=50).

Mulheres sem evidências de neoplasia maligna ovariana, submetidas a ooforectomia por solicitação ou indicação em mulheres na pós menopausa ou por alguma intercorrência cirúrgica que necessitou a retirada do órgão (por exemplo sangramento ativo), consideradas como controle neste estudo. Não será realizada retirada de parte ou todo o ovário apenas para fins de estudo. (n=10).

O recrutamento das pacientes será realizado nos Ambulatórios de Ginecologia Oncológica do Hospital das Clínicas/UFMG (HC-UFMG) e as coletas das amostras teciduais serão realizadas nos blocos cirúrgicos do referido hospital, durante o tratamento cirúrgico. O processamento das amostras teciduais, o cultivo de células, incubação e análises moleculares serão realizadas no Serviço de Biologia Celular da Fundação Ezequiel Dias.

De acordo com os pesquisadores serão incluídas no estudo mulheres com diagnóstico de CEO, acima de 18 anos, com proposta de tratamento cirúrgico ou quimioterápico (casos); mulheres com diagnóstico de CEO, com ou sem história familiar de câncer de mama e/ou câncer de ovário mulheres, acima de 18 anos, submetidas à histerectomia abdominal para mioma uterino (controle); Termo de consentimento livre e esclarecido assinado. Entre os critérios de exclusão estão: não concordância em participar do estudo; evidência à laparotomia de algum processo infeccioso agudo peritoneal; mulheres com história de uso de imunossuppressores, corticosteroides e/ou anti-inflamatórios não esteroides há pelo menos três meses.

As amostras clínicas serão submetidas à extração de DNA e RNA utilizando o reagente TRIzol® (Invitrogen) seguindo o protocolo descrito pelo fabricante. Após a extração, as amostras de DNA e RNA serão quantificadas por espectrometria e avaliadas quanto a sua integridade no sistema 2200 TapeStation. As amostras de RNA seguirão, então, para etapas de síntese de cDNA utilizando o kit SuperScript™ II Reverse Transcriptase, Invitrogen e o cDNA será submetido à amplificação pela reação em cadeia da polimerase em tempo real (qRT-PCR) para quantificação relativa da expressão dos genes alvo. Resumidamente, os ensaios serão realizados usando a tecnologia de arranjos de sondas lineares marcadas com fluorocromos (TaqManArray® -Applied Biosystems). Serão usados ensaios inventariados, seguindo o protocolo descrito pelo fabricante. O sequenciamento dos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2) será executado pelo método de Sanger seguindo protocolo padronizado para detectar a

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad Si 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

presença de SNPs e outras pequenas variações de sequências DNA das regiões codificantes desses genes nas amostras clínicas. Para a exclusão de grandes rearranjos nesses genes não detectáveis pelo sequenciamento será utilizado o método MLPA. Essa técnica é uma alternativa para estudar variações de número de cópias (CNV's) menores e baseia-se em única análise por sondas compostas em kits determinados pelo fabricante (MRC-HOLLAND®) para regiões que apresentam mutações tipo microdeleções/microduplicações. O serviço de sequenciamento e do MLPA serão terceirizados, contudo, sua execução será acompanhada por membros do grupo de pesquisa. Para determinação do perfil de metilação das ilhas CpGs da região promotora dos genes BRCA1 e BRCA2, o gDNA extraído das amostras será tratado com bissulfito de sódio (Methylseqr™ Bisulfite Conversion Kit - Applied Biosystems) e submetido a PCR metilação específica (MSP) em tempo real.

Para análise dos dados, os pesquisadores informam que os dados serão armazenados no programa SPSS 18.0 para Windows e Bionumerics 7.0 (Applied Maths) e analisados por meio de ferramentas estatísticas apropriadas dos referidos programas, a fim de avaliar e comparar os resultados obtidos com dados clínico-patológicos das pacientes com câncer ovariano. Os dados obtidos poderão permitir um melhor entendimento do cenário de mutação e da expressão dos alvos ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2) em estudo no CO e poderão elucidar marcadores moleculares, candidatos à predição de susceptibilidade/resistência deste tumor ao tratamento quimioterápico. Além dessas análises estatísticas serão utilizados outros métodos de análise a fim de mensurar e visualizar graficamente as diferenças na expressão gênica entre os grupos de tumores ovarianos estudados (CO primário, CO metastático e cistoadenoma seroso ovariano). Serão utilizados métodos de construção de árvore de decisão e heatmaps. Na construção da árvore de decisão, foi utilizado o método J48 que é uma implementação do C4.5 (QUINLAN, 1993), presente no software Weka (FRANK, et al., 2010). O J48 será executado utilizando parâmetros de entrada padrão. A árvore de decisão construída para cada um dos marcadores estudados utilizando como registros os grupos de pacientes descritos na casuística, visando segregar os registros em dois grupos distintos responsivos e não responsivos ao tratamento. Para validação e generalização do método de classificação será utilizada a validação cruzada leave one out cross validation (LOOCV). O heatmap será construído utilizando o pacote "gplots" com parâmetros default para clusterização e a função "heatmap.2" no software R (R DEVELOPMENT CORE TEAM, 2011).

No referencial teórico, os pesquisadores afirmam que o câncer de ovário (CO) é o tumor

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005**Bairro:** Unidade Administrativa II **CEP:** 31.270-901**UF:** MG **Município:** BELO HORIZONTE**Telefone:** (31)3409-4592**E-mail:** coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

ginecológico mais difícil de ser diagnosticado e o de menor chance de cura, pois cerca de três quartos dos cânceres desse órgão apresentam-se em estágio avançado no momento do diagnóstico. O número de mortes de mulheres com CO em 2013 foi de 3.283, de acordo com o Sistema de Informação sobre Mortalidade, do Ministério da Saúde. A ultrassonografia endovaginal e a dosagem de CA 125 são o esteio da detecção do CO, mas possuem baixa especificidade e sensibilidade para detecção da doença em estágios iniciais. Um importante fator de risco para o diagnóstico e tratamento do CO é a história familiar. Cerca de 10 a 15% dos casos de CO e quase 20% dos CO de alto grau são decorrentes de mutações nos genes BRCA1 e BRCA2. Além das mutações germinativas, mutações somáticas nestes genes podem produzir tumores que agem como deficientes em BRCA. Alterações genéticas específicas tumorais se correlacionam com a presença de CO, assim, elas são importantes marcadores com potenciais clínicos para avaliar risco, diagnóstico, detecção, prognóstico e resposta ao tratamento do câncer.

Ainda, segundo os pesquisadores, a identificação da predisposição de um câncer hereditário, como o CO associado a mutação genética, é de suma importância na saúde pública, pois muda o manejo da doença e permite que os pacientes e seus familiares em risco se beneficiem da profilaxia e diagnóstico precoce do câncer, o que reduz seu risco à população em geral. Dentro de uma perspectiva da medicina translacional, um teste molecular de predição da resposta terapêutica individualizada poderá contribuir para o aprimoramento da terapia do CO, resultando aumento da expectativa e qualidade de vida das mulheres afetadas.

Objetivo da Pesquisa:

Os pesquisadores descrevem como objetivos.

Primário:

Avaliar uma coorte de mulheres diagnosticadas com Câncer Epitelial de Ovário (CEO) tipo seroso para detectar e caracterizar variantes genética presentes germinativas e somáticas dos genes em ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), bem como associar a presença dessas variantes com a classificação desses tumores como tipo I e tipo II, as características clínicas das pacientes e o desfecho primário do tratamento.

Específicos:

1. Avaliar a presença de mutações germinativas nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), em mulheres de câncer

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005
Bairro: Unidade Administrativa II **CEP:** 31.270-901
UF: MG **Município:** BELO HORIZONTE
Telefone: (31)3409-4592 **E-mail:** coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

epitelial de ovário.

2. Avaliar a presença de mutações somáticas nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), em mulheres de câncer epitelial de ovário.
3. Associar a presença de mutações nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), com as características clínicas das pacientes e o desfecho primário do tratamento.
4. Caracterizar variantes genética detectadas nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), entre as pacientes dos diferentes grupos de tumores ovarianos e o grupo controle.
5. Comparar as variantes genética detectadas nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), entre as pacientes dos diferentes grupos de tumores ovarianos com o grupo controle.
6. Determinar o perfil de expressão dos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), entre as pacientes dos diferentes grupos de tumores ovarianos e o grupo controle.
7. Comparar a variabilidade de expressão dos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), entre as pacientes dos diferentes grupos de tumores ovarianos com o grupo controle.
8. Associar as variantes genética nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), e o perfil de expressão desses genes entre os grupos de mulheres com CEO primário e metastático.
9. Sequenciar os genes TP53 e PTEN nas amostras de tumores ovarianos para classificação dos tumores malignos como Tipo I ou II.

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005

Bairro: Unidade Administrativa II

CEP: 31.270-901

UF: MG

Município: BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

10. Avaliar a associação das variantes genética presentes nos genes ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1, BRCA2 e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), o perfil de expressão desses genes com a classificação dos tumores malignos como Tipo I ou II e os dados clínico-patológicos, tais como dosagem sérica do CA-125, estadiamento FIGO, sobrevida livre da doença e sobrevida geral, das pacientes com CEO.

11. Identificar uma assinatura genética dos tumores ovarianos que permitam prever a efetividade do tratamento à base de PARPI, como Lynparza™.

Avaliação dos Riscos e Benefícios:

Os riscos descritos no projeto pelos autores são os que se seguem:

“Os riscos do estudo proposto são aqueles inerentes ao procedimento cirúrgico primário que motivou a cirurgia, sem riscos adicionais às pacientes.”

Os benefícios descritos no projeto pelos autores são os que se seguem:

Benefícios

“Não haverá benefício direto para as pacientes, exceto nos casos em que forem identificados miRNAs específicos, que já estão associados, na literatura, às características específicas da progressão da doença ou da resposta terapêutica. A longo prazo, o maior benefício é o do conhecimento sobre o Câncer de Ovário (CO), proporcionando melhor entendimento a relação de expressão de miRNAs envolvidos na regulação das vias de morte celular (apoptose e anoikis), reparo de DNA e metabolismo de RNA e as características clínicas observadas. Tais resultados podem apresentar implicações clínicas relevantes, permitindo construir uma plataforma para o diagnóstico precoce e de recidivas da doença, avaliação de resposta terapêutica, quimioprevenção adequada e prognóstico, além de auxiliar no desenvolvimento de novas modalidades de tratamento e promover aconselhamento genético e manejo preventivo para os pacientes e seus familiares.”

Comentários e Considerações sobre a Pesquisa:

Trata-se de uma coorte prospectiva de mulheres diagnosticadas com CEO tipo seroso para detectar e caracterizar variantes genéticas presentes germinativas e somáticas dos genes em ATM, ATR, PALB2, BRIP, RAD51, FANC, BRCA1 e BRCA2, e genes MMR (MLH1, MSH2, MSH6, EPCAM e PMS2), bem como associar a presença destas variantes com a classificação desses tumores como

Endereço: Av. Presidente Antônio Carlos, 6627 2ª Ad SI 2005

Bairro: Unidade Administrativa II

CEP: 31.270-901

UF: MG

Município: BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

tipo I e II, as características clínicas das pacientes e o desfecho primário do tratamento.

Os pesquisadores relatam que, de acordo com a Resolução nº306/2004 da Agência Nacional de Vigilância Sanitária - ANVISA - e a Resolução nº358/2005 do Conselho Nacional do Meio Ambiente – CONAMA (que dispõem sobre o tratamento e descarte final de resíduos de serviços de saúde), os resíduos biológicos que serão gerados (amostras de câncer de ovário humano), classificados como A3, serão acondicionados em sacos brancos contendo o símbolo universal de risco biológico, sem quaisquer outros tipos de resíduos misturados, e encaminhados para descarte pela Fundação Ezequiel Dias – FUNED.

Os pesquisadores também informam que são participantes do projeto:

- Membros do Departamento de Ginecologia e Obstetrícia da Faculdade de Medicina da UFMG
- Membro do Departamento de Patologia e Medicina Legal da Faculdade de Medicina da UFMG
- Membros do Departamento de Biologia Geral do Instituto de Ciências Biológicas da UFMG
- Membros do Serviço de Biologia Celular, Laboratório de Pesquisa e Gestão de Desenvolvimento da Divisão de Pesquisa e Desenvolvimento da Fundação Ezequiel Dias-Funed
- Membros da Universidade Federal de Uberlândia (UFU)

O projeto de pesquisa foi aprovado pelo Departamento de Ginecologia e Obstetrícia da Faculdade de Medicina sem restrições em 24 de fevereiro de 2017.

Este projeto teve o orçamento previamente estimado de R\$460.270.,77 (quatrocentos e sessenta mil e duzentos e setenta reais e setenta e sete centavos). Os pesquisadores informam que o serviço hospitalar e laboratorial envolvido na execução desta proposta possuem toda a infraestrutura e os equipamentos necessários ao desenvolvimento do projeto proposto, bem como contam com as equipes técnicas qualificadas em seu quadro de servidores públicos/funcionários para o alcance dos objetivos e metas propostos neste projeto. Além disso, alguns reagentes e insumos necessários à coleta das amostras e execução da proposta (não incluídos no orçamento) estão disponíveis no laboratório. Além da infraestrutura e de apoio técnico oferecido pela instituição, o projeto deverá contar com os recursos captados em linhas de fomento pelo coordenador (Fapemig e CNPq) e apoio a pesquisa por iniciativa do investigador oferecido pela AstraZeneca, através de termo de parceria específico a ser firmado.

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad S/C 2005	
Bairro: Unidade Administrativa II	CEP: 31.270-901
UF: MG	Município: BELO HORIZONTE
Telefone: (31)3409-4592	E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

O cronograma apresentado prevê início de coleta de dados a partir da aprovação do Comitê de Ética (entre 01/11/2017 a 01/05/2018). O término da pesquisa está previsto 01 de novembro de 2019.

As solicitações do Comitê de Ética foram atendidas:

1. Os pesquisadores definiram como critério de inclusão tanto para caso como para controles participantes com idade acima de 18 anos, desta forma, torna-se dispensável o Termo de Assentimento Livre e Esclarecido.
2. Os pesquisadores acrescentaram no projeto de pesquisa que haverá uso de fontes secundárias.
3. Os pesquisadores fizeram as modificações em ambos os TCLE: páginas foram numeradas; campos de rubrica foram acrescidos; foi solicitada autorização ao participante para consulta ao prontuário.

Considerações sobre os Termos de apresentação obrigatória:

Na segunda versão, os pesquisadores apresentaram um total de quatro (4) documentos que estão descritos abaixo.

PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_971006.pdf

Carta_resposta_ao_parecer.pdf

TCLE_controle.doc

TCLE_caso.doc

Projeto.docx

Na primeira versão, os documentos apresentados, em um total de oito (8), são os que estão descritos abaixo.

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad S1 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_971006.pdf

cartasubmissao_GEP_HC.pdf (acusa recebimento; datada e assinada em 26 de novembro de 2017)

carta_departamento.pdf (projeto aprovado; data e assinada em 24 de fevereiro de 2017)

TCLE_controle.doc (de acordo com a Resolução nº 466/12 do Conselho Nacional de Saúde; necessários pequenos ajustes)

TCLE_caso.doc (de acordo com a Resolução nº 466/12 do Conselho Nacional de Saúde; necessários pequenos ajustes)

FOLHA_DE_ROSTO.pdf (aprovada, data e assinada em 06 de novembro de 2017)

Projeto.docx (datado em novembro de 2016)

Recomendações:

Não há recomendações.

Conclusões ou Pendências e Lista de Inadequações:

Sou S.M.J. pela aprovação do projeto de pesquisa.

Considerações Finais a critério do CEP:

Tendo em vista a legislação vigente (Resolução CNS 466/12), o COEP-UFMG recomenda aos Pesquisadores: comunicar toda e qualquer alteração do projeto e do termo de consentimento via emenda na Plataforma Brasil, informar imediatamente qualquer evento adverso ocorrido durante o desenvolvimento da pesquisa (via documental encaminhada em papel), apresentar na forma de notificação relatórios parciais do andamento do mesmo a cada 06 (seis) meses e ao término da pesquisa encaminhar a este Comitê um sumário dos resultados do projeto (relatório final).

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_P ROJETO_971006.pdf	05/05/2018 17:25:00		Aceito

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad Sl 2005

Bairro: Unidade Administrativa II **CEP:** 31.270-901

UF: MG **Município:** BELO HORIZONTE

Telefone: (31)3409-4592

E-mail: coep@prpq.ufmg.br

Continuação do Parecer: 2.667.131

Outros	Carta_resposta_ao_parecer.pdf	05/05/2018 17:24:32	CAMILA MARTINS DE CARVALHO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_controle.doc	05/05/2018 17:15:58	CAMILA MARTINS DE CARVALHO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_caso.doc	05/05/2018 17:15:40	CAMILA MARTINS DE CARVALHO	Aceito
Projeto Detalhado / Brochura Investigador	Projeto.docx	05/05/2018 17:15:12	CAMILA MARTINS DE CARVALHO	Aceito
Outros	cartasubmissao_GEP_HC.pdf	14/03/2018 20:16:37	CAMILA MARTINS DE CARVALHO	Aceito
Outros	carta_departamento.pdf	18/12/2017 20:36:08	CAMILA MARTINS DE CARVALHO	Aceito
Outros	carta_unidade_funcional.pdf	18/12/2017 20:35:34	CAMILA MARTINS DE CARVALHO	Aceito
Folha de Rosto	FOLHA_DE_ROSTO.pdf	06/12/2017 10:10:05	CAMILA MARTINS DE CARVALHO	Aceito
Outros	853753170aprovacaoassinada.pdf	22/05/2018 09:36:10	Vivian Resende	Aceito
Outros	853753170parecerassinado.pdf	22/05/2018 09:36:19	Vivian Resende	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

BELO HORIZONTE, 22 de Maio de 2018

Assinado por:
Vivian Resende
(Coordenador)

Endereço: Av. Presidente Antônio Carlos, 6627 2º Ad SI 2005
Bairro: Unidade Administrativa II **CEP:** 31.270-901
UF: MG **Município:** BELO HORIZONTE
Telefone: (31)3409-4592 **E-mail:** coep@prpq.ufmg.br