

INVESTIGAÇÃO DE CAUSAS GENÉTICAS ASSOCIADAS AO DESENVOLVIMENTO DE SARCOMAS EM IDADE JOVEM

NATHÁLIA DE ANGELIS DE CARVALHO

Tese apresentada à Fundação Antônio Prudente para obtenção
do Título de Doutor em Ciências.

Área de Concentração: Oncologia

Orientadora: Dra. Giovana Tardin Torrezan

Coorientadora: Dra. Dirce Maria Carraro

São Paulo

2022

FICHA CATALOGRÁFICA

Preparada pelo Ensino/Apoio ao aluno da Fundação Antônio Prudente

de Angelis de Carvalho, Nathália.

Investigação de causas genéticas associadas ao desenvolvimento de sarcomas em idade jovem. / Nathália de Angelis de Carvalho. São Paulo, 2022.

158f.

Tese de Doutorado - Fundação Antônio Prudente. Curso de Pós-Graduação em Ciências - Área de concentração: Oncologia.

Orientador: Giovana Tardin Torrezan.

1. Sarcomas, 2. Alterações genéticas, 3. Sequenciamento de Nova-Geração

CDU 616

Nome: Nathália de Angelis de Carvalho

Título: Investigação de causas genéticas associadas ao desenvolvimento de sarcomas em idade jovem

Aprovado em: ____/____/____

Banca Examinadora

Orientador: Dra. Giovana Tardin Torrezan

Instituição: Fundação Antônio Prudente

Membro da banca: Dr. Benedito Mauro Rossi

Instituição: Hospital Sírio Libanês

Membro da banca: Dra. Patrícia Ashton Prolla

Instituição: Universidade Federal do Rio Grande do Sul

Membro da banca: Dr. José Andres Yunes

Instituição: Centro Infantil de Investigações Hematológicas Dr. Domingos Ademar Boldrini

Membro da banca: Dr. Victor Evangelista de Faria Ferraz

Instituição: Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo

SUPORTE À PESQUISA POR AGÊNCIA DE FOMENTO

Este trabalho recebeu apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), através de auxílio à Pesquisa Regular, processo número - 2018/06269-5.

Este trabalho recebeu apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), através do auxílio à Pesquisa – Bolsa de mestrado, inicialmente de 01/07/2018 a 30/06/2019 e depois com a vigência de 01/07/2020 - 31/10/2022 (Nesse momento, como bolsa de doutorado direto), processo número - 88887.509818/2020-00.

Ainda, o trabalho foi contemplado com uma bolsa do Programa de Doutorado Sanduíche no Exterior (PDSE-CAPES), processo número - 88881.623036/2021-01, entre o período de 01/10/2021 a 31/01/2022.

Entre o período de 01/07/2019 a 30/06/2020, esse projeto (em nível mestrado) recebeu apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), através de auxílio à Pesquisa, processo número - 2018/17027-2.

“A revolução é a harmonia da forma e da cor e tudo está e se move sob uma única lei: a vida.

Ninguém se aparta de ninguém. Ninguém luta por si mesmo. Tudo é tudo e um.”

Frida Khalo

Dedico esse trabalho a todos que continuam persistindo, resistindo e não desistindo – na ciência, na luta individual e coletiva, e em todos os âmbitos da vida.

AGRADECIMENTOS

Primeiramente, agradeço a todos os pacientes que concordaram em participar desse trabalho. Vocês são muito fortes e corajosos nessa luta! Espero poder retribuir de alguma forma.

À minha orientadora, Dra. Giovana Tardin Torrezan, agradeço imensamente. Por todos esses anos de descobertas, de crescimento pessoal e profissional. Por ser apoio e acreditar em mim, muitas vezes mais do que eu mesma, sendo a minha maior força em toda essa trajetória na pesquisa. Você é a minha inspiração como pessoa e cientista!

À minha coorientadora no doutorado-sanduíche, Dra. Barbara Rivera, agradeço por aceitar me orientar e fazer isso da melhor forma, nunca me deixando desamparada. Agradeço pela paciência, por me trazer desafios e por enriquecer esse estudo com todo seu conhecimento.

À minha coorientadora, Dra. Dirce Maria Carraro e a todos do LGBM pela oportunidade em fazer parte desse grupo de pesquisa, responsável por todo meu aprendizado durante esses anos. Em especial, à Karina Miranda Santiago e ao Marcelo Moreno.

À Dra. Thais Bartelli e ao Dr. Benedito Mauro Rossi por fazerem parte da minha banca de acompanhamento, complementando o trabalho com diversas sugestões valorosas.

A todos os colaboradores envolvidos no desenvolvimento desse trabalho, ao A.C. Camargo Cancer Center, aos departamentos de Oncogenética, de Oncologia Clínica (em especial à Dra Joyce Lisboa pelo apoio no recrutamento de pacientes) e Oncologia Pediátrica, ao Biobanco, à Bioinformática, aos funcionários do CIPE e à pós-graduação por todo o suporte.

À Deus, por sempre se fazer presente em todos os detalhes da minha vida, cuidando e protegendo.

Aos meus pais, Cristina de Angelis e Renato Augusto de Carvalho, que sempre confiam e abraçam minhas decisões. E à toda minha família, em especial às minhas avós, que também sempre me apoiaram em tudo.

Aos meus companheiros caninos, Pitchuco, Lilica e Bidu, por serem colo e minha motivação em todos esses anos, especialmente na decisão de estudar e persistir na medicina veterinária. E à felina Phoebe, que é uma companheirinha essencial.

Ao meu amigo-irmão Cícero Júlio Silva Costa, meu grande presente da graduação, onde fomos descobrindo e explorando juntos o meio da pesquisa, experimentando diversos desafios e obstáculos. Vivemos tantas coisas, que eu não seria nada sem a sua companhia e seu suporte nesse trajeto.

À minha amiga-irmã Sara Iolanda Oliveira da Silva por ser um dos meus maiores apoios na vida, dividindo alegrias, conquistas e angústias. Seu apoio foi e é essencial em todo esse processo.

À Dayane Schmidt, uma amiga que me inspira pela sua inteligência, bondade, amor e acolhimento. Obrigada por me motivar e por estar sempre por perto, mesmo a kms de distância.

Aos meus presentes do laboratório de pesquisa, Ana Carolina Araújo, Ana Carolina Kerekes, Débora Andrade e Julia Salles, que foram espaço de acolhimento, amparo e parceria nesse caminho todo.

Aos amigos de Barcelona, os quais se tornaram minha família no exterior. Lá compartilhamos experiências e criamos histórias das quais eu nunca esquecerei. À Jenniffer Hauschildt, Luana Homma, Marina Soncini, Mirele Corrêa, Franciany Braga, Will, Lucca Tori, Vando França, Jose Luiz Jr, Willy Sandino, Yovani Villanueva e Oscar Rondon, pela grande parceria, acolhida e união. A experiência foi imensamente melhor com vocês. Gracias por todo, chicos!

Em especial, à Elisabet Munté e ao Lluís Salvador do IDIBELL, em Barcelona, que sempre foram pacientes, preocupados e acolhedores, vocês foram fundamentais nesse período. À Clara, Carla e Demetria, e a todos os colegas da instituição em que pude compartilhar conhecimentos e experiências.

Por fim, às agências de fomento CAPES e FAPESP, que possibilitaram a realização desse trabalho e continuam persistindo no apoio e incentivo a diversos outros estudos.

RESUMO

de Carvalho N.d.A. **Investigação de causas genéticas associadas ao desenvolvimento de sarcomas em idade jovem** [Tese]. São Paulo, Fundação Antônio Prudente, 2022.

Os avanços recentes na genômica permitiram o reconhecimento de quase uma centena de novos genes de predisposição ao câncer (GPC). Embora o rastreamento genético de uma parte destes genes esteja atualmente bem estabelecido para os tumores hereditários mais comuns, há uma série de tumores raros, como sarcomas, que podem estar associados a síndromes de câncer hereditário, mas cuja relação genótipo-fenótipo ainda é desconhecida. Dessa forma, o presente estudo propôs-se a: definir a frequência de variantes raras germinativas patogênicas em GPCs em crianças, adolescentes e adultos jovens com sarcomas; avaliar as características moleculares dos tumores para confirmar sua causa genética em casos selecionados, realizando análise de perda de heterozigose (LOH); identificar novos genes e vias relevantes associados ao desenvolvimento de condrossarcomas e sarcomas ultrarraros. Para isso, foram rastreadas variantes germinativas em 177 pacientes jovens diagnosticados com sarcoma abaixo de 40 anos, através da análise de sequenciamento de nova geração, com painéis customizados de 113/126 GPCs. Primeiramente, foi rastreada a variante fundadora do gene *TP53* mais prevalente na população brasileira (p.Arg337His) e 5/177 pacientes foram detectados como portadores e excluídos da análise de painel. No painel, foram avaliados 172 pacientes e em 33 foram detectadas variantes patogênicas (P) ou provavelmente patogênicas (PP) (18,6%), sendo estas em genes previamente associados ao risco no desenvolvimento de sarcomas, como *CHEK2*, *EXT1*, *EXT2*, *NF1*, *RB1* e *TP53*, mas também em genes no qual esse risco ainda é desconhecido (*AKT1*, *ERCC3*, *ERCC4*, *FANCM*, *MITF*, *MUTYH*, *NTHL1*, *SLX4* e *TSC2*) ou é emergente (*BRCA2*, *ERCC2*, *PALB2*). Cento e trinta e nove (78,5%) pacientes tiveram variantes de significado incerto ou foram negativos. Portadores de variantes P/PP foram mais propensos a apresentar mais de um tumor primário do que não portadores (21,1% X 6,5%; p=0,012). Dos 25 tumores avaliados para LOH, a perda do alelo selvagem foi detectada em nove (36%), sendo um dos genes de associação desconhecida a sarcomas (*MITF*). Para explorar novos genes candidatos, 9 pacientes com condrossarcomas ou sarcomas ultrarraros negativos para variantes P/PP foram incluídos em análise de sequenciamento de exoma tumoral e germinativo. Foram priorizadas na análise germinativa variantes em 10 GPCs conhecidos e 22 genes com variantes de perda de função e *missenses* em que a relação com câncer hereditário é desconhecida. Foi encontrada somente uma variante em um gene associado a sarcomas (*RECQL4*) e em um gene emergente (*BRCA2*). Uma variante patogênica no gene *RAD50* foi detectada em um paciente com sarcoma fibromixóide. Em dois tumores distintos, foi observada alteração no gene *COL7A1*. Na análise somática, em três genes (*TP53*, *PTCH1*, *CREBBP*) foram encontradas alterações potencialmente significantes clinicamente, visto que são genes de associação conhecida a sarcomas. Também foram identificados outros genes mutados, alguns deles incluídos em vias biológicas que podem ser interessantes para sarcomas, como *VPS16* e *MYF6*. Contudo, para que seja possível realizar associações genótipo-fenótipo dos dados de exoma, tanto germinativo quanto somático, outras evidências são necessárias, como análise de transcriptoma e co-segregação. Com os dados do painel, nossos resultados destacam uma alta taxa de variantes P/PP em GPCs em pacientes jovens brasileiros

com sarcoma (21,5% incluindo os pacientes portadores da variante *TP53* p.Arg337His), mesmo em pacientes sem história familiar de câncer. As taxas de variantes P/PP variaram de 0% a 33% entre os subtipos de sarcoma e exibiram associações específicas entre subtipos e genes. Isso aponta a urgência de implementar estratégias de triagem genética adequadas para esses indivíduos e suas famílias.

Palavras-chave: 1. Sarcomas. 2. Alterações genéticas. 3. Sequenciamento de Nova-Geração.

ABSTRACT

de Carvalho N.d.A. [Investigation of genetic causes associated with the development of sarcomas at young age] [Tese]. São Paulo, Fundação Antônio Prudente, 2022.

Recent advances in genomics have allowed the recognition of nearly a hundred new cancer predisposing genes (CPGs). Although genetic screening in some of these genes is currently well established for the most common hereditary tumors, there are a number of rare tumors, such as sarcomas, that may be associated with hereditary cancer syndromes, but whose genotype-phenotype relationship is still unknown. Thus, the present study aimed to: define the frequency of rare pathogenic germline variants in CPGs in children, adolescents and young adults with sarcomas; assess the molecular characteristics of tumors to confirm their genetic cause in selected cases, performing loss of heterozygosity (LOH) analysis; identify novel genes and relevant pathways associated with the development of chondrosarcomas and ultra-rare sarcomas. For this, we searched for germline variants in 177 young patients diagnosed with sarcoma under 40 years old, through next-generation sequencing analysis, with customized panels of 113/126 CPGs. First, we searched for the founder variant in *TP53* gene (p.Arg337His) - most prevalent in the Brazilian population - and found that 5/177 patients were detected as carriers, being excluded from the panel analysis. In the panel, 172 patients were evaluated and in 33 (18.6%) there were detected germline pathogenic variants (GPVs), occurring in genes previously associated with the risk of developing sarcomas (*TP53*, *RB1*, *NF1*, *CHEK2*, *EXT1* and *EXT2*), but also in genes where that risk is still unknown (*ERCC2/3*, *TSC2*, *RAD50*, *FANCM*, and others) or is emerging (*PALB2*, *BRCA2*). GPVs carriers were more likely to present more than one primary tumor than non-carriers (21.1% X 6.5%; $p=0.012$). One hundred and thirty-nine (78.5%) patients had variants of uncertain significance or were negative. Of the 25 tumors evaluated for LOH, loss of the wild-type allele was detected in nine (36%), in which one of the genes has an unknown association with sarcomas (*MITF*). To explore new candidate genes, 9 patients with chondrosarcomas or ultra-rare sarcomas negative for GPVs were included in tumor exome and germline sequencing analysis. In the germline analysis, there were prioritized variants in 10 known CPGs and 22 genes with loss-of-function and missense variants in which the relationship with hereditary cancer is unknown. Only one variant was found in a sarcoma-associated gene (*RECQL4*) and in an emergent gene (*BRCA2*). A pathogenic variant in the *RAD50* gene was detected in a patient with fibromyxoid sarcoma. An alteration in *COL7A1* gene was observed in two different tumors. In the somatic analysis, there were found potentially clinically significant alterations in three genes (*TP53*, *PTCH1*, *CREBBP*), since they are genes with a known association with sarcomas. Other mutated genes were also identified, some of them included in biological pathways that may be interesting for sarcomas, such as *VPS16* and *MYF6*. However, in order to be able to perform genotype-phenotype associations of exome data, other evidences are needed for both germline and somatic evaluation, such as transcriptome analysis and co-segregation. With panel data, our results highlight a high rate of GPVs in CPGs in young Brazilian patients with sarcoma (21.5% including patients carrying the *TP53*:p.Arg337His variant), even in patients with no family history of cancer. Our findings reveal that 1 in each 5 patients with sarcomas under 40 years old presented a GPV, even in patients without family history of cancer. GPVs rates varied from 0%

to 33% across sarcoma subtypes and exhibited specific gene-subtypes associations. This pinpoints the urgency of implementing appropriated genetic screening strategies for these individuals and their families.

Keywords: 1. Sarcomas. 2. Genetic alterations. 3. Next-Generation Sequencing.

LISTA DE FIGURAS

Figura 1 - Representação topográfica do desenvolvimento de sarcomas no corpo humano.	23
Quadro 1 - Critérios clínicos utilizados para o diagnóstico de SLF e LFL.....	25
Quadro 2 - Descrição de SPHCs e outras síndromes hereditárias associadas ao desenvolvimento de sarcomas e os genes envolvidos nesse processo.	29
Quadro 3 - Critérios recomendados pela ACMG/AMP para classificação de variantes germinativas.	34
Quadro 4 - Critérios combinados para classificação de variantes germinativas de acordo com as recomendações do ACMG/AMP	35
Quadro 5 - Categorização em Tier das variantes somáticas de acordo com os níveis de evidência.....	36
Quadro 6 - Detalhamento das evidências para classificação de mutações somáticas entre Tier I, II, III e IV.	38
Figura 2 - Esquema dos Níveis de Evidência Terapêutica do OncoKB™ V2	39
Quadro 7 - Genes incluídos nos painéis de 113 e 126 GPCs e sua associação de risco com sarcomas.....	44
Gráfico 1 - Distribuição dos subtipos histológicos de sarcoma de acordo com o diagnóstico desses tumores por faixa etária dos pacientes.	57
Figura 3 - Panorama geral do desenho experimental do estudo, metodologia e achados principais.....	58
Quadro 8 - Dados clínicos e moleculares dos pacientes portadores da variante patogênica <i>TP53</i> :p.Arg337His.	60

Figura 4 - Panorama geral das alterações germinativas patogênicas e provavelmente patogênicas encontradas em nossa coorte (38/177).....	63
Quadro 9 - Síndromes e tumores associados aos genes na qual a relação com o desenvolvimento de sarcomas é ainda incerta ou não descrita.	66
Gráfico 2 - Frequência de detecção de GPVs dentre os subtipos de sarcomas.....	68
Figura 5 - Avaliação do efeito da variante intrônica c.1722+1G>C no transcrito do gene <i>EXT1</i>	73
Gráfico 3 - Frequência de variantes germinativas patogênicas em genes encontrados nessa coorte (177 pacientes com sarcomas) comparado com a coorte de indivíduos idosos saudáveis (1.171 indivíduos brasileiros do banco de dados ABraOM).	76
Quadro 10 - Dados clínicos dos 9 pacientes avaliados por ES.	77
Figura 6 - Fluxograma da anotação e filtragem de variantes na chamada germinativa do sequenciamento de exoma de 9 pacientes.....	78
Figura 7 - Fluxograma da anotação e filtragem de variantes na chamada somática do sequenciamento de exoma de 9 pacientes.....	81
Gráfico 4 - Carga mutacional (TMB) dos 9 tumores avaliados por ES.	88

LISTA DE TABELAS

Tabela 1 - Aspectos clínicos dos 177 pacientes incluídos, comparados de acordo com os grupos de faixa etária (Crianças e Adolescentes e Adultos Jovens).....	56
Tabela 2 - Comparação das características clínicas entre pacientes portadores de GPVs e VUS/Negativos.....	58
Tabela 3 - Descrição de dados clínicos e variantes P e PP detectadas no sequenciamento de painel multigene em 33 pacientes.	64
Tabela 4 - Alterações clínicas de pacientes com 21 VUS de interesse, classificadas como VUS seguindo as recomendações do ACMG ⁸⁶	69
Tabela 5 - Casos avaliados para perda de heterozigose em tecido tumoral.	74
Tabela 6 - Panorama geral das informações clínicas e moleculares dos pacientes avaliados por ES na etapa de priorização de genes.	79
Tabela 7 - Informações gerais das 27 variantes (27 genes) priorizadas utilizando o Franklin Genoox.....	82
Tabela 8 - Informações gerais das variantes somáticas selecionadas no VarSeq.....	85
Tabela 9 - Detalhamento das variantes encontradas em genes priorizados através do VarElect.	86
Tabela 10 - VarElect top 20 genes com associação ao termo sarcoma.	87
Tabela 11 - Variantes somáticas encontradas em genes de associação estabelecida aos sarcomas anotadas pelo Franklin Genoox, em 2 pacientes (n = 9) avaliados por ES.	88

LISTA DE ABREVIATURAS

AAJ	Adolescentes e Adultos Jovens
AD	<i>Allele Depth</i>
AF	<i>Allele frequency</i>
C	Criança
CAC	Carcinoma Adrenocortical
CCR	Câncer colorretal
CGH-array	<i>Comparative genomic hybridization</i>
CMMRD	<i>Constitutional Mismatch Repair Deficiency</i>
CNV	Variação de número de cópias (do inglês <i>Copy number variation</i>)
CPC	Carcinoma de Plexo Coroide
CS	Condrossarcoma
DDR	Resposta ao dano do DNA (do inglês <i>DNA damage response</i>)
DFSP	Dermatofibrossarcoma protuberans
DNA	Ácido Desoxirribonucleico
DP	<i>Read Depth</i>
Dx	Diagnóstico
ES	Sequenciamento de exoma (do inglês <i>Exome sequencing</i>)
ET	Esclerose tuberosa
F	Feminino
FFPE	Fixado em formalina e embebido em parafina (do inglês <i>formalin fixed paraffin embedded</i>)
GIST	Tumor estromal gastrointestinal (do inglês <i>Gastrointestinal stromal tumor</i>)
GPC	Gene de predisposição ao câncer
GPV	<i>Germline pathogenic variant</i>
GST	Gene supressor tumoral
HCEPR	Hipertrofia congênita do epitélio pigmentar da retina
HF	História familiar
HLRCC	<i>Hereditary Leiomyomatosis and Renal Cell Cancer</i>
HMO	Osteocondromas múltiplos hereditários
ID	Identificação da amostra
LFL	Li-Fraumeni-like
LMS	Leiomiossarcoma
LoF	<i>Loss of Function</i>
LOH	Perda de heterozigose (do inglês <i>Loss of Heterozygosity</i>)
LPS	Lipossarcoma
M	Masculino
MAF	<i>Minor allele frequency</i>
MEN2	Neoplasia múltipla endócrina do tipo II
MFS	Mixofibrossarcoma
MLPA	<i>Multiplex Ligation-dependent Probe Amplification</i>

MINAS	Síndrome da Neoplasia Hereditária Multilocus (do inglês <i>Multilocus Inherited Neoplasia Alleles Syndrome</i>)
MMR	<i>Mismatch repair</i>
MPNST	Tumor maligno da bainha do nervo periférico (do inglês <i>Malignant peripheral nerve sheath tumor</i>)
MVA	<i>Mosaic Variegated Aneuploidy</i>
NBS	<i>Nijmegen Breakage Syndrome</i>
ND	Não descrito/não disponível
NE	Não especificado
NF1	Neurofibromatose Tipo 1
NGS	Sequenciamento de Nova Geração (do inglês <i>Next-Generation Sequencing</i>)
OS	Osteossarcoma
P	Patogênica
PCR	Reação em cadeia de polimerase (do inglês <i>Polymerase chain reaction</i>)
PP	Provavelmente Patogênica
RB	Retinoblastoma
RMS	Rabdomiossarcoma
RNA	Ácido ribonucleico
RT-PCR	<i>Reverse transcription-polymerase chain reaction</i>
SAPM	Sarcoma alveolar de partes moles
SCDF	Sarcoma de células dendríticas foliculares
SE	Sarcoma de Ewing
SL	Síndrome de Lynch
SLF	Síndrome de Li-Fraumeni
SNC	Sistema Nervoso Central
SNP-array	<i>Single nucleotide polymorphism</i>
SPHC	Síndrome de Predisposição Hereditária ao Câncer
SS	Sarcoma sinovial
STS	Sarcoma de partes moles (do inglês <i>Soft Tissue Sarcomas</i>)
TFS	Tumor fibroso solitário
TMB	<i>Tumor Mutation Burden</i>
VADS	Carcinomas epidermoides de vias aerodigestivas superiores
VAF	<i>Variant allele frequency</i>
VUS	Variante de significado incerto (do inglês <i>Variant of uncertain significance</i>)
WT	<i>Wild-type</i>

LISTA DE SIGLAS

ABraOM	Arquivo Brasileiro Online de Mutações
ACCCC	A.C. Camargo Cancer Center
ACMG/AMP	American College of Medical Genetics and Genomics
AMP	Association for Molecular Pathology
BWA	Burrows-Wheeler Aligner
BY	Bavária
CA	Califórnia
CEP	Conselho de Ética em Pesquisa
COSMIC	Catalog of Somatic Mutations in Cancer
dbSNP	The Database of Short Genetic Variation
ExAC	Exome Aggregation Consortium
FDA	Food And Drug Administration
GATK	Genome Analysis Toolkit
HGMD	Human Gene Mutation Database
IA	Iowa
IDIBELL	L'Institut d'Investigació Biomèdica de Bellvitge
IIRC	International Intraocular Retinoblastoma Classification
IGV	Integrative Genomics Viewer
MA	Massachusetts
NRW	Renânia do Norte-Vestfália
ON	Ontário
PG	Professional Guideline
PLSD	Prospective Lynch Syndrome Database
SP	São Paulo
TALE	Termo de Assentimento Livre e Esclarecido
TCGA	The Cancer Genome Atlas
TCLE	Termo de Consentimento Livre e Esclarecido
VQSR	Variant Quality Score Recalibration
WHO	World Health Organization

SUMÁRIO

1. INTRODUÇÃO	21
1.1 SARCOMAS.....	22
1.1 SÍNDROMES DE PREDISPOSIÇÃO HEREDITÁRIA ASSOCIADAS A SARCOMAS.....	24
1.1.1 Síndrome de Li-Fraumeni	24
1.1.2 Retinoblastoma Hereditário	26
1.1.3 Neurofibromatose Tipo 1	27
1.1.4 Outras síndromes hereditárias associadas a sarcomas.....	29
1.2 ANÁLISES MOLECULARES PARA A IDENTIFICAÇÃO DE ALTERAÇÕES GENÉTICAS	31
1.3 ALTERAÇÕES GENÉTICAS E SUAS CLASSIFICAÇÕES	32
2. OBJETIVOS.....	40
2.1 OBJETIVOS GERAIS	40
2.1 OBJETIVOS ESPECÍFICOS	40
3. MATERIAIS E MÉTODOS.....	41
3.1 INCLUSÃO DOS PACIENTES	41
3.1 ANÁLISES MOLECULARES.....	42
3.1.1 Extração de DNA germinativo	42
3.1.2 Sequenciamento Sanger ou de amplicon.....	42
3.1.3 Sequenciamento de Nova Geração do painel multigênico.....	43
3.1.1 Testes genéticos em pools de amostras.....	47
3.1.2 Análise e interpretação das variantes	48
3.1.3 Validação das variantes provavelmente patogênicas e patogênicas.....	49
3.1.4 Avaliação de perda de heterozigose (LOH)	49
3.1.5 Análise de transcrito.....	50
3.2 COMPARAÇÃO DA FREQUÊNCIA DAS VARIANTES ENCONTRADAS COM BANCO DE DADOS DE INDIVÍDUOS SAUDÁVEIS.....	50
3.3 ANÁLISE ESTATÍSTICA.....	51
3.4 SEQUENCIAMENTO DE EXOMA	51
4. RESULTADOS	55
4.1 DADOS CLÍNICOS DOS PACIENTES INCLUÍDOS	55
4.2 ANÁLISES MOLECULARES.....	57
4.2.1 Sequenciamento Sanger ou amplicon.....	59
4.2.2 Análise de painel multigênico por NGS	61

4.2.3	Casos com fenótipos não-câncer esperados.....	71
4.2.4	Avaliação de perda de heterozigose (LOH)	73
4.3	COMPARAÇÃO DAS VARIANTES ENCONTRADAS COM BANCOS DE DADOS DE INDIVÍDUOS SAUDÁVEIS.....	76
4.4	SEQUENCIAMENTO DE EXOMA GERMINATIVO E TUMORAL DE PACIENTES SELECIONADOS	77
4.4.1	Análise Germinativa.....	77
4.4.2	Análise Somática.....	81
5.	DISCUSSÃO.....	89
6.	CONCLUSÕES.....	100
7.	REFERÊNCIAS.....	102
8.	ANEXOS.....	115
8.1	ANEXO I: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	115
8.2	ANEXO II: TERMO DE ASSENTIMENTO PARA CRIANÇAS NÃO ALFABETIZADAS E PARA CRIANÇAS E ADOLESCENTES.....	122
8.3	ANEXO III: TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO.....	123
9.	APÊNDICES	124
9.1	MÉTRICAS DE COBERTURA DO ES DE LEUCÓCITO E TUMOR.	124
9.2	INFORMAÇÕES GERAIS SOBRE AS 212 VARIANTES GERMINATIVAS DE PERDA DE FUNÇÃO EM GENES NÃO GPCs ENCONTRADAS EM UMA DAS ETAPAS DE FILTRAGEM E PRIORIZAÇÃO.....	124
9.3	INFORMAÇÕES GERAIS DAS 76 VARIANTES (75 GENES) ENCONTRADAS NO GRY_383 E SELECIONADAS NO VARSEQ.	132
9.4	ARTIGO DOS ACHADOS DO SEQUENCIAMENTO DE PAINEL: EM PREPARO PARA SUBMISSÃO.....	135

1. INTRODUÇÃO

Estima-se que de todos os casos de câncer, 5 a 10% sejam decorrentes de fatores hereditários, o que envolve variantes germinativas patogênicas causadoras das chamadas Síndromes de Predisposição Hereditária ao Câncer (SPHCs). Já os tumores esporádicos, resultantes apenas de alterações somáticas, compreendem a maioria dos casos^{1,2}. Grande parte das SPHCs possui padrão autossômico dominante, no qual se observa que parentes de primeiro grau de um indivíduo afetado possuem um risco de 50% de carregar a variante germinativa causadora da predisposição ao câncer³. Considerando os fatores hereditários, estudos com os mais diversos tipos de câncer buscam fornecer informações que auxiliem na determinação de estratégias para rastreamento, detecção precoce e prevenção do aparecimento da doença em famílias afetadas⁴. Contudo, a possibilidade da existência de aspectos hereditários ainda é um tema muito negligenciado no momento do diagnóstico dos pacientes, provavelmente devido à infrequência de alguns tipos de tumores e riscos pouco estabelecidos de hereditariedade em diversas neoplasias⁵.

Neoplasias malignas raras podem se originar devido a herança genética e, com isso, a busca e o rastreamento para fatores hereditários são indispensáveis, uma vez que a identificação de indivíduos portadores de SPHC permitiria acompanhar clinicamente esses pacientes, objetivando diagnóstico precoce de neoplasias metacrônicas, além de promover medidas de redução de risco para o indivíduo e familiares, e mais recentemente, a indicar terapias alvo específicas para estes tumores⁶. Nesse sentido, aliados ao diagnóstico clínico, são empregadas técnicas moleculares para o rastreamento de variantes genéticas germinativas, as quais existe uma série de genes conhecidos associados as SPHCs. Em sua maioria os genes responsáveis pelas SPHCs são genes supressores de tumor ou proto-oncogenes que atuam no controle do ciclo celular, como: *TP53*, *APC*, *RET*, *RB1*, *NF1* ou em vias de reparo do DNA, como: *BRCA1*, *BRCA2*, *MSH2*, *MLH1*, *MSH6*, *PMS2*, entre outros. No entanto, existem genes mais recentemente descritos, na qual a relação com tumores pertencentes ao espectro das SPHCs ou outros tumores raros ainda é desconhecida³.

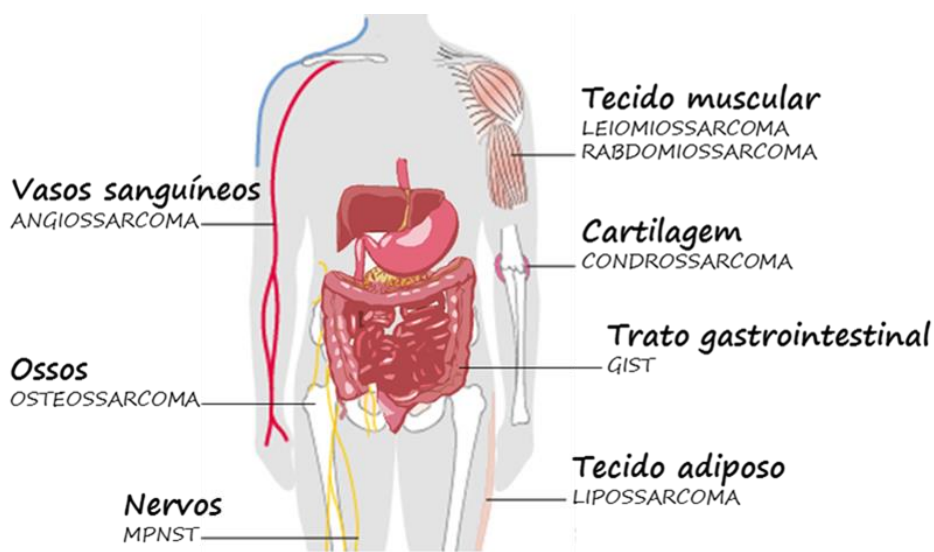
Uma das principais características das SPHCs é a idade precoce ao desenvolvimento de tumores (quando comparada a tumores esporádicos), nas quais alguns cânceres podem se desenvolver predominantemente em adultos jovens (como por exemplo, câncer de ovário hereditário e ou tumores da Síndrome de Lynch); enquanto outros ocorrem no período da

infância (como por exemplo, retinoblastoma hereditário). Ainda, em algumas síndromes, tumores específicos e múltiplos tumores primários podem se desenvolver em crianças e em adultos jovens (como por exemplo, Síndrome de Li-Fraumeni)⁷.

1.1 SARCOMAS

O grupo de tumores raros de origem mesenquimal denominado por sarcoma afeta principalmente a população mais jovem, compreendendo 20% dos tumores em crianças, 10% em adolescentes e adultos jovens^{8,9} e apenas 1% de tumores malignos em adultos^{10,11}. Nesse contexto, devido a sua maior frequência em pacientes jovens, esse grupo de neoplasias malignas tem sido associados às SPHCs. Sua incidência é de < 6 casos por 100.000 indivíduos e, dependendo de características moleculares, ainda pode ser dividido em outro subgrupo, designado sarcomas ultrarraros, que compreendem 20% de todos os tumores de partes moles e ósseos, com uma incidência de ≤ 1 por 1.000.000 de indivíduos por ano. Com isso, existem poucos dados relacionados a biologia, epidemiologia, prognóstico e tratamento dessas neoplasias¹².

O grupo dos sarcomas abrange mais de 100 subtipos histológicos e essa grande heterogeneidade (Figura 1) foi considerada na classificação de tumores de tecidos moles e ósseos publicada pela Organização Mundial da Saúde (WHO, do inglês *World Health Organization*), que se baseou em sua diversidade morfológica e molecular¹³. Nesse trabalho, esses tumores (benignos e malignos) foram divididos de acordo com a característica morfológica do tecido acometido e subdivididos em subtipos, considerando particularidades moleculares. No grupo principal foram nomeados: tumores adipocíticos, fibroblásticos/miofibroblásticos, pseudo fibro-histiocíticos, vasculares, pericíticos (perivasculares), tumores do músculo liso, tumores do músculo esquelético, tumores estromais gastrointestinais (GISTs), tumores condro-ósseos, tumores da bainha do nervo periférico, tumores de diferenciação incerta e sarcomas indiferenciados de pequenas células redondas do osso e tecidos moles¹⁴.



Fonte: Ilustração adaptada de Damerell, Pepper e Prince (2021)¹⁵.

Legenda: MPNST – tumor maligno da bainha do nervo periférico; GIST – tumor estromal gastrointestinal.

Figura 1 - Representação topográfica do desenvolvimento de sarcomas no corpo humano.

A etiologia da maior parte dos sarcomas é desconhecida, porém sabe-se que a exposição prolongada à radiação ionizante aumenta o risco do aparecimento de tumores ósseos^{8,16}. Dentre os fatores hereditários, indivíduos portadores de SPHC parecem apresentar maior susceptibilidade ao aparecimento de câncer quando expostos a radiação (como em pacientes jovens submetidos ao tratamento de radioterapia), devido a alterações em genes supressores de tumor (GST)¹⁷⁻¹⁹. Ainda, existem estudos que associam algumas SPHCs diretamente ao desenvolvimento de sarcomas, sendo as mais conhecidas: Síndrome de Li-Fraumeni, Retinoblastoma Hereditário e Neurofibromatose Tipo I²⁰.

O diagnóstico histopatológico desses tumores ainda é desafiador, dado o seu perfil heterogêneo e sua raridade, com uma discordância diagnóstica de aproximadamente 30%²¹. No entanto, análises de imuno-histoquímica, de testes genéticos moleculares e de citogenética são técnicas importantes para avaliação e descrição dos sarcomas²². Ainda, a compreensão de mecanismos envolvidos no desenvolvimento desses tumores, além de determinar o diagnóstico, atualmente é importante para a definição terapêutica específica²³. Alterações citogenéticas são comuns em sarcomas²⁴ e, nesse contexto, esses tumores são divididos em dois grupos: sarcomas com alterações genéticas específicas e de cariótipo simples; e sarcomas com alterações genéticas desconhecidas e de cariótipo complexo. Dentre as alterações genéticas específicas, a maior parte é decorrente de translocações

cromossômicas (na qual geram genes de fusão), sendo observadas em um terço de todos os sarcomas^{25,26}.

O tratamento preconizado para os sarcomas consiste no controle local da doença com ressecção da neoplasia com uma adequada margem de segurança, associada ou não a radioterapia, e no uso de drogas de forma sistêmica (quimioterapia e imunoterapia)¹¹. A doxorubicina, é, na maioria dos casos, o quimioterápico de escolha, podendo ser combinado a outros agentes ou não²⁷. Atualmente, existem estudos que se baseiam no uso de terapias-alvo, como observado no GIST, no qual são utilizados inibidores de tirosina quinase²⁸, evidenciando a relevância de se conhecer as bases moleculares destes tumores. A sobrevida dos pacientes varia de acordo com os fatores prognósticos, como o grau de diferenciação da neoplasia, tamanho e a localização da lesão primária no momento do diagnóstico. Pacientes com tumores localizados, de baixo grau e que permitem procedimentos de ressecção cirúrgica completa (com margens livres e amplas), combinados ou não a terapias adjuvantes, possuem maior sobrevida em cinco anos, quando comparados aos pacientes que apresentam neoplasias com características biológicas mais agressivas^{29,30}.

1.1 SÍNDROMES DE PREDISPOSIÇÃO HEREDITÁRIA ASSOCIADAS A SARCOMAS

Existem SPHCs nas quais o aumento do risco para sarcomas é conhecido e critérios clínicos para diagnóstico e rastreamento genético de famílias afetadas são usualmente empregados, como, por exemplo, para pacientes portadores de Síndrome de Li-Fraumeni (SLF). Entretanto, dada à raridade da maioria das SPHCs, a relação com o diagnóstico de sarcomas, na maioria das vezes, é desconsiderada, fazendo com que a frequência dessa associação ainda seja desconhecida.

1.1.1 Síndrome de Li-Fraumeni

A SLF é uma síndrome autossômica dominante rara, caracterizada por alterações germinativas detectadas no gene *TP53* (GST), que determinam alto risco de desenvolvimento precoce de tumores malignos primários. Os tumores predominantemente observados são sarcomas, neoplasias malignas de mama e sistema nervoso central, leucemia, câncer de laringe e pulmão, e câncer adrenocortical^{31,32}. Entretanto, diversos outros tumores já foram

relatados em portadores dessa síndrome, como melanoma, câncer colorretal e de pâncreas^{3,33}. Em famílias portadoras de SLF, a realização de exames periódicos é extremamente importante, assim como o manejo ambiental, com a diminuição do contato com fatores carcinógenos e a vigilância familiar, sendo condutas que auxiliam na diminuição da morbidade e mortalidade³⁴. Os critérios clínicos para diagnóstico de SLF tem sido modificados ao longo do tempo, entretanto, o critério clínico clássico descrito por Li e Fraumeni (1988)³⁵ ainda é amplamente utilizado e foram surgindo outros estudos complementares que abordavam critérios para pacientes sem o fenótipo clássico da doença, nomeados de Li-Fraumeni-like (LFL), mas que apresentavam alteração no gene *TP53*³⁶⁻³⁹ (Quadro 1).

Quadro 1 - Critérios clínicos utilizados para o diagnóstico de SLF e LFL.

CRITÉRIO CLÁSSICO – SLF
Paciente diagnosticado com sarcoma antes de 45 anos E
Parente de 1º grau com qualquer tipo de câncer antes de 45 anos E
Parente de 1º ou 2º grau com qualquer tipo de câncer antes de 45 anos OU sarcoma em qualquer idade
LFL (Birch et al., 1994)³⁶
Câncer pediátrico ou sarcoma, tumor SNC ou CAC antes de 45 anos E
Parente de 1º ou 2º grau com tumores do espectro da SLF* em qualquer idade E
Parente de 1º ou 2º grau com qualquer tipo de câncer antes de 60 anos
CRITÉRIO DE CHOMPRET MODIFICADO – LFL
Probando com fenótipo clássico de SLF* antes de 46 anos E
Parente de 1º ou 2º grau com fenótipo clássico de SLF antes de 56 anos (exceto câncer de mama, se o probando possui diagnóstico desse câncer) OU
Múltiplos tumores em qualquer idade OU
Probando com múltiplos tumores primários (exceto múltiplos tumores de mama), sendo ao menos dois deles pertencentes ao espectro de tumores da SLF e o primeiro tumor diagnosticado antes de 46 anos OU
Probando com CAC ou CPC ou RMS pleomórfico em qualquer idade e independente de HF OU
Câncer de mama antes de 31 anos sem alterações em <i>BRCA1/2</i>

Fonte: adaptado de Alcalde et al., 2021⁷.

Nota: * sarcoma, câncer de mama, tumor do sistema nervoso central, carcinoma adrenocortical, leucemia, carcinoma broncoalveolar. SNC – sistema nervoso central; CAC – carcinoma adrenocortical; CPC – carcinoma de plexo coroide; RMS – rabiomiossarcoma; HF – história familiar.

No Brasil, existe uma variante no gene *TP53* considerada como uma variante fundadora, conhecida como *TP53:p.Arg337His* e quando comparado a população mundial (frequência alélica menor – MAF = 0,000012 no gnomAD_v2 1.1), é mais frequentemente detectada na população brasileira (MAF = 0,001, observada no estudo de Custódio et al. (2013)⁴⁰ através de rastreamento em recém-nascidos do Paraná). Ela pode ser encontrada em

pacientes que não preenchem os critérios clínicos clássicos da SLF, sendo que há um o variado risco e espectro de penetrância dessa variante para diferentes tumores. Dos tumores observados, enquanto alguns indivíduos podem apresentar somente tumores adrenocorticais, outros podem desenvolver variados tipos de tumores, como carcinoma de plexo coroide, câncer de mama, pulmão, tireoide e sarcomas⁴¹⁻⁴³. Em relação a sarcomas, Volc et al. (2020)⁴⁴ avaliaram 502 indivíduos brasileiros portadores dessas neoplasias não selecionados para história familiar e verificaram que 8% deles apresentavam a variante *TP53*:p.Arg337His. Além disso, observaram que o subtipo histológico de sarcoma de partes moles mais frequente nesse grupo foi o leiomisossarcoma (52,5%).

Recentemente, Pinto et al. (2020)⁴⁵ identificaram uma variante rara no gene *XAF1* como um modificador de fenótipo em pacientes portadores da variante p.Arg337His. A variante *XAF1*:p.Glu134Ter (MAF = 0,0035, gnomAD_v2 1.1) compartilha o mesmo haplótipo da variante *TP53*:p.Arg337His e portadores de ambas as variantes parecem desenvolver um fenótipo mais agressivo da SLF, apresentando maior risco para o desenvolvimento de sarcomas e de múltiplas neoplasias.

1.1.2 Retinoblastoma Hereditário

O Retinoblastoma (RB) em sua forma hereditária é uma doença autossômica dominante causada por mutações germinativas no gene *RB1* (GST), que atinge a região da retina (podendo ser bilateral ou unilateral – mais comum) e acomete principalmente crianças. O RB hereditário compreende 30-40% dos casos, atingindo pacientes com menos de um ano de idade, enquanto a forma esporádica compreende 60-70%, ocorrendo em crianças com idade entre dois e cinco anos. No RB hereditário, um alelo apresenta alteração germinativa e o outro alelo, mutação somática. Já no RB esporádico, ambos os alelos são afetados por mutações somáticas. Geralmente, ainda, alterações oculares bilaterais ocorrem na RB hereditária, enquanto as unilaterais têm uma maior frequência na doença esporádica^{46,47}.

Os pacientes pediátricos têm boa resposta terapêutica em relação ao RB, entretanto é observado que, ao longo da vida, podem desenvolver outras neoplasias primárias como osteossarcoma e sarcoma de partes moles (STS, do inglês *soft tissue sarcomas*)^{24,48}. No trabalho de Kleinerman et al. (2005)⁴⁹ foi observado que dentre a coorte de pacientes com idade inferior a 25 anos que receberam radioterapia no tratamento de RB hereditário, 76%

dos tumores secundários que apareceram pós-tratamento radioterápico eram sarcomas. Já entre os pacientes com a mesma faixa etária e RB hereditário não submetidos à radioterapia, 33% dos tumores secundários diagnosticados eram sarcomas. Ketteler et al. (2020)⁵⁰ observaram que após o tratamento radioterápico, os tumores secundários mais frequentes foram os sarcomas, compreendendo 62,5% (25/40) dos tumores avaliados.

O RB esporádico não parece aumentar o risco de desenvolvimento de outros tumores. Além de alteração genética identificada no gene *RB1*, outro fator predisponente para o aparecimento de outros tumores é o tratamento radioterápico para RB, que pode induzir ao câncer através de mutações somáticas em pacientes portadores de RB hereditário. Alguns quimioterápicos também parecem aumentar o risco para novas neoplasias^{48,51}. O risco aumentado para o desenvolvimento de tumores epiteliais como câncer de bexiga, pulmão e mama também tem sido observado em pacientes com diagnóstico prévio de RB⁵².

Para definir a severidade da doença e implementar condutas terapêuticas menos nocivas, foi criado o IIRC (*International Intraocular Retinoblastoma Classification*) que, também, descreve respostas clínicas à quimioterapia sistêmica, auxiliando na definição da forma do tratamento inicial e de decisões posteriores⁵³⁻⁵⁵.

1.1.3 Neurofibromatose Tipo 1

A Neurofibromatose Tipo 1 (NF1) é uma doença autossômica dominante causada por mutação no gene *NF1* (GST), que afeta o crescimento celular de tecidos neurais. A doença possui incidência de 1 em 3.000 indivíduos no mundo, com uma penetrância de quase 100%, na qual os pacientes podem apresentar fenótipos variados⁵⁶. As características da NF1 são definidas por máculas café-com-leite (*café au lait*), neurofibromas, sardas presentes em regiões de dobra cutânea, nódulos de Lisch (hamartomas melanocíticos da íris) e displasias ósseas. Os indivíduos afetados possuem alto risco de desenvolvimento de tumores benignos e malignos, sendo os mais comuns, respectivamente: neurofibroma e tumor maligno da bainha do nervo periférico (MPNST, do inglês *Malignant peripheral nerve sheath tumor*). O risco do desenvolvimento de MPNST em pacientes portadores de NF1 varia de 7 a 13%. Os MPNST são de difícil detecção e podem apresentar metástase pulmonar, hepática, cerebral, em partes moles, óssea, em linfonodos regionais, cutânea e de retroperitônio; e geralmente associados a mau prognóstico⁵⁷⁻⁵⁹.

Em 1988, o *National Institute of Health Consensus Development Conference* determinou critérios para diagnóstico clínico da neurofibromatose⁶⁰. Mais recentemente, esse critério foi revisado por Legius et al. (2021)⁶¹, que estipulou: A) o critério diagnóstico para NF1 é dado por um indivíduo que não tem um parente diagnosticado previamente com NF1 e que possui ao menos duas das seguintes características: ≥ 6 máculas café-au-lait > 5 mm em maior diâmetro em indivíduos pré-puberal e > 15 mm em indivíduos pós-puberal; sardas em região axilar ou inguinal; ≥ 2 neurofibromas de qualquer tipo ou um neurofibroma plexiforme; glioma de via óptica; ≥ 2 nódulos de Lisch identificados através do exame por lâmpada de fenda ou ≥ 2 anomalias coroides definidas por nódulos de manchas brilhantes identificados por tomografia de coerência óptica ou espectroscopia de reflectância no infravermelho próximo; uma lesão óssea distinta como displasia de asa esfenoidal, arqueamento ântero-lateral da tíbia ou pseudoartrose de osso longo; uma variante patogênica heterozigota no gene NF1 com uma frequência do alelo variante de 50% em tecido normal, como as células brancas do sangue. B) o filho de um parente que preenche os critérios diagnósticos especificados no item A recebe o diagnóstico de NF1 se um ou mais critérios do item A estiverem presentes.

Pacientes diagnosticados com NF1 podem apresentar problemas neurológicos, psicológicos, cognitivos, comportamentais, ortopédicos, perda de visão, problemas cardiovasculares e gastrointestinais⁶¹. Ainda não há um tratamento efetivo ou cura para NF1, sendo utilizadas condutas cirúrgicas quando necessário, mas os pacientes vivem com a doença, realizando exames periódicos e sendo alertados sobre a investigação de sinais e sintomas que possam ser sugestivos de outras alterações⁶². Ainda, recentemente foi aprovada uma terapia-alvo com a classe dos inibidores de MEK, como o Selumetinib, que demonstrou eficácia no tratamento de doenças relacionadas ao NF1, como neurofibromas plexiformes inoperáveis. Em outros tumores, como gliomas de baixo grau, também foram observados resultados promissores através de testes clínicos. Apesar disso, os desafios no tratamento dos pacientes com NF1 ainda são grandes, visto a prevalência de diversas neoplasias em crianças, como leucemia e rabdomyosarcomas, e adultos jovens, como MPNST, GIST, câncer de mama, entre outros, o que evidencia ainda mais a importância de estudos com modelos pré-clínicos e clínicos que possibilitam uma maior compreensão da biologia da NF1 para o desenvolvimento de terapias-alvo e outras abordagens combinadas⁶³.

1.1.4 Outras síndromes hereditárias associadas a sarcomas

Além das SPHCs anteriormente descritas, existe a relação do aparecimento precoce de sarcomas a outros tipos de síndromes (Quadro 2), estas principalmente relacionadas a dismorfismos e malformações⁶. Contudo, ainda existe uma escassez de estudos que atribuem aspectos hereditários ao desenvolvimento de sarcomas e o papel dos genes envolvidos nesse contexto.

Quadro 2 - Descrição de SPHCs e outras síndromes hereditárias associadas ao desenvolvimento de sarcomas e os genes envolvidos nesse processo.

Síndrome	Herança	Gene(s)	Sarcoma
Beckwith-Wiedemann	Autossômica dominante	<i>CDKNQC, NSD1, P57Kip</i>	RMS
Bloom	Autossômica recessiva	<i>BLM</i>	RMS, OS
Carney-Stratakis	Autossômica dominante	<i>SDHB, SDHC, SDHD</i>	GIST
CMMRD	Autossômica recessiva	<i>PMS2</i>	RMS
Costello	Autossômica dominante	<i>HRAS</i>	RMS
DICER1	Autossômica dominante	<i>DICER1</i>	RMS
FAP	Autossômica dominante	<i>APC</i>	Tumor desmoide
GIST familiar	Autossômica dominante	<i>KIT, PDGFRA</i>	GIST
Gorlin-Goltz	Autossômica dominante	<i>PTCH1</i>	RMS
SLF	Autossômica dominante	<i>TP53</i>	RMS, OS, STS
HLRCC	Autossômica dominante	<i>FH</i>	LMS
MVA	Autossômica recessiva	<i>BUB1B</i>	RMS
NBS	Autossômica recessiva	<i>NBS1</i>	RMS
Neurofibromatose Tipo 1	Autossômica dominante	<i>NF1</i>	MPNST
Noonan	Autossômica dominante	<i>PTPN11, SOS1</i>	RMS
Retinoblastoma	Autossômica dominante	<i>RB1</i>	OS
Rothmund-Thomson	Autossômica recessiva	<i>REQL4, RECQL4</i>	OS
Rubinstein-Taybi	Autossômica dominante	<i>CREBBP, EP300</i>	RMS, OS
Werner	Autossômica recessiva	<i>WNR</i>	OS

Fonte: Quadro adaptado de Rahner, Steinke (2008)³ e Farid, Ngeow (2016)¹¹.

Legenda: CMMRD – Constitutional Mismatch Repair Deficiency; GIST – Gastrointestinal Stromal Tumor; SLF – Síndrome de Li-Fraumeni; HLRCC – Hereditary Leiomyomatosis and Renal Cell Cancer; MVA – Mosaic Variegated Aneuploidy; NBS – Nijmegen Breakage Syndrome; RMS – rabdmiossarcoma; OS – osteossarcoma; GIST – tumor estromal gastrointestinal (do inglês *gastrointestinal stromal tumor*); STS – sarcoma de partes moles (do inglês *soft tissue sarcomas*); LMS – leiomiiossarcoma; MPNST – tumor maligno da bainhado nervo periférico (do inglês *malignant peripheral nerve sheath tumor*).

Neste sentido, Ballinger e colaboradores (2016)⁹ investigaram a causa genética de sarcomas em 1.162 pacientes maiores de 15 anos e de quatro coortes diferentes, independente do histórico familiar. Neste estudo, foram avaliadas características como a idade média de acometimento – 46 anos (3-93), a presença de múltiplos tumores primários, observada em 15% dos indivíduos, e a relação dos tumores com SPHCs conhecidas, através do

sequenciamento de um painel gênico. Aproximadamente 20% dos pacientes portadores de sarcoma eram portadores de variantes patogênicas em um dos genes pertencentes ao painel de 72 GPCs utilizado no estudo. Destes, foram detectadas alterações patogênicas em genes na qual a relação com aumento de risco para sarcomas já é bem estabelecida, como *TP53* (12/1.162) e *NF1* (6/1.162). Contudo, outros genes também foram identificados com frequência elevada de alterações quando comparados aos indivíduos controle, como *ERCC2*, *ATR*, *BRCA2* e *ATM*, os quais sua relação com o aparecimento de sarcomas era previamente desconhecida. Indivíduos com variantes raras em dois genes ou mais também tinham um risco aumentado comparado a indivíduos com variante patogênica em um único gene.

Em outro estudo, Chan et al. (2017)⁶⁴ avaliaram uma coorte de 66 indivíduos asiáticos com idade inferior a 50 anos, acometidos com sarcomas classificados clinicamente como esporádicos, no qual foi observado que 13,6% dos pacientes carregavam pelo menos uma variante germinativa patogênica em genes como *ATM*, *BRCA2*, *ERCC4*, *FANCC*, *FANCE*, *FANCI*, *MSH6*, *POLE*, *SDHA* e *TP53*. Diferentemente de ambos os estudos, Mirabello et al. (2020)⁶⁵ avaliou 1.244 indivíduos com sarcoma, considerando desde 2 até 80 anos de idade, mas incluindo somente osteossarcomas e realizando análise de sequenciamento completo do exoma nesse grupo de tumores, encontrando 28% dos pacientes com variantes germinativas patogênicas. Nesse estudo, eles também identificaram novos genes candidatos no aparecimento de osteossarcomas, como *CDKN2A*, *MEN1*, *VHL*, *POT1*, *APC* e *ATRX*.

Ainda, um trabalho recente do nosso grupo avaliou indivíduos com diagnóstico de sarcoma e outro tumor primário do espectro da Síndrome de Lynch (SL), a qual é uma SPHC causada por variantes germinativas nos genes de reparo de malpareamento do DNA (MMR). Os sarcomas não são considerados como tumores pertencentes ao espectro da SL, entretanto, existem alguns relatos que os associam a esta síndrome⁶⁶⁻⁶⁸. Nesse trabalho, foram identificados cinco pacientes nos quais o sarcoma foi associado à SL e, além disso, evidenciamos um papel principal do gene *MSH2* como possivelmente associado ao risco de sarcomas⁶⁹. Ainda, essa associação foi posteriormente confirmada em um estudo do *Prospective Lynch Syndrome Database* (PLSD), que identificou que portadores de variantes patogênicas em *MSH2* apresentam risco 42 vezes (5,4-79,3) maior de desenvolverem osteossarcomas comparado a população em geral⁷⁰.

Além dos genes conhecidos de SPHCs, existem relatos de novos genes associados ao risco de desenvolvimento de sarcomas e outros tumores, como *POT1*⁷¹ e *BAP1*⁷². Com isso,

alguns sarcomas podem ser decorrentes de variantes patogênicas em genes não classicamente relacionados a esse grupo de neoplasias e, dessa forma, novas associações ainda não descritas podem existir, especialmente em populações menos estudadas, como a Brasileira.

1.2 ANÁLISES MOLECULARES PARA A IDENTIFICAÇÃO DE ALTERAÇÕES GENÉTICAS

Existe uma variedade de técnicas moleculares que visam o rastreamento de alterações genéticas, conhecidas por: sequenciamento Sanger, MLPA (*Multiplex Ligation-dependent Probe Amplification*), microarranjos genômicos (CGH-array e SNP-array), sequenciamento de Nova Geração (NGS), entre outras⁷³.

Com o avanço tecnológico, o desenvolvimento da técnica de NGS possibilitou uma revolução nos estudos em larga escala genômica⁷⁴. O NGS permite a realização de sequenciamentos do genoma, de éxons ou de transcritos específicos⁷³ e, com isso, é uma técnica bastante utilizada para a detecção de variantes patogênicas conhecidas ou não, relacionadas às SPHC, como observado em diversos estudos⁷⁵⁻⁷⁸. Além disso, o NGS tem mostrado grande eficiência com relação ao tempo de resposta de suas análises e redução de custos, visto que realiza um sequenciamento rápido de milhões de pares de bases de múltiplos indivíduos em uma mesma reação. Por esses motivos, o NGS tem sido amplamente utilizado na medicina translacional, facilitando o diagnóstico e o direcionamento para uma melhor definição terapêutica, assim como o prognóstico. Entretanto, os laboratórios devem trabalhar com a técnica atentando-se às limitações relacionadas à qualidade analítica, como a interpretação das variantes genéticas e possíveis achados incidentais^{76,79}.

Outras abordagens genômicas também podem ser consideradas para pacientes com tumores de início precoce, como o sequenciamento de exoma (ES, do inglês *exome sequencing*) e do genoma completo, os quais há uma maior amplitude de regiões avaliadas, o que possibilita a detecção de alterações que podem ser deletérias ao indivíduo e que não foram encontradas através de outras estratégias. Muitos trabalhos têm utilizado o ES para identificação de novos genes de predisposição a diversos tipos de câncer. Essa estratégia tem sido considerada efetiva para elucidar a causa genética em alguns tumores. Apesar disso, a identificação de novos genes de predisposição ao câncer ainda é complexa, visto a grande dificuldade na interpretação do impacto funcional e clínico das variantes encontradas^{80,81}.

Com isso, a utilização do ES em tecidos somáticos e sequenciamento de RNA pode ser uma alternativa para auxiliar na compreensão do efeito dessas variantes no desenvolvimento de tumores e, ainda, essas abordagens podem permitir a realização de análises como perda de heterozigose (LOH, do inglês *loss of heterozygosity*)⁸⁰ e de alterações transcricionais que modificam a função proteica. Os eventos de LOH podem desempenhar um papel biológico importante em genes supressores tumorais, aos quais ocorre a deleção do alelo selvagem e são frequentemente encontrados em tumores com características hereditárias^{82,83}. Dessa forma, essas estratégias são necessárias na compreensão da participação conjunta de fenômenos somáticos e germinativos no processo de tumorigênese e a combinação de diversas ferramentas pode permitir a elucidação da relevância clínica e funcional das variantes para o risco no desenvolvimento de tumores⁸⁴, assim como o reconhecimento de novos alvos pode possibilitar o aprimoramento e personalização de terapias, e na predição do risco de câncer dos pacientes⁸⁵.

1.3 ALTERAÇÕES GENÉTICAS E SUAS CLASSIFICAÇÕES

A interpretação dos dados gerados do sequenciamento genético é um desafio. Com isso, para facilitar e padronizar a classificação de variantes genéticas encontradas a partir do NGS, foram desenvolvidos critérios que auxiliam na definição da patogenicidade e relevância clínica das alterações detectadas, sejam elas germinativas ou somáticas, sendo que as recomendações são reconhecidas e amplamente utilizadas por laboratórios de diagnóstico molecular.

As alterações germinativas são geralmente classificadas seguindo as recomendações da ACMG/AMP (*American College of Medical Genetics and Genomics/Association for Molecular Pathology*)⁸⁶, na qual as variantes são divididas entre cinco categorias, sendo as Patogênicas (P) e Provavelmente Patogênicas (PP), consideradas de relevância clínica e que podem impactar diretamente no risco de câncer, manejo e tratamento; as Variantes de Significado Incerto (VUS, do inglês *Variant of Uncertain Significance*), as quais não são utilizadas para decisões médicas; e as Provavelmente Benignas e Benignas, variantes que não possuem relevância clínica, em sua maioria são polimorfismos e geralmente não são reportadas em laudos de testes genéticos. Essas classificações são baseadas nos tipos de evidência das variantes, sendo considerados dados populacionais (frequência dessas variantes

na população controle ou indivíduos saudáveis), computacionais (programas de predição *in silico*), funcionais (*in vitro* e *in vivo*), de segregação, dentre outros. No Quadro 3 estão descritos os critérios e suas respectivas pontuações determinadas pela ACMG/AMP. Esses critérios, quando combinados, definem a classificação de cada variante, que envolve as cinco categorias citadas, sendo desde patogênica a benigna. As variantes são classificadas como VUS quando não há a combinação definida entre os critérios ou os critérios para benigna e patogênica forem contraditórios (Quadro 4).

Quadro 3 - Critérios recomendados pela ACMG/AMP para classificação de variantes germinativas.

	Evidência	Dados populacionais	Dados computacionais e preditivos	Dados funcionais	Dados de segregação	Dados <i>de novo</i>	Dados alélicos	Outros bancos de dados	Outros dados
Benigna	Forte (<i>Strong</i>)	MAF na população controle é maior que 5% no Exome Sequencing Project, 1000 Genomes ou ExAC BA1/BS1 OU a variante encontrada em um indivíduo adulto saudável para uma doença recessiva (homozigoto), dominante (heterozigoto) ou ligada ao X (hemizigoto), com penetrância total esperada em idade precoce BS2	ND	Estudos funcionais bem estabelecidos <i>in vitro</i> ou <i>in vivo</i> não demonstram efeitos deletérios sobre a função da proteína ou <i>splicing</i> BS3	Ausência de segregação da doença BS4	ND	ND	ND	ND
	Apoio (<i>Supporting</i>)	ND	Múltiplas linhas de evidência computacional sugerem que não há impacto no gene/produto do gene BP4 ; Missenses em genes onde somente variantes truncadas causam a doença BP1 ; Variante sinônima (silenciosa) sem predição de impacto em regiões de <i>splicing</i> e o nucleotídeo não é altamente conservado BP7	ND	ND	ND	Variante identificada em <i>trans</i> com uma variante patogênica para um gene/doença dominante totalmente penetrante ou observada em <i>cis</i> com uma variante patogênica em qualquer padrão de herança BP2 ; Deleções/inserções <i>in frame</i> em uma região repetitiva sem uma função conhecida BP3	Fontes conceituadas sem os dados compartilhados, mas variante relatada como benigna BP6	Variante encontrada em um caso com causa alternativa para a doença BP5
Patogênica	Apoio (<i>Supporting</i>)	ND	Múltiplas linhas de evidência computacional suportam o efeito deletério no gene/produto do gene PP3	Missense no gene com taxa baixa de variantes missenses benignas e missenses P comuns PP2	Co-segregação com doença em múltiplos membros da família afetados PP1	ND	ND	Fontes conceituadas = patogênica PP5	Fenótipo do paciente ou HF altamente específica para o gene PP4
	Moderada (<i>Moderate</i>)	Ausente em dados populacionais (ou em uma frequência extremamente baixa se recessivo) PM2	Nova alteração do tipo missense em um aminoácido no qual uma alteração missense patogênica diferente já foi descrita PM5 ; Variante que altera o comprimento da proteína, como ocorre em deleções/inserções <i>in frame</i> em regiões não repetitivas ou no caso de variantes de <i>stop loss</i> PM4	Variante em <i>hot spot</i> mutacional e/ou em domínio funcional bem estudado sem variação benigna PM1	Maiores dados de segregação*	Variante <i>de novo</i> (sem confirmação paterna/materna) PM6	Para doenças recessivas, variante detectada em <i>trans</i> com uma variante patogênica conhecida PM3	ND	ND
	Forte (<i>Strong</i>)	Prevalência em afetados estatisticamente mais aumentada do que os controles PS4	Uma mesma alteração no aminoácido já estabelecida como variante patogênica PS1	Estudos funcionais bem estabelecidos mostrando efeito deletério	*	De novo (com confirmação paterna/materna) PS2	ND	ND	ND
	Muito forte (<i>Very Strong</i>)	ND	Variante nula (<i>nonsense</i> , <i>frameshift</i> , <i>splice site</i> , códon de iniciação, deleção de um ou mais éxons) predita em um gene onde LOF é um mecanismo conhecido da doença PVS1	ND	*	ND	ND	ND	ND

Fonte: Retirado e traduzido de Richards et al. (2015)⁸⁶.

Nota: * dados mais fortes de segregação aumentam a evidência de patogenicidade da variante. MAF – *minor allele frequency*; ND – não descrito; HF – história familiar; LOF – *loss of function*.

Quadro 4 - Critérios combinados para classificação de variantes germinativas de acordo com as recomendações do ACMG/AMP.

Patogênica
(i) 1 Muito forte (PVS1) E (a) ≥ 1 Forte (PS1 – PS4) OU (b) ≥ 2 Moderado (PM1 – PM6) OU (c) 1 Moderado (PM1 – PM6) e 1 Apoio (PP1 – PP5) OU (d) ≥ 2 Apoio (PP1 – PP5) (ii) ≥ 2 Forte (PS1 – PS4) OU (iii) 1 Forte (PS1 – PS4) E (a) ≥ 3 Moderado (PM1 – PM6) OU (b) 2 Moderado (PM1 – PM6) E ≥ 2 Apoio (PP1 – PP5) OU (c) 1 Moderado (PM1 – PM6) E ≥ 4 Apoio (PP1 – PP5)
Provavelmente patogênica
(i) 1 Muito forte (PVS1) E 1 Moderado (PM1 – PM6) OU (ii) 1 Forte (PS1 – PS4) E 1-2 Moderado (PM1 – PM6) OU (iii) 1 Forte (PS1 – PS4) E ≥ 2 Apoio (PP1 – PP5) OU (iv) ≥ 3 Moderado (PM1 – PM6) OU (v) 2 Moderado (PM1 – PM6) E ≥ 2 Apoio (PP1 – PP5) OU (vi) 1 Moderado (PM1 – PM6) E ≥ 4 Apoio (PP1 – PP5)
Benigna
(i) 1 Autônomo (BA1) OU (ii) ≥ 2 Forte (BS1 – BS4)
Provavelmente benigna
(i) 1 Forte (BS1 – BS4) e 1 Apoio (BP1 – BP7) OU (ii) ≥ 2 Apoio (BP1 – BP7)

Fonte: Retirado e traduzido de Richards et al. (2015)⁸⁶.

A classificação ACMG/AMP de 2015 tem sido amplamente utilizada em laboratórios clínicos e de pesquisa, e após essa publicação, diversos trabalhos de grupos de especialistas da própria ACMG e do grupo ClinGen foram sendo publicados para promover ajustes e melhorias na interpretação de cada critério, como ajustes para os critérios PVS1⁸⁷, BA1⁸⁸, PS3/BS3⁸⁹, entre outros, e também para tipos específicos de alterações, como alterações de número de cópias⁹⁰ ou para genes específicos como *CDH1*⁹¹, *TP53*⁹² e *PTEN*⁹³.

Para a classificação de mutações somáticas, foram criadas recomendações pela AMP, juntamente com a Sociedade Americana de Oncologia Clínica (do inglês *American Society of Clinical Oncology*) e o Colégio Americano de Patologistas (do inglês *College of American Pathologists*)⁹⁴. Essas recomendações são importantes para que a interpretação das variantes reportadas seja uniforme entre os laboratórios, visto que o resultado do NGS geralmente evidencia um grande número de variantes genéticas detectadas no DNA tumoral e a informação que muitas delas carrega é fundamental para definir uma abordagem mais

personalizada ao paciente, assim como possibilitar o desenvolvimento de diretrizes de tratamento com base em achados mais específicos. Nesse sentido, são considerados níveis de evidência para as variantes somáticas, que irão determinar se essas variantes possuem impacto clínico relevante ou não, baseado em sua significância no diagnóstico do câncer, no prognóstico e/ou terapia, na qual a classificação é dividida entre quatro categorias Tier: Tier I, variantes com forte significância clínica (níveis de evidência A e B); Tier II, variantes com potencial significado clínico (níveis de evidência C ou D); Tier III, variantes com significado clínico desconhecido; e Tier IV, variantes que são benignas ou provavelmente benignas (Quadros 5 e 6).

Quadro 5 - Categorização em Tier das variantes somáticas de acordo com os níveis de evidência.

Tier I: Variantes com Forte Significância Clínica (<i>Terapêutica, prognóstico e diagnóstico</i>)
Evidência nível A: Terapia com aprovação pelo FDA, inclusa em diretrizes profissionais
Evidência nível B: Estudos bem estabelecidos com consenso de experts na área
Tier II: Variantes com Potencial Significado Clínico (<i>Terapêutica, prognóstico e diagnóstico</i>)
Evidência nível C: Terapias aprovadas pelo FDA para diferentes tipos de tumores ou terapias investigativas; Múltiplos estudos menores publicados com algum consenso
Evidência nível D: Testes pré-clínicos ou alguns relatos de caso sem consenso
Tier III: Variantes com Significado Clínico Desconhecido
Não observada em uma frequência alélica significativa em bancos de dados da população geral ou subpopulações específicas, ou pan-cancer, ou bancos de dados com variantes tumor-específicas; Ausência de evidências convincentes de associação com câncer publicadas
Tier IV: Variantes Benignas ou Provavelmente Benignas
Observada em uma frequência alélica significativa na população geral ou subpopulações específicas; Ausência de evidências de associação com câncer publicadas

Fonte: Retirado e traduzido de Li et al. (2017)⁹⁴.

Legenda: FDA – *Food and Drug Administration*, Agência Federal.

Além do modelo de Tier para classificação de mutações somáticas, existem algumas outras ferramentas que são curadas e utilizadas para definir o impacto clínico das alterações detectadas. O banco de dados OncoKB™ é um banco de oncologia de precisão reconhecido pelo FDA, no qual possui informações biológicas e clínicas sobre as alterações genômicas no

câncer. Foi desenvolvido por Chakravarty et al. (2017)⁹⁵, no qual os eventos somáticos são classificados utilizando um sistema de Níveis de Evidência, o qual determina o desfecho clínico das alterações (Figura 2). Resumidamente, dados sobre os genes e as variantes encontradas em cada gene respectivo são obtidos através de relatos das variantes em bancos de dados confiáveis e informações estatísticas, além de considerar diretrizes de tratamento e de literatura científica. Com isso, é possível estabelecer o efeito mutacional e oncogênico das variantes, além da sua correlação com o gene afetado e com o tipo de tumor envolvido. Ainda, as implicações clínicas dessas correlações também são descritas, para as quais são avaliadas as implicações preditivas e prognósticas e a existência de terapias convencionais e/ou investigativas. Todos esses dados são revisados por cientistas e ao final dessa curadoria, os dados se tornam publicamente disponíveis no *website*. As informações são frequentemente revisadas e atualizadas.

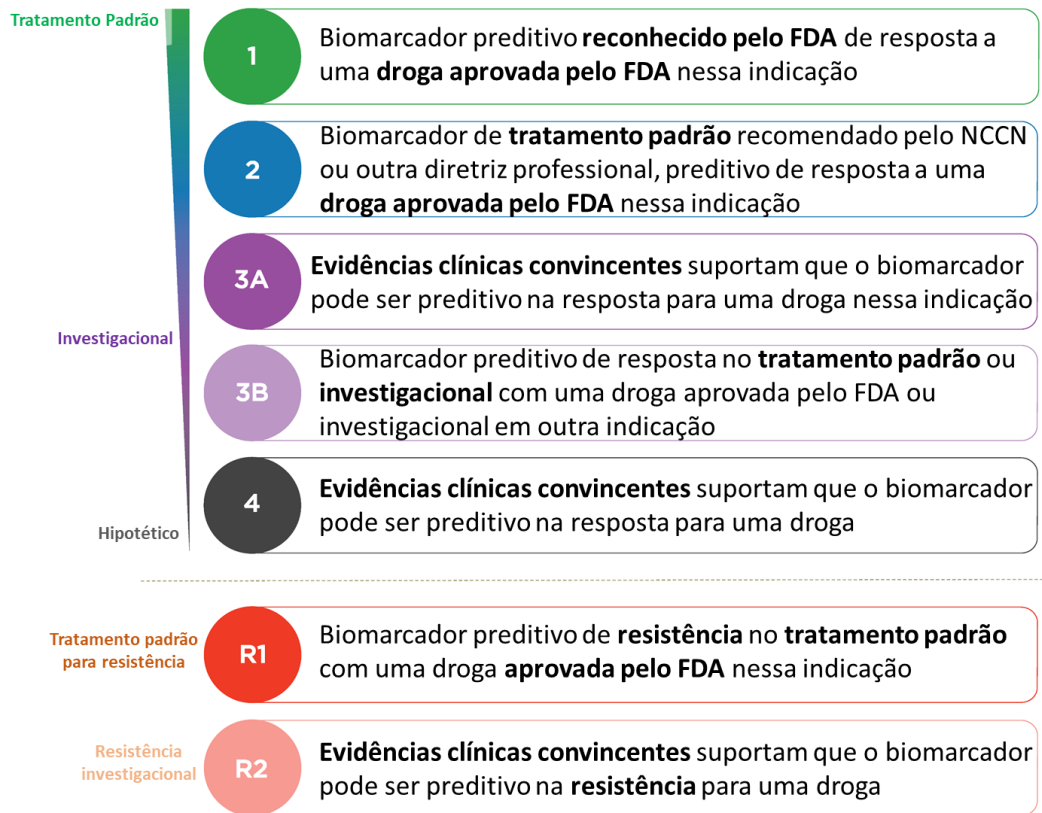
Quadro 6 - Detalhamento das evidências para classificação de mutações somáticas entre Tier I, II, III e IV.

		Fontes/tipo de evidência									
Terapias aprovadas pelo FDA, PG, terapias investigacionais		Tipo de mutação	Frequência das variantes	Potencial germinativo	Banco de dados populacional: <i>ESP, dbSNP, 100Genome, ExAC</i>	Banco de dados germinativo: <i>HGMD, ClinVar</i>	Banco de dados somático: <i>COSMIC, My Cancer Genome, TCGA</i>	Programas preditivos: <i>SIFT, PolyPhen2, MutTaster, CADD</i>	Envolvimento de vias	Publicações: estudo funcional, estudo populacional, outros	
Evidência disponível	Tier I	Terapêutica: Aprovada pelo FDA ou investigacional com forte evidência; Diagnóstica: em PG ou evidência reportada com consenso; Prognóstica: em PG ou evidência reportada com consenso	Ativadora, LOF (missense, nonsense, indel, splicing), CNVs, fusões	Maioria mosaico	Maioria não mosaico (VAF de aproximadamente 50% ou 100%)	MAF ausente ou extremamente baixa	Pode ou não estar presente	Presença muito provável	Maioria deletéria: informação para ser usada apenas como referência	Vias associadas a doenças	Terapêutica: evidência reportada com consenso; Diagnóstica: evidência reportada com consenso; Prognóstica: evidência reportada com consenso
	Tier II	Terapêutica: Aprovada pelo FDA para diferentes tipos tumorais; terapias investigacionais com alguma evidência; Diagnóstica: Ausente em PG, mas com dados de publicação acurados; Prognóstica: Ausente em PG, mas com dados de publicação acurados	Ativadora, LOF (missense, nonsense, indel, splicing), CNVs, fusões	Maioria mosaico	Maioria não mosaico (VAF de aproximadamente 50% ou 100%)	MAF ausente ou extremamente baixa	Pode ou não estar presente	Presença provável	Maioria deletéria: informação para ser usada apenas como referência	Envolve vias associadas a doenças ou vias patogênicas	Terapêutica: evidência do uso de terapias aprovadas pelo FDA para diferentes tipos de tumor; testes clínicos de fase 2 ou 3 para terapias investigacionais; Diagnóstica: múltiplas linhas de evidências reportadas sem consenso; Prognóstica: múltiplas linhas de evidência reportadas sem consenso
	Tier III	Genes associados a câncer: nenhum; Genes não associados a câncer (aplicado ao sequenciamento de exoma no câncer): nenhum	Funcionalmente desconhecida; maioria missense, in-frame indels; outros tipos, menos comum	Mosaico ou não mosaico	Maioria não mosaico (VAF de aproximadamente 50% ou 100%)	MAF ausente ou extremamente baixa	Ausente ou reclassificada de patogênica para VUS	Ausente ou presente sem associação para tumores específicos (VUS potencialmente germinativa); presente, mas em raros casos	Variável; informação para ser usada somente como referência	Pode ou não envolver vias associadas a doenças	Nenhuma ou evidências não acuradas para determinar significância clínica/biológica
	Tier IV	Nenhuma	Funcionalmente benigna ou desconhecida; maioria missense; outros tipos, menos comum	Maioria não mosaico (VAF de aproximadamente 50% ou 100%)	Maioria não mosaico (VAF de aproximadamente 50% ou 100%)	MAF >= 1% na população geral; ou MAF maior em algumas etnias	Ausente ou presente, mas reclassificada para benigna/provavelmente benigna	Ausente ou presente sem associação para tumores específicos (potencial polimorfismo germinativo raro)	Maioria benigna; informação para ser usada somente como referência	Pode ou não envolver vias associadas a doenças	Evidência reportada suporte para benigna/provavelmente benigna; ou nenhuma

Fonte: Retirado e traduzido de Li et al. (2017)⁹⁴.

Notas: os bancos de dados e programas preditivos são apenas alguns exemplos dentre vários outros amplamente utilizados. Para as variantes com potencial germinativo testadas somente no tumor, deve-se ter a confirmação no tecido normal e o aconselhamento genético deve ser recomendado. CNV – *copy number variation*; COSMIC – Catalog of Somatic Mutations in Cancer; dbSNP – The Database of Short Genetic Variation; ExAC – Exome Aggregation Consortium; FDA – Food and Drug Administration; HGMD – Human Gene Mutation Database; indel – inserção e deleção; LOF – *loss of function*; MAF – *minor allele frequency*; PG – *professional guideline*; TCGA – The Cancer Genome Atlas; VAF – *variant allele frequency*.

Figura 2 - Esquema dos Níveis de Evidência Terapêutica do OncoKB™ V2.



Fonte: Retirado e traduzido de Chakravarty et al. (2017)⁹⁵.

2. OBJETIVOS

2.1 OBJETIVOS GERAIS

Identificar as causas genéticas e vias moleculares associadas ao desenvolvimento de sarcomas em crianças, adolescentes e adultos jovens.

2.1 OBJETIVOS ESPECÍFICOS

- Identificar variantes genéticas germinativas pontuais e variações no número de cópias em genes de predisposição ao câncer em pacientes jovens portadores de sarcoma;
- Investigar a relação genótipo-fenótipo de pacientes com variantes germinativas patogênicas em genes de associação incerta com sarcomas através da análise de perda de heterozigose;
- Comparar a frequência de variantes patogênicas detectadas na coorte com sarcomas com a frequência dessas variantes em bancos de dados de indivíduos saudáveis;
- Buscar novos genes de predisposição a sarcomas em pacientes com condrossarcomas ou sarcomas de histologia ultrarraras através da análise integrada de sequenciamento de exoma germinativo e tumoral de pacientes selecionados negativos para variantes patogênicas.

3. MATERIAIS E MÉTODOS

3.1 INCLUSÃO DOS PACIENTES

Nesse estudo, foram incluídos 177 indivíduos que apresentaram diagnóstico de sarcoma no A.C. Camargo Cancer Center entre os anos 2005 e 2021, considerando os seguintes critérios: 1. pacientes com idade inferior a 40 anos; 2. pacientes sem diagnóstico prévio de síndromes genéticas associadas à dismorfismos e malformações.

Com isso, 141 pacientes foram incluídos de maneira retrospectiva, cujas amostras de sangue periférico e tecido tumoral já se encontravam armazenadas no Biobanco do A.C. Camargo, e foram priorizados pacientes com osteossarcomas e condrossarcomas. Adicionalmente, de 2018 a 2021, trinta e seis pacientes foram incluídos prospectivamente, contatando o departamento de Oncogenética, o departamento de Oncologia Pediátrica e demais departamentos clínicos do ACCCC que realizam a assistência de pacientes com sarcomas. Os pacientes elegíveis prospectivamente foram encaminhados ao Departamento de Oncogenética para aconselhamento genético pré-teste e assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) (Anexo I). Os TCLEs para pacientes menores de 18 anos foram assinados por um dos pais ou responsável e os pacientes expressaram sua concordância em participar do estudo pela apresentação do Termo de Assentimento Livre e Esclarecido (TALE) (Anexo II). Uma vez que este projeto é parte do projeto Departamental intitulado “Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens” (CEP 2505/18), foram utilizados o TCLE e TALE referentes a este projeto. O projeto possui aprovação do CEP 2546/18, CAAE 84725318.4.0000.5432. Os resultados do teste estão sendo retornados aos pacientes em consulta de aconselhamento pós-teste. Para os pacientes retrospectivos de amostras obtidas no Biobanco, os quais já assinaram o TCLE padrão do Biobanco, apenas os resultados em que foram identificadas variantes patogênicas (P) ou provavelmente patogênicas (PP) estão sendo devolvidos aos pacientes em consulta de aconselhamento pós-teste.

Para coleta de dados clínicos dos pacientes incluídos, foram priorizadas as seguintes informações extraídas de prontuários médicos eletrônicos: histórico familiar de neoplasias, se

há ou não presença de diagnóstico clínico de alguma SPHC, histologia tumoral (sarcoma), se há ou não presença de outros tumores, identificação e topografia tumoral.

3.1 ANÁLISES MOLECULARES

3.1.1 Extração de DNA germinativo

O DNA genômico foi extraído a partir do leucócito de sangue periférico das amostras armazenadas no Biobanco do A.C. Camargo, de acordo com seu procedimento padrão no equipamento QiaSymphony (Qiagen, NRW, Alemanha). Para os casos incluídos prospectivamente, o DNA foi obtido a partir de saliva, seguindo o protocolo de purificação do kit PrepIT-L2P (PT-L2P) (DNA Genotek Inc., ON, Canadá). Os DNAs genômicos foram quantificados através do espectrofotômetro NanoDrop® ND-1000 (ThermoFisher Scientific, MA, Estados Unidos) para obtenção de valores de pureza e quantidade de DNA presente na amostra e Qubit 4 Fluorometer (ThermoFisher Scientific) seguindo os protocolos dos fabricantes.

3.1.2 Sequenciamento Sanger ou de amplicon

Inicialmente, foi realizado sequenciamento Sanger para identificação da variante germinativa patogênica p.Arg337His no gene *TP53*. Para isso, foi realizada a amplificação do DNA genômico dos pacientes selecionados através da reação em cadeia da polimerase (PCR, do inglês *polymerase chain reaction*) com iniciadores específicos para a região do éxon 10 do gene *TP53*, sendo o *forward* CCGTCATAAAGTCAAACAATT e o *reverse* GAATGGAATCCTATGGCTTTC. Posteriormente, o produto amplificado foi sequenciado com o kit BigDye®Terminator v3.1 (Applied Biosystems, SP, Brasil) e sequenciador Automático de DNA, ABI Prism 3130 XL (Applied Biosystems), seguindo as instruções do fabricante do aparelho. O resultado do sequenciamento foi analisado utilizando a ferramenta CLCbioMain Workbench (CLC Bio, Aarhus, Dinamarca). Para uma parte das amostras, ao invés de Sanger, foi realizado NGS de *amplicon* (método descrito no item 3.2.6). Adicionalmente, para os pacientes em que foi detectada a variante *TP53*:p.Arg337His, foi investigada também a presença da variante germinativa *XAF1*:p.Glu134Ter, utilizando a técnica de NGS de *amplicon*.

3.1.3 Sequenciamento de Nova Geração do painel multigênico

Posteriormente, amostras negativas para a variante p.Arg337His no *TP53* foram avaliadas por sequenciamento NGS utilizando um painel customizado de 113 ou 126 genes de predisposição ao câncer, no qual os GPCs foram selecionados baseados no estudo de Rahman (2014)⁹⁶, de painéis comerciais de laboratórios clínicos e através do banco de dados PanelApp⁹⁷⁻⁹⁹. O painel de 126 genes compreendeu os genes incluídos no painel de 113 GPCs para tumores sólidos somado a mais 13 outros GPCs envolvidos em tumores hematológicos. Nessa etapa, foram utilizadas duas estratégias para a captura e sequenciamento dos genes alvo, sendo a primeira realizada com amostras individuais e a segunda com amostras organizadas em *pools* (detalhado no item 4.5). Em ambas, o DNA genômico (individual ou em *pool*) foi fragmentado por digestão enzimática e ligado a adaptadores e barcodes específicos com o kit QIAseq FX DNA Library Kit (Qiagen). A captura das regiões alvo foi realizada através de hibridização com sondas biotinizadas e foi utilizado o painel de 113 ou 126 genes (xGen Gene Capture Pools - IDT Technologies), seguindo as recomendações do fabricante. O sequenciamento foi realizado na plataforma NextSeq 500 (Illumina, CA, Estados Unidos), utilizando os kits NextSeq 500/550 High Output Kit v2 (300 cycles) ou NextSeq 500/550 Mid Output Kit v2 (300 cycles), visando obter uma cobertura mínima de 30X (amostras individuais) ou 200X (*pools* de amostras) em 99 a 100% das bases alvo.

Quadro 7 - Genes incluídos nos painéis de 113 e 126 GPCs e sua associação de risco com sarcomas.

Gene	Painel	Relação com sarcomas	Referência	Gene	Painel	Relação com sarcomas	Referência
<i>ACD</i>	126	Não descrita	Não disponível	<i>MEN1</i>	113/126	Não descrita	Não disponível
<i>AIP</i>	126	Não descrita	Não disponível	<i>MET</i>	113/126	Não descrita	Não disponível
<i>AKT1</i>	113/126	Não descrita	Não disponível	<i>MITF</i>	113/126	Não descrita	Não disponível
<i>ALK</i>	113/126	Não descrita	Não disponível	<i>MLH1</i>	113/126	Moderada (recessivo)	PanelApp ⁹⁷
<i>APC</i>	113/126	Emergente (dominante)	Mirabello et al. (2020) ⁶⁵	<i>MLH3</i>	126	Não descrita	Não disponível
<i>ATM</i>	113/126	Emergente (dominante)	Ballinger et al. (2016) ⁹ ; Chan et al. (2017) ⁶⁴	<i>MRE11A</i>	126	Não descrita	Não disponível
<i>AXIN2</i>	113/126	Não descrita	Não disponível	<i>MSH2</i>	113/126	Moderada (recessivo)	PanelApp ⁹⁷
<i>BAP1</i>	113/126	Emergente (dominante)	Kaszuba et al. (2018) ⁷²	<i>MSH3</i>	113/126	Não descrita	Não disponível
<i>BARD1</i>	113/126	Não descrita	Não disponível	<i>MSH6</i>	113/126	Moderada (recessivo)	PanelApp ⁹⁷
<i>BLM</i>	113/126	Moderada (recessivo)	PanelApp ⁹⁷	<i>MSR1</i>	113/126	Não descrita	Não disponível
<i>BMPR1A</i>	113/126	Não descrita	Não disponível	<i>MTAP</i>	113/126	Não descrita	Não disponível
<i>BRCA1</i>	113/126	Não descrita	Não disponível	<i>MUTYH</i>	113/126	Não descrita	Não disponível
<i>BRCA2</i>	113/126	Emergente (dominante)	Ballinger et al. (2016) ⁹ ; Chan et al. (2017) ⁶⁴	<i>NBN</i>	113/126	Moderada (recessiva)	PanelApp ⁹⁷
<i>BRIP1</i>	113/126	Não descrita	Não disponível	<i>NF1</i>	113/126	Definitiva (dominante)	GeneReviews ⁹⁸
<i>BUB1</i>	113/126	Não descrita	Não disponível	<i>NF2</i>	113/126	Não descrita	Não disponível
<i>BUB1B</i>	113/126	Moderada (recessivo)	PanelApp ⁹⁷	<i>NSD1</i>	113/126	Não descrita	Não disponível
<i>BUB3</i>	113/126	Não descrita	Não disponível	<i>NTHL1</i>	113/126	Não descrita	Não disponível
<i>CASR</i>	113/126	Não descrita	Não disponível	<i>PALB2</i>	113/126	Não descrita	Não disponível
<i>CDC73</i>	113/126	Não descrita	Não disponível	<i>PDGFRA</i>	113/126	Definitiva (dominante, GIST)	Gene Curation Coalition ⁹⁹
<i>CDH1</i>	113/126	Não descrita	Não disponível	<i>PHOX2B</i>	113/126	Não descrita	Não disponível
<i>CDK4</i>	113/126	Não descrita	Não disponível	<i>PMS2</i>	113/126	Moderada (recessivo)	PanelApp ⁹⁷
<i>CDKN1B</i>	113/126	Não descrita	Não disponível	<i>POLD1</i>	113/126	Não descrita	Não disponível
<i>CDKN1C</i>	113/126	Moderada (dominante)	PanelApp ⁹⁷	<i>POLE</i>	113/126	Não descrita	Não disponível
<i>CDKN2A</i>	113/126	Emergente	Mirabello et al. (2020) ⁶⁵	<i>POLH</i>	113/126	Não descrita	Não disponível

CEBPA	126	Não descrita	Não disponível	POT1	113/126	Emergente (dominante)	Calvet et al. (2015) ⁷¹
CHEK2	113/126	Definitiva (dominante)	GeneReviews ⁹⁸	PRKAR1A	113/126	Não descrita	Não disponível
CREBBP	113/126	Emergente (dominante)	GeneReviews ⁹⁸	PRSS1	126	Não descrita	Não disponível
CTNNA1	113/126	Não descrita	Não disponível	PTCH1	113/126	Não descrita	Não disponível
CTR9	113/126	Não descrita	Não disponível	PTEN	113/126	Moderada (recessivo)	PanelApp ⁹⁷
CYLD	113/126	Não descrita	Não disponível	RAD50	113/126	Não descrita	Não disponível
DDB2	113/126	Não descrita	Não disponível	RAD51C	113/126	Não descrita	Não disponível
DICER1	113/126	Emergente (dominante)	PanelApp ⁹⁷	RAD51D	113/126	Não descrita	Não disponível
DIS3L2	113/126	Não descrita	Não disponível	RB1	113/126	Definitiva (dominante)	GeneReview ⁹⁸
EGFR	126	Não descrita	Não disponível	RECQL4	113/126	Forte (recessivo)	PanelApp ⁹⁷
EGLN1	113/126	Não descrita	Não disponível	REST	113/126	Não descrita	Não disponível
ENG	113/126	Não descrita	Não disponível	RET	113/126	Não descrita	Não disponível
EP300	113/126	Emergente (dominante)	GeneReviews ⁹⁸	RHBDF2	113/126	Não descrita	Não disponível
EPCAM	113/126	Não descrita	Não disponível	RUNX1	126	Não descrita	Não disponível
ERBB2	126	Não descrita	Não disponível	SDHA	113/126	Não descrita	Não disponível
ERCC2	113/126	Emergente (recessivo)	Ballinger et al. (2016) ⁹	SDHAF2	113/126	Não descrita	Não disponível
ERCC3	113/126	Não descrita	Não disponível	SDHB	113/126	Forte (dominante)	PanelApp ⁹⁷
ERCC4	113/126	Emergente (recessivo)	Chan et al. (2017) ⁶⁴	SDHC	113/126	Forte (dominante)	PanelApp ⁹⁷
ERCC5	113/126	Não descrita	Não disponível	SDHD	113/126	Não descrita	Não disponível
EXT1	113/126	Definitiva (dominante)	PanelApp ⁹⁷	SLX4	113/126	Não descrita	Não disponível
EXT2	113/126	Definitiva (dominante)	PanelApp ⁹⁷	SMAD4	113/126	Não descrita	Não disponível
EZH2	113/126	Não descrita	Não disponível	SMARCA4	113/126	Moderada (recessivo, sarcoma de corpo uterino); Definitiva (dominante)	PanelApp ⁹⁷ ; GeneReviews ⁹⁸
FANCC	113/126	Emergente (recessivo)	Chan et al. (2017) ⁶⁴	SMARCB1	113/126	Não descrita	Não disponível
FANCG	113/126	Não descrita	Não disponível	SMARCE1	113/126	Não descrita	Não disponível
FANCM	113/126	Não descrita	Não disponível	STK11	113/126	Não descrita	Não disponível
FH	113/126	Moderada (dominante)	PanelApp ⁹⁷	SUFU	113/126	Não descrita	Não disponível
FLCN	113/126	Não descrita	Não disponível	TERF2IP	126	Não descrita	Não disponível
GALNT12	113/126	Não descrita	Não disponível	TERT	113/126	Não descrita	Não disponível
GATA2	126	Não descrita	Não disponível	TMEM127	113/126	Não descrita	Não disponível

GPC3	113/126	Não descrita	Não disponível	TP53	113/126	Definitiva (dominante)	GeneReviews ⁹⁸
GREM1	113/126	Não descrita	Não disponível	TSC1	113/126	Não descrita	Não disponível
HNF1A	113/126	Não descrita	Não disponível	TSC2	113/126	Não descrita	Não disponível
HOXB13	113/126	Não descrita	Não disponível	VHL	113/126	Não descrita	Não disponível
HRAS	113/126	Moderada (dominante)	PanelApp ⁹⁷	WRN	113/126	Moderada (recessivo)	PanelApp ⁹⁷
KIF1B	113/126	Não descrita	Não disponível	WT1	113/126	Não descrita	Não disponível
KIT	113/126	Definitiva (dominante, GIST)	Gene Curation Coalition ⁹⁹	XPA	113/126	Não descrita	Não disponível
LZTR1	113/126	Não descrita	Não disponível	XPC	113/126	Não descrita	Não disponível
MAX	113/126	Não descrita	Não disponível	XRCC2	126	Não descrita	Não disponível
MC1R	113/126	Não descrita	Não disponível	YAP1	126	Não descrita	Não disponível

Fonte: Autoria própria.

3.1.1 Testes genéticos em pools de amostras

Com o objetivo de diminuir os custos do sequenciamento e aumentar o número de amostras avaliadas, foi utilizada uma estratégia implementada no grupo de Genômica Clínica e Funcional para realização de teste genético em *pools* de amostras que possibilitasse o rastreamento de variantes germinativas raras com alta qualidade dos dados gerados usando como base a estratégia bi-dimensional de amostras em *pools*, semelhante ao descrito previamente por Zuzarte et al. (2014)¹⁰⁰.

De forma resumida nesse método, os DNAs dos pacientes são individualmente distribuídos em placas de 96 poços e combinados em 2 dimensões, uma considerando as linhas da placa e gerando 8 *pools* e outra considerando as colunas, gerando 12 *pools*, totalizando 20 *pools* a serem sequenciados para as 96 amostras. Os *pools* de amostras são posteriormente processados para a captura e sequenciamento dos genes alvo e o sequenciamento é realizado na plataforma NextSeq 500 (Illumina), visando obter uma cobertura mínima de 200X para as bases alvo. Esta cobertura representa que, se considerarmos um *pool* equimolar, pelo menos 25 leituras (para os *pools* de 8 amostras) ou 16,6 leituras (para *pools* de 12 amostras) estariam presentes para cada indivíduo representado no *pool*.

Esta estratégia bidimensional permite que sejam identificados exatamente em quais pacientes estão presentes variantes raras, uma vez que a variante será detectada em 2 *pools* de linha/coluna podendo ser facilmente validadas em relação às amostras carreadoras, através do sequenciamento direcionado destas variantes em poucas amostras. Com essa abordagem, foram sequenciadas amostras retrospectivas utilizando o painel de 113/126 GPCs, seguindo a mesma estratégia descrita no item 4.4 para as amostras individuais. As variantes encontradas foram consideradas se o alelo variante detectado fosse > 0,5% nas leituras, considerando uma profundidade de cobertura mínima de 100X, em pelo menos um pool de coluna e um pool de linha. Todas as variantes P/PP detectadas por esta análise foram validadas por metodologia de sequenciamento de *amplicon*.

3.1.2 Análise e interpretação das variantes

Após a realização do NGS, foram realizadas análises de bioinformática e as variantes encontradas foram interpretadas de acordo com os protocolos do Grupo de Genômica Clínica e Funcional e Laboratório de Diagnóstico Genômico do A.C. Camargo Cancer Center. Em resumo, a ferramenta Isaac Enrichment v2.1, disponível no aplicativo BaseSpace (Illumina), foi utilizada para realizar o alinhamento e as chamadas variantes. As variantes identificadas foram anotadas e filtradas no *software* VarSeq (Golden Helix, Bozeman, Estados Unidos). Foram aplicados filtros específicos para seleção de variantes germinativas raras (Qualidade ≥ 30 , cobertura da base $\geq 30X$, frequência do alelo variante $\geq 0,25$, frequência do alelo menor [MAF] $\leq 0,02$ ou ausente nos bancos de dados populacionais (gnomAD, 1000genomes e ABraOM). Foram selecionadas apenas variantes em regiões codificantes ou em regiões adjacentes aos limites éxon-intron (até 10pb intrônicos). As variantes detectadas no estudo foram avaliadas quanto à sua classificação em bancos de dados e ferramentas de classificação clínica de variantes (ClinVar, VarSome e Franklin Genoox) e foram classificadas segundo os critérios sugeridos pelo American College of Medical Genetics and Genomics (ACMG)⁸⁶. Foram considerados informações de programas preditores de patogenicidade *in silico*, que são fornecidas pelo *software* VarSeq: *SIFT*, *Polyphen2 HVAR*, *MutationTaster*, *MutationAssessor*, *FATHMM* e *FATHMM*, *REVEL*, *dbsnpSNP*, *MaxEntScan* e *SpliceAI*, e foram consideradas as predições de plataformas digitais como o VarSome e o Franklin Genoox, que também são preditores *in silico*, no qual compilam diversas informações já conhecidas das variantes descritas e as classificam seguindo os critérios da ACMG. Além dessas ferramentas, foi utilizado o ClinVar, que apresenta o registro de variantes previamente descritas e classificadas por outros trabalhos e laboratórios. Os pacientes nos quais foram identificadas variantes patogênicas ou provavelmente patogênicas foram considerados como portadores de alterações germinativas patogênicas (GPVs, do inglês *germline pathogenic variants*).

A análise de alterações de número de cópias (CNV, do inglês *copy number variation*) foi realizada através do *software* VarSeq (Golden Helix). Resumidamente, o algoritmo de chamada de CNV do VarSeq realiza uma avaliação comparativa da cobertura normalizada das regiões alvo entre as amostras caso e amostras controle, para identificar regiões de menor cobertura (deleções) e maior cobertura (duplicações). Foram utilizados parâmetros padrão do

software para uma chamada estrigente de CNVs, em que a variação da cobertura da região deve ter Z-score maior que 3 (duplicações) ou menor que -3 (deleções).

3.1.3 Validação das variantes provavelmente patogênicas e patogênicas

As variantes genéticas pontuais classificadas como provavelmente patogênicas e patogênicas foram validadas através do sequenciamento pelo método de Sanger ou NGS de *amplicon*. Para o sequenciamento Sanger foi utilizado o kit BigDye® Terminator v3.1 (Applied Biosystems) e sequenciador Automático de DNA, ABI Prism 3130 XL (Applied Biosystems), seguindo as instruções do fabricante aparelho. O resultado do sequenciamento foi analisado utilizando a ferramenta CLCbioMain Workbench (CLC Bio). Para o sequenciamento de *amplicons* foi utilizado o kit Ion Plus Fragment Library (ThermoFisher Scientific) na construção de biblioteca e o sequenciamento foi realizado através da plataforma IonProton ou IonS5 (ThermoFisher Scientific), seguindo as instruções do fabricante. A validação das variantes foi realizada utilizando o *software* IGV (BroadInstitute e UC San Diego).

3.1.4 Avaliação de perda de heterozigose (LOH)

Para a avaliação de LOH foi utilizado um método estabelecido pelo grupo e que tem sido usado em várias publicações^{101,102}. Resumidamente foram desenhados iniciadores específicos para as regiões das variantes de interesse, visando à amplificação de produtos de 80 a 180 pares de base. Os tecidos tumorais, congelados a fresco ou fixados em formalina e emblocados em parafina (FFPE, do inglês *formalin fixed paraffin embedded*), utilizados nesta análise foram macrodissecados para selecionar áreas contendo um mínimo de 40% de células tumorais. A amplificação por PCR, preparação e construção de biblioteca de *amplicons* e análise foram realizadas conforme o protocolo dos *kits* Ion Plus Fragment Library kit (ThermoFisher Scientific), Ion PI™ Hi-Q™ OT2 200 Kit e Ion PI™ Hi-Q™ Sequencing 200 Kit, do sistema Ion Proton (ThermoFisher Scientific). A frequência alélica da variante foi verificada visualmente com o *software* IGV e frações alélicas $\geq 65\%$ foram consideradas como LOH (perda do alelo selvagem)⁸³.

3.1.5 Análise de transcrito

Foi realizada a análise de transcrito de um gene específico para um caso com uma variante intrônica provavelmente patogênica no gene *EXT1*. A análise de transcrito foi realizada utilizando os transcritos do gene *EXT1* através da extração de RNA de amostra de tumor fresco congelado e leucócito do probando (portador da variante *EXT1*:c.1722+1G>C) e, como controle positivo, foi utilizado leucócito de um indivíduo não detectado com essa alteração. Os cDNAs foram sintetizados através da reação de transcriptase reversa (RT) utilizando o kit SuperScript™ VILO™ cDNA Synthesis (Invitrogen, MA, EUA), seguindo as instruções do fabricante. Para a PCR, foram desenhados *primers* entre os éxons 6 e 10, a fim de amplificar os fragmentos que poderiam conter as variantes de splicing. Os produtos da RT-PCR foram visualizados posteriormente à reação de eletroforese em gel de agarose 2% (corado com SYBR Safe (Invitrogen), no qual o resultado foi observado em um transiluminador de luz ultravioleta (Alpha Innotec, BY, Alemanha). Os produtos amplificados foram sequenciados na plataforma Ion Proton (ThermoFisher Scientific) e as leituras foram mapeadas utilizando o *software* CLC Main Workbench (CLC Bio versão 6) e STAR. O gráfico do tipo *sashimi plot* foi gerado após alinhamento com a ferramenta STAR¹⁰³ e visualização utilizando o *software* IGV.

3.2 COMPARAÇÃO DA FREQUÊNCIA DAS VARIANTES ENCONTRADAS COM BANCO DE DADOS DE INDIVÍDUOS SAUDÁVEIS

As frequências de variantes P/PP detectadas nessa coorte de pacientes foram agrupadas e comparadas com a frequência dessas alterações nos mesmos genes descritas no ABraOM (Arquivo Brasileiro Online de Mutações), um banco de dados disponível publicamente, que descreve alterações genéticas em indivíduos brasileiros. O conjunto de dados do ABraOM SABE-WGS-1171 contém o sequenciamento de genoma de 1.171 indivíduos não relacionados obtidos de uma amostra baseada no censo de indivíduos idosos de São Paulo, Brasil¹⁰⁴. Dessa forma, foram avaliados 19 genes nos quais foram detectadas GPVs, e as variantes do ABraOM foram classificadas seguindo os mesmos critérios descritos para os pacientes.

3.3 ANÁLISE ESTATÍSTICA

A estatística descritiva foi utilizada para resumir os resultados. A comparação das características clínicas entre os pacientes com ou sem GPVs foi realizada por meio do teste Qui-quadrado de Pearson ou exato de Fisher para variáveis binárias e Mann-Whitney para variáveis contínuas. O teste bicaudal com $p < 0,05$ foi considerado significativo. Para comparação de pacientes e coorte de controle, foi utilizado o teste exato de Fisher.

3.4 SEQUENCIAMENTO DE EXOMA

Nesta etapa do projeto, foram selecionados pacientes já anteriormente avaliados pelas outras técnicas, mas considerando os seguintes critérios: 1 – indivíduos com resultado negativo para variantes provavelmente patogênica ou patogênica no sequenciamento do painel de genes; E 2 – com diagnóstico histopatológico de condrossarcoma ou sarcomas ultrarraros. Em virtude da complexidade dos sarcomas, foram definidos os subtipos a serem avaliados, a fim de permitir uma maior elucidação dos mecanismos tumorigênicos envolvidos a esses tipos histológicos mais específicos. Foram priorizados pacientes com história pessoal de múltiplas neoplasias e/ou história familiar positiva para câncer, inespecífica para SPHCs conhecidas. O DNA de células normais (leucócito ou saliva) e de células tumorais foram extraídos e processados no Biobanco do A.C. Camargo, seguindo seus procedimentos padrão. O ES do DNA de células normais e tumorais foi realizado utilizando o kit xGen® Exome Research Panel (IDT Technologies, IA, Estados Unidos) e o equipamento NextSeq500 (Illumina).

A análise e interpretação integrada das variantes identificadas nos dados de ES foram realizadas de acordo com os protocolos estabelecidos nos Grupo de Genômica Clínica e Funcional e Laboratório de Bioinformática do A.C. Camargo Cancer Center. Para essa análise, o *pipeline* de chamada de variante germinativa utilizado segue as boas práticas definidas pelo *software* do GATK (*Genome Analysis Toolkit*)^{105–107}. Dessa forma, a partir dos arquivos FASTQs, os dados gerados são alinhados contra a referência do genoma humano GRCh37 (hg19) utilizando o BWA (*Burrows-Wheeler Aligner* software). Depois é realizado o pré-processamento com o GATK e a partir do BAM processado, é utilizado o GATK HaplotypeCaller para a chamada de variante germinativa por amostra. Os VCFs por amostra são unidos com o

CombineGVCFs e esse arquivo de VCF combinado é recalibrado. As variantes identificadas foram anotadas com os *softwares* snpEff e SnpSift (v4.3) utilizando vários bancos de polimorfismos, como dbNSFP (v3.5a), dbSNP (v150) e ABRaOM (v1.0.2). Em seguida, as variantes foram filtradas aplicando o método VariantRecalibration (VQSR, do inglês *Variant Quality Score Recalibration*) do GATK com vários *resources*, como HapMap e dbSNP, e com os *tranches* 100.0, 99.90, 99.0 e 90.0. Somente as variantes que passaram em todos os filtros seguiram para as análises posteriores.

O *pipeline* de chamada somática utilizado também foi o do GATK, no alinhamento, pré-processamento e chamada de variante. Seguindo o GATK, foi feito o alinhamento dos arquivos FASTQ com o *software* BWA no modo *MEM* contra a sequência referência do genoma humano GRCh37 (hg19). Os BAMs foram pré-processados a partir do GATK, foi utilizado o Mutect2 de forma pareada (amostras de tumor e normal) para identificação das mutações somáticas e as variantes foram anotadas com os *softwares* snpEff e SnpSift (v4.3) utilizando os bancos de dados dbnsfp (v3.5) e ABRaOM (v1.0.2). Por fim, foram aplicados os *hard filters* descritos na documentação do GATK Best Practices, incluindo filtro de *strand bias* (OxOG). Somente as variantes que passaram em todos os filtros foram utilizadas nas análises posteriores. Adicionalmente, foi realizada a análise de Tumor Mutation Burden (TMB), na qual foi utilizado o *software* maftools no R 3.6 através do método *tmb* disponível no pacote. O parâmetro *captureSize* foi alterado para 39.08646 para refletir o número de milhões de bases sequenciadas no experimento de NGS das amostras.

Os arquivos VCFs da chamada germinativa e somática foram importados para o *software* VarSeq (Golden Helix) para anotações adicionais e filtragem de variantes. As variantes identificadas foram filtradas seguindo os seguintes critérios: qualidade >30, PASS (variantes que passam pelos filtros de qualidade); DP (*Read Depth*) > 10; AD (*Allele Depth*) > 5; AF (*Allele Frequency*) <= 0,01 ou ausente; Frequência na corrida < 6 ou ausente; base variante presente em pelo menos 25% das leituras (exoma germinativo) ou 5% (exoma somático).

Na filtragem de variantes germinativas foi criada uma lista de 184 GPCs incluída no *template* do VarSeq, na qual se o GPC dos dados de ES estivesse presente nessa lista, seria sinalizado na planilha. Ainda, também foram filtrados genes de relação desconhecida/incerta com câncer, sendo priorizados com base nos seguintes parâmetros: efeito de perda de função ou *missense*; classificação do ClinVar: outros, conflitante, VUS, PP, P, sem classificação;

preditores de patogenicidade (SIFT, Polyphen-2 HVAR, MutationTaster, MutationAssessor, FATHMM, FATHMM MKL): mínimo 4 de 6 predizendo como danosa; REVEL $\geq 0,5$. Destes, foram consideradas as variantes das quais havia perda de heterozigose no tumor ou classificadas como P/PP pelo VarSome. Dos genes incluídos na lista de 184 GPCs, os seguintes fatores foram considerados para a priorização, sendo eles: classificação do ClinVar: outros, conflitante, VUS, PP, P, sem classificação; alterações: *missense*, *frameshift*, *splice site* e com GnomAD POPMAX Alt Allele Freq (AF): $< 0,01$. Destes, foram consideradas as variantes em que havia pelo menos 4 de 6 preditores de patogenicidade assinalados e o REVEL $\geq 0,5$.

Na avaliação somática, investigando outros genes e vias candidatas possivelmente relevantes no desenvolvimento de sarcomas, os dados do ES também foram anotados utilizando o VarSeq, e as variantes foram filtradas considerando parâmetros de qualidade (conforme descritos nos parágrafos anteriores deste item), além de consequência da variante (LoF e *missense*) e validação por inspeção visual através da ferramenta IGV. Para priorização, esses genes foram avaliados quanto a: frequência de alterações encontradas em sarcomas (acima de 5%) reportadas no cBioPortal^{108,109} e/ou OncoKb⁹⁵, ou alteração no mesmo gene em pelo menos dois pacientes dessa coorte; GPC já conhecido e genes com funções e vias gênicas relevantes. Também foram consideradas informações de preditores como o REVEL ($\geq 0,5$) e classificação das alterações no banco de dados ClinVar.

As variantes somáticas encontradas também foram priorizadas de acordo a função do gene afetado e seu efeito na proteína, através da utilização de outros programas de predição de patogenicidade *in silico*, como o VarElect. O VarElect¹¹⁰ é uma ferramenta que contribui na busca por correlações genótipo-fenótipo e mostra vias e redes de interações que genes diferentes compartilham, gerando parâmetros que embasam essas correlações, como o score e a probabilidade média de causar doença. Os scores do VarElect são baseados na ligação direta ou indireta entre os genes sequenciados com as palavras-chave de doenças, sintomas e fenótipos. Com isso, as pontuações são determinadas pela frequência do aparecimento dessas informações/termos no banco GeneCards, e são comparadas quando descritas em cada gene individualmente e entre todos os genes. Dessa forma, o programa ainda consegue classificar os genes, apontando sua probabilidade de estarem relacionados à doença sinalizada. O termo “sarcoma” foi utilizado para busca no programa VarElect, assim, foram considerados somente os genes de associação direta com o fenótipo sarcoma.

O Franklin Genoox¹¹¹ – modalidade para variantes somáticas - também foi utilizado na anotação e priorização, no qual as variantes foram filtradas com os seguintes parâmetros: as classificadas entre Tier 1, 2 e 3; de região exônica e *splice*; com os tipos de alteração: *missense*, *stop gain*, *stop loss*, *start gain*, *start loss*, *frameshift*; com a frequência não disponível ou muito rara; e confiança baixa, média e alta. Depois, dentre essas, foram consideradas somente as variantes de classificação Tier 1 e 2. Esse modelo de classificação de variantes baseado em Tier considera pontuações de 1 a 4, sendo que dentro de 1 e 2 estão incluídas as variantes de possível relevância clínica e tidas como potenciais alvos terapêuticos, na 3 as de significância clínica não compreendida e na 4 as variantes sem importância clínica conhecida. O detalhamento dessa classificação está descrito no item 1.4.

Essa etapa de priorização de variantes germinativas e somáticas foi realizada em colaboração com o grupo da pesquisadora Dra. Barbara Rivera, do L'Institut d'Investigació Biomèdica de Bellvitge (IDIBELL) localizado em Barcelona, Espanha, a qual fui contemplada com uma bolsa de estágio no exterior, do Programa de Doutorado-Sanduiche no Exterior da CAPES (PDSE-CAPES, processo 88881.623036/2021-01), no período de 4 de outubro de 2021 a 18 de fevereiro de 2022.

4. RESULTADOS

4.1 DADOS CLÍNICOS DOS PACIENTES INCLUÍDOS

Foram identificados 436 pacientes diagnosticados com sarcomas e com material de DNA de leucócito disponível no Biobanco da instituição. Os prontuários eletrônicos desses pacientes foram avaliados para confirmação de dados clínicos como: idade ao diagnóstico, histologia, presença de outros tumores, histórico familiar de neoplasias e diagnóstico genético prévio. Seguindo os critérios de inclusão e exclusão do estudo, foram excluídos pacientes com diagnóstico genético prévio de Síndrome de Li-Fraumeni (6 pacientes), e pacientes diagnosticados com sarcoma com idade superior a 40 anos (244 pacientes). Após atender todos os critérios, foram selecionados retrospectivamente 141 pacientes. Adicionalmente, 36 pacientes foram incluídos prospectivamente entre 2018 e 2021, em colaboração com os departamentos de Oncogenética, Oncologia Clínica e Oncologia Pediátrica da instituição, totalizando, por fim, 177 pacientes com amostras disponíveis para avaliação.

A média de idade dessa coorte foi de 26 anos (0-40, exceto por um paciente com sarcoma diagnosticado aos 54 anos, que incluído prospectivamente considerando seu histórico familiar, no qual o pai já falecido foi diagnosticado com tumor ósseo aos 25). A coorte foi dividida em dois grupos considerando as faixas etárias, sendo de 0 a 14 anos (Crianças – C) e de 15 a 40 anos (Adolescentes e Adultos Jovens - AAJ), compreendendo 20,3% (36/177) e 79,7% (141/177) dos pacientes, respectivamente. Na Tabela 1 estão descritas as variáveis clínicas dessa coorte. Quase 10% dos pacientes tiveram mais do que um tumor primário, e 40,7% tiveram história familiar positiva para câncer relatada nos prontuários médicos. Não houve relevância estatística quando foram comparadas as características clínicas entre os grupos C e AAJ. Dentre os pacientes diagnosticados com mais do que um tumor primário (17/177), sete fizeram tratamento radioterápico, o qual provavelmente não apresenta relação com o desenvolvimento do sarcoma por ter sido realizado em outras localizações.

Tabela 1 - Aspectos clínicos dos 177 pacientes incluídos, comparados de acordo com os grupos de faixa etária (Crianças e Adolescentes e Adultos Jovens).

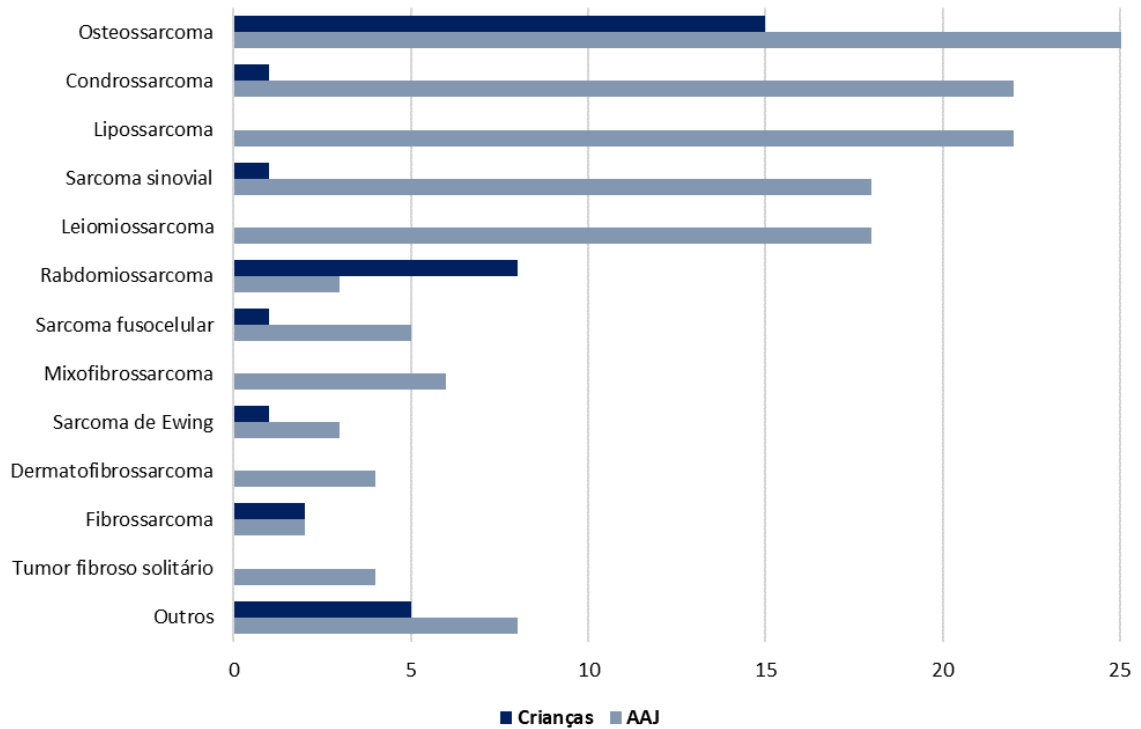
Características	Pacientes N (%)	C N (%)	AAJ N (%)	P-valor
Total de indivíduos	177 (100%)	36 (20,3%)	141 (79,7%)	
Sexo				1
F	97 (54,8%)	20 (55,6%)	77 (54,6%)	
M	80 (45,2%)	16 (44,4%)	64 (45,4%)	
Média de idade (faixa)	26 (0 – 54)*	9 (0 – 14)	30 (14 – 54)	
Outros tumores primários				0,12
Positivo	17 (9,6%)	6 (16,7%)	11 (7,8%)	
Negativo	160 (90,3%)	30 (83,3%)	130 (92,2%)	
História familiar de câncer				1
Positiva	72 (40,7%)	7 (19,4%)	62 (44,0%)	
Negativa	28 (15,8%)	2 (5,6%)	26 (18,4%)	
Não informado	77 (43,5%)	27 (75,0%)	53 (37,6%)	

Fonte: Autoria própria.

Nota: * foi incluído um paciente com diagnóstico de sarcoma (condrossarcoma) aos 54 anos. O caso foi incluído prospectivamente considerando seu histórico familiar, no qual o pai foi diagnosticado com tumor ósseo aos 25 anos e não estava disponível para coleta de material germinativo (falecido). N – número de pacientes da coorte; C - Crianças; AAJ – Adolescentes e Adultos Jovens; F – feminino; M – masculino.

Dentre os subtipos histológicos de sarcomas, foram observadas diferenças esperadas entre os grupos etários, aos quais os condrossarcomas, lipossarcomas, sarcomas sinoviais e leiomiossarcomas foram mais comuns nos AAJs, enquanto o rabdomiossarcoma foi mais frequente em crianças (Gráfico 1). É importante destacar a seleção das amostras foi enriquecida com indivíduos diagnosticados com condrossarcomas e osteossarcomas do nosso Biobanco, consequentemente aumentando a frequência de ambos os tumores em nossa coorte, comparado a outros subtipos histológicos de sarcomas.

Gráfico 1 - Distribuição dos subtipos histológicos de sarcoma de acordo com o diagnóstico desses tumores por faixa etária dos pacientes.

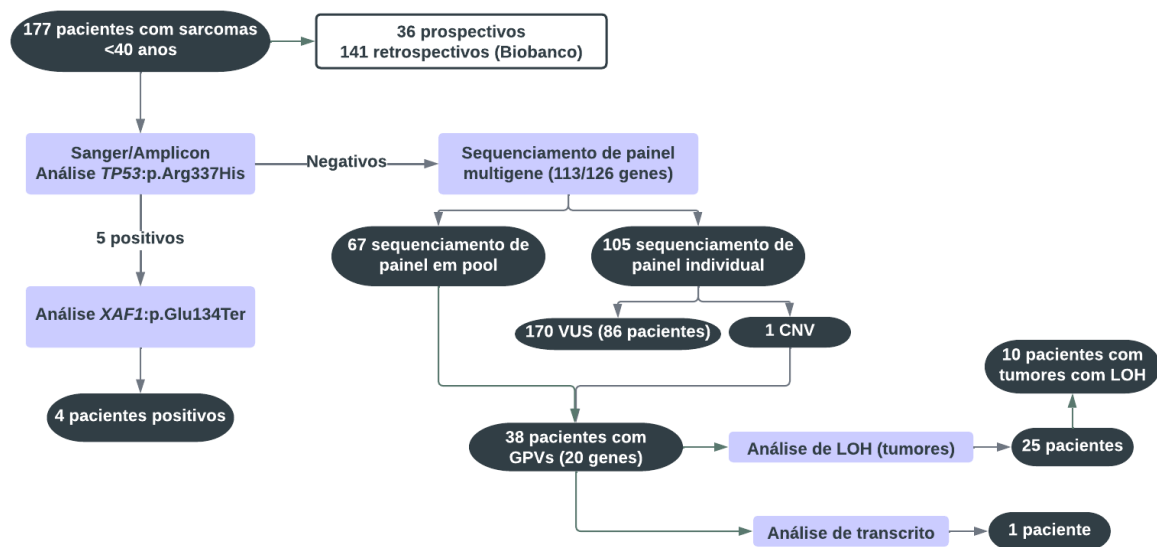


Fonte: Autoria própria.

Notas: AAJ – Adolescentes e Adultos Jovens.

4.2 ANÁLISES MOLECULARES

A Figura 3 descreve o delineamento experimental do estudo e as etapas de análises germinativas e moleculares do projeto.



Fonte: Fluxograma criado no Lucidchart <www.lucidchart.com>.

Legenda: VUS – *variant of uncertain clinical significance*; CNV – *copy number variants*; GPVs – *germline pathogenic variants*; LOH – *loss of heterozygosity*.

Figura 3 - Panorama geral do desenho experimental do estudo, metodologia e achados principais.

Na comparação das características clínicas entre os pacientes portadores de GPVs e VUS/Negativos, foi observado que pacientes que apresentaram diagnóstico de outros tumores primários tiveram uma maior detecção de variantes patogênicas e provavelmente patogênicas (P-valor = 0,012). Dentre as outras informações, como sexo, faixa etária e história familiar de câncer, não houve relevância estatística (Tabela 2).

Tabela 2 - Comparação das características clínicas entre pacientes portadores de GPVs e VUS/Negativos.

Características	Pacientes N (%)	GPVs N (%)	VUS/Negativos N (%)	P-valor
Total de indivíduos	177 (100%)	38 (21,4%)	139 (78,6%)	0,32
F	97 (54,8%)	24 (13,5%)	73 (41,2%)	
M	80 (45,2%)	14 (36,8%)	66 (32,3%)	
Média de idade (anos)	26 (0 – 54*)	27 (6 – 54*)	26 (0 – 41)	0,94
Faixa etária				0,58
C (0-14 anos)	36 (20,3%)	6 (15,8%)	30 (21,6%)	
AAJ (15-40* anos)	141 (79,7%)	32 (84,2%)	109 (78,4%)	
Outros tumores primários				0,012
Positivo	17 (9,6%)	8 (21,1%)	9 (6,5%)	
Negativo	160 (90,3%)	30 (78,9%)	130 (93,5%)	

História familiar de câncer				0,64
Positiva	72 (40,7%)	15 (39,5%)	57 (41,0%)	
Negativa	28 (15,8%)	4 (10,5%)	24 (17,3%)	
Não informado	77 (43,5%)	19 (50,0%)	58 (41,7%)	

Fonte: Autoria própria.

Legenda: N – número de pacientes da coorte; C – crianças; AAJ – Adolescentes e Adultos Jovens.

4.2.1 Sequenciamento Sanger ou amplicon

Inicialmente, foi realizado sequenciamento Sanger ou de *amplicon* para pesquisa da variante p.Arg337His no gene *TP53*, ao qual as amostras positivas para esta alteração não foram incluídas nas próximas etapas do projeto. O resultado do sequenciamento Sanger ou de *amplicon* evidenciou a presença da variante p.Arg337His em cinco amostras. No Quadro 8 estão descritos os dados clínicos dos pacientes portadores desta alteração e a presença da alteração modificadora de risco *XAF1*:p.Glu134Ter. A maior parte dos pacientes apresentou o subtipo histológico de leiomiossarcoma (3/5), 2/5 pacientes apresentaram histórico pessoal e familiar compatível com os critérios modificados de Chompret³⁸ somados ao descritos por Birch et al. (1994)³⁶ (LFL) e nenhum paciente preencheu os critérios clínicos clássicos de Li-Fraumeni.

Quadro 8 - Dados clínicos e moleculares dos pacientes portadores da variante patogênica *TP53:p.Arg337His*.

ID	Idade	Sexo	Histologia (sarcoma)	História familiar de câncer	Presença de <i>XAF1:p.Glu134Ter</i>	Critério clínico
GRY_7	30	M	LMS	Avô: câncer gástrico aos 93 anos.	Presente	Nenhum
GRY_22	29	F	LMS	Tia paterna com câncer de mama aos 39 anos; tia paterna com câncer de SNC aos 35 anos; avô materno com câncer gástrico (idade não informada).	Presente	Nenhum
GRY_23	33	M	LMS	Pai com neoplasia de próstata aos 59 anos e sarcoma aos 62 anos; avó materna com câncer de laringe aos 72 anos (tabagista); prima paterna com glioblastoma aos 35 anos; primo paterno com linfoma de Hodgkin aos 22 anos; tio paterno com câncer de cólon aos 50 anos; tio paterno com leucemia aos 5 anos; avó paterno câncer de cólon aos 60 anos; avô paterno câncer de cólon aos 60 anos.	Ausente	Chompret (2015)
GRY_327	33	F	LS	Mãe: câncer de mama aos 44 anos e câncer de tireoide aos 46 anos; avó materna: câncer de pulmão aos 70 anos; tia-avó materna: câncer de mama aos 63 anos; tia-avó materna: câncer colorretal aos 50 anos; tia-avó materna: câncer de útero aos 50 anos; primo materno: câncer de SNC aos 5 anos.	Presente	Chompret (2015)
GRY_369	23	M	OS	Tia materna: câncer de pulmão.	Presente	Nenhum

Fonte: Autoria própria.

Legenda: ID – identificação (da amostra); M – masculino; F – feminino; LMS – leiomiossarcoma; LS – lipossarcoma; OS – osteossarcoma; SNC – Sistema Nervoso Central.

4.2.2 Análise de painel multigênico por NGS

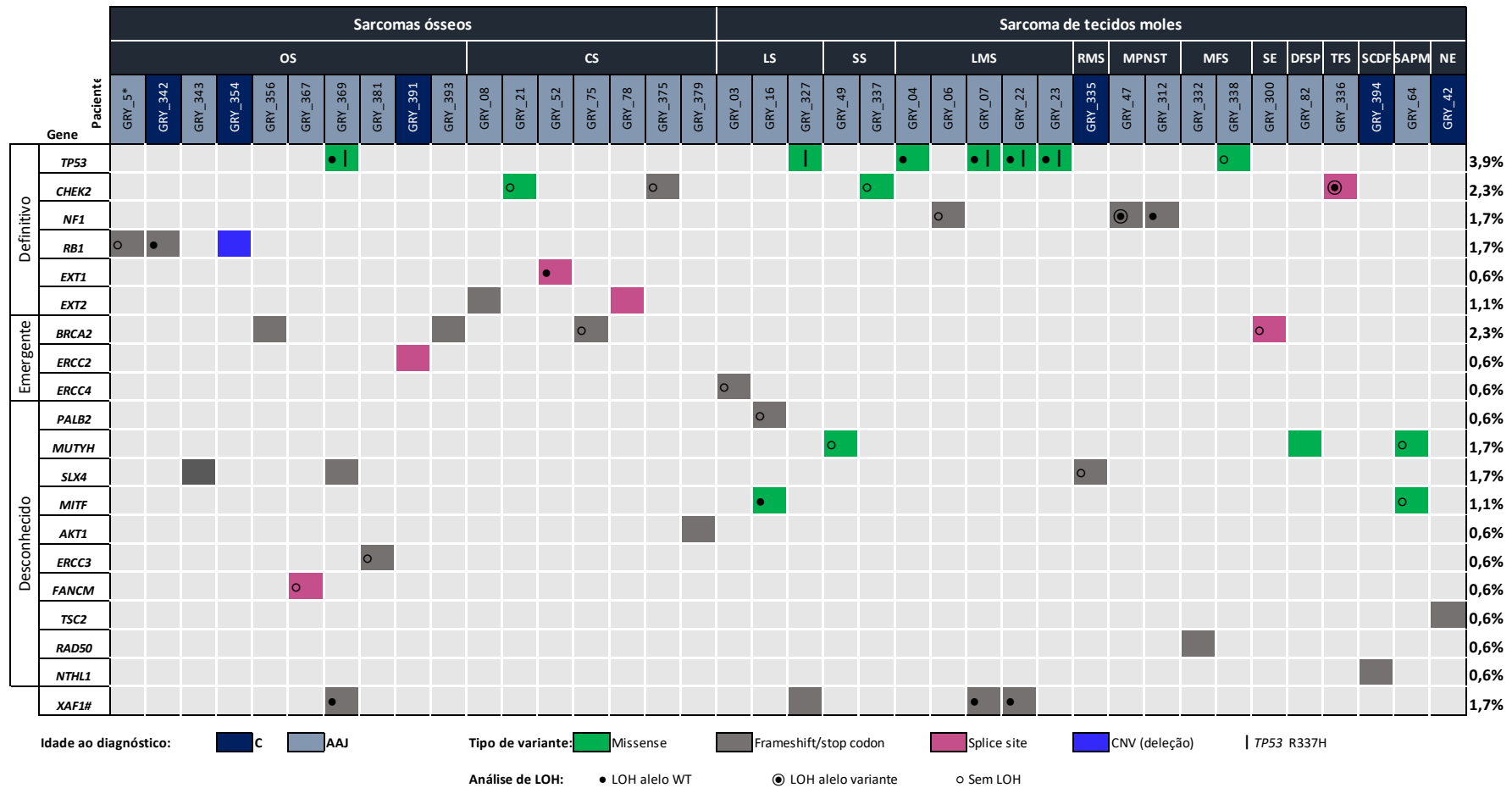
Das 177 amostras incluídas neste projeto, 172 foram sequenciadas por NGS com o painel de 113 ou 126 genes (excluindo os pacientes portadores da variante *TP53*:p.Arg337His). A cobertura média das 172 amostras foi de 706 vezes, e 98% das bases do painel apresentaram cobertura maior do que 50X. O sequenciamento individual foi realizado em 105 amostras e 67 foram incluídas no sequenciamento de painel por pool. As variantes foram classificadas seguindo as recomendações da ACMG/AMP, detalhadas no item 1.4, sendo que 33 pacientes apresentaram variantes P/PP (18,6%) e 139 pacientes apresentaram VUS ou foram negativos (WT, do inglês *wild-type*) (78,5%). As VUS foram avaliadas apenas nos 105 pacientes que fizeram sequenciamento individual, e desses, 81,9% apresentaram VUS.

No geral, considerando os cinco portadores de *TP53*:p.Arg337His, 21,5% (38/177) dos pacientes apresentaram GPVs. Os genes mais frequentemente alterados foram *TP53* (3,9%), *CHEK2* (2,3%), *BRCA2* (2,3%), e *NF1*, *RB1*, *MUTYH*, *SLX4* (1,7% cada) (Figura 4). Além da variante *TP53*:p.Arg337His, outras três GPVs foram detectadas em mais de um paciente: três pacientes foram detectados com *MUTYH*:p.Gly396Asp; dois pacientes com *RB1*:p.Arg251Ter e dois pacientes com *MITF*:p.Glu318Lys.

Os genes de predisposição ao câncer avaliados foram classificados de acordo com sua evidência de relação com predisposição a sarcomas, como descrito no banco Gene Curation Coalition⁹⁹ e através de revisão de literatura (Quadro 7, na metodologia). Foram identificadas variantes em genes previamente descritos em aumentar o risco para o desenvolvimento de sarcomas, como *TP53*, *CHEK2*, *NF1*, *RB1*, *EXT1* e *EXT2*, mas também em genes aos quais essa associação é emergente, como *BRCA2*, *ERCC2* e *ERCC4*, ou é desconhecida, como *MUTYH*, *SLX4*, *MITF*, *AKT1*, *ERCC3*, *FANCM*, *NTLH1*, *RAD50*, *PALB2* e *TSC2*. Foi detectada uma deleção no gene *RB1*, sendo a única CNV encontrada (Figura 4 e Tabela 3). A descrição das síndromes, tumores e outros fenótipos associados aos genes de associação desconhecida aos sarcomas está apresentada no Quadro 9.

Considerando somente os genes classificados como definitivos e emergentes para risco de sarcomas, 14,7% (26/177) dos pacientes apresentaram GPVs. A maior parte dos genes com relação desconhecida para sarcomas representam genes associados com síndromes de predisposição ao câncer recessivas. De todos os pacientes, três apresentaram alterações P/PP em mais de um GPC, sendo genes de relação conhecida e desconhecida para sarcomas: um

paciente com lipossarcoma foi detectado com GPVs em *PALB2* e *MITF*, um paciente com sarcoma alveolar apresentou alterações em *MUTYH* e *MITF* e um paciente com osteossarcoma apresentou GPVs em *TP53* e *SLX4* (Figura 4).



Fonte: Autoria própria.

Legenda: À direita, as porcentagens representam a quantidade de amostras detectadas com alterações P/PP no gene indicado e, na parte inferior estão descritos os tipos de alterações, assim como sinalizando quais amostras apresentaram ou não LOH na análise tumoral. As amostras sem essa indicação não foram submetidas a essa análise. Os pacientes foram coloridos de acordo com o grupo etário. * Paciente com osteossarcoma, leiomiossarcoma e lipossarcoma (portador de Retinoblastoma). # Alterações em *XAF1* foram investigadas em pacientes com *TP53*:p.Arg337His OS – osteossarcoma; CS – condrossarcoma; LS – lipossarcoma; SS – sarcoma sinovial; LMS – leiomiossarcoma; RMS – rabiomiossarcoma; MPNST - tumor maligno da bainha do nervo periférico (do inglês *malignant peripheral nerve sheath tumour*); MFS – mixofibrossarcoma; SE – sarcoma de Ewing; DFSP – dermatofibrossarcoma protuberans; TFS – tumor fibroso solitário; SCDF – sarcoma de células dendríticas foliculares; SAPM – sarcoma alveolar de partes moles; NE – não especificado; C – crianças; AAJ – adolescentes e adultos jovens; CNV – variação de número de cópias (do inglês *copy number variation*); LOH – perda de heterozigose (do inglês *loss of heterozygosity*).

Figura 4 - Panorama geral das alterações germinativas patogênicas e provavelmente patogênicas encontradas em nossa coorte (38/177).

Tabela 3 - Descrição de dados clínicos e variantes P e PP detectadas no sequenciamento de painel multigene em 33 pacientes.

ID da amostra	Idade ao diagnóstico	Sexo	História pessoal e Histologia	História Familiar	Gene	Variante c. p.	Classificação
GRY_3	26	F	Lipossarcoma mixóide	ND	<i>ERCC4</i>	c.1349G>A; p.(Trp450Ter)	PP
GRY_4	38	F	Leiomiossarcoma	Avó: câncer de pâncreas	<i>TP53</i>	c.844C>T; p.(Arg282Trp)	P
GRY_5	23	M	Retinoblastoma; Osteossarcoma; Lipossarcoma; Leiomiossarcoma	ND	<i>RB1</i>	c.751C>T; p.(Arg251Ter)	P
GRY_6	37	M	Leiomiossarcoma	ND	<i>NF1</i>	c.1155delT; p.(Ile386fs)	PP
GRY_8	27	F	Condrossarcoma	ND	<i>EXT2</i>	c.2044C>T; p.(Arg682Ter)	PP
GRY_16	29	F	Lipossarcoma	ND	<i>MITF</i> <i>PALB2</i>	c.952G>A; p.(Glu318Lys) c.355delC; p.(Gln119Lysfs*58)	PP P
GRY_21	28	F	Condrossarcoma	ND	<i>CHEK2</i>	c.470T>C; p.(Ile157Thr)	PP
GRY_42	34	M	Sarcoma de partes moles	Avô paterno: câncer de pâncreas; pai: câncer de reto; avô materno: câncer de bexiga; avó materna: câncer em intestino	<i>TSC2</i>	c.570C>A; p.(Tyr190Ter)	P
GRY_47	19	F	Sarcoma fusocelular	Mãe: câncer em cólon	<i>NF1</i>	c.1756_1759delACTA; p.(Thr586Valfs*18)	P
GRY_49	30	F	Sarcoma sinovial	Bisavó materna: câncer em esôfago; avô materno: câncer em estômago; mãe: câncer de mama	<i>MUTYH</i>	c.1187G>A; p.(Gly396Asp)	P
GRY_52*	54	F	Condrossarcoma	Pai: câncer ósseo NE; tia paterna: carcinomatose; avó materna: câncer de mama; avô materno: câncer de próstata	<i>EXT1</i>	c.1722+1G>C; p.(Met546_Val575del)	PP
GRY_64	30	M	Sarcoma alveolar	Tia paterna: câncer de garganta	<i>MUTYH</i> ; <i>MITIF</i>	c.1187G>A; p.(Gly396Asp) c.952G>A; p.(Glu318Lys)	P PP
GRY_75	36	F	Condrossarcoma	Tio paterno: câncer renal aos 65 anos	<i>BRCA2</i>	c.5112_5119delinsG; p.(Ile1705Leufs*5)	P

GRY_78	36	M	Condrossarcoma	Pai: câncer gástrico aos 65 anos	<i>EXT2</i>	c.1079+1G>A	P
GRY_82	37	F	Dermatofibrossarcoma	Paciente negro	<i>MUTYH</i>	c.1187G>A; p.Gly396Asp	P
GRY_300	30	F	Sarcoma de Ewing	ND	<i>BRCA2</i>	c.7617+1G>A; p.(spl?)	P
GRY_312	31	M	Sarcoma fusocelular; Rabdomiossarcoma embrionário	ND	<i>NF1</i>	c.1818C>A; p.(Cys606Ter)	P
GRY_332	M	24	Sarcoma fibromixóide de baixo grau	ND	<i>RAD50</i>	c.2165dupA; p.Glu723Glyfs*5	P
GRY_335	12	M	Rabdomiossarcoma alveolar	ND	<i>SLX4</i>	c.5197G>T; p.(Glu1733Ter)	P
GRY_336	31	F	Tumor fibroso solitário	Paciente negro	<i>CHEK2</i> ; <i>SLX4</i>	c.319+2T>A; p.(spl?) c.5197G>T; p.(Glu1733Ter)	P P
GRY_337	37	F	Sarcoma sinovial do tipo monofásico fibroso	Avô paterno: câncer VADS (tabagista); avô materno: câncer de próstata; tio paterno: câncer de pulmão	<i>CHEK2</i>	c.349A>G; p.(Arg117Gly)	P
GRY_338	33	M	Mixofibrossarcoma	Pai: câncer de próstata	<i>TP53</i>	c.524G>A; p.(Arg175His)	P
GRY_342	8	M	Retinoblastoma; Osteossarcoma	ND	<i>RB1</i>	c.751C>T; p.(Arg251Ter)	P
GRY_343	18	F	Osteossarcoma condroblástico	ND	<i>SLX4</i>	c.4828delT; p.(Ser1610Profs*105)	PP
GRY_354	12	F	Retinoblastoma; Osteossarcoma osteoblástico	ND	<i>RB1</i>	g.(?_48916672)_(48923185_?)	Deleção (CNV)
GRY_356	39	F	Osteossarcoma	ND	<i>BRCA2</i>	c.2808_2811delACAA; p.(Ala938Profs*21)	P
GRY_367	16	M	Osteossarcoma ostecondroblástico	ND	<i>FANCM</i>	c.1397-3_1397-2dupTA; p.(spl?)	P
GRY_375	29	M	Condrossarcoma	ND	<i>CHEK2</i>	c.277delT; p.(Trp93Glyfs*17)	P

GRY_379	35	F	Condrossarcoma	ND	<i>AKT1</i>	c.698dupG; p.(Glu234Argfs*23)	P
GRY_381	37	F	Osteossarcoma osteoblástico	Duas tias maternas com câncer de mama (uma aos 45 anos), sem informações adicionais	<i>ERCC3</i>	c.659_660insA; p.(Ser221Phefs*2)	P
GRY_391	12	F	Osteossarcoma	ND	<i>ERCC2</i>	c.361-1G>A; p.(spl?)	PP
GRY_393	19	F	Osteossarcoma	ND	<i>BRCA2</i>	c.3680_3681delTG; p.(Leu1227Glnfs*5)	P
GRY_394	6	F	Sarcoma de células dendríticas foliculares interdigitantes	ND	<i>NTHL1</i>	c.268C>T;p.(Gln90Ter)	PP

Fonte: Autoria própria.

Nota: * Paciente com pais consanguíneos e história familiar de parente de primeiro grau (pai) com câncer ósseo aos 25 anos. ID – identificação (da amostra); F – feminino; M – masculino; ND – não disponível; NE – não especificado; VADS – carcinomas epidermóides de vias aerodigestivas superiores; PP – provavelmente patogênica; P – patogênica.

Quadro 9 - Síndromes e tumores associados aos genes na qual a relação com o desenvolvimento de sarcomas é ainda incerta ou não descrita.

Genes	Síndromes/doenças caracterizadas	Principais tumores/mais comuns	Principais tumores raros/menos comuns	Outras características
AKT1	Síndrome de Proteus	Câncer de ovário	Carcinoma de mama ductal invasivo, adenocarcinoma de pulmão e cólon, carcinoma de mama invasivo	ND
BRCA2	Anemia de Fanconi, câncer de mama e ovário hereditários	Câncer de mama e de ovário	Câncer de trompa de falópio, câncer primário peritoneal, câncer de próstata, câncer de pâncreas	Aumenta a susceptibilidade para gliomas
ERCC2	Xeroderma pigmentoso, Síndrome cérebro-óculo-fácio-esquelética, tricotiodistrofia	Adenocarcinoma de próstata, tumores sólidos malignos, carcinoma colorretal e carcinoma de ovário	ND	ND
ERCC3	Tricotiodistrofia, Xeroderma pigmentoso	Tumores de pele, cólon e estômago	ND	ND
ERCC4	Anemia de Fanconi, Xeroderma pigmentoso grupo F e tipo F, síndrome progeroide	Homozigosidade: carcinoma de células basais, carcinoma de células escamosas, melanoma	Incerto: a associação de risco com outros tumores para portadores homo e heterozigotos ainda está em andamento	Fenótipo complexo e variável

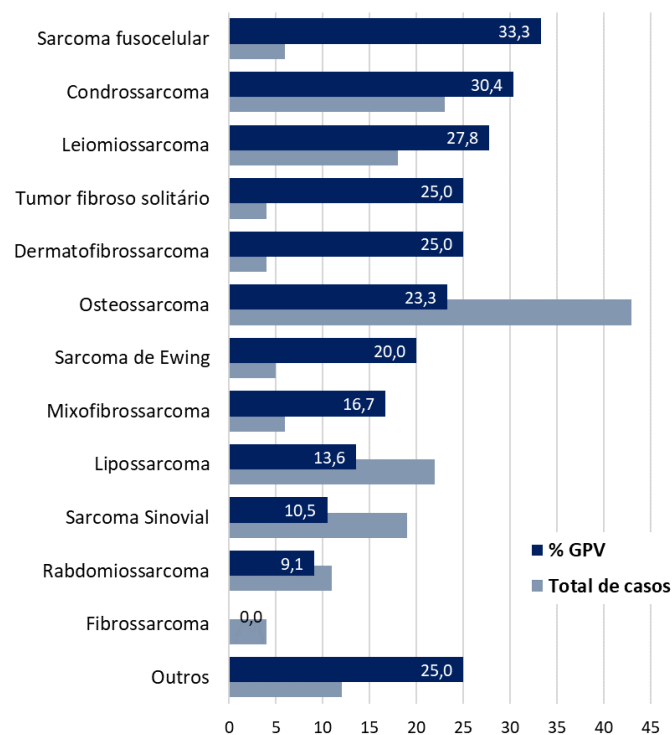
FANCM	Anemia de Fanconi, alterações na espermatogênese e parada prematura da função do ovário	Adenocarcinoma de pulmão, melanoma, glioblastoma, carcinoma ductal invasivo, câncer de cólon	ND	ND
MITF	ND	Melanoma, carcinoma de células renais	Paragangliomas, feocromocitomas	Aumento do número de nevos
MUTYH	Polipose associada ao <i>MUTYH</i> (recessiva)	Bialélicos: CCR, carcinoma papilífero de tireoide, endométrio, ovário, bexiga, mama; Monoalélicos: CCR, mama, gástrico	Câncer de duodeno, tumores benignos e malignos de glândula sebácea	Pólipos intestinais, cistos mandibulares, HCEPR, alteração em tireoide
NTHL1	Polipose Adenomatosa Familiar 3, Síndrome da telangiectasia cutânea familiar e predisposição ao câncer	CCR	ND	Pólipos intestinais
PALB2	Anemia de Fanconi, síndrome de predisposição ao câncer hereditário	Câncer de pâncreas, de mama e ovário	ND	ND
RAD50	Síndrome de quebras de Nijmegen-like	Câncer de mama	Adenocarcinoma de cólon e pulmão, adenocarcinoma endometriode, carcinoma urotelial de bexiga e melanoma cutâneo	ND
SLX4	Anemia de Fanconi	Adenocarcinoma de cólon e pulmão, adenocarcinoma endometriode, carcinoma de mama ductal invasivo, adenocarcinoma prostático	Melanoma, adenocarcinoma de ovário seroso de alto grau, glioblastoma, adenocarcinoma pancreático	ND
TSC2	Esclerose tuberosa 2 Displasia focal cortical tipo 2 (somática) Linfangioleiomiomatose (somática)	ND	ND	ND

Fonte: Informações compiladas de bancos de dados como GeneCards (<https://www.genecards.org/>), Omim (<https://www.omim.org/>) e My Cancer Genome (<https://www.mycancergenome.org/>).

Legenda: ND – não descrito; CCR – câncer colorretal; HCEPR – hipertrofia congênita do epitélio pigmentar da retina.

A detecção de GPVs foi comparada dentre os subtipos de sarcomas, revelando que alguns subtipos tiveram maiores frequências de GPVs, como o grupo de sarcomas fusocelulares (33,3%), condrossarcoma (30,4%) e leiomiossarcoma (27,8%). Os subtipos com uma menor detecção de GPVs foram fibrossarcoma (0% - de 4 casos), rabdomiossarcoma (9,1%) e sarcoma sinovial (10,5%) (Gráfico 2). Também foram observadas algumas associações gene-subtipo de sarcoma esperadas. Por exemplo, GPVs em *EXT1* e *EXT2* foram detectadas somente em condrossarcomas e GPVs em *RB1* somente em pacientes com osteossarcomas (sendo um deles diagnosticado com osteossarcoma, leiomiossarcoma e lipossarcoma). Além disso, houve um enriquecimento de alterações em *TP53* em leiomiossarcomas e de *NF1* em MPNST.

Gráfico 2 - Frequência de detecção de GPVs dentre os subtipos de sarcomas.



Fonte: Autoria própria.

Nota: O grupo dos sarcomas fusocelulares inclui dois tumores que foram reclassificados como tumores malignos da bainha do nervo periférico, ambos com GPV em *NF1*.

Dentre as VUS, 81,9% dos pacientes (86 de 105) foram detectados com essas variantes no sequenciamento de painel individual. De 170 variantes pontuais, 21 foram consideradas VUS de maior interesse através de preditores *in silico* de patogenicidade (Revel > 0,6) ou pela classificação do Franklin Genoox ser P/PP (Tabela 4).

Tabela 4 - Alterações clínicas de pacientes com 21 VUS de interesse, classificadas como VUS seguindo as recomendações do ACMG⁸⁶.

ID da amostra	Idade	Sexo	Histologia (sarcoma)	HF de câncer	Gene	RefSeq (NM)	HGVS c.	HGVS p.	VAF (gnomAD/ABraOM)	Franklin	REVEL
GRY_3	26	F	LPS	ND	<i>CHEK2</i>	NM_001005735.1	c.1046G>C	p.Gly349Ala	0,000029/ND	PP	0,536
GRY_11	36	M	LPS	Avó paterna: câncer de mama (60 anos)	<i>MLH1</i>	NM_000249.3	c.2221C>A	p.Leu741Met	ND	VUS	0,683
GRY_11	36	M	LPS	Avó paterna: câncer de mama (60 anos)	<i>SDHA</i>	NM_004168.3	c.205G>A	p.Ala69Thr	0,000032/0,000427	VUS	0,724
GRY_12	13	F	OS	Bisavó maternal: câncer de mama; prima materna: câncer de mama (47 anos); primo: leucemia (infância); primo: tumor SNC (infância); avô: câncer de pulmão	<i>PTCH1</i>	NM_000264.3	c.4240G>A	p.Val1414Met	0,000143/ND	VUS	0,606
GRY_13	34	F	LPS	ND	<i>PDGFRA</i>	NM_001347828.1	c.1748G>A	p.Arg583His	0,000028/0,000854	VUS	0,767
GRY_15	38	F	LPS	Pai: câncer de língua (fumante); avó materna: câncer de útero; primo materno de 2º grau: sarcoma NE	<i>BRCA2</i>	NM_000059.3	c.8351G>A	p.Arg2784Gln	0,000012/ND	P	0,77
GRY_17	38	F	LMS	Tia avó maternal: câncer de mama	<i>ATM</i>	NM_000051.3	c.4148C>T	p.Ser1383Leu	0,000036/ND	VUS	0,797
GRY_27	39	F	Sarcoma fusocelular	Mãe: câncer de mama	<i>NBN</i>	NM_002485.4	c.666C>G	p.Phe222Leu	ND	VUS	0,838
GRY_31	40	F	LPS	Negativa	<i>MSH2</i>	NM_000251.2	c.1571G>A	p.Arg524His	0,000008/0,000427	VUS	0,918
GRY_32	39	F	LMS	Pai: câncer de próstata (68 anos); avó materna: câncer de pulmão (80 anos)	<i>CDKN2A</i>	NM_000077.4	c.383G>C	p.Arg128Pro	0,000004/ND	VUS	0,66
GRY_35	35	F	LPS	Avô paterno: câncer de próstata (70 anos); avó paterna: leucemia (65 anos)	<i>SDHB</i>	NM_003000.2	c.112C>T	p.Arg38Cys	0,000020/ND	VUS	0,873
GRY_35	35	F	LPS	Avô paterno: câncer de próstata (70 anos); avó paterna: leucemia (65 anos)	<i>RET</i>	NM_020975.4	c.1699G>A	p.Asp567Asn	0,000171/0,000854	VUS	0,643
GRY_37	29	F	LPS	Negativa	<i>EXT2</i>	NM_207122.1	c.749G>A	p.Arg250Gln	0,00002/ND	VUS	0,83
GRY_43	38	M	LPS	ND	<i>PMS2</i>	NM_000535.6	c.857A>G	p.Asp286Gly	0,000119/ND	VUS	0,626
GRY_49	30	F	SS	Mãe: câncer de mama; tio avó materno: câncer de esôfago; avó materno: câncer gástrico	<i>EXT1</i>	NM_000127.2	c.700A>T	p.Ile234Phe	0,000004/ND	VUS	0,783

GRY_62	37	F	LPS	Avô materno: câncer de próstata; tio: câncer de cólon	<i>PTCH1</i>	NM_001083602.1	c.-169delG	-	ND	P	ND
GRY_69	39	F	CS	ND	<i>TP53</i>	NM_000546.5	c.467G>A	p.Arg156His	0,000016/ND	PP	0,627
GRY_80	20	M	RMS	ND	<i>BRCA1</i>	NM_007294.4	c.3691T>C	p.Phe1231Leu	0,000032/0,000427	VUS	0,642
GRY_309	12	M	FS	ND	<i>EXT2</i>	NM_207122.1	c.238C>T	p.Arg80Trp	0,000048/ND	VUS	0,6
GRY_326	14	F	SE/PNET	Bisavô materno: câncer de mama; avó maternal: melanoma; avô paterno: câncer gástrico; avô materno: câncer gástrico	<i>ATM</i>	NM_000051.3	c.7475T>C	p.Leu2492Pro	ND	PP	0,93
GRY_384	26	F	CS	Mãe: câncer de mama; tio: câncer de cólon; tio: câncer de fígado	<i>CDKN2A</i>	NM_000077.5	c.160A>C	p.Met54Leu	0,000050/ND	VUS	0,696

Fonte: Autoria própria.

Legenda: HF – história familiar; VAF – *variant allele frequency*; F – feminino; M – masculino; ND – não disponível; LPS – lipossarcoma; OS – osteossarcoma; LMS – leiomiossarcoma; SS – sarcoma sinovial; CS – condrossarcoma; RMS – rabdomiossarcoma; FS – fibrossarcoma; SE – sarcoma de Ewing; P – patogênica; PP – provavelmente patogênica; VUS – variante de significado incerto (do inglês *variant of uncertain significance*).

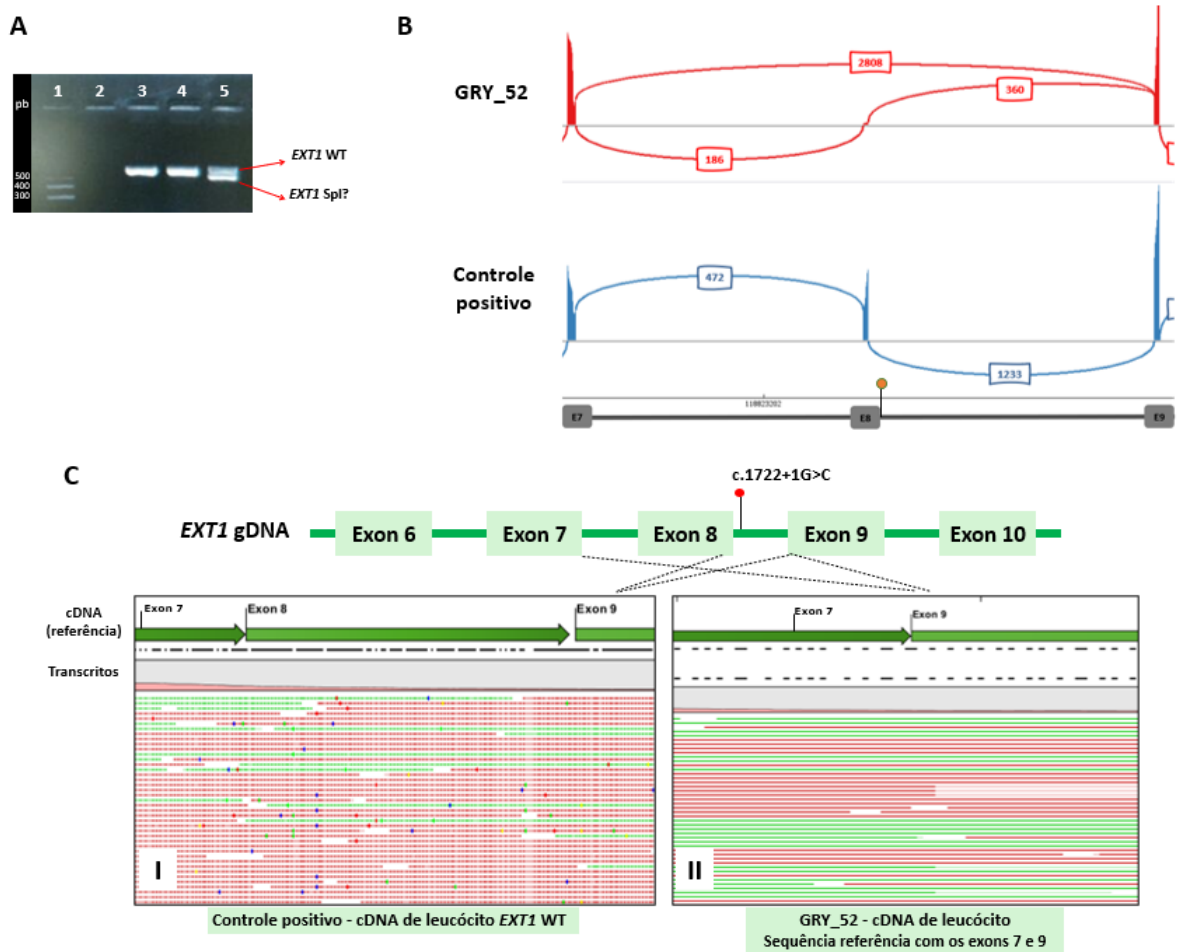
4.2.3 Casos com fenótipos não-câncer esperados

Alguns dos genes em que GPVs foram identificadas levam a fenótipos não neoplásicos e a tumores benignos que representam um critério diagnóstico para as síndromes associadas. Por exemplo, portadores de alterações em *NF1* apresentam múltiplos neurofibromas cutâneos, múltiplas máculas café-ao-leite (*café au lait*), sardas em áreas focais, dentre outras características. Pacientes com *TSC2* alterado podem apresentar múltiplas anormalidades em pele, cérebro, rins, coração e pulmões. Já em pacientes com GPVs em *EXT1/2* é esperado encontrar múltiplos osteocondromas (exostoses), os quais são tumores benignos caracterizados por protuberâncias ósseas envolvidas por uma camada cartilaginosa. Foram identificados 7 pacientes com GPVs nesses genes e os fenótipos clínicos desses indivíduos foram investigados.

Para os três portadores de *NF1* identificados nessa coorte, apenas um dos pacientes apresentou diagnóstico prévio da síndrome Neurofibromatose Tipo I (GRY_47), com múltiplos neurofibromas detectados desde os 11 anos de idade. Os outros dois pacientes (GRY_06 e GRY_312) foram acompanhados em nossa instituição por no mínimo 4 anos, e nenhum outro sinal clínico de NF1 foi relatado nos prontuários médicos (embora nenhum deles tenha sido avaliado por um geneticista durante o período, e ambos tiveram perda de seguimento em 2011). O paciente GRY_06 apresentou diagnóstico de leiomiossarcoma, não sendo um tumor associado ao espectro da NF1. Já os outros dois pacientes, GRY_47 e GRY_312 apresentaram MPNST, o qual é sabido que indivíduos portadores dessa síndrome possuem risco aumentado para essa neoplasia. No entanto, o diagnóstico inicial desses dois tumores era de sarcoma fusocelular (GRY_47) e rabdomiossarcoma embrionário (GRY_312), e foram classificados como MPNST após revisão de lâminas realizada nesse estudo, em combinação com o resultado do teste germinativo.

Um paciente com GPV em *TSC2* teve diagnóstico tardio de esclerose tuberosa (ET) (GRY_42) e apresentou três tumores primários: um sarcoma de partes moles grau III não especificado, aos 13 anos; um adenocarcinoma colorretal aos 29 anos e um carcinoma de células renais com rearranjo-TFE3 aos 38 anos. Ainda, ele foi diagnosticado com a síndrome de West na infância e polipose colônica aos 35 anos. O diagnóstico clínico de ET foi confirmado somente aos 37 anos de idade, a partir de alterações dermatológicas, com lesões em região de face.

Com relação aos genes *EXT1/2*, três pacientes com condrossarcoma apresentaram GPVs nesses genes e nenhum deles foi previamente diagnosticado com osteocondromas múltiplos hereditários (HMO), anteriormente denominados de exostoses múltiplas hereditárias. Um caso (GRY_78) relatou a retirada de um osteocondroma aos 10 anos, e um paciente (GRY_52) foi diagnosticado com múltiplos osteocondromas em membros inferiores em exames de imagem solicitados após os resultados do teste genético desse estudo mostrarem uma variante PP em *EXT1*. Esse paciente era portador de uma variante localizada na região doadora de splice do íntron 8 de *EXT1* (c.1722+1G>C). Para confirmar o efeito dessa variante no splicing, foi realizado sequenciamento dos transcritos de *EXT1* do sangue, o qual foi evidenciado o *skipping* total do éxon 8, causando uma deleção in frame de 30 aminoácidos (c.1633_1722 del; p.(Met546_Val575del)) no domínio de Glycosyl transferase (Figura 5). Baseado nesse achado e no fenótipo específico do paciente, a variante foi reclassificada como patogênica.



Fonte: Autoria própria.

Nota: A. Produtos da RT-PCR de transcritos do gene *EXT1* amplificados dos éxons 6 ao 10, demonstrados em gel de agarose 2%, sendo 1) marcador de peso molecular (100 pb); 2) controle negativo; 3 e 4) controles positivos utilizando cDNA de leucócito e tumor de um paciente sem alteração nesse gene, respectivamente; 5) cDNA de leucócito do paciente GRY_52, evidenciando o transcrito esperado *EXT1* WT (wild-type) de 516 pb e o transcrito aberrante (*EXT1* Spl?) de 426 pb. **B.** *Sashimi plot* demonstrando as junções de *splice*, no qual abaixo, em azul, está representado o controle positivo com todas as junções esperadas dos éxons 7 ao 9 e acima, em vermelho, o cDNA de leucócito do paciente GRY_52 evidencia ambas as junções, sendo a junção esperada entre os éxons 8 e 9 e a junção aberrante entre os éxons 7 e 9, levando a deleção do éxon 8. **C.** Resultado do sequenciamento de *amplicon* de cDNA de leucócito do gene *EXT1* analisado pelo software CLC Genomics Workbench, no qual I - controle positivo, evidenciando o alinhamento dos éxons avaliados e a presença de mRNA com todos os éxons (éxons 7, 8 e 9) e II - cDNA de leucócito do paciente GRY_52, mostrando o alinhamento com a sequência referência com a exclusão do éxon 8 e o ponto de junção éxon 7 e 9. A alteração evidenciada no sítio doador de *splice* do íntron 8 leva à deleção completa do éxon 8 (90 pares de bases).

Figura 5 - Avaliação do efeito da variante intrônica c.1722+1G>C no transcrito do gene *EXT1*.

4.2.4 Avaliação de perda de heterozigose (LOH)

Para investigar a inativação somática através da deleção do alelo selvagem de portadores de GPsV como um possível segundo hit, foi avaliado LOH em tumores frescos congelados ou FFPE de 25 pacientes. Foi observada evidência de LOH do alelo selvagem em 36% desses casos, em pacientes com GPsV relacionados a sarcomas, como *TP53* (5/5), *NF1* (1/3), *RB1* (1/2) e *EXT1* (1/1), e também em *MITF* (1/1), gene não associado a sarcomas (Figura 4, Tabela 5). Em relação à associação dos genes ao risco de sarcomas, LOH do alelo selvagem foi detectada em 50% dos genes com conhecida associação (8/16), em 0% dos genes com associação emergente (0/3) e em 14,3% dos genes de associação desconhecida (1/8). Além disso, também foi rastreado LOH em *XAF1*, em três tumores com o haplótipo *TP53-XAF1*, sendo detectado LOH em todos eles. Para 2 genes, *NF1* e *CHEK2*, foi observado LOH do alelo variante.

Tabela 5 - Casos avaliados para perda de heterozigose em tecido tumoral.

ID do paciente	Histologia do sarcoma	% Celularidade do tumor	Gene	Variante (HGVS c., p.)	Tumor (VAF)	LOH
GRY_03	Lipossarcoma	ND	<i>ERCC4</i>	c.1349G>A, p.Trp450Ter	49%	Sem LOH
GRY_04	Sarcoma NE	90%	<i>TP53</i>	c.844C>T, p.Arg282Trp	96%	LOH
GRY_05	Leiomiossarcoma	80%	<i>RB1</i>	c.751C>T, p.Arg251Ter	54%	Sem LOH
GRY_06	Leiomiossarcoma	100%	<i>NF1</i>	c.1155delT, p.Ile386fs	59%	Sem LOH
GRY_07*	Leiomiossarcoma	90%	<i>TP53</i>	c.1010G>A, p.Arg337His	92%	LOH
GRY_07	Leiomiossarcoma	90%	<i>XAF1</i>	c.400G>T, p.Glu134Ter	97%	LOH
GRY_07	Leiomiossarcoma	90%	<i>TP53</i>	c.1010G>A, p.Arg337His	94%	LOH
GRY_07	Leiomiossarcoma	90%	<i>XAF1</i>	c.400G>T, p.Glu134Ter	98%	LOH
GRY_16	Lipossarcoma	60%	<i>MITF</i>	c.952G>A, p.Glu318Lys	72%	LOH
GRY_16	Lipossarcoma	60%	<i>PALB2</i>	c.355delC, p.Gln119Lysfs*58	47%	Sem LOH
GRY_21*	Condrossarcoma	95%	<i>CHEK2</i>	c.470T>C, p.Ile157Thr	41%	Sem LOH
GRY_21	Condrossarcoma	95%	<i>CHEK2</i>	c.470T>C, p.Ile157Thr	40%	Sem LOH
GRY_22	Leiomiossarcoma	100%	<i>TP53</i>	c.1010G>A, p.Arg337His	96%	LOH
GRY_22	Leiomiossarcoma	100%	<i>XAF1</i>	c.400G>T, p.Glu134Ter	96%	LOH
GRY_23	Leiomiossarcoma	60%	<i>TP53</i>	c.1010G>A, p.Arg337His	72%	LOH
GRY_47	Sarcoma fusocelular	70%	<i>NF1</i>	c.1756_1759delACTA, p.Thr586Valfs*18	3%	Sem LOH
GRY_49	Sarcoma sinovial	90%	<i>MUTYH</i>	c.1187G>A, p.Gly396Asp	47%	Sem LOH
GRY_52	Condrossarcoma	60%	<i>EXT1</i>	c.1633_1722 del, p.Met546_Val575 del	74%	LOH
GRY_64	Sarcoma alveolar	ND	<i>MUTYH</i>	c.1187G>A, p.Gly396Asp	47%	Sem LOH
GRY_64	Sarcoma alveolar	ND	<i>MITF</i>	c.952G>A, p.Glu318Lys	58%	Sem LOH
GRY_75	Condrossarcoma	ND	<i>BRCA2</i>	c.5112_5119delinsG, p.Ile1705Leufs*5	61%	Sem LOH
GRY_300*	Tumor desmoplásico	70%	<i>BRCA2</i>	c.7617+1G>A, p.spl?	53%	Sem LOH
GRY_300	Tumor desmoplásico	70%	<i>BRCA2</i>	c.7617+1G>A, p.spl?	53%	Sem LOH
GRY_312	MPNST	70%	<i>NF1</i>	c.1818C>A, p.Cys606Ter	99%	LOH
GRY_335	Rabdomiossarcoma alveolar	80%	<i>SLX4</i>	c.5197G>T, p.Glu1733Ter	45%	Sem LOH
GRY_336	Tumor fibroso solitário	70%	<i>CHEK2</i>	c.319+2T>A, p.spl?	8%	Sem LOH

GRY_337	Sarcoma sinovial	60%	<i>CHEK2</i>	c.349A>G, p.Arg117Gly	45%	Sem LOH
GRY_338	Mixofibrossarcoma	ND	<i>TP53</i>	c.524G>A, p.Arg175His	55%	Sem LOH
GRY_342	Osteossarcoma	80%	<i>RB1</i>	c.751C>T, p.Arg251Ter	97%	LOH
GRY_367	Osteossarcoma	70%	<i>FANCM</i>	c.1397-3_1397-2dupTA, p.spl?	5,20%	Sem LOH
GRY_369	Osteossarcoma	80%	<i>SLX4</i>	c.2133T>A, p.Tyr711Ter	56%	Sem LOH
GRY_369	Osteossarcoma	80%	<i>TP53</i>	c.1010G>A; p.Arg337His	65%	LOH
GRY_369	Osteossarcoma	80%	<i>XAF1</i>	c.400G>T; p. Glu134Ter	65%	LOH
GRY_375	Osteossarcoma	50%	<i>CHEK2</i>	c.277delT, p.Trp93Glyfs*17	37%	Sem LOH
GRY_381	Osteossarcoma	50%	<i>ERCC3</i>	c.659_660insA, p.Ser221Phefs*2	48%	Sem LOH

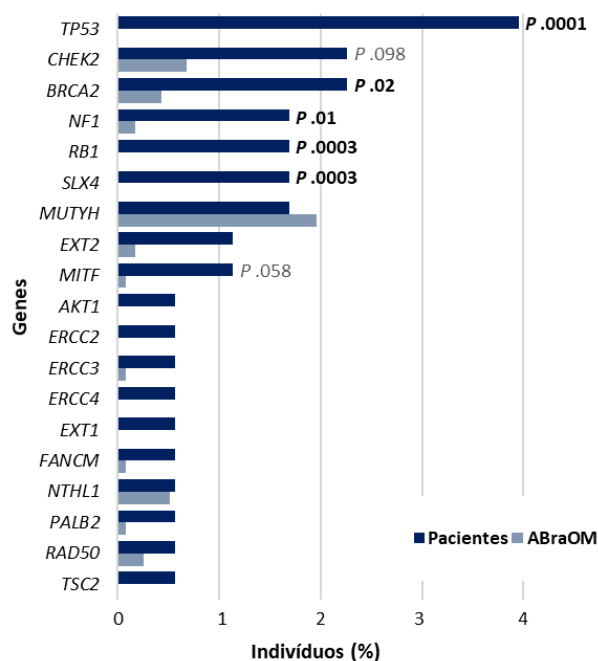
Fonte: Autoria própria.

Nota: * os pacientes sinalizados tiveram 2 áreas diferentes do mesmo tumor avaliadas. VAF – *Variant Allele Frequency*; LOH – *Loss of Heterozygosity*; ND – não descrita.

4.3 COMPARAÇÃO DAS VARIANTES ENCONTRADAS COM BANCOS DE DADOS DE INDIVÍDUOS SAUDÁVEIS

Para a análise de comparação da frequência das variantes P e PP encontradas nos pacientes com sarcomas com a de variantes encontradas em indivíduos saudáveis, foram coletados dados do banco ABraOM (Arquivo Brasileiro Online de Mutações), filtrando variantes descritas somente nos genes que apresentaram GPVs em nosso estudo. Foram encontradas 151 variantes raras, que não apresentam classificação como benignas no banco ClinVar e foram classificadas de acordo com os critérios do ACMG. Os resultados foram submetidos à análise estatística descrita no item 3.4. Alguns genes apresentaram uma maior porcentagem de variantes P/PP quando comparada a frequência dessas alterações em nossa coorte e na coorte de indivíduos coletados do banco de dados ABraOM (Gráfico 3). A frequência das variantes patogênicas detectadas em nossa coorte foi significativamente maior em genes como *TP53*, *RB1* e *NF1* – estes de relação conhecida aos sarcomas, e outros genes sem associação estabelecida a esses tumores, como *BRCA2* e *SLX4* (teste de qui-quadrado $\leq 0,05$ para significância estatística). Interessantemente, alterações patogênicas germinativas em *MITF* tiveram um $P = 0,058$, próximo do limiar de significância.

Gráfico 3 - Frequência de variantes germinativas patogênicas em genes encontrados nessa coorte (177 pacientes com sarcomas) comparado com a coorte de indivíduos idosos saudáveis (1.171 indivíduos brasileiros do banco de dados ABraOM).



Fonte: A autoria própria.

4.4 SEQUENCIAMENTO DE EXOMA GERMINATIVO E TUMORAL DE PACIENTES SELECIONADOS

Foram selecionados nove pacientes para a análise de exoma germinativo e tumoral considerando os critérios descritos no item 3.5. Dessa coorte, cinco pacientes apresentaram condrossarcomas e quatro possuíam diagnóstico de sarcomas ultrarraros. Ainda, três pacientes mencionaram histórico de familiares com câncer (Quadro 10). As métricas de cobertura do sequenciamento de leucócito e tumor estão descritas no Apêndice 9.1, as quais foram consideradas as informações de média de cobertura e cobertura maior do que 10X.

Quadro 10 - Dados clínicos dos 9 pacientes avaliados por ES.

ID da amostra	Sexo	Idade (dx)	Histologia (sarcoma)	Outras neoplasias primárias	História familiar de câncer
GRY_02	F	33	Condrossarcoma	ND	ND
GRY_28	M	39	Condrossarcoma	Neoplasia urotelial papilar	ND
GRY_41	M	31	Condrossarcoma	ND	Avô materno: câncer não especificado
GRY_383	M	35	Condrossarcoma	ND	Mãe: câncer de mama aos 50 anos
GRY_388	F	18	Condrossarcoma	ND	ND
GRY_38	M	30	Angiossarcoma epitelióide	ND	Avó paterna: câncer de pulmão (tabagista); avó maternal: câncer de útero aos 30 anos
GRY_68	M	39	Mixofibrossarcoma	ND	ND
GRY_332	M	24	Mixofibrossarcoma	ND	ND
GRY_365	M	8	Osteossarcoma fibroblástico	ND	ND

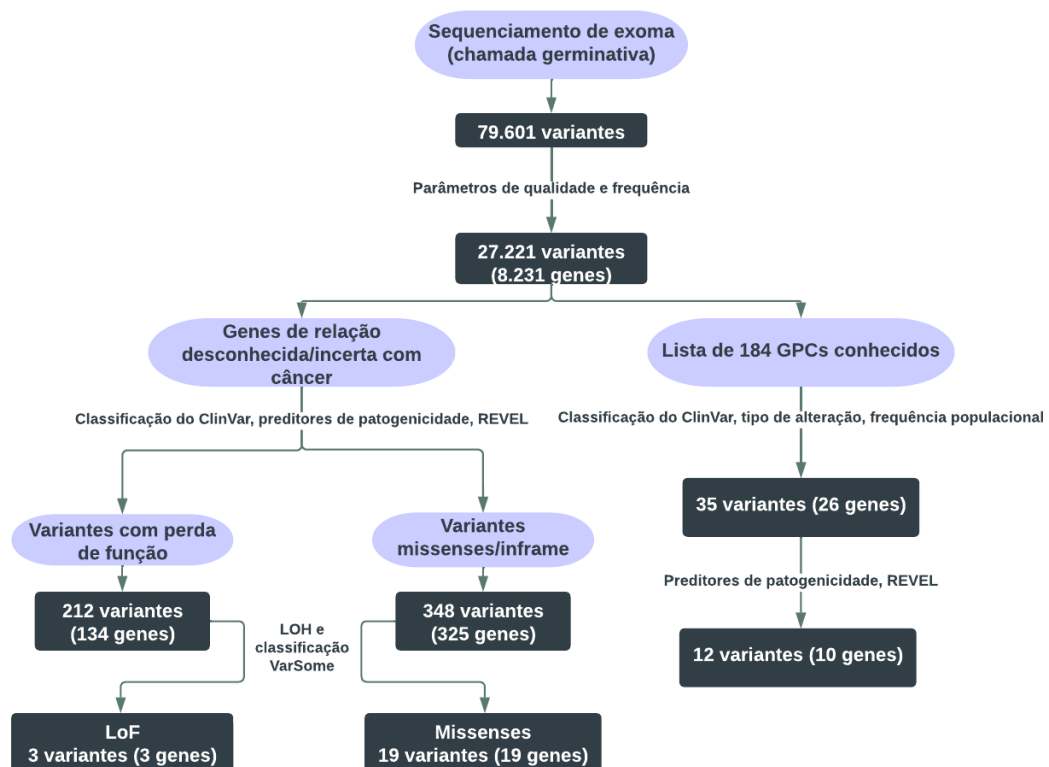
Fonte: Autoria própria.

Legenda: F – feminino; M – masculino; ID – identificação; dx – diagnóstico; ND – não disponível.

4.4.1 Análise Germinativa

Na chamada germinativa e anotação das variantes, foi encontrado um total de 79.601 variantes. Considerando parâmetros de qualidade e frequência alélica populacional descritos no item 3.5 da metodologia, esse número diminuiu para 27.221 variantes (8.231 genes). Nesse momento, foi aplicado um filtro para a identificação de genes conhecidamente relacionados a câncer hereditário, utilizando como base uma lista de 184 genes de predisposição ao câncer. Nessa etapa de filtragem, dois grupos foram gerados: os GPCs conhecidos e os genes de

associação desconhecida/incerta com câncer. Com isso, foram aplicados filtros para a priorização de variantes considerando o tipo da alteração, a classificação em bancos de dados, predição de patogenicidade das variantes encontradas e presença de LOH no tumor, no qual, por fim, foram identificadas 12 variantes em GPCs (10 genes) e, nos genes com essa relação desconhecida, 3 variantes de LoF (3 genes) e 19 *missenses* (19 genes) (Figura 6). No Apêndice 9.2 estão detalhadas as 212 variantes (134 genes) com perda de função selecionadas na etapa anterior à priorização final desse grupo de variantes.



Fonte: Fluxograma criado no Lucidchart <www.lucidchart.com>.

Figura 6 - Fluxograma da anotação e filtragem de variantes na chamada germinativa do sequenciamento de exoma de 9 pacientes.

Na Tabela 6 estão descritas as variantes encontradas em cada gene, além da análise de LOH nos tumores observada através do IGV e informações como as classificações das variantes com base no preditor *in silico* VarSome, do banco de dados ClinVar e do preditor de patogenicidade REVEL. Um mesmo gene (*COL7A1*) apareceu alterado em dois sarcomas distintos (condrossarcoma e sarcoma fibromixóide). O gene *BMPR1A* foi classificado como benigno pelo VarSome, mas foi considerado na priorização por apresentar classificação conflitante pelo ClinVar e LOH no tumor.

Tabela 6 - Panorama geral das informações clínicas e moleculares dos pacientes avaliados por ES na etapa de priorização de genes.

ID da amostra	Histologia (sarcoma)	Chr:Pos	Gene	HGVS c.; p.	LOH no tumor	Classificação do VarSome	Banco de dados ClinVar	REVEL
GRY_02	Condrossarcoma	5:1293676	<i>TERT</i> *	c.1323_1325delGGA; p.Glu441del	sem LOH	VUS	Conflitante	ND
		3:48608578	<i>COL7A1</i> *	c.7120G>A; p.Gly2374Arg	sem LOH	PP	VUS	0,749
		6:29573464	<i>GABBR1</i>	c.2321A>G; p.Tyr774Cys	LOH (65%)	VUS	ND	0,817
		18:28934363	<i>DSG1</i>	c.2204G>T; p.Cys735Phe	sem LOH (61%)	VUS/P	ND	0,712
		2:26684692	<i>OTOF</i>	c.5405C>T; p.Ala1802Val	LOH (70%)	PP	VUS	0,729
GRY_28	Condrossarcoma	14:81574762	<i>TSHR</i> *	c.658G>A; p.Ala220Thr	LOH (88%)	VUS/PP	ND	0,472
		3:19938252	<i>EFHB</i>	c.1651_1652dupGT; p.Arg552fs	LOH (73%)	VUS	VUS	ND
		8:39494831	<i>ADAM18</i>	c.637C>T; p.Gln213Ter	LOH (87%)	VUS/PP	ND	ND
		7:117254733	<i>CFTR</i>	c.3434G>A; p.Trp1145Ter	LOH (95%)	P	ND	ND
		16:2158419	<i>PKD1</i>	c.6749C>T; p.Thr2250Met	LOH (80%)	VUS	Conflitante	0,579
		1:24192072	<i>FUCA1</i>	c.433T>C; p.Trp145Arg	LOH (88%)	VUS/PP	Conflitante	0,803
		22:40757526	<i>ADSL</i>	c.897G>A; p.Met299Ile	LOH (91%)	VUS/PP	ND	0,650
		16:78064560	<i>CLEC3A</i>	c.416G>A; p.Gly139Asp	LOH (77%)	VUS/PP	ND	0812
		10:105658695	<i>STN1</i>	c.521A>G; p.Tyr174Cys	LOH (71%)	VUS	VUS	0,602
		7:100173865	<i>LRCH4</i>	c.1634G>A; p.Arg545His	LOH (90%)	VUS	ND	0,874
		1:28863367	<i>RCC1</i>	c.1046C>T; p.Ser349Leu	LOH (85%)	VUS	ND	0,573
Y:6931939	<i>TBL1Y</i>	c.206A>T; p.Asp69Val	LOH (100%) homozigoto	VUS/PP	P	ND		
GRY_41	Condrossarcoma	X:18923932	<i>PHKA2</i>	c.2852T>C; p.Met951Thr	LOH (89%) homozigoto	PP	ND	0,910
GRY_383	Condrossarcoma	16:89842176	<i>FANCA</i> *	c.1874G>C; p.Cys625Ser	sem LOH	VUS/PP	Conflitante	0,806
		10:88681437	<i>BMPRI1</i> *	c.1327C>T; p.Arg443Cys	LOH (78%)	B	Conflitante	0,888
		10:43622062	<i>RET</i> *	c.3079C>G; p.Leu1027Val	sem LOH	VUS/PP	VUS	0,758
		2:215633977	<i>BARD1</i> *	c.1374C>A; p.Asp458Glu	sem LOH	VUS/PP	VUS	0,729
		X:57146524	<i>SPIN2B</i>	c.539A>G; p.Asp180Gly	LOH (71%) homozigoto	VUS	ND	0,492
GRY_388	Condrossarcoma	1:94517254	<i>ABCA4</i>	c.2588G>C; p.Gly863Ala	LOH (69%)	P	Conflitante	0,798
		12:12312831	<i>LRP6</i>	c.2347C>T; p.Arg783Cys	LOH (65%)	VUS/PP	ND	0,729

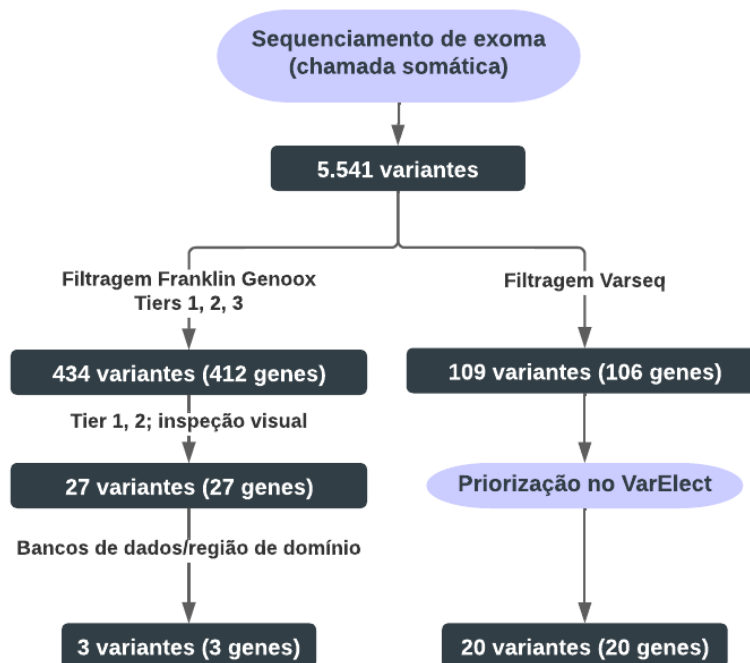
		6:43488700	<i>POLR1C</i>	c.836G>A; p.Arg279Gln	LOH (69%)	P	P	0,859
GRY_38	Angiossarcoma epitelióide	13:32893426	<i>BRCA2</i> *	c.280C>T; p.Pro94Ser	sem LOH	VUS	Conflitante	ND
		22:37465142	<i>TMPRSS6</i>	c.2084T>C; p.Ile695Thr	sem LOH (61%)	VUS/PP	ND	0,860
GRY_332	Sarcoma fibromixóide	5:131931451	<i>RAD50</i> *	c.2165dupA; p.Glu723Glyfs*5	sem LOH	P	P	ND
		3:48631804	<i>COL7A1</i> *	c.263C>T; p.Pro88Leu	sem LOH	PP	ND	0,643
		3:48602358	<i>COL7A1</i> *	c.8676G>T; p.Trp2892Cys	sem LOH	VUS/P	VUS	0,509
GRY_365	Osteossarcoma fibroblástico	8:145738410	<i>RECQL4</i> *	c.2569_2574dupTGCACC; p.Cys857_Thr858dup	sem LOH	VUS/P	Conflitante	ND
		17:7948640	<i>ALOX15B</i>	c.934G>T; p.Ala312Ser	LOH (73%)	VUS	ND	0,702
		10:71332433	<i>NEUROG3</i>	c.367G>C; p.Glu123Gln	LOH (86%)	PP	VUS	0,854

Fonte: Autoria própria.

Legenda: * Os genes destacados com o asterisco são GPCs e os genes não destacados são genes de relação desconhecida com câncer hereditário. ND – não disponível; B – benigna; P – patogênica; PP – provavelmente patogênica; VUS – variante de significado incerto; LOH – perda de heterozigose.

4.4.2 Análise Somática

Adicionalmente, foi realizada a chamada somática do ES, na qual inicialmente foram identificadas 5.541 variantes. Uma das plataformas utilizadas para anotação foi o Franklin Genoox, ao qual foram considerados diversos filtros, como o do modelo de Tier: 1, 2 e 3, tipo de alteração encontrada, frequência alélica e confiança, restando 227 variantes. Destas, foram mantidas as variantes com Tier 1 e 2 (48 variantes), das quais 27 foram consideradas verdadeiras e somáticas após inspeção visual comparativa do alinhamento germinativo e somático de cada amostra (Figura 7, Tabela 7).



Fonte: Fluxograma criado no Lucidchart <www.lucidchart.com>.

Figura 7 - Fluxograma da anotação e filtragem de variantes na chamada somática do sequenciamento de exoma de 9 pacientes.

Tabela 7 - Informações gerais das 27 variantes (27 genes) priorizadas utilizando o Franklin Genoox.

Amostra	Chr:Pos	Gene	HGVS p.	VAF	Classificação pelo Franklin	Oncokb	cBioPortal/Domínio	Gene descrito em sarcomas
GRY_02	1:173,952,597	<i>RC3H1</i>	p.Trp184fs	29.76	PP	NE	ND	Sim
GRY_02	9: 43,844,264	<i>CNTNAP3B</i>	p.Leu534fs	7.55	VUS/P	NE	ND	Não
GRY_02	1: 33,476,435	<i>AK2</i>	c.671-2_671-1insAGCACCTAGCAGCCTTCTCCAAA	7.24	VUS/P	NE	ND	Sim
GRY_02	19: 38,948,809	<i>RYR1</i>	p.Arg682Trp	11.65	VUS/P	NE	ND, domínio SPRY	Sim
GRY_02	5: 127,645,681	<i>FBN2</i>	p.Cys1732Gly	16.67	VUS/P	NE	ND, domínio Calcium-binding EGF	Sim
GRY_28	7: 144,059,763	<i>ARHGEF5</i>	p.Ala10fs	33.73	VUS/P	NE	ND, Rho guanine nucleotide exchange factor	Sim
GRY_28	19: 24,115,978	<i>ZNF726</i>	p.Gly354*	9.52	VUS/P	NE	ND	Sim
GRY_28	10: 46,999,596	<i>GPRIN2</i>	p.Glu240_Val241insMetLysGlu	8.97	VUS/P	NE	ND	Sim
GRY_28	6:109764549	<i>SMPD2</i>	p.Ser270Cys	15.73	VUS/P	NE	ND, Endonuclease/Exonuclease/phosphatase family	Sim
GRY_28	14: 52,521,004	<i>NID2</i>	p.Phe268Ser	74.55	VUS/P	NE	ND, Nidogen-like	Sim
GRY_28	20: 39,316,902	<i>MAFB</i>	p.Ser197Ala	38.14	VUS/P	ND	ND	Sim
GRY_28	19: 8,155,130	<i>FBN3</i>	p.Arg2013Trp	49.02	VUS/P	NE	ND	Sim
GRY_41	16: 2,164,922	<i>PKD1</i>	p.Thr701Asn	13.56	VUS/P	NE	ND	Sim
GRY_68	1: 152,681,680	<i>LCE4A</i>	p.Gly47_Cys48insTyrCysSerSerGlyGly	8.06	VUS/P	NE	ND, Late cornified envelope	Sim
GRY_365	2: 166,532,867	<i>CSRNP3</i>	p.Thr152fs	27.91	VUS/P	NE	ND	Sim
GRY_365	11: 71,238,675	<i>KRTAP5-7</i>	p.Cys111_Cys112insCysGlnSerSerCysTerLysProCysCys	8.96	VUS/P	NE	ND	Sim
GRY_365	17: 60,766,288	<i>MRC2</i>	p.Glu1101*	25.00	VUS/P	NE	ND	Sim
GRY_365	11: 17,419,310	<i>ABCC8</i>	p.Ala1263Val	48.62	VUS/P	NE	ND, ABC transporter transmembrane region	Sim
GRY_365	2: 179,398,695	<i>TTN</i>	p.Gly34216Asp	23.86	VUS/P	NE	ND, domínio Immunoglobulin I-set	Sim
GRY_383	17: 7,578,479	<i>TP53</i>	p.Pro151Ser	26.32	P	Descrita	Descrita, domínio P53 DNA-binding	Sim
GRY_383	9: 98,231,380	<i>PTCH1</i>	p.His634fs	69.70	PP	ND	ND	Sim
GRY_383	18: 42,531,093	<i>SETBP1</i>	p.Leu596Leu	13.92	VUS/B	ND	ND	Sim

GRY_383	20: 2,841,438	<i>VPS16</i>	p.Arg187*	30.30	P	NE	ND, Vps16, região N-terminal	Sim
GRY_383	3: 193,081,961	<i>ATP13A5</i>	p.Trp57fs	31.08	VUS/P	NE	ND, P5-type ATPase cation	Sim
GRY_388	6: 30,997,594	<i>MUC22</i>	p.Ser1466_Gly1475del	27.82	VUS/P	NE	ND	Sim
GRY_388	16: 3,778,098	<i>CREBBP</i>	p.Gln2317Leu	6.04	VUS/P	ND	ND	Sim
GRY_388	4: 190,878,577	<i>FRG1</i>	p.Asn153Asp	6.48	VUS	NE	ND, domínio FRG1-like	Sim

Fonte: Autoria própria.

Legenda: VAF – variant allele frequency; VUS – variant of uncertain significance; PP – provavelmente patogênica; P – patogênica; NE – não especificado; ND – não descrita e não encontrada em região de domínio.

Utilizando o VarSeq para anotação e filtragem das variantes, considerando os parâmetros descritos no item 3.5 da metodologia, foram selecionadas 109 variantes em 106 genes distintos (Tabela 8, Apêndice 9.3). Quanto ao tipo de alteração, foram observadas 47 mutações com perda de função e 62 *missenses*. Quatro variantes foram mantidas na priorização inicial, mas apresentaram REVEL < 0,5 e 41 variantes de LoF não tiveram essa classificação disponível. A ferramenta VarElect foi utilizada para buscar as associações genótipo-fenótipo e, dos 106 genes selecionados, 20 foram discriminados no programa como de associação direta com o fenótipo sarcoma (Tabela 9 e 10).

Posteriormente, as variantes somáticas priorizadas pelas duas estratégias foram analisadas em bancos de dados como OncoKB e cBioPortal, aos quais apenas uma delas estava descrita em ambos os bancos como uma alteração *driver* (*TP53*:p.Pro151Ser). Apesar disso, outros genes com alterações somáticas previamente relatadas em sarcomas também foram encontrados, como *PTCH1* e *CREBBP* (Tabela 11). Ainda, comparando os genes priorizados utilizando as diferentes ferramentas, Franklin e VarElect, os genes que se repetiram foram *TP53*, *PTCH1*, *CREBBP* e *TTN*. Considerando a filtragem do VarSeq, os genes *FBN2* e o *RHPN2* apareceram mutados dois pacientes diferentes, com sarcomas distintos (Tabela 8).

Tabela 8 - Informações gerais das variantes somáticas selecionadas no VarSeq.

ID da amostra	Chr:Pos	Gene	Efeito	HGVS c.	p.	Éxon	REVEL
GRY_28	1:108786278	<i>NBPF4</i>	LoF	NM_001143989.2:c.106C>T	p.Arg36Ter	2	ND
	5:175395058	<i>THOC3</i>	LoF	NM_032361.4:c.154delC	p.Leu52Trpfs*19	1	ND
	7:144059764	<i>ARHGEF5</i>	LoF	NM_005435.4:c.14_15ins(22)	p.Ala10Glyfs*2	2	ND
	22:43930637	<i>EFCAB6</i>	LoF	NM_022785.4:c.4164T>A	p.Cys1388Ter	30	ND
	6:109764549	<i>SMPD2</i>	Missense	NM_003080.3:c.809C>G	p.Ser270Cys	9	0,22;0,936
	10:94250269	<i>IDE</i>	Missense	NM_004969.4:c.1514T>C	p.Ile505Thr	12	0,55
	14:52521004	<i>NID2</i>	Missense	NM_007361.4:c.803T>C	p.Phe268Ser	4	0,872
	15:78882515	<i>CHRNA5</i>	Missense	NM_000745.3:c.782G>A	p.Cys261Tyr	5	0,915
19:8155130	<i>FBN3</i>	Missense	NM_032447.5:c.6037C>T	p.Arg2013Trp	49	0,798	
GRY_38	1:144917828	<i>PDE4DIP</i>	LoF	NM_001350520.1:c.1947delT	p.Ala650Leufs*6	7	ND
	15:22382609	<i>OR4N4</i>	LoF	NM_001005241.3:c.140dupT	p.Thr48Hisfs*41	1	ND
GRY_41	11:89608809	<i>TRIM64B</i>	LoF	NM_001164397.2:c.377delG	p.Ser126Thrfs*3	1	ND
	13:21746601	<i>SKA3</i>	LoF	NM_145061.6:c.208delC	p.Gln70Lysfs*7	3	ND
	17:38955859	<i>KRT28</i>	Missense	NM_181535.3:c.287G>A	p.Arg96His	1	0,937
	18:42532372	<i>SETBP1</i>	Missense	NM_015559.3:c.3067G>A	p.Ala1023Thr	4	0,546
GRY_68	22:50943373	<i>LMF2</i>	Missense	NM_033200.3:c.1295T>C	p.Leu432Pro	10	0,525
	7:55565340	<i>VOPP1</i>	Missense	NM_030796.5:c.157C>T	p.Arg53Trp	3	0,732
	7:102182021	<i>UPK3BL2</i>	Missense	NM_001097615.2:c.*143G>T,NM_001363506.1:c.201G>T	p.Gln67His	5,3	0,022;0,586
GRY_332	1:243328887	<i>CEP170</i>	LoF	NM_014812.3:c.2375C>A	p.Ser792Ter	13	0,023
	9:70176769	<i>FOXD4L5</i>	LoF	NM_001126334.1:c.1215delC	p.Trp406Glyfs	1	ND
	19:33490585	<i>RHPN2</i>	LoF	NM_033103.5:c.1132C>T	p.Gln378Ter	10	ND
	9:75431107	<i>TMC1</i>	Missense	NM_138691.3:c.1744T>C	p.Phe582Leu	19	0,637
22:50599441	<i>MOV10L1</i>	Missense	NM_018995.3:c.3511T>C	p.Tyr1171His	26	0,775	
GRY_365	2:166532868	<i>CSRNP3</i>	LoF	NM_001172173.2:c.455delC	p.Thr152Asnfs*17	6	ND
	6:29912029	<i>HLA-A</i>	LoF	NM_002116.8:c.751delG	p.Asp251Thrfs*46	4	ND
	17:60766288	<i>MRC2</i>	LoF	NM_006039.5:c.3301G>T	p.Glu1101Ter	23	0,241

	19:33490585	<i>RHPN2</i>	LoF	NM_033103.5:c.1132C>T	p.Gln378Ter	10	ND
	2:27609125	<i>PPM1G</i>	Missense	NM_177983.3:c.212G>T	p.Cys71Phe	3	0,714
	2:179398695	<i>TTN</i>	Missense	NM_001267550.2:c.102647G>A	p.Gly34216Asp	358	0,803
	11:17419310	<i>ABCC8</i>	Missense	NM_000352.6:c.3788C>T	p.Ala1263Val	31	0,725
	17:8025275	<i>HES7</i>	Missense	NM_032580.4:c.292G>A	p.Glu98Lys	4	0,558
GRY_383*	9:98231381	<i>PTCH1</i>	LoF	NM_000264.5:c.1899_1902delACAC	p.His634Thrfs*58	14	ND
	17:7578479	<i>TP53</i>	Missense	NM_000546.6:c.451C>T	p.Pro151Ser	5	0,898
GRY_388	8:42036566	<i>PLAT</i>	Missense	NM_000930.5:c.1379C>T	p.Ser460Leu	13	0,392;0,874
	16:3778098	<i>CREBBP</i>	Missense	NM_004380.3:c.6950A>T	p.Gln2317Leu	31	0,726

Fonte: Autoria própria.

Nota: * O GRY_383 foi detectado com 76 variantes somáticas, sendo essas melhor detalhadas no item Apêndice 9.3. ND – não descrito.

Tabela 9 - Detalhamento das variantes encontradas em genes priorizados através do VarElect.

ID da amostra	Chr:Pos	Gene	Efeito	HGVS c.	HGVS p.	Éxon	REVEL
GRY_28	5:175395058	<i>THOC3</i>	LoF	NM_032361.4:c.154delC	p.Leu52Trpfs*19	1	ND
GRY_28	10:94250269	<i>IDE</i>	Missense	NM_004969.4:c.1514T>C	p.Ile505Thr	12	0,55
GRY_332	1:243328887	<i>CEP170</i>	LoF	NM_014812.3:c.2375C>A	p.Ser792Ter	13	0,023
GRY_365	6:29912029	<i>HLA-A</i>	LoF	NM_002116.8:c.751delG	p.Asp251Thrfs*46	4	ND
GRY_365	17:60766288	<i>MRC2</i>	LoF	NM_006039.5:c.3301G>T	p.Glu1101Ter	23	0,241
GRY_365	2:27609125	<i>PPM1G</i>	Missense	NM_177983.3:c.212G>T	p.Cys71Phe	3	0,714
GRY_365	2:179398695	<i>TTN</i>	Missense	NM_001267550.2:c.102647G>A	p.Gly34216Asp	358	0,803
GRY_383	3:49210205	<i>KLHDC8B</i>	LoF	NM_173546.3:c.3G>T	p.Met1Ile	2	0,628
GRY_383	9:98231381	<i>PTCH1</i>	LoF	NM_000264.5:c.1899_1902delACAC	p.His634Thrfs*58	14	ND
GRY_383	8:57026229	<i>MOS</i>	Missense	NM_005372.1:c.313G>T	p.Ala105Ser	1	0,531
GRY_383	8:144945191	<i>EPPK1</i>	Missense	NM_031308.4:c.2231A>G	p.Tyr744Cys	2	0,611
GRY_383	12:81101790	<i>MYF6</i>	Missense	NM_002469.3:c.292G>A	p.Ala98Thr	1	0,872
GRY_383	14:20876282	<i>TEP1</i>	Missense	NM_007110.5:c.317C>T	p.Ser106Phe	2	0,535
GRY_383	14:24976574	<i>CMA1</i>	Missense	NM_001836.5:c.197A>G	p.His66Arg	2	0,627

GRY_383	15:40707168	<i>IVD</i>	Missense	NM_002225.5:c.865G>C	p.Gly289Arg	8	0,888
GRY_383	16:30992394	<i>SETD1A</i>	Missense	NM_014712.3:c.4705A>C	p.Lys1569Gln	17	0,584
GRY_383	17:7578479	<i>TP53</i>	Missense	NM_000546.6:c.451C>T	p.Pro151Ser	5	0,898
GRY_383	22:46614274	<i>PPARA</i>	Missense	NM_005036.6:c.484C>G	p.Leu162Val	6	0,668
GRY_388	8:42036566	<i>PLAT</i>	Missense	NM_000930.5:c.1379C>T	p.Ser460Leu	13	0,392;0,874
	16:3778098	<i>CREBBP</i>	Missense	NM_004380.3:c.6950A>T	p.Gln2317Leu	31	0,726

Fonte: Autoria própria.

Legenda: ND – não descrito; LoF – *loss of function*.

Tabela 10 - VarElect top 20 genes com associação ao termo sarcoma.

Gene	Descrição	Score (VarElect)	Probabilidade média de causar doença
<i>TP53</i>	Tumor Protein P53	52,57	72,85%
<i>MOS</i>	MOS Proto-Oncogene, Serine/Threonine Kinase	7,27	47,11%
<i>PTCH1</i>	Patched 1	3,74	44,93%
<i>MYF6</i>	Myogenic Factor 6	3,00	44,87%
<i>CREBBP</i>	CREB Binding Protein	2,20	69,94%
<i>CMA1</i>	Chymase 1	2,18	32,01%
<i>MRC2</i>	Mannose Receptor C Type 2	1,66	55,57%
<i>KLHDC8B</i>	Kelch Domain Containing 8B	1,59	85,73%
<i>HLA-A</i>	Major Histocompatibility Complex, Class I, A	1,52	0,05%
<i>THOC3</i>	THO Complex 3	1,48	34,55%
<i>PLAT</i>	Plasminogen Activator, Tissue Type	0,61	53,97%
<i>PPARA</i>	Peroxisome Proliferator Activated Receptor Alpha	0,52	36,99%
<i>IVD</i>	Isovaleryl-CoA Dehydrogenase	0,52	68,51%
<i>SETD1A</i>	SET Domain Containing 1A, Histone Lysine Methyltransferase	0,52	69,43%
<i>EPPK1</i>	Epiplakin 1	0,39	0,00%
<i>TTN</i>	Titin	0,28	0,25%
<i>PPM1G</i>	Protein Phosphatase, Mg2+/Mn2+ Dependent 1G	0,22	78,15%
<i>TEP1</i>	Telomerase Associated Protein 1	0,14	0,53%
<i>CEP170</i>	Centrosomal Protein 170	0,06	34,50%
<i>IDE</i>	Insulin Degrading Enzyme	0,06	81,57%

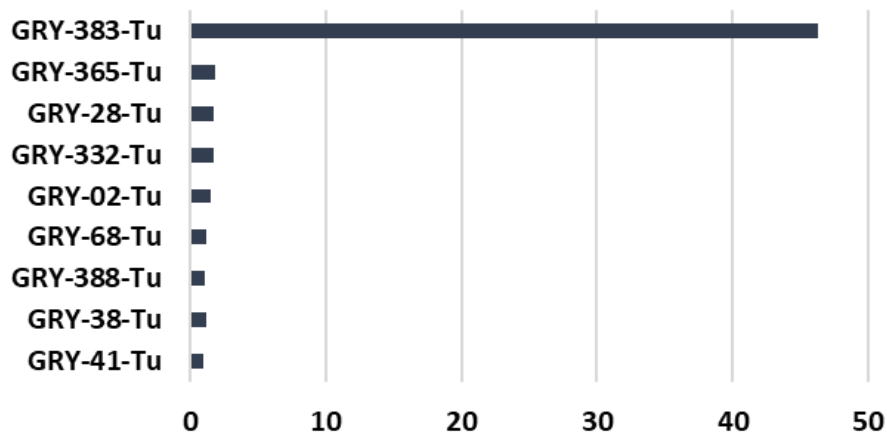
Tabela 11 - Variantes somáticas encontradas em genes de associação estabelecida aos sarcomas anotadas pelo Franklin Genoox, em 2 pacientes (n = 9) avaliados por ES.

ID da amostra	Gene	p.	Chr:Pos/Troca	Tipo de alteração	Região de domínio	Oncokb/cBioPortal
GRY_383	<i>TP53</i>	p.Pro151Ser	17:7578479 G>A	Missense	Sim	Descrita em ambos
	<i>PTCH1</i>	p.His634*	9:98231381- 98231381 GTGT>-	Frameshift	Não	Não descrita
GRY_388	<i>CREBBP</i>	p.Gln2317Leu	16:3778098 T>A	Missense	Não	Não descrita

Fonte: Autoria própria.

A carga mutacional dos tumores (TMB, do inglês *Tumor Mutation Burden*) também foi avaliada, no qual 8 tumores apresentaram uma média de 65 mutações e uma média de TMB de 1,4. Um condrossarcoma (GRY_383) apresentou um número de mutações somáticas e TMB expressivamente maior, sendo 4.517 mutações somáticas e TMB de 46 (Gráfico 4).

Gráfico 4 - Carga mutacional (TMB) dos 9 tumores avaliados por ES.



Fonte: Autoria própria.

5. DISCUSSÃO

Nosso estudo evidenciou que 21,5% dos pacientes jovens brasileiros não selecionados e diagnosticados com sarcomas são portadores de GPVs em genes de predisposição ao câncer. Foram encontradas GPVs em genes de relação já conhecida com sarcomas, mas também em genes em que essa associação é desconhecida, como *PALB2*, *MUTYH*, *SLX4*, *MITF*, *AKT1*, *ERCC3*, *FANCM*, *RAD50* e *TSC2*. Considerando somente os genes de associação bem estabelecida ou emergente ao risco de sarcomas, 14,7% dos pacientes tiveram GPVs. Além disso, 11,3% dos pacientes tiveram VUSs consideradas deletérias por programas de predição *in silico* em genes autossômicos dominantes de predisposição ao câncer. Na comparação entre nossa coorte e uma população controle Brasileira, o enriquecimento na coorte de pacientes para variantes GPVs em genes conhecidos de predisposição a sarcomas foi confirmado para *TP53*, *NF1* e *RB1*, assim como para *BRCA2*, um gene com evidência emergente para o risco de sarcoma, e *SLX4*, um gene sem evidência de aumento de risco para esse tipo de tumor.

Nossos achados corroboram com outros estudos recentes, nos quais detectaram uma média de 20% de GPVs em pacientes com sarcomas, com diferentes critérios de inclusão. Em seu estudo pioneiro, Ballinger et al. (2016)⁹ observaram que 19% dos 1.162 pacientes não selecionados com sarcomas (3-93 anos) eram portadores de GPVs em 72 GPCs e quase 50% foram detectados com variantes possivelmente patogênicas (preditas *in silico* de serem danosas). Além disso, comparado a população controle, houve uma maior frequência de variantes possivelmente patogênicas em genes conhecidos em sarcomas, como *TP53*, mas também em *ATM*, *ATR*, *BRCA2* e *ERCC2*. Em outro estudo de painel multigene avaliando 52 GPCs, 13,6% de 66 pacientes com sarcomas esporádicos abaixo de 50 anos eram portadores de uma GPV, igualmente em genes de associação conhecida e não conhecida com sarcomas (*ATM*, *BRCA2*, *ERCC4*, *FANCC*, *FANCE*, *FANCI*, *MSH6*, *POLE*, *SDHA* e *TP53*)⁶⁴. Mirabello et al. (2020)⁶⁵ realizaram sequenciamento de exoma de 1.244 indivíduos (2-80 anos) com sarcomas, encontrando GPVs em 28% deles. Nesse estudo, eles também realizaram uma comparação da frequência de alterações patogênicas por gene com coortes controle e foi possível identificar novos genes candidatos para o risco aumentado de sarcomas, como *CDKN2A*, *MEN1*, *VHL*, *POT1*, *APC* e *ATRX*. Por último, em um estudo caso-controle mais recente, Gillani et al (2022)¹¹² avaliaram 141 GPCs utilizando sequenciamento de genoma e de exoma germinativo de 1.147

pacientes com sarcomas pediátricos (coorte enriquecida para sarcoma de Ewing, osteossarcoma e rabdmiossarcoma) e mais de 10.000 controles. A partir da testagem de diferentes tipos de sarcomas e associações subtipo-específicas, eles confirmaram o papel de genes previamente implicados na patogênese de desses tumores (*TP53*, *DICER1*, *RB1*, *RECQL4*, *NF1*) e também observaram associação em genes sem uma evidência prévia substancial no risco de sarcomas (*FANCC*, *FANCA*, *PTPN11*, *RECQL*, *MUTYH*, *BRCA2*, *SDHD*). Adicionalmente, outros relatos de casos também descreveram a possível associação de sarcomas com outros GPCs, como *POT1*⁷¹ e *BAP1*⁷².

Embora nosso estudo não tenha fortes evidências para confirmar as diferenças na predisposição genética dentre os vários subtipos de sarcoma, foram observadas associações qualitativas gene-subtipos e uma variada taxa de detecção de GPCs, indo de 0% a 33% entre os subtipos de sarcomas. Por exemplo, associações gene-subtipos bem estabelecidas foram observadas em nossa coorte, visto que todos os três pacientes com *RB1* apresentaram osteossarcomas (um deles com leiomiossarcoma e lipossarcoma metacrônicos), todos os três portadores de GPCs em *EXT1/2* desenvolveram condrossarcomas e 2 de 3 pacientes com GPC em *NF1* foram diagnosticados com MPNST. Adicionalmente, a alteração *TP53*:p.Arg337His foi mais detectada em leiomiossarcomas (60% - 3/5), o que corrobora com outros relatos anteriores na população Brasileira⁴⁴, que também mostram o comportamento distinto dessa alteração quando comparada com outras variantes em *TP53*, uma vez que os leiomiossarcomas correspondem a apenas 9,1% dos sarcomas em pacientes com outras variantes patogênicas em *TP53*¹¹³.

A variante *TP53*:p.Arg337His é conhecida como uma variante causadora da síndrome de Li-Fraumeni associada a um variado risco e espectro de penetrância para diferentes tumores, geralmente levando a um fenótipo diferente da SLF clássica⁴⁰. Essa variante é considerada fundadora no Brasil, visto que sua frequência nessa população é expressivamente maior (MAF = 0,001)⁴⁰ do que comparada a população mundial (MAF gnomAD_v2 1.1 = 0,000012)^{41,42}. Um estudo brasileiro investigou a frequência da *TP53*:p.Arg337His em 502 pacientes não selecionados com sarcomas (1-91 anos, com média de idade de 40 anos) e observou uma taxa de detecção de 8%, sendo que a maioria dos afetados apresentaram leiomiossarcomas (52,5%) com idade de diagnóstico inferior a 40 anos de idade (82,5%)⁴³. Em nosso estudo, a taxa de detecção dessa variante foi 3% (5/177), provavelmente devido a coorte apresentar uma idade mais jovem ao diagnóstico. Além disso, em nosso estudo os

fenótipos dos probandos e seus familiares foram representativos da penetrância e espectro de tumores variáveis dessa variante, uma vez que apenas dois de cinco probandos preencheram os critérios de Chompret (2015)³⁸ e nenhum preencheu os critérios clássicos de Li-Fraumeni. Recentemente, uma variante modificadora, *XAF1*:p.Glu134Ter, que ocorre em haplótipo com a *TP53*:p.Arg337His foi descrita como uma variante que aumenta o risco de sarcomas e outras neoplasias subsequentes⁴⁵. Nesse estudo, a variante p.Glu134Ter foi detectada em 79% dos 203 pacientes carreadores da *TP53*:p.Arg337His⁴⁵, e em nosso estudo, 80% (4/5) dos pacientes foram detectados com esse haplótipo mutante composto.

Nós identificamos que 1,7% da nossa coorte tiveram mais de uma GPV em genes distintos. Esse fenômeno foi recentemente descrito na literatura como Síndrome da Neoplasia Hereditária Multilocus (MINAS, do inglês *Multilocus Inherited Neoplasia Alleles Syndrome*)^{114,115}. Existem alguns poucos estudos descrevendo MINAS em pacientes com sarcomas, como descrito na revisão de McGuigan et al. (2022)¹¹⁵, no qual foram descritos casos com GPVs em dois genes autossômicos dominantes e altamente penetrantes, como *TP53* com *PTEN* e *NF1* com *BRCA2*. Em nossos casos, apenas um dos três indivíduos detectados com MINAS teve uma GPV em um gene dominante com alta penetrância associado a sarcoma (*TP53*) combinado com uma GPV monoalélica em um gene de risco desconhecido para sarcomas (e de herança recessiva – *SLX4*). Os outros dois pacientes apresentaram ambas as GPVs em genes sem correlação prévia a esses tumores (*PALB2* com *MITF* e *MUTYH* com *MITF*). Interessantemente, dois indivíduos com MINAS tiveram a mesma GPV (p.Glu318Lys - MAF ABraOM = 0,000427) em *MITF*, um fator transcricional relacionado a regulação do ciclo celular, que desempenha um papel importante na homeostasia de melanócitos. Essa variante rara, de moderada penetrância, é conhecida por sua associação com um risco aumentado de carcinoma de células renais e melanoma e foi recentemente associada a um maior risco de carcinosarcoma¹¹⁶. Ainda, um estudo utilizando um Registro de Câncer Hereditário, com uma coorte Brasileira, descreveu a prevalência de 0,9% dessa variante (10/1.056 pacientes), sendo que o câncer de mama foi o mais comum dentre os tumores nos probandos e seus familiares, não havendo descrição de sarcomas¹¹⁷.

Além da ocorrência de variantes monogênicas de alto risco que levam a uma maior susceptibilidade para sarcomas, estudos recentes têm proposto um efeito combinado de múltiplas variantes menos penetrantes contribuindo para o risco de sarcomas através do modelo de risco poligênico⁹. Ainda, efeitos poligênicos e variantes germinativas monoalélicas

em genes de resposta ao dano do DNA (DDR, do inglês *DNA damage response*) foram sugestivos de predispor, em idade mais jovem, a subtipos de sarcomas associados a translocações⁹, como foi posteriormente mostrado no caso de variantes no *FANCC* em sarcomas de Ewing¹¹². Nossos pacientes com múltiplas GPVs e pacientes com GPVs monoalélicas em genes DDR associados a síndromes recessivas de predisposição ao câncer (como nos genes das famílias *FANC* e *ERCC*) poderiam representar exemplos desses efeitos poligênicos. Para confirmar a associação desses outros efeitos é necessário ampliar o número de pacientes e controles e, será ainda mais desafiador gerar evidências que justifiquem propor um manejo clínico específico para esses casos.

Existem inúmeras estratégias moleculares que podem ser implementadas para a compreensão do papel de GPVs distintas e genes candidatos na carcinogênese de um tumor específico. A análise somática de LOH considera o modelo dos dois hits que, no geral, ocorre quando genes supressores tumorais são inativados, no qual o alelo selvagem pode ser perdido no tecido tumoral devido a grandes deleções. Para sarcomas, é previsível que ocorra LOH em *GST* quando essa relação é bem definida, como em *NF1*, *RB1* e *TP53*, os quais apresentaram esse evento em vários casos desse estudo (8 de 16 casos avaliados com GPVs em genes de sarcoma). Ainda, alguns genes podem apresentar outros mecanismos mais complexos de iniciação tumoral, como já foi descrito para os genes *EXT1/2* em condrossarcomas, no qual uma perda em mosaico dessas proteínas é necessária para instaurar a tumorigênese^{118,119}. Dos três casos de *EXT1/2* desse estudo, foi avaliada LOH em apenas um caso, o qual apresentou perda do alelo selvagem. Além disso, dentre os genes de associação desconhecida para sarcomas, um paciente com lipossarcoma e GPVs em *MITF* e *PALB2* apresentou LOH na variante de *MITF*. Esse caso ilustra que os dados de LOH precisam ser interpretados com ressalva, especialmente em tumores com grande instabilidade genômica e alterações de número de cópias frequentes, como os sarcomas, uma vez que o *MITF* é um gene conhecido por influenciar na oncogênese através de atividades de ganho de função e um segundo hit não é esperado.

O sequenciamento de exoma é outra abordagem utilizada em diversos estudos, o qual também visa a compreensão de mecanismos que possam estar envolvidos no desenvolvimento de várias doenças. No câncer, pode possibilitar o reconhecimento de novos genes e vias associadas ao aparecimento de determinados tumores e também pode contribuir para a identificação de biomarcadores preditivos e prognósticos. Em sarcomas existem alguns

estudos de ES, sendo que a maior parte deles é centralizada em histologias como osteossarcoma^{65,120}, rabdomyossarcoma^{121,122} e sarcoma de Ewing¹²³. A realização de estudos focados em histologias específicas dos sarcomas auxilia no melhor entendimento desses tumores, visto sua grande heterogeneidade e complexidade, abrangendo características particulares dentre os subtipos. Pensando nisso, o ES foi realizado para 9 pacientes negativos para GPVs nos dados de painel da nossa coorte, e foram selecionados somente pacientes com diagnóstico de condrossarcomas e sarcomas ultrarraros, visto que na literatura há uma carência de estudos de ES com esses tumores, o que revela ainda mais a premência da elaboração de novos trabalhos voltados para histologias distintas.

O condrossarcoma é um dos sarcomas ósseos mais comuns¹²⁴ e a maior parte dos estudos de testes genéticos nesses tumores é realizada para determinação de marcadores tumorais que possam influenciar em decisões terapêuticas¹²⁵⁻¹²⁷. Com isso, poucos trabalhos tem explorado esses tumores a nível germinativo, assim como os sarcomas ultrarraros, os quais existe uma dificuldade ainda maior no seu entendimento, uma vez que são tumores pouco caracterizados e sem informações epidemiológicas, biológicas, preditivas e prognósticas bem definidas¹². Na análise germinativa do ES de pacientes com condrossarcomas e sarcomas ultrarraros, foi possível identificar variantes de significado incerto priorizadas em genes de predisposição ao câncer conhecidos, como *TERT*, *COL7A1*, *TSHR*, *FANCA*, *BMPR1A*, *RET*, *BARD1*, *BRCA2*, *RAD50* e *RECQL4*. Desses GPCs, somente *BRCA2* é um gene de relação emergente com sarcomas e *RECQL4* é um gene com relação bem estabelecida, sendo relacionado a síndrome autossômica recessiva Rothmund-Thomson, a qual pacientes portadores dessa síndrome apresentam maior risco para osteossarcomas. Em nosso estudo, foi identificada uma alteração germinativa monoalélica em *RECQL4* (p.Cys857_Thr858dup) em um paciente com osteossarcoma fibroblástico. A variante encontrada no indivíduo de nossa coorte é do tipo duplicação *in frame*, localizada em uma região que não é de domínio funcional conhecido, porém foi classificada pelo preditor VarSome como VUS/P e no ClinVar com classificação conflitante, o que sugere a necessidade de mais evidências para classificá-la definitivamente. O gene *RECQL4* pertence à família RecQ helicase que codifica a proteína DNA helicase e pode participar da modulação de segregação cromossômica. Um trabalho recente¹²⁸ avaliou dados de sequenciamento germinativo de 5.562 pacientes com tumores pediátricos, no qual 0,43% dos pacientes apresentaram variantes monoalélicas de perda de função no *RECQL4*, sendo que dentre os osteossarcomas,

2% apresentaram alterações heterozigotas nesse gene, estando significativamente enriquecida nesse grupo de pacientes em relação aos controles ($P = 0,00087$, $OR = 7,1$, 95% CI, 2,9–17). Ainda, 38% dos pacientes detectados com variantes em *RECQL4* tiveram essas alterações descritas no domínio DNA helicase, que quando alterado pode levar a instabilidade genômica e impactar no reparo do DNA.

O gene *TERT*, codificante da proteína telomerase transcriptase reversa, que atua na manutenção das extremidades do telômero, foi encontrado alterado em um indivíduo com condrossarcoma nesse estudo. Mutações somáticas em *TERT*, especialmente as mutações *hotspots* na região promotora do gene, são encontradas em até 45% dos condrossarcomas, e foram recentemente descritas como um fator de pior prognóstico¹²⁶. Ainda, mutações somáticas nesse gene também foram descritas em outros sarcomas como lipossarcoma mixóide^{129,130}, e leiomiossarcoma uterino¹³¹. Alterações germinativas em condrossarcomas ou outros sarcomas ainda não foram descritas, sendo que segundo o banco Gene Curation Coalition⁹⁹ foram relatadas associação para risco de câncer apenas para leucemia (evidência moderada) e melanoma (evidência limitada). Apesar disso, no mesmo paciente foi detectada uma variante germinativa em outro GPC, o *COL7A1*, que também foi identificado em outro indivíduo, diagnosticado com sarcoma fibromixóide. Esse gene pode desempenhar um papel importante na tumorigênese dos sarcomas, ainda não esclarecido, visto que quando alterado pode causar um distúrbio no processo de maturação de colágeno através da produção de cadeias pro-colágeno alfa, podendo prejudicar o processo de osteogênese. Mutações somáticas de *COL7A1* em condrossarcomas já foram descritas^{124,132}, apesar de não serem vistos valores preditivos e prognósticos desse achado até o momento¹³³. Alterações germinativas no *COL7A1A* são conhecidas por causar as síndromes de malformações esqueléticas, como as collagenopatias tipo II¹³⁴. Apesar disso, alterações germinativas candidatas em sarcomas foram somente descritas no presente estudo.

O paciente com sarcoma fibromixóide e alteração em *COL7A1A* também apresentou uma outra variante germinativa em *RAD50*, esta de perda de função e previamente classificada no ClinVar como patogênica. O *RAD50* codifica uma proteína de reparo de quebra de DNA dupla fita, e é considerado de associação desconhecida aos sarcomas. Dentre os outros GPCs encontrados em condrossarcomas nesse trabalho (*TSHR*, *FANCA*, *BMPR1A*, *RET* e *BARD1*), o gene *RET*, que se comporta como um oncogene e pode ativar receptores de tirosina quinase, já foi descrito em um estudo de co-segregação, sendo que uma variante germinativa

nesse gene foi detectada no probando, de 35 anos, e em sua mãe, ambos portadores da síndrome da neoplasia múltipla endócrina do tipo II (MEN2), no qual o paciente também apresentou osteossarcoma¹³⁵. No mesmo paciente em que foi detectada a variante em *RET*, também foi visto alteração em *BMPR1A*, gene associado a codificação do receptor de proteína morfogenética óssea 1^a, no qual a proteína BMPR1A se liga a outros ligantes da via do fator de crescimento TGF- β , estando envolvida na transmissão de sinais químicos da membrana celular para o núcleo. O *BMPR1A* é associado a formação de polipose juvenil e pode ser um fator de predisposição para tumores gastrointestinais e colorretais¹³⁶. Em lipossarcomas esporádicos pode atuar como um marcador prognóstico¹³⁷, entretanto, não há descrição de alterações germinativas nesse gene em sarcomas. Dentre os genes *TSHR*, *FANCA* e *BARD1*, nenhum foi anteriormente descrito em condrossarcomas.

Em relação aos GPCs encontrados em sarcomas ultrarraros, somente o *BRCA2* – gene que atua na via de recombinação homóloga para reparo de DNA de fita dupla – já foi descrito em sarcomas, sendo um candidato potencial no aumento de risco para esses tumores^{9,122}. Dentre os genes não GPCs com variantes germinativas priorizadas, para a maioria há descrição de variantes – somente somáticas - em sarcomas, citadas em bancos de dados como TCGA, PanCancer Atlas¹³⁸, The Angiosarcoma Project¹³⁹ e o banco Pediatric Ewing Sarcoma¹¹², e todas classificadas como variantes de significado clínico desconhecido. Um único gene (*SPIN2B*) não foi descrito com mutações somáticas em sarcomas nesses bancos, e nem com alterações germinativas na literatura. De modo geral, apesar de alguns desses genes apresentarem LOH na análise pareada do tumor e classificação patogênica ou provavelmente patogênica através de preditores *in silico*, não é possível afirmar que eles podem de fato aumentar o risco no desenvolvimento de sarcomas, visto que outras análises, como as transcriptoma e de co-segregação familiar, são cruciais para compreender melhor o papel deles nesses tumores.

Na análise somática do ES, foram encontradas variantes de possível relevância clínica em genes relacionados ao câncer e com associação bem estabelecida para sarcomas, como *TP53*, *PTCH1* e *CREBBP*. O *TP53* é um gene supressor tumoral localizado no braço curto do cromossomo 17, o qual codifica a p53, uma fosfoproteína nuclear de 53 kDa, que atua como fator de transcrição através da regulação de diversas atividades celulares, como no controle da progressão do ciclo celular, na diferenciação celular, nos processos de angiogênese, apoptose, na senescência e no reparo do DNA. Com isso, a p53 participa da ativação

transcricional de uma série de outros genes, predominantemente relacionados a parada do ciclo celular, e esse evento pode acontecer em condições de estresse genotóxico, como hipóxia e danos no DNA¹⁴⁰. Embora existam muitos relatos de alterações LoF em GSTs, as alterações genéticas mais encontradas no *TP53* são do tipo *missense*, localizadas em região de domínio de ligação ao DNA, que podem conferir um efeito “dominante negativo”, ou seja, o qual o produto do único alelo mutado interfere na atuação do alelo normal – o que também pode, em alguns casos, conferir um efeito oncogênico do gene, papel descrito mais recentemente^{141,142}. A variante *missense* em *TP53* encontrada no sequenciamento de exoma, p.Pro151Ser, é considerada *driver* em bancos de dados como o TCGA e já foi relatada em leiomiomas no mesmo banco¹³⁸. Nesse estudo, a variante foi detectada em um paciente diagnosticado com condrossarcoma.

O gene *PTCH1* codifica a proteína patched-1 e um componente da via de sinalização de hedgehog, a qual é importante para o desenvolvimento embrionário e tumorigênese. Alterações *driver* no *PTCH1* em pacientes com osteossarcoma também já foram relatadas¹⁴³ e a variante encontrada nesse gene (p.His634Ter) em um indivíduo de nossa coorte é considerada de perda de função, por levar a um códon de parada prematuro, sendo também classificada como potencialmente deletéria em preditores *in silico*, como VarSome e Franklin Genoox.

O gene *CREBBP*, envolvido em diversas atividades transcricionais e com um importante papel na embriogênese, crescimento celular e homeostase, atuando no remodelamento de cromatina, já foi associado ao desenvolvimento de diferentes histologias de sarcomas, no qual em alguns estudos foram relatadas alterações como fusão gênica¹⁴⁴. Vias modificadas que afetam a regulação do gene¹⁴⁵ e alterações pontuais encontradas em pacientes com diagnóstico de sarcoma histiocítico e neoplasias linfoides¹⁴⁶ parecem contribuir no desenvolvimento desses tumores. No presente trabalho, a variante encontrada nesse gene (p.Gln2317Leu), embora também não esteja localizada em região de domínio, é classificada como de possível relevância clínica nos preditores *in silico* anteriormente citados.

Igualmente, outros genes ainda não relatados em sarcomas podem desempenhar um papel importante na tumorigênese dessa condição. Através da ferramenta VarElect e da priorização dos genes baseado em outros parâmetros citados no item 3.4, foram evidenciadas em nossa coorte variantes somáticas de possível significado clínico em genes como *VPS16*, *MOS*, *MYF6*, *CMA1*, *MRC2*, *KLHDC8B* e *THOC3*. Apesar de muito raro, já foram detectadas

variantes nesses genes em pacientes com diferentes subtipos de sarcomas^{138,139}, contudo, os genes *VPS16*, *MOS*, *MYF6*, *CMA1* e *KLHDC8B* foram encontrados alterados em apenas um indivíduo de nossa coorte, com condrossarcoma, ainda detectado com a mutação *driver* em *TP53* e variante em *PTCH1*. As alterações nesses genes podem representar eventos “*passengers*”, uma vez que este caso apresentou uma grande disparidade nos dados de TMB comparado aos outros pacientes (TMB 42 comparado a um TMB médio de 1,4 dos demais tumores), indicando a existência de algum fenômeno que desencadeou mutações pontuais de forma global no genoma, evento que será melhor explorado em um estudo futuro.

Ainda considerando os top 20 genes pontuados pelo VarElect como relacionados a sarcomas, considerando genes com mutações e vias interessantes, no *TTN* foi encontrada uma mutação *missense* em região de domínio, em um paciente de nossa coorte. Esse gene é o maior gene humano e pode aparecer mutado em muitos tumores devido ao acaso sendo que, em sarcomas, ao menos 132 *missenses* já foram relatadas nos bancos de dados presentes no cBioPortal¹⁰⁸, muitas delas localizadas em regiões de domínio, como por exemplo uma encontrada em região de domínio de proteína quinase (p.Glu32222Lys), no subtipo mixofibrossarcoma, e também sendo relatadas outras mutações como as do tipo *frameshift* e de *splice*. Apesar disso, é um gene que desempenha um papel importante na manutenção do tecido muscular, principalmente estriado. No paciente desse estudo, o gene apareceu mutado em um osteossarcoma fibroblástico e, visto que os fibroblastos auxiliam na nutrição e sustentação da musculatura, essa mutação pode ser relevante nesse caso.

Da priorização no Franklin Genoox, foi observado que dois genes (*FBN3* e *RHPN2*) apresentaram mutações somáticas em dois pacientes diferentes, envolvendo condrossarcoma (nos dois indivíduos com alteração no gene *FBN3*), sarcoma fibromixóide e osteossarcoma fibroblástico (no gene *RHPN2*). O *FBN3* codifica a fibrilina, molécula da matriz extracelular que se agrupa em microfibrilas em muitos tecidos conjuntivos, já o *RHPN2* atua como um transdutor de sinais. Ambos são genes citados em sarcomas através de estudos como TCGA, PanCancer Atlas¹³⁸ e The Angiosarcoma Project¹³⁹, apresentando alterações pontuais ou CNVs que podem representar alguma importância nesses tumores.

Ainda, genes de predisposição ao câncer, como *HLA-A*, *PDE4DIP* e *SETBP1* também foram detectados com mutações somáticas que podem desempenhar algum papel importante para esses pacientes. O *HLA-A* (complexo principal de histocompatibilidade classe I, A) atua na ligação de RNA e de antígenos peptídeos, sendo mais relatado em tumores como

o colorretal e de mama. Nesse estudo, foi detectada a variante *HLA-A:p.(Asp251Thrfs*46)* em um paciente com osteossarcoma fibroblástico, ainda não relatada em outros trabalhos, mas, nesses tumores, já foram descritas alterações em pacientes com sarcoma sinovial e osteossarcoma, atuando como marcadores preditivos para decisões terapêuticas, principalmente no uso de imunoterápicos^{147,148}. Alterações no gene *PDE4DIP* podem levar a distúrbios mieloproliferativos associados a eosinofilia, sendo observadas em pacientes com câncer de mama, de órgãos do sistema reprodutor feminino e do trato urinário. Em sarcomas, poucos trabalhos observaram alterações nesse gene, sendo relatadas tanto variantes pontuais quanto fusões gênicas^{149,150}. Nessa coorte, foi detectada a variante *PDE4DIP:p.(Ala650Leufs*6)* em um paciente com angiossarcoma, variante ainda não descrita nessa histologia, embora outras variantes nesse gene já tenham sido encontradas em angiossarcomas, leiomiossarcoma, sarcoma sinovial, dentre outros^{138,139}. Ainda dentre os GPCs, também foi detectada a variante *SETBP1:p.(Ala1023Thr)* em um condrossarcoma. Esse gene atua na replicação de DNA como um oncogene, através da codificação da proteína SET, sendo conhecido por apresentar alterações em neoplasias mieloides^{146,151,152}. Em sarcomas, não há relatos de variantes patogênicas associadas a esse gene.

Nós reconhecemos que esse estudo possui algumas limitações. Primeiramente, em relação ao sequenciamento de painel, variantes causais em regiões não codificantes de GPCs e de outros genes que poderiam estar envolvidos na tumorigênese de sarcomas não foram avaliadas. Segundo, para algumas amostras foi realizado o sequenciamento de painel utilizando a estratégia de pool, o que consequentemente inviabilizou a detecção de CNVs. Por fim, nossos resultados evidenciaram muitas VUSs interessantes e a maioria delas não foi explorada, mesmo que os pacientes apresentassem uma forte história familiar de câncer. Ainda assim, em comparação com estudos anteriores, nós realizamos NGS utilizando um painel multigene amplo, incluímos indivíduos de todos os subtipos de sarcoma e com início mais precoce (tanto pediátricos quanto adultos jovens), de uma população geneticamente mista e sub-representada, e realizamos análise de LOH tumoral para variantes patogênicas, fornecendo *insights* sobre a predisposição genética do sarcoma.

Em relação aos dados de exoma, uma grande quantidade de variantes e genes candidatos foram identificados através dessa abordagem, o que dificulta a compreensão do impacto dessas variantes para definição de relações genótipo-fenótipo, evidenciando a necessidade de outras análises mais abrangentes para que essas correlações possam ser

exploradas, como a análise de transcriptoma e estudos de co-segregação. Nesse estudo, não foram realizadas análises de co-segregação e essa investigação poderia ser inviabilizada para muitos indivíduos devido à ausência de familiares afetados, seja por perda de seguimento ou óbito dos pacientes. Ademais, esses achados ainda são muito preliminares e foi utilizada uma coorte pequena de pacientes com condrossarcomas e sarcomas ultrarraros, sendo necessário um número maior de indivíduos com esses tumores para que mecanismos tumorigênicos sejam melhor elucidados e para que sejam possibilitadas análises comparativas mais adequadas dos achados entre os tumores de uma mesma histologia.

No geral, nossos achados revelaram que 1 entre 5 pacientes com sarcomas abaixo de 40 anos apresentaram uma variante patogênica germinativa em um gene de predisposição ao câncer, incluindo pacientes sem história familiar de câncer e sem sinais clínicos de alterações não tumorais esperadas para uma síndrome subjacente. Atualmente existe apenas uma recomendação de rastreamento genético em pacientes com sarcomas¹⁵³, que compreende particularmente sarcomas de partes moles e sarcomas viscerais em indivíduos abaixo de 46 anos e com um parente de primeiro ou segundo grau com tumores do espectro de *TP53* abaixo de 56 anos, expandindo também para pacientes com STS principalmente diagnosticados em locais aos quais foi realizado radioterapia e ainda com outro tumor relacionado ao *TP53*. Embora este seja um grande passo para melhorar o manejo clínico desses pacientes, ainda um número substancial de pacientes com GPs não seria detectado¹⁵³⁻¹⁵⁵, apontando a urgência na implementação de estratégias de rastreamento genético apropriadas para esses indivíduos e suas famílias.

6. CONCLUSÕES

No presente estudo, pudemos concluir:

- Foi evidenciado que 21,5% dos pacientes jovens brasileiros não selecionados e com sarcomas apresentaram variantes patogênicas em genes de predisposição ao câncer, sendo encontradas GPs tanto em genes aos quais o risco no desenvolvimento de sarcomas é conhecido ou emergente, quanto em genes aos quais esse risco é desconhecido. Ainda, foi possível identificar uma grande deleção no gene *RB1* utilizando a ferramenta do VarSeq. A variante *TP53:p.Arg337His* foi detectada em 2,8% dos pacientes dessa coorte, sendo que 80% deles também apresentaram a variante modificadora de risco *XAF1:p.(Glu134Ter)*;

- Através da análise de LOH do alelo selvagem, foi possível detectar esse evento mais frequentemente em genes reconhecidamente associados a sarcomas, além de um gene não associado, *MITF*, que também apareceu alterado em outro paciente com a mesma variante (*p.Glu318Lys*). Através da análise de *splicing*, foi possível definir o efeito a nível proteico da variante avaliada no *EXT1*, que foi reclassificada para patogênica;

- Na análise comparativa da frequência das variantes P e PP encontradas nos pacientes com sarcomas com a frequência dessas variantes encontradas em indivíduos saudáveis, foi possível identificar que em nossa coorte a frequência foi significativamente maior em genes de relação conhecida com sarcomas (*TP53*, *RB1* e *NF1*), sendo também estatisticamente significativa em um gene emergente (*BRCA2*) e um gene de associação desconhecida para sarcomas (*SLX4*). O gene *MITF* apresentou um P-valor limítrofe;

- Na análise germinativa do ES, foi possível identificar alterações VUS candidatas em GPCs, sendo que desses achados o *RECQL4* foi o único gene de relação estabelecida com sarcomas, além do *BRCA2*, de associação emergente. Foram vistas alterações no gene *COL7A1* em dois sarcomas distintos e, apesar de não haver mais evidências para uma associação genótipo-fenótipo, é um gene que atua em uma via interessante. Dentre os genes não GPs, foram encontradas diversas variantes candidatas, algumas delas classificadas como P/PP e com evidência de LOH no tumor. Na análise somática, foi possível identificar uma mutação *driver* no gene *TP53* e outros genes relacionados a sarcomas também foram detectados com mutações somáticas, como *PTCH1* e *CREBBP*. Utilizando o Franklin Genoox e o VarSeq para a priorização, foi possível evidenciar tanto genes potencialmente acionáveis quanto genes ainda

não relatados nos tumores avaliados, no qual a relevância clínica dessas alterações ainda é incompreendida.

No geral, com base nos achados do painel, as taxas de GPVs variaram entre os subtipos de sarcoma e exibiram algumas associações específicas entre subtipos-genes. No exoma foi possível identificar alguns genes de possível relevância clínica. Com isso, é ainda mais explicitado a necessidade de análises mais amplas, as quais podem permitir uma maior elucidação do comportamento desses tumores, para criar novas estratégias que possam definir condutas mais personalizadas aos pacientes afetados. Dada a grande heterogeneidade dos sarcomas, a elaboração de diretrizes específicas baseadas nos subtipos e em dados populacionais pode ser mais apropriada. No entanto, até que estudos maiores e com estratégias mais abrangentes possam identificar um melhor planejamento de rastreamento, não seria irracional, considerando os achados mais recentes, recomendar que seja realizado o encaminhamento para um especialista em oncogenética e possivelmente testes genéticos para todos os pacientes com sarcoma.

7. REFERÊNCIAS

1. Lodish H, Berk A, Matsudaira P, A. Kaiser C, Krieger M, P. Scott M, et al. *Molecular Cell Biology*. 5th ed. W. H. Freeman; 2003. 973 p.
2. de Freitas Saito R, Vinícius Gomes Lana M, F.V Medrano R, Chammas R. *Fundamentos da Oncologia Molecular*. 1st ed. Atheneu; 2015. 516 p.
3. Rahner N, Steinke V. Hereditary cancer syndromes. *Dtsch Arzteblatt Int*. 2008 Oct;105(41):706–14.
4. Hereditary Cancer Syndromes and Risk Assessment: ACOG COMMITTEE OPINION, Number 793. *Obstet Gynecol*. 2019 Dec;134(6):e143–9.
5. McBride KA, Schlub TE, Ballinger ML, Thomas DM, Tattersall MH. International survey of awareness of genetic risk in the clinical sarcoma community. *Asia Pac J Clin Oncol*. 2016 Jun;12(2):133–42.
6. Banks KC, Moline JJ, Marvin ML, Newlin AC, Vogel KJ. 10 rare tumors that warrant a genetics referral. *Fam Cancer*. 2013 Mar;12(1):1–18.
7. Miranda Alcalde B, Villa Alcázar M, Martínez Romera I, López Ibor B. The importance of Li-Fraumeni syndrome, a hereditary cancer predisposition disorder. *Arch Argent Pediatr*. 2021 Feb;119(1):e11–7.
8. Thomas DM, Ballinger ML. Etiologic, environmental and inherited risk factors in sarcomas. *J Surg Oncol*. 2015 Apr;111(5):490–5.
9. Ballinger ML, Goode DL, Ray-Coquard I, James PA, Mitchell G, Niedermayr E, et al. Monogenic and polygenic determinants of sarcoma risk: an international genetic study. *Lancet Oncol*. 2016 Sep 1;17(9):1261–71.
10. Schiavi A, Lavigne J, Turcotte R, Kasprzak L, Dumas N, Chong G, et al. Using a family history questionnaire to identify adult patients with increased genetic risk for sarcoma. *Curr Oncol Tor Ont*. 2015 Oct;22(5):317–25.
11. Farid M, Ngeow J. Sarcomas Associated With Genetic Cancer Predisposition Syndromes: A Review. *The Oncologist*. 2016 Aug;21(8):1002–13.
12. Stacchiotti S, Frezza AM, Blay JY, Baldini EH, Bonvalot S, Bovée JVMG, et al. Ultra-rare sarcomas: A consensus paper from the Connective Tissue Oncology Society community of experts on the incidence threshold and the list of entities. *Cancer*. 2021 Aug 15;127(16):2934–42.
13. Kallen ME, Hornick JL. The 2020 WHO Classification: What's New in Soft Tissue Tumor Pathology? *Am J Surg Pathol*. 2021 Jan;45(1):e1–23.

14. Sbaraglia M, Bellan E, Dei Tos AP. The 2020 WHO Classification of Soft Tissue Tumours: news and perspectives. *Pathologica*. 2021 Apr;113(2):70–84.
15. Damerell V, Pepper MS, Prince S. Molecular mechanisms underpinning sarcomas and implications for current and future therapy. *Signal Transduct Target Ther*. 2021 Jun 30;6(1):246.
16. Wakeford R. Radiation in the workplace-a review of studies of the risks of occupational exposure to ionising radiation. *J Radiol Prot Off J Soc Radiol Prot*. 2009 Jun;29(2A):A61–79.
17. Cohen RJ, Curtis RE, Inskip PD, Fraumeni JF. The risk of developing second cancers among survivors of childhood soft tissue sarcoma. *Cancer*. 2005 Jun 1;103(11):2391–6.
18. Kleinerman RA, Tucker MA, Abramson DH, Seddon JM, Tarone RE, Fraumeni JF. Risk of soft tissue sarcomas by individual subtype in survivors of hereditary retinoblastoma. *J Natl Cancer Inst*. 2007 Jan 3;99(1):24–31.
19. Kleinerman RA. Radiation-sensitive genetically susceptible pediatric sub-populations. *Pediatr Radiol*. 2009 Feb;39 Suppl 1:S27–31.
20. Benjamin RS, Futreal A. Are sarcomas hereditary? *Lancet Oncol*. 2016 Sep;17(9):1179–81.
21. Thway K, Wang J, Mubako T, Fisher C. Histopathological diagnostic discrepancies in soft tissue tumours referred to a specialist centre: reassessment in the era of ancillary molecular diagnosis. *Sarcoma*. 2014;2014:686902.
22. Schaefer IM, Hong K, Kalbasi A. How Technology Is Improving the Multidisciplinary Care of Sarcoma. *Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Annu Meet*. 2020 May;40:445–62.
23. Baird K, Davis S, Antonescu CR, Harper UL, Walker RL, Chen Y, et al. Gene expression profiling of human sarcomas: insights into sarcoma biology. *Cancer Res*. 2005 Oct 15;65(20):9226–35.
24. Skubitz KM, D'Adamo DR. Sarcoma. *Mayo Clin Proc*. 2007 Nov;82(11):1409–32.
25. Liu W, Wang R, Zhang Y, Wang H, Huang Z, Jin T, et al. Whole-exome sequencing in osteosarcoma with distinct prognosis reveals disparate genetic heterogeneity. *Cancer Genet*. 2021 Aug;256–257:149–57.
26. Krishnan B, Khanna G, Clohisy D. Gene translocations in musculoskeletal neoplasms. *Clin Orthop*. 2008 Sep;466(9):2131–46.
27. Ryan CW, Desai J. The past, present, and future of cytotoxic chemotherapy and pathway-directed targeted agents for soft tissue sarcoma. *Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Annu Meet*. 2013;

28. Shoushtari AN, Van Tine BA, Schwartz GK. Novel treatment targets in sarcoma: more than just the GIST. *Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Annu Meet.* 2014;e488-495.
29. Woll PJ, Reichardt P, Le Cesne A, Bonvalot S, Azzarelli A, Hoekstra HJ, et al. Adjuvant chemotherapy with doxorubicin, ifosfamide, and lenograstim for resected soft-tissue sarcoma (EORTC 62931): a multicentre randomised controlled trial. *Lancet Oncol.* 2012 Oct;13(10):1045–54.
30. Judson I, Verweij J, Gelderblom H, Hartmann JT, Schöffski P, Blay JY, et al. Doxorubicin alone versus intensified doxorubicin plus ifosfamide for first-line treatment of advanced or metastatic soft-tissue sarcoma: a randomised controlled phase 3 trial. *Lancet Oncol.* 2014 Apr;15(4):415–23.
31. Li FP, Fraumeni JF. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann Intern Med.* 1969 Oct;71(4):747–52.
32. Strong LC, Williams WR, Tainsky MA. The Li-Fraumeni syndrome: from clinical epidemiology to molecular genetics. *Am J Epidemiol.* 1992 Jan 15;135(2):190–9.
33. Etzold A, Schröder JC, Bartsch O, Zechner U, Galetzka D. Further evidence for pathogenicity of the TP53 tetramerization domain mutation p.Arg342Pro in Li-Fraumeni syndrome. *Fam Cancer.* 2015 Mar;14(1):161–5.
34. Villani A, Tabori U, Schiffman J, Shlien A, Beyene J, Druker H, et al. Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: a prospective observational study. *Lancet Oncol.* 2011 Jun;12(6):559–67.
35. Li FP, Fraumeni JF, Mulvihill JJ, Blattner WA, Dreyfus MG, Tucker MA, et al. A cancer family syndrome in twenty-four kindreds. *Cancer Res.* 1988 Sep 15;48(18):5358–62.
36. Birch JM, Hartley AL, Tricker KJ, Prosser J, Condie A, Kelsey AM, et al. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res.* 1994 Mar 1;54(5):1298–304.
37. Eeles RA. Germline mutations in the TP53 gene. *Cancer Surv.* 1995;25:101–24.
38. Chompret A, Abel A, Stoppa-Lyonnet D, Brugieres L, Pages S, Feunteun J, et al. Sensitivity and predictive value of criteria for p53 germline mutation screening. *J Med Genet.* 2001 Jan;38(1):43–7.
39. Malkin D. Li-Fraumeni Syndrome. *Genes Cancer.* 2011 Apr;2(4):475–84.
40. Custódio G, Parise GA, Kiesel Filho N, Komechen H, Sabbaga CC, Rosati R, et al. Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. *J Clin Oncol Off J Am Soc Clin Oncol.* 2013 Jul 10;31(20):2619–26.

41. Achatz MI, Zambetti GP. The Inherited p53 Mutation in the Brazilian Population. *Cold Spring Harb Perspect Med*. 2016 Dec;6(12):a026195.
42. Formiga MN da C, de Andrade KC, Kowalski LP, Achatz MI. Frequency of Thyroid Carcinoma in Brazilian TP53 p.R337H Carriers With Li Fraumeni Syndrome. *JAMA Oncol*. 2017 Oct 1;3(10):1400–2.
43. Couto PP, Bastos-Rodrigues L, Schayek H, Melo FM, Lisboa RGC, Miranda DM, et al. Spectrum of germline mutations in smokers and non-smokers in Brazilian non-small-cell lung cancer (NSCLC) patients. *Carcinogenesis*. 2017 Oct 26;38(11):1112–8.
44. Volc SM, Ramos CRN, Galvão H de CR, Felicio PS, Coelho AS, Berardineli GN, et al. The Brazilian TP53 mutation (R337H) and sarcomas. *PLoS One*. 2020;15(1):e0227260.
45. Pinto EM, Figueiredo BC, Chen W, Galvao HCR, Formiga MN, Fragoso MCBV, et al. XAF1 as a modifier of p53 function and cancer susceptibility. *Sci Adv*. 2020 Jun;6(26):eaba3231.
46. Hansen MF, Koufos A, Gallie BL, Phillips RA, Fodstad O, Brøgger A, et al. Osteosarcoma and retinoblastoma: a shared chromosomal mechanism revealing recessive predisposition. *Proc Natl Acad Sci U S A*. 1985 Sep;82(18):6216–20.
47. Kleinerman RA, Schonfeld SJ, Tucker MA. Sarcomas in hereditary retinoblastoma. *Clin Sarcoma Res*. 2012 Oct 4;2(1):15.
48. Wong FL, Boice JD, Abramson DH, Tarone RE, Kleinerman RA, Stovall M, et al. Cancer incidence after retinoblastoma. Radiation dose and sarcoma risk. *JAMA*. 1997 Oct 15;278(15):1262–7.
49. Kleinerman RA, Tucker MA, Tarone RE, Abramson DH, Seddon JM, Stovall M, et al. Risk of new cancers after radiotherapy in long-term survivors of retinoblastoma: an extended follow-up. *J Clin Oncol Off J Am Soc Clin Oncol*. 2005 Apr 1;23(10):2272–9.
50. Ketteler P, Hülsenbeck I, Frank M, Schmidt B, Jöckel KH, Lohmann DR. The impact of RB1 genotype on incidence of second tumours in heritable retinoblastoma. *Eur J Cancer Oxf Engl* 1990. 2020 Jul;133:47–55.
51. Wong JR, Morton LM, Tucker MA, Abramson DH, Seddon JM, Sampson JN, et al. Risk of subsequent malignant neoplasms in long-term hereditary retinoblastoma survivors after chemotherapy and radiotherapy. *J Clin Oncol Off J Am Soc Clin Oncol*. 2014 Oct 10;32(29):3284–90.
52. Marees T, Moll AC, Imhof SM, de Boer MR, Ringens PJ, van Leeuwen FE. Risk of second malignancies in survivors of retinoblastoma: more than 40 years of follow-up. *J Natl Cancer Inst*. 2008 Dec 17;100(24):1771–9.
53. Linn Murphree A. Intraocular retinoblastoma: the case for a new group classification. *Ophthalmol Clin N Am*. 2005 Mar;18(1):41–53, viii.

54. Mallipatna AC, Sutherland JE, Gallie BL, Chan H, Héon E. Management and outcome of unilateral retinoblastoma. *J AAPOS Off Publ Am Assoc Pediatr Ophthalmol Strabismus*. 2009 Dec;13(6):546–50.
55. Scelfo C, Francis JH, Khetan V, Jenkins T, Marr B, Abramson DH, et al. An international survey of classification and treatment choices for group D retinoblastoma. *Int J Ophthalmol*. 2017;10(6):961–7.
56. Wang W, Wei CJ, Cui XW, Li YH, Gu YH, Gu B, et al. Impacts of NF1 Gene Mutations and Genetic Modifiers in Neurofibromatosis Type 1. *Front Neurol*. 2021 Sep 8;12:704639.
57. Evans DGR, Baser ME, McGaughan J, Sharif S, Howard E, Moran A. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J Med Genet*. 2002 May;39(5):311–4.
58. Gutmann DH, Ferner RE, Listernick RH, Korf BR, Wolters PL, Johnson KJ. Neurofibromatosis type 1. *Nat Rev Dis Primer*. 2017 Feb 23;3:17004.
59. Musella A, Palaia I, Domenici L, Casorelli A, Martoccia A, Benedetti Panici P. Vulvar malignancy in neurofibromatosis syndrome. *Case Rep Obstet Gynecol*. 2013;2013:217924.
60. Boyd KP, Korf BR, Theos A. Neurofibromatosis type 1. *J Am Acad Dermatol*. 2009 Jul;61(1):1–14; quiz 15–6.
61. Legius E, Messiaen L, Wolkenstein P, Pancza P, Avery RA, Berman Y, et al. Revised diagnostic criteria for neurofibromatosis type 1 and Legius syndrome: an international consensus recommendation. *Genet Med*. 2021;23(8):1506–13.
62. Ferner RE, Huson SM, Thomas N, Moss C, Willshaw H, Evans DG, et al. Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. *J Med Genet*. 2007 Feb;44(2):81–8.
63. Farschtschi S, Mautner VF, McLean ACL, Schulz A, Friedrich RE, Rosahl SK. The Neurofibromatoses. *Dtsch Arzteblatt Int*. 2020 May 15;117(20):354–60.
64. Galvin R, Watson AL, Largaespada DA, Ratner N, Osum S, Moertel CL. Neurofibromatosis in the Era of Precision Medicine: Development of MEK Inhibitors and Recent Successes with Selumetinib. *Curr Oncol Rep*. 2021 Mar 15;23(4):45.
65. Chan SH, Lim WK, Ishak NDB, Li ST, Goh WL, Tan GS, et al. Germline Mutations in Cancer Predisposition Genes are Frequent in Sporadic Sarcomas. *Sci Rep*. 2017 Sep 6;7(1):10660.
66. Mirabello L, Zhu B, Koster R, Karlins E, Dean M, Yeager M, et al. Frequency of Pathogenic Germline Variants in Cancer-Susceptibility Genes in Patients With Osteosarcoma. *JAMA Oncol*. 2020 May;6(5):724–34.

67. Nilbert M, Therkildsen C, Nissen A, Akerman M, Bernstein I. Sarcomas associated with hereditary nonpolyposis colorectal cancer: broad anatomical and morphological spectrum. *Fam Cancer*. 2009;8(3):209–13.
68. Lee N, Luthra R, Lopez-Terrada D, Wang WL, Lazar AJ. Retroperitoneal undifferentiated pleomorphic sarcoma having microsatellite instability associated with Muir-Torre syndrome: case report and review of literature. *J Cutan Pathol*. 2013 Aug;40(8):730–3.
69. Urso E, Agostini M, Pucciarelli S, Bedin C, D'angelo E, Mescoli C, et al. Soft tissue sarcoma and the hereditary non-polyposis colorectal cancer (HNPCC) syndrome: formulation of an hypothesis. *Mol Biol Rep*. 2012 Oct;39(10):9307–10.
70. de Angelis de Carvalho N, Niitsuma BN, Kozak VN, Costa FD, de Macedo MP, Kupper BEC, et al. Clinical and Molecular Assessment of Patients with Lynch Syndrome and Sarcomas Underpinning the Association with MSH2 Germline Pathogenic Variants. *Cancers*. 2020 Jul 9;12(7):E1848.
71. Dominguez-Valentin M, Sampson JR, Møller P, Seppälä TT, PLSD Collaborators. Analysis in the Prospective Lynch Syndrome Database identifies sarcoma as part of the Lynch syndrome tumor spectrum. *Int J Cancer*. 2021 Jan 15;148(2):512–3.
72. Calvete O, Garcia-Pavia P, Domínguez F, Bougeard G, Kunze K, Braeuninger A, et al. The wide spectrum of POT1 gene variants correlates with multiple cancer types. *Eur J Hum Genet EJHG*. 2017 Nov;25(11):1278–81.
73. Kaszuba MC, Pulido JS, Folpe AL, Pichurin PN, Goodenberger ML, Spinner RJ. Malignant Peripheral Nerve Sheath Tumor in a Patient With BAP1 Tumor Predisposition Syndrome. *World Neurosurg*. 2018 Jan;109:362–4.
74. Encinas G, K. F. Mazzotti T, A. Nagai M. Alterações Genéticas no Câncer. In: *Fundamentos da Oncologia Molecular*. 1st ed. Atheneu; 2015. p. 81–95.
75. Garcia FOT, Mazzotti TKF, Pasini FS. Controle da expressão gênica e suas alterações no câncer. In: *Fundamentos da Oncologia*. 1st ed. Atheneu; 2015. p. 117–31.
76. Kamihara J, Rana HQ, Garber JE. Germline TP53 mutations and the changing landscape of Li-Fraumeni syndrome. *Hum Mutat*. 2014 Jun;35(6):654–62.
77. Soto JL, Blanco I, Díez O, Planells JG, Sánchez IL, Matthijs G, et al. Documento de consenso sobre la implementación de la secuenciación masiva de nueva generación en el diagnóstico genético de la predisposición hereditaria al cáncer. *Med Clínica*. 2018;151(2):80–80.
78. O'Shea R, Clarke R, Berkley E, Giffney C, Farrell M, O'Donovan E, et al. Next generation sequencing is informing phenotype: a TP53 example. *Fam Cancer*. 2018 Jan;17(1):123–8.

79. Soares BL, Brant AC, Gomes R, Pastor T, Schneider NB, Ribeiro-Dos-Santos Â, et al. Screening for germline mutations in mismatch repair genes in patients with Lynch syndrome by next generation sequencing. *Fam Cancer*. 2018 Jul;17(3):387–94.
80. Kadri S, Long BC, Mujacic I, Zhen CJ, Wurst MN, Sharma S, et al. Clinical Validation of a Next-Generation Sequencing Genomic Oncology Panel via Cross-Platform Benchmarking against Established Amplicon Sequencing Assays. *J Mol Diagn JMD*. 2017 Jan;19(1):43–56.
81. Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med*. 2015 Dec 10;373(24):2336–46.
82. Jones RM, Melton PE, Pinese M, Rea AJ, Ingley E, Ballinger ML, et al. Identification of novel sarcoma risk genes using a two-stage genome wide DNA sequencing strategy in cancer cluster families and population case and control cohorts. *BMC Med Genet*. 2019 May 3;20(1):69.
83. Jouenne F, Chauvot de Beauchene I, Bollaert E, Avril MF, Caron O, Ingster O, et al. Germline CDKN2A/P16INK4A mutations contribute to genetic determinism of sarcoma. *J Med Genet*. 2017 Sep;54(9):607–12.
84. Huang KL, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell*. 2018 Apr 5;173(2):355-370.e14.
85. Lu C, Xie M, Wendl MC, Wang J, McLellan MD, Leiserson MDM, et al. Patterns and functional implications of rare germline variants across 12 cancer types. *Nat Commun*. 2015 Dec 22;6:10086.
86. Meyer UA, Zanger UM, Schwab M. Omics and drug response. *Annu Rev Pharmacol Toxicol*. 2013;53:475–502.
87. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med Off J Am Coll Med Genet*. 2015 May;17(5):405–24.
88. Abou Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, Biesecker LG, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat*. 2018 Nov;39(11):1517–24.
89. Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG, ClinGen Sequence Variant Interpretation Working Group. Updated recommendation for the benign stand-alone ACMG/AMP criterion. *Hum Mutat*. 2018 Nov;39(11):1525–30.
90. Brnich SE, Abou Tayoun AN, Couch FJ, Cutting GR, Greenblatt MS, Heinen CD, et al. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med*. 2019 Dec 31;12(1):3.

91. Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med Off J Am Coll Med Genet.* 2020 Feb;22(2):245–57.
92. Lee K, Krempely K, Roberts ME, Anderson MJ, Carneiro F, Chao E, et al. Specifications of the ACMG/AMP variant curation guidelines for the analysis of germline CDH1 sequence variants. *Hum Mutat.* 2018 Nov;39(11):1553–68.
93. Fortuno C, Lee K, Olivier M, Pesaran T, Mai PL, de Andrade KC, et al. Specifications of the ACMG/AMP variant interpretation guidelines for germline TP53 variants. *Hum Mutat.* 2021 Mar;42(3):223–36.
94. Mester JL, Ghosh R, Pesaran T, Huether R, Karam R, Hruska KS, et al. Gene-specific criteria for PTEN variant curation: Recommendations from the ClinGen PTEN Expert Panel. *Hum Mutat.* 2018 Nov;39(11):1581–92.
95. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer. *J Mol Diagn JMD.* 2017 Jan;19(1):4–23.
96. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol.* 2017 May 16;1:PO.17.00011.
97. Rahman N. Realizing the promise of cancer predisposition genes. *Nature.* 2014 Jan 16;505(7483):302–8.
98. Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet.* 2019 Nov;51(11):1560–5.
99. Adam MP, Everman DB, Mirzaa GM, Pagon RA, Wallace SE, Bean LJ, et al., editors. *GeneReviews*[®] [Internet]. Seattle (WA): University of Washington, Seattle; 1993 [cited 2022 Oct 14]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1116/>
100. DiStefano MT, Goehringer S, Babb L, Alkuraya FS, Amberger J, Amin M, et al. The Gene Curation Coalition: A global effort to harmonize gene-disease evidence resources. *Genet Med Off J Am Coll Med Genet.* 2022 Aug;24(8):1732–42.
101. Zuzarte PC, Denroche RE, Fehringer G, Katzov-Eckert H, Hung RJ, McPherson JD. A two-dimensional pooling strategy for rare variant detection on next-generation sequencing platforms. *PLoS One.* 2014;9(4):e93455.
102. Grasel RS, Felicio PS, de Paula AE, Campacci N, Garcia FA de O, de Andrade ES, et al. Using Co-segregation and Loss of Heterozygosity Analysis to Define the Pathogenicity of Unclassified Variants in Hereditary Breast Cancer Patients. *Front Oncol.* 2020 Oct 2;10:571330.

103. Santana Dos Santos E, Spurdle AB, Carraro DM, Briaux A, Southey M, Torrezan G, et al. Value of the loss of heterozygosity to BRCA1 variant classification. *NPJ Breast Cancer*. 2022 Jan 17;8(1):9.
104. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013 Jan;29(1):15–21.
105. Naslavsky MS, Scliar MO, Yamamoto GL, Wang JYT, Zverinova S, Karp T, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. *Nat Commun*. 2022 Dec;13(1):1004.
106. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010 Sep;20(9):1297–303.
107. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011 May;43(5):491–8.
108. Auwera GV der, O'Connor B. *Genomics in the Cloud*. 1st ed. O'Reilly Media; 2020. 493 p.
109. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012 May;2(5):401–4.
110. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013 Apr 2;6(269):p11.
111. Stelzer G, Plaschkes I, Oz-Levi D, Alkelai A, Olender T, Zimmerman S, et al. VarElect: the phenotype-based variation prioritizer of the GeneCards Suite. *BMC Genomics*. 2016 Jun 23;17 Suppl 2:444.
112. Franklin by Genoox. Available from: <https://franklin.genoox.com>
113. Gillani R, Camp SY, Han S, Jones JK, Chu H, O'Brien S, et al. Germline predisposition to pediatric Ewing sarcoma is characterized by inherited pathogenic variants in DNA damage repair genes. *Am J Hum Genet*. 2022 Jun;109(6):1026–37.
114. Ognjanovic S, Olivier M, Bergemann TL, Hainaut P. Sarcomas in TP53 germline mutation carriers: A review of the IARC TP53 database. *Cancer*. 2012 Mar 1;118(5):1387–96.
115. Whitworth J, Skytte AB, Sunde L, Lim DH, Arends MJ, Happerfield L, et al. Multilocus Inherited Neoplasia Alleles Syndrome: A Case Series and Review. *JAMA Oncol*. 2016 Mar 1;2(3):373.

116. McGuigan A, Whitworth J, Andreou A, Hearn T, Tischkowitz M, Maher ER. Multilocus Inherited Neoplasia Allele Syndrome (MINAS): an update. *Eur J Hum Genet.* 2022 Mar;30(3):265–70.
117. Guhan SM, Artomov M, McCormick S, Njauw CN, Stratigos AJ, Shannon K, et al. Cancer risks associated with the germline MITF(E318K) variant. *Sci Rep.* 2020 Dec;10(1):17051.
118. Oliveira LJC, Gongora ABL, Lima FAS, Canedo FSNA, Quirino CV, Pisani JP, et al. Expanding the phenotype of E318K (c.952G > A) MITF germline mutation carriers: case series and review of the literature. *Hered Cancer Clin Pract.* 2021 Jul 21;19:32.
119. Reijnders CMA, Waaijer CJF, Hamilton A, Buddingh EP, Dijkstra SPD, Ham J, et al. No Haploinsufficiency but Loss of Heterozygosity for EXT in Multiple Osteochondromas. *Am J Pathol.* 2010 Oct;177(4):1946–57.
120. de Andrea CE, Zhu JF, Jin H, Bovée JV, Jones KB. Cell cycle deregulation and mosaic loss of *Ext1* drive peripheral chondrosarcomagenesis in the mouse and reveal an intrinsic cilia deficiency: Polarity in peripheral chondrosarcomagenesis. *J Pathol.* 2015 Jun;236(2):210–8.
121. Reimann E, Kõks S, Ho XD, Maasalu K, Märtson A. Whole exome sequencing of a single osteosarcoma case—integrative analysis with whole transcriptome RNA-seq data. *Hum Genomics.* 2014 Dec 11;8:20.
122. Tlemsani C, Leroy K, Gimenez-Roqueplo AP, Mansuet-Lupo A, Pasmant E, Larousserie F, et al. Chemoresistant pleomorphic rhabdomyosarcoma: whole exome sequencing reveals underlying cancer predisposition and therapeutic options. *J Med Genet.* 2020 Feb;57(2):104–8.
123. Kim J, Light N, Subasri V, Young EL, Wegman-Ostrosky T, Barkauskas DA, et al. Pathogenic Germline Variants in Cancer Susceptibility Genes in Children and Young Adults With Rhabdomyosarcoma. *JCO Precis Oncol.* 2021 Jan 11;5:PO.20.00218.
124. Capasso M, Montella A, Tirelli M, Maiorino T, Cantalupo S, Iolascon A. Genetic Predisposition to Solid Pediatric Cancers. *Front Oncol.* 2020 Oct 28;10:590033.
125. Tarpey PS, Behjati S, Cooke S, Van Loo P, Wedge DC, Pillay N, et al. Frequent mutation of the major cartilage collagen gene, COL2A1, in chondrosarcoma. *Nat Genet.* 2013 Aug;45(8):923–6.
126. Davis EJ, Wu YM, Robinson D, Schuetze SM, Baker LH, Athanikar J, et al. Next generation sequencing of extraskeletal myxoid chondrosarcoma. *Oncotarget.* 2017 Feb 21;8(13):21770–7.
127. Zhang Y, Chen Y, Yang C, Seger N, Hesla AC, Tsagkozis P, et al. TERT promoter mutation is an objective clinical marker for disease progression in chondrosarcoma. *Mod Pathol Off J U S Can Acad Pathol Inc.* 2021 Nov;34(11):2020–7.

128. Tap WD, Villalobos VM, Cote GM, Burris H, Janku F, Mir O, et al. Phase I Study of the Mutant IDH1 Inhibitor Ivosidenib: Safety and Clinical Activity in Patients With Advanced Chondrosarcoma. *J Clin Oncol Off J Am Soc Clin Oncol*. 2020 May 20;38(15):1693–701.
129. Maciaszek JL, Oak N, Chen W, Hamilton KV, McGee RB, Nuccio R, et al. Enrichment of heterozygous germline RECQL4 loss-of-function variants in pediatric osteosarcoma. *Cold Spring Harb Mol Case Stud*. 2019 Oct;5(5):a004218.
130. Koelsche C, Renner M, Hartmann W, Brandt R, Lehner B, Waldburger N, et al. TERT promoter hotspot mutations are recurrent in myxoid liposarcomas but rare in other soft tissue sarcoma entities. *J Exp Clin Cancer Res CR*. 2014 Apr 11;33(1):33.
131. Kunieda J, Yamashita K, Togashi Y, Baba S, Sakata S, Inamura K, et al. High prevalence of TERT aberrations in myxoid liposarcoma: TERT reactivation may play a crucial role in tumorigenesis. *Cancer Sci*. 2022 Mar;113(3):1078–89.
132. Choi J, Manzano A, Dong W, Bellone S, Bonazzoli E, Zammataro L, et al. Integrated mutational landscape analysis of uterine leiomyosarcomas. *Proc Natl Acad Sci U S A*. 2021 Apr 13;118(15):e2025182118.
133. Totoki Y, Yoshida A, Hosoda F, Nakamura H, Hama N, Ogura K, et al. Unique mutation portraits and frequent COL2A1 gene alteration in chondrosarcoma. *Genome Res*. 2014 Sep;24(9):1411–20.
134. Nicolle R, Ayadi M, Gomez-Brouchet A, Armenoult L, Banneau G, Elarouci N, et al. Integrated molecular characterization of chondrosarcoma reveals critical determinants of disease progression. *Nat Commun*. 2019 Oct 11;10(1):4622.
135. Nishimura G, Haga N, Kitoh H, Tanaka Y, Sonoda T, Kitamura M, et al. The phenotypic spectrum of COL2A1 mutations. *Hum Mutat*. 2005 Jul;26(1):36–43.
136. Kovac M, Woolley C, Ribi S, Blattmann C, Roth E, Morini M, et al. Germline RET variants underlie a subset of paediatric osteosarcoma. *J Med Genet*. 2021 Jan;58(1):20–4.
137. Blatter R, Tschupp B, Aretz S, Bernstein I, Colas C, Evans DG, et al. Disease expression in juvenile polyposis syndrome: a retrospective survey on a cohort of 221 European patients and comparison with a literature-derived cohort of 473 SMAD4/BMPR1A pathogenic variant carriers. *Genet Med*. 2020;22(9):1524–32.
138. O'Neill HL, Cassidy AP, Harris OB, Cassidy JW. BMP2/BMPR1A is linked to tumour progression in dedifferentiated liposarcomas. *PeerJ*. 2016 Apr 19;4:e1957.
139. Cancer Genome Atlas Research Network. Electronic address: elizabeth.demicco@sinaihealthsystem.ca, Cancer Genome Atlas Research Network. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell*. 2017 Nov 2;171(4):950-965.e28.

140. Painter CA, Jain E, Tomson BN, Dunphy M, Stoddard RE, Thomas BS, et al. The Angiosarcoma Project: enabling genomic and clinical discoveries in a rare cancer through patient-partnered research. *Nat Med*. 2020 Feb;26(2):181–7.
141. Tokino T, Nakamura Y. The role of p53-target genes in human cancer. *Crit Rev Oncol Hematol*. 2000 Jan;33(1):1–6.
142. Thoenen E, Curl A, Iwakuma T. TP53 in bone and soft tissue sarcomas. *Pharmacol Ther*. 2019 Oct;202:149–64.
143. Pinto EM, Zambetti GP. What 20 years of research has taught us about the TP53 p.R337H mutation. *Cancer*. 2020 Nov 1;126(21):4678–86.
144. Czarnecka AM, Synoradzki K, Firlej W, Bartnik E, Sobczuk P, Fiedorowicz M, et al. Molecular Biology of Osteosarcoma. *Cancers*. 2020 Jul 31;12(8):E2130.
145. Micci F, Gorunova L, Agostini A, Johannessen LE, Brunetti M, Davidson B, et al. Cytogenetic and molecular profile of endometrial stromal sarcoma. *Genes Chromosomes Cancer*. 2016 Nov;55(11):834–46.
146. Feng D, Li Z, Yang L, Liang H, He H, Liu L, et al. BMSC-EV-derived lncRNA NORAD Facilitates Migration, Invasion, and Angiogenesis in Osteosarcoma Cells by Regulating CREBBP via Delivery of miR-877-3p. *Oxid Med Cell Longev*. 2022;2022:8825784.
147. Egan C, Lack J, Skarshaug S, Pham TA, Abdullaev Z, Xi L, et al. The mutational landscape of histiocytic sarcoma associated with lymphoid malignancy. *Mod Pathol Off J U S Can Acad Pathol Inc*. 2021 Feb;34(2):336–47.
148. Sundara YT, Kostine M, Cleven AHG, Bovée JVMG, Schilham MW, Cleton-Jansen AM. Increased PD-L1 and T-cell infiltration in the presence of HLA class I expression in metastatic high-grade osteosarcoma: a rationale for T-cell-based immunotherapy. *Cancer Immunol Immunother Cll*. 2017 Jan;66(1):119–28.
149. Rosenbaum E, Seier K, Bandlamudi C, Dickson M, Gounder M, Keohan ML, et al. HLA Genotyping in Synovial Sarcoma: Identifying HLA-A*02 and Its Association with Clinical Outcome. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2020 Oct 15;26(20):5448–55.
150. Roszik J, Khan A, Conley AP, Livingston JA, Groisberg R, Ravi V, et al. Unique Aberrations in Intimal Sarcoma Identified by Next-Generation Sequencing as Potential Therapy Targets. *Cancers*. 2019 Aug 31;11(9):E1283.
151. da Costa LT, Dos Anjos LG, Kagohara LT, Torrezan GT, De Paula CAA, Baracat EC, et al. The mutational repertoire of uterine sarcomas and carcinosarcomas in a Brazilian cohort: A preliminary study. *Clin Sao Paulo Braz*. 2021;76:e2324.
152. Makishima H. Somatic SETBP1 mutations in myeloid neoplasms. *Int J Hematol*. 2017 Jun;105(6):732–42.

153. Carratt SA, Braun TP, Coblenz C, Schonrock Z, Callahan R, Curtiss BM, et al. Mutant SETBP1 enhances NRAS-driven MAPK pathway activation to promote aggressive leukemia. *Leukemia*. 2021 Dec;35(12):3594–9.
154. Gronchi A, Miah AB, Dei Tos AP, Abecassis N, Bajpai J, Bauer S, et al. Soft tissue and visceral sarcomas: ESMO-EURACAN-GENTURIS Clinical Practice Guidelines for diagnosis, treatment and follow-up☆. *Ann Oncol Off J Eur Soc Med Oncol*. 2021 Nov;32(11):1348–65.
155. National Comprehensive Cancer Network. Bone Cancer (Version 1.2023 [Internet]. [cited 2022 Sep 25]. Available from:
https://www.nccn.org/professionals/physician_gls/pdf/bone.pdf
156. National Comprehensive Cancer Network. Soft tissue Sarcoma (Version 2.2022). [Internet]. [cited 2022 Sep 25]. Available from:
https://www.nccn.org/professionals/physician_gls/pdf/sarcoma.pdf

8. ANEXOS

8.1 ANEXO I: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO



TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)
RESOLUÇÃO 466/12 CNS/MS

I - DADOS DE IDENTIFICAÇÃO DO PARTICIPANTE DA PESQUISA OU RESPONSÁVEL LEGAL

1. NOME DO PARTICIPANTE:
 DOCUMENTO DE IDENTIDADE Nº: SEXO : M F
 DATA NASCIMENTO:/...../..... ENDEREÇO:
 Nº APTO:
 BAIRRO: CIDADE: CEP:.....
 TELEFONE: (.....) (.....)
 EMAIL:

2. RESPONSÁVEL LEGAL:
 NATUREZA (grau de parentesco, tutor, curador etc.):
 DOCUMENTO DE IDENTIDADE Nº: SEXO : M F
 DATA NASCIMENTO:/...../..... ENDEREÇO:
 Nº APTO:
 BAIRRO: CIDADE: CEP:.....
 TELEFONE: (.....) (.....)
 EMAIL:

II - DADOS SOBRE A PESQUISA CIENTÍFICA

1. TÍTULO DO PROTOCOLO DE PESQUISA:
 Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens.

2. PESQUISADOR RESPONSÁVEL:
 Dra. Giovana Tardin Torrezan

3. PESQUISADORES ENVOLVIDOS:
 Dra Dirce Maria Carraro – Laboratório de Genômica e Biologia Molecular

Rúbrica do
participante:

Rúbrica do pesquisador
responsável:



Dra Maria Nirvana e Dr Diogo Cordeiro de Queiroz Soares – Departamento de Oncogenética
 Dr Samuel Aguiar – Núcleo de Tumores Colorretais
 Dr Stênio de Cássio Zequi – Núcleo de Urologia
 Dra Cecilia Maria Lima Da Costa e Dra Viviane Sonaglio – Departamento de Oncologia Pediátrica
 Dra Isabela Werneck da Cunha e Dra Mariana Petaccia de Macedo – Departamento de Anatomia Patológica

4. DURAÇÃO DA PESQUISA:

36 meses

III - INFORMAÇÕES AO PARTICIPANTE

O(A) senhor(a) está sendo convidado(a) a participar do projeto de pesquisa (“Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens”), que será realizado no A.C. Camargo Cancer Center, e que tem por objetivo entender as causas genéticas ou herdáveis do aumento de risco para pacientes que desenvolveram tumores raros em idade jovem (crianças, adolescentes e adultos com menos de 40 anos).

Este documento fornece informações sobre o estudo para o qual o(a) senhor(a) está sendo convidado(a) a participar.

IV – OBJETIVOS DA PESQUISA

Em nosso estudo, iremos buscar as possíveis causas genéticas e alterações celulares envolvidas no surgimento de tumores raros de pacientes jovens.

V- JUSTIFICATIVA PARA A PROPOSTA DA PESQUISA

A maioria dos tumores é causada por uma combinação do nosso envelhecimento, fatores ambientais e estilo de vida. No entanto, em alguns casos, certos tipos de câncer parecem ocorrer com maior frequência em algumas famílias e em indivíduos mais jovens. Nesta porção de pacientes, o risco aumentado de desenvolver tumores é causado por alterações herdadas em certos genes (partes específicas do nosso material genético, o DNA), e podem ser transmitidas de pais para filhos. Essas condições genéticas são conhecidas como Síndromes de Predisposição Hereditária ao Câncer.

Pesquisas anteriores identificaram um grupo de genes que, quando alterados, conferem um maior risco de desenvolver diversos tipos de tumores, e estes genes podem ser analisados através de um teste genético para entender as causas genéticas do câncer. A identificação das causas genéticas dos tumores raros em pacientes jovens pode ser importante para que os médicos modifiquem recomendações de exames de rastreamento e acompanhamento de outros familiares.

Rúbrica do
participante:

Rúbrica do pesquisador
responsável:



VI – DESENHO DA PESQUISA

Para realizar este estudo, os dados clínicos e de história familiar de câncer dos pacientes selecionados serão coletados a partir da revisão do prontuário hospitalar e de comunicação pessoal em consulta de aconselhamento genético. O DNA será obtido de amostra de sangue ou saliva e será usado para a análise de alterações genéticas em genes já conhecidos de predisposição hereditária ao câncer (painel gênico) e em outros genes (exoma) através de sequenciamento de nova geração. Para alguns casos, também iremos avaliar o DNA de amostra do tumor que já está armazenada no Biobanco da instituição. Esses dados serão analisados e interpretados para contribuir com a compreensão das bases genéticas do surgimento desses tumores.

VII – DESCRIÇÃO DOS PROCEDIMENTOS

A sua aceitação em participar desse estudo será formalizada pela assinatura deste documento. Se você decidir participar, deverá preencher e assinar esse Termo de Consentimento Livre e Esclarecido onde você autoriza:

- Coletar 2 mL de saliva (1 tubo) para o exame genético. O(a) senhor(a) não poderá ter comido, bebido, fumado ou mastigado chiclete nos 30 minutos que antecedem a coleta;

OU

- Coletar 8 mL de sangue em frasco contendo anticoagulante EDTA (2 tubos) para o exame genético. Não há necessidade de preparo para essa coleta;
- Fornecer informação clínica sobre os tumores que você desenvolveu e/ou seus familiares desenvolveram, caso seja necessário informações adicionais que não estão em seu prontuário hospitalar.

Caso você concorde em participar do estudo, esta será uma colaboração voluntária, não havendo nenhuma forma de remuneração por isto. Do mesmo modo, esta decisão não afeta de nenhum modo o que será cobrado de você ou de sua seguradora de saúde. Informamos que não haverá nenhum custo relacionados a sua participação neste estudo. A coleta de sangue ou saliva poderá ser realizada em qualquer dia em que você estiver no A.C. Camargo Cancer Center. Preferencialmente, iremos contatá-lo(a) nas datas que você tiver exames, consultas ou qualquer procedimento médico. Desta forma, você não precisará vir ao hospital apenas para a coleta desta pesquisa e, portanto, não haverá gastos adicionais referentes ao transporte e alimentação.

Rúbrica do
participante:

Rúbrica do pesquisador
responsável:



As amostras de sangue ou saliva coletadas serão processadas e o DNA extraído será armazenado pelo Biobanco do A.C. Camargo Cancer Center. Toda identificação do seu material será feita por códigos para preservar sua confidencialidade. O DNA será utilizado para a realização dos testes genéticos (análise de painel de genes e/ou exoma). Após o teste, se estiver de acordo, você receberá um laudo com o resultado durante consulta de aconselhamento genético com o médico geneticista colaborador deste projeto, que irá explicar detalhadamente as implicações dos resultados genéticos encontrados.

VIII - DESCONFORTOS E RISCOS ESPERADOS DECORRENTES DO PROCEDIMENTO

No caso de ser realizada a coleta de sangue, poderá haver alguma dor em decorrência da punção. Complicações de coleta de sangue são raras e geralmente são de pequeno porte. Se houver extravasamento de sangue da veia no local da punção geralmente há uma mancha roxa (hematoma) e um pequeno desconforto que desaparece em poucos dias. As coletas de sangue serão realizadas por profissionais especialmente treinados, o que diminui as chances de complicações.

Nenhum dano imediato ou tardio, que comprometa a sua saúde, poderá ser decorrente deste estudo, uma vez que você não será submetido a nenhum procedimento adicional aqueles realizados durante o seu tratamento neste hospital.

Existe a possibilidade de identificarmos alterações no seu DNA que podem influenciar o seu risco de desenvolver câncer. A detecção destas alterações pode resultar em mudanças de conduta clínica para o participante e seus familiares. O resultado do teste genético pode causar angústias e ansiedade, e para minimizar estes riscos ele é realizado em consulta de aconselhamento genético.

IX - BENEFÍCIOS QUE PODERÃO SER OBTIDOS

A maioria dos participantes não se beneficiará diretamente com a participação nesse estudo. Contudo, o estudo contribuirá para um melhor entendimento dos fatores que podem causar o aparecimento de tumores raros em idade jovem, desta forma, pode ajudar no desenvolvimento de métodos de diagnóstico e tratamento mais eficazes. Alguns participantes poderão ser beneficiados a partir da identificação de variantes genéticas de aumento de risco para o desenvolvimento de tumores, o que pode resultar em mudanças de conduta clínica apropriadas para estes pacientes.

X – CONFIDENCIALIDADE

Rúbrica do
participante:

Rúbrica do pesquisador
responsável:



A confidencialidade de suas informações será mantida e sua identidade será preservada, sendo que somente os membros da equipe médica e do Comitê de Ética em Pesquisa terão acesso aos registros. A sua participação neste estudo é voluntária, tendo o direito de retirar-se a qualquer momento. A recusa ou desistência da participação nesse estudo não irá prejudicar seu acompanhamento médico e tratamento.

XI - DANOS RELACIONADOS À PESQUISA

Qualquer dano resultante da sua participação no estudo será avaliado e tratado de acordo com os benefícios e cuidados a que você tem direito. Ao assinar este formulário de consentimento você não está abrindo mão de qualquer um dos seus direitos legais.

XII - ACOMPANHAMENTO, ASSISTÊNCIA E RESPONSÁVEIS

Os pesquisadores, médicos e a equipe envolvida na pesquisa se comprometem a dar informação atualizada ao longo do estudo, caso este seja o seu desejo. **TELEFONES PARA CONTATO EM CASO DE INTERCORRÊNCIAS CLÍNICAS OU QUALQUER DÚVIDA SOBRE O ESTUDO:** (011) 2189-5000 – ramal 2953, entrar em contato com Dra. Giovana Tardin Torrezan (pesquisadora).

XIII - OBSERVAÇÕES COMPLEMENTARES

As amostras coletadas serão processadas pelo Biobanco do A.C. Camargo Cancer Center. O excedente, não utilizado por este projeto de pesquisa, poderá ser descartado ou, se autorizado pelo participante, ser armazenado e utilizado em outros projetos, após a aprovação dos mesmos pelo Comitê de Ética em Pesquisa desta Instituição, e quando necessário, da Comissão Nacional de Ética em Pesquisa. Dependendo da escolha do participante, poderá ser ou não necessária a assinatura de novo TCLE. Para isso, pedimos que O(A) senhor (a) escolha uma das seguintes opções quanto ao armazenamento de sua amostra biológica:

1. () **Autorizo** o armazenamento da minha amostra biológica excedente ao final deste estudo para que a mesma possa ser utilizada em outros protocolos de pesquisa.
2. () **Não autorizo** o armazenamento da minha amostra biológica excedente ao final deste estudo para que a mesma possa ser utilizada em outros protocolos de pesquisa. Desta forma, compreendo que o excedente será inutilizado pelo Biobanco da instituição.

Se você escolheu a resposta do item 1, favor assinalar uma das seguintes opções:

- () **Quero receber** um novo TCLE a cada nova pesquisa realizada no A.C. Camargo Cancer Center.
- () **Não quero receber** um novo TCLE a cada nova pesquisa no A.C. Camargo Cancer Center. Portanto, concordo que este TCLE, uma vez assinado, valerá também para os demais protocolos de pesquisa.

Rúbrica do participante:

Rúbrica do pesquisador responsável:



Em casos de óbito ou condições incapacitantes, o(A) senhor (a) tem a possibilidade de ceder os direitos sobre sua amostra biológica armazenada aos seus sucessores ou outras pessoas de sua confiança. Para isso, pedimos que expresse a sua vontade com base nas opções abaixo:

() Cedo os direitos sobre minha amostra biológica armazenada nas situações acima referidas.

Favor indicar ao menos 1 (um) nome de pessoa autorizada:

Nome: _____

Grau de Parentesco: _____

Telefone: () _____

Nome: _____

Grau de Parentesco: _____

Telefone: () _____

() Não cedo os direitos sobre minha amostra biológica armazenada nas situações acima referidas.

Os dados fornecidos, coletados e obtidos neste protocolo de pesquisa, poderão ser utilizados em estudos futuros. A partir de seu material biológico armazenado, futuras pesquisas poderão realizar novas abordagens científicas com o intuito de, por exemplo, investigar a ocorrência de outras alterações em seu DNA.

Em qualquer ocasião e a qualquer tempo, você terá o direito de acesso às informações e resultados obtidos com a utilização de seu material biológico armazenado, recebendo orientações quanto as suas implicações, incluindo aconselhamento genético quando necessário.

Esta pesquisa poderá ser interrompida durante a sua realização e após o aval do CEP da instituição, em casos onde: a) O pesquisador responsável decida que algum motivo ou situação possa pôr em risco a segurança do participante ou; b) O CEP julgue que o estudo esteja sendo conduzido de maneira eticamente inaceitável.

XIV - QUEM DEVO CONTATAR EM CASO DE DÚVIDAS:

Pesquisador Responsável: Dra. Giovana Tardin Torrezan

Laboratório de Genômica e Biologia Molecular do A.C. Camargo Cancer Center - São Paulo.

Telefones para contato: (11) 2189-5000 ramal 5023

Endereço: Rua Professor Antônio Prudente, 211 – Liberdade – São Paulo.

Se o pesquisador responsável não fornecer as informações/ esclarecimentos suficientes, por favor, entre em contato com o Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antônio Prudente – A.C. Camargo Cancer Center pelo telefone (11) 2189-5000, ramal 5020 de segunda-feira à quinta-feira das 8 horas às 18 horas e sexta-feira das 8 horas às 17 horas.

Na impossibilidade de contato com as referências acima, o senhor (a) pode solicitar informações do estudo para a Comissão Nacional de Ética em Pesquisa (CONEP) através do telefone (61) 3315-2951 / fax (61) 3226-6453 / ou e-mail conep@saude.gov.br.

Rúbrica do
participante:

Rúbrica do pesquisador
responsável:



Este documento será elaborado em 2 (duas) vias. O(A) senhor (a) receberá uma das vias originais e a outra será arquivada pelo pesquisador em seu arquivo de pesquisa.

Eu,

_____ declaro ter lido, compreendido e discutido o conteúdo do presente Termo de Consentimento e **concordo em participar desse estudo de forma livre e esclarecida** autorizando os procedimentos acima relacionados:

Assinatura do participante ou responsável legal

_____/_____/_____
Data

Assinatura do responsável pela pesquisa
Dra. Giovana Tardin Torrezan

_____/_____/_____
Data

8.2 ANEXO II: TERMO DE ASSENTIMENTO PARA CRIANÇAS NÃO ALFABETIZADAS E PARA CRIANÇAS E ADOLESCENTES



Termo de Assentimento para crianças não alfabetizadas.

Título do projeto: Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens.

Este documento deve ser lido para a criança por um adulto.

Os seres vivos possuem **um montão** de estruturas **pequeninhas**, chamadas de **células**, que nós não conseguimos ver, e que trabalham juntas para manter nosso corpo funcionando bem.



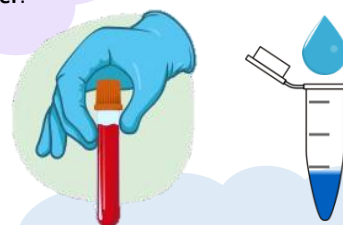
Todos os dias, as nossas células **conseguem lutar** contra vários vilões que tentam deixar nosso corpo doente.



Mas, às vezes, elas não conseguem lutar sozinhas e consertar os problemas que aparecem, e nisso, o **câncer pode aparecer**.



Para buscar entender porquê o câncer aparece, cientistas estudam as células dos pacientes. E para isso, **precisamos da sua ajuda!**



É bem simples!

Pedimos para coletar um pouquinho de sangue do seu braço ou de saliva de sua boca, pois eles serão **muito importantes** para o nosso trabalho!

Você gostaria de nos ajudar? [] SIM [] NÃO

Nome da criança : _____

Assinatura _____ Data: _____

Nome da pessoa que obteve o consentimento: _____

Assinatura _____ Data: _____

Rúbrica do pesquisador responsável:

Versão 1.0, Julho de 2018

Página 1 de 1

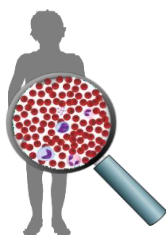
8.3 ANEXO III: TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO

**Termo de Assentimento Livre e Esclarecido.**

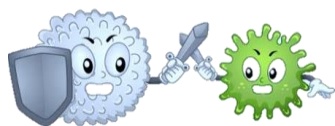
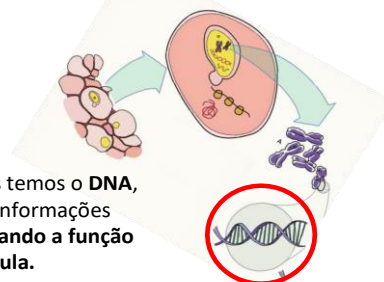
Título do projeto: Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens.

Este documento deve ser lido para a criança/adolescente por um adulto.

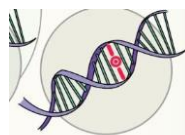
Nosso corpo é composto por **trilhões de partes - chamadas células**, que trabalham em conjunto para manter nosso corpo em **equilíbrio**.



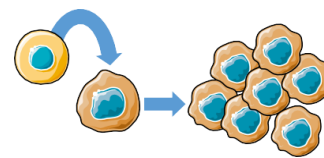
Dentro das células nós temos o **DNA**, que carrega várias informações importantes, **comandando a função de cada célula**.



Dessa forma, diariamente, nosso corpo **consegue lutar** contra vários vilões que tentam causar o desequilíbrio de nossas células e gerar doenças.



Mas, às vezes, ele não consegue lutar sozinho e não corrige os **erros** que podem ocorrer no DNA e em **regiões específicas** dele, chamadas de **genes**.



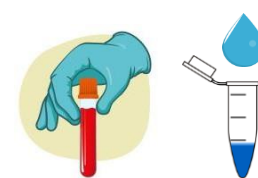
Então, **células defeituosas** se multiplicam, podendo causar o **câncer**.



Algumas famílias podem carregar esses **genes alterados, transmitindo de pais para filhos**. Com isso, às vezes câncer aparece em pessoas mais jovens.



Assim, precisamos entender quais genes estão alterados e podem causar o câncer. Para isso, **precisamos da sua ajuda!**



Para isso, pedimos para coletar um pouquinho de sangue de seu braço ou de saliva de sua boca. Eles serão **muito importantes** para o nosso trabalho!

Você gostaria de nos ajudar? SIM NÃO

Nome da criança: _____

Assinatura _____ Data: _____

Nome da pessoa que obteve o consentimento: _____

Assinatura _____ Data: _____

Rúbrica do pesquisador responsável:

9. APÊNDICES

9.1 MÉTRICAS DE COBERTURA DO ES DE LEUCÓCITO E TUMOR.

ID da amostra	Média de cobertura (X)		% Cobertura maior do que 10X	
	Leucócito	Tumor	Leucócito	Tumor
GRY_02	50,9	92,01	97,2%	98,3%
GRY_28	94,10	80,23	98,4%	98,2%
GRY_41	56,26	85,95	97,4%	98,2%
GRY_383	34,88	62,76	94,1%	96,8%
GRY_388	38,39	97,20	95,5%	98,2%
GRY_38	56,01	110,75	97,4%	98,5%
GRY_68	56,16	104,36	97,5%	98,4%
GRY_332	34,29	93,62	93,8%	98,2%
GRY_365	38,35	98,16	94,8%	98,4%

Fonte: Autoria própria.

9.2 INFORMAÇÕES GERAIS SOBRE AS 212 VARIANTES GERMINATIVAS DE PERDA DE FUNÇÃO EM GENES NÃO GPCS ENCONTRADAS EM UMA DAS ETAPAS DE FILTRAGEM E PRIORIZAÇÃO

ID da amostra	Chr:Pos	Zigosidade	Gene	Alteração na sequência	HGVS c.	HGVS p.	Éxon	REVEL
GRY_02	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND
GRY_02	17:3374973	HMZ	SPATA22	splice_donor	NM_001170698.2:c.-74+2_-74+3delTG	ND	1	ND
GRY_02	16:81194329	HT	PKD1L2	frameshift	NM_001278425.2:c.1600_1604delAGC AG	p.Ser534Leufs*15	11	ND
GRY_02	11:8942889	HT	C11orf16	frameshift	NM_020643.3:c.1377_1378delAT	p.Ile459Metfs*11	6	ND
GRY_02	7:99270539	HT	CYP3A5	splice_acceptor	NM_000777.5:c.219-237G>A	ND	4	ND

GRY_02	7:64169019	HT	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_02	3:167083684	HT	ZBBX	frameshift	NM_001199201.2:c.263delA	p.Asn88Metfs*10	6	ND
GRY_02	2:233712226	HT	GIGYF2	frameshift	NM_001103146.3:c.3629_3632delCAC A	p.Pro1210Argfs*24	27	ND
GRY_02	2:25016051	HT	PTRHD1	frameshift	NM_001013663.2:c.169_196del	p.Ala57Argfs*26	1	ND
GRY_02	1:159923112	HT	SLAMF9	frameshift	NM_033438.4:c.378_379insCTACA	p.Ile127Leufs*16	2	ND
GRY_02	20:35807791	HT	MROH8,RPN 2	frameshift	NM_152503.7:c.93- 1_93ins(29),NM_002951.5:c.13+21_13 +22ins(29)	p.Asn31Lysfs*10,	2,1	ND
GRY_02	19:501702	HT	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_02	18:47311569	HT	ACAA2	frameshift	NM_006111.3:c.1085_1107del	p.Thr362Lysfs*37	9	ND
GRY_02	12:55820903	HT	OR6C76	frameshift	NM_001005183.1:c.866dupT	p.Arg290Lysfs*10	1	ND
GRY_02	11:125818408	HT	VSIG10L2	stop_gained	NM_001365077.1:c.642G>A	p.Trp214Ter	3	ND
GRY_02	10:6067844	HT	IL2RA	frameshift	NM_000417.3:c.209delA	p.Asn70Thrfs*25	2	ND
GRY_02	7:100609662	HT	MUC3A	frameshift	NM_005960.2:c.4216_4217dupGG	p.Val1407Alafs*64	12	ND
GRY_02	6:170871055	HT	TBP	frameshift	NM_003194.5:c.231_237delGCAGCAG	p.Gln77Hisfs*65	3	ND
GRY_02	6:32551955	HMZ	HLA-DRB1	frameshift	NM_002124.4:c.301delC	p.Arg101Glyfs*28	2	ND
GRY_02	3:88040823	HT	HTR1F	frameshift	NM_001322209.2:c.930dupT	p.Val311Cysfs*7	3	ND
GRY_02	2:233712225	HT	GIGYF2	frameshift	NM_001103146.3:c.3629delC	p.Pro1210Hisfs*25	27	ND
GRY_02	1:159923110	HT	SLAMF9	frameshift	NM_033438.4:c.380_381insCT	p.Cys128Tyrfs*14	2	ND
GRY_28	22:37906309	HT	CARD10	frameshift	NM_014550.4:c.819delG	p.Glu273Aspfs*18	4	ND
GRY_28	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND
GRY_28	19:15132683	HT	CCDC105	frameshift	NM_173482.3:c.1204delC	p.Gln402Argfs	6	ND
GRY_28	19:10224842	HT	P2RY11,PPA N-P2RY11	frameshift	NM_002566.5:c.554_561delCCGAGGC C,NM_001040664.3:c.1814_1821delCC GAGGCC	p.Pro185Leufs*19,NP _001035754.1	2,13	ND
GRY_28	11:4566460	HT	OR52M1	frameshift	NM_001004137.1:c.41delT	p.Phe14Serfs*32	1	ND
GRY_28	8:17503646	HT	MTUS1	splice_acceptor	NM_001363057.2:c.3600-2A>G	ND	15	0.661
GRY_28	7:64169019	HMZ	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_28	6:31324892	HT	HLA-B	frameshift	NM_005514.8:c.44delC	p.Ala15Glyfs*5	1	ND
GRY_28	22:37906311	HT	CARD10	frameshift	NM_014550.4:c.813_817delGAAGG	p.Lys272Alafs*9	4	ND
GRY_28	20:18287451	HT	ZNF133	frameshift	NM_001352452.2:c.217+421delA	ND	6	ND
GRY_28	20:5948146	HT	MCM8	frameshift	NM_032485.6:c.940delG	p.Glu314Serfs*16	9	ND

GRY_28	19:53740407	HT	ZNF677	frameshift	NM_182609.4:c.1573dupA	p.Thr525Asnfs*8	5	ND
GRY_28	19:49560921	HT	CGB7	splice_donor	NM_001385261.1:c.-1551delG	ND	1	ND
GRY_28	19:501702	HMZ	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_28	12:100631797	HT	DEPDC4	stop_lost	NM_001387201.1:c.*126T>C	ND	11	ND
GRY_28	12:40830451	HT	MUC19	frameshift	NM_173600.2:c.3129_3132delTCAG	p.Ser1043Argfs*8	27	ND
GRY_28	11:65422265	HT	RELA	splice_acceptor	NM_021975.4:c.1235_1240delTCCTAG	p.Val412_Leu413del	11	ND
GRY_28	7:39379282	HT	POU6F2	frameshift	NM_007252.4:c.552_553insAG	p.Gln185Serfs*118	6	ND
GRY_28	7:39379280	HT	POU6F2	frameshift	NM_007252.4:c.550_551insA	p.Leu184Hisfs*94	6	ND
GRY_28	6:168376926	HT	HGC6.3	frameshift	NM_001129895.3:c.407_408insA	p.Thr137Hisfs*7	1	ND
GRY_28	6:168376883	HMZ	HGC6.3	frameshift	NM_001129895.3:c.450_451insA	p.Gln151Thrfs*7	1	ND
GRY_28	6:168376881	HMZ	HGC6.3	frameshift	NM_001129895.3:c.452_453insC	p.Gln151Hisfs*7	1	ND
GRY_28	6:131156072	HT	SMLR1	frameshift	NM_001195597.2:c.301delT	p.Tyr101Thrfs*4	2	ND
GRY_28	3:128634099	HT	CFAP92	frameshift	NM_001348520.2:c.2269_2270delCT	p.Leu757Valfs*3	14	ND
GRY_28	3:97888323	HT	OR5H15	frameshift	NM_001005515.2:c.782dupG	p.Ala263Cysfs*33	2	ND
GRY_28	3:38163607	HT	DLEC1	frameshift	NM_007337.4:c.5021-171_5021- 168dupAACA	ND	36	ND
GRY_28	3:19938252	HT	EFHB	frameshift	NM_144715.4:c.1651_1652dupGT	p.Arg552Phefs*5	9	ND
GRY_28	2:201760050	HT	NIF3L1	frameshift	NM_001369441.2:c.662_663insA	p.Val222Glyfs*37	4	ND
GRY_28	3:19938251	HT	EFHB	frameshift	NM_001330688.2:c.1651_1652dupGT	p.Arg552fs	13	ND
GRY_38	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND
GRY_38	19:46543528	HT	IGFL4	stop_gained	NM_001002923.3:c.206_217delGGCCC TGCTTCC	p.Trp69Ter	3	ND
GRY_38	17:3374973	HT	SPATA22	splice_donor	NM_001170698.2:c.-74+2_-74+3delITG	ND	1	ND
GRY_38	16:19729016	HT	IQCK,KNOP1	splice_acceptor	NM_153208.2:c.-612- 1G>C,NM_001012991.3:c.-3+464C>G	ND	2,1	ND
GRY_38	7:64169019	HMZ	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_38	7:45123805	HT	NACAD	frameshift	NM_001146334.2:c.1974delG	p.Lys658Asnfs*25	2	ND
GRY_38	4:3076654	HT	HTT	frameshift	NM_002111.8:c.102delG	p.Gln34Hisfs*67	1	ND
GRY_38	22:19766883	HT	TBX1	frameshift	NM_080646.2:c.1149_1150insTG	p.Thr384Ter	9	ND
GRY_38	22:19766879	HT	TBX1	frameshift	NM_080646.2:c.1146_1147delCA	p.Phe382Leufs*13	9	ND
GRY_38	22:19766874	HT	TBX1	frameshift	NM_080646.2:c.1143delC	p.Phe382Serfs	9	ND
GRY_38	20:18287451	HT	ZNF133	frameshift	NM_001352452.2:c.217+421delA	ND	6	ND

GRY_38	19:501702	HMZ	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_38	16:89788941	HT	ZNF276	frameshift	NM_001113525.2:c.210_213delAGGC	p.Arg72Leufs*116	2	ND
GRY_38	16:56913033	HT	SLC12A3	frameshift	NM_000339.3:c.1232delC	p.Pro411Leufs*37	10	ND
GRY_38	12:130839090	HT	PIWIL1	frameshift	NM_004764.5:c.1054delC	p.Gln352Lysfs*6	10	ND
GRY_38	11:125818408	HT	VSIG10L2	stop_gained	NM_001365077.1:c.642G>A	p.Trp214Ter	3	ND
GRY_38	11:64757664	HT	BATF2	splice_donor	NM_138456.4:c.142-380G>A	ND	3	ND
GRY_38	7:45123817	HT	NACAD	frameshift	NM_001146334.2:c.1842_1962del	p.Ser615Glnfs*28	2	ND
GRY_38	2:206641240	HT	NRP2	frameshift	NM_201266.2:c.2440+9699_2440+970 2delGCAC	ND	15	ND
GRY_41	X:135961586	HT	RBMX	initiator_codon	NM_002139.4:c.-1_1delCA	p.Met1Trpfs*34	2	ND
GRY_41	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND
GRY_41	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53dupG	ND	6	ND
GRY_41	17:39305506	HT	KRTAP4-5	frameshift	NM_033188.4:c.514delC	p.Arg172Alafs	1	ND
GRY_41	17:33520792	HT	SLC35G3	frameshift	NM_152462.2:c.535delA	p.Thr179Hisfs*27	1	ND
GRY_41	17:15406369	HT	TVP23C,TVP2 3C-CDRT4	frameshift	NM_145301.3:c.640delT,NM_0012044 78.2:c.462+42730delT	p.Tyr214Thrfs,	6,5	ND
GRY_41	17:4936628	HT	SLC52A1	frameshift	NM_017986.4:c.1062delC	p.Tyr355Thrfs*2	4	ND
GRY_41	17:3336401	HT	OR1E2	frameshift	NM_003554.2:c.726_735delTATCTGCA AG	p.Ile243Profs*33	1	ND
GRY_41	17:263387	HT	C17orf97	frameshift	NM_001013672.5:c.753delT	p.Leu252Serfs*166	2	ND
GRY_41	17:263286	HT	C17orf97	frameshift	NM_001013672.5:c.652delG	p.Asp218Thrfs*200	2	ND
GRY_41	16:19729016	HT	IQCK,KNOP1	splice_acceptor	NM_153208.2:c.-612- 1G>C,NM_001012991.3:c.-3+464C>G	ND	2,1	ND
GRY_41	8:11628978	HT	NEIL2	frameshift	NM_145043.4:c.22delA	p.Arg8Glyfs*51	2	ND
GRY_41	7:99270539	HT	CYP3A5	splice_acceptor	NM_000777.5:c.219-237G>A	ND	4	ND
GRY_41	7:64169019	HMZ	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_41	6:31380160	HT	MICA	frameshift	NM_001177519.3:c.951_952delTG	p.Gly318Leufs	5	ND
GRY_41	6:26368279	HT	BTN3A2	splice_donor	NM_007047.5:c.-6+1G>A	ND	2	ND
GRY_41	2:115219702	HT	DPP10	stop_gained	NM_020868.6:c.60+19287T>G	ND	1	ND
GRY_41	1:201356001	HT	LAD1	frameshift	NM_005558.4:c.485_488delTGGG	p.Val162Alafs*56	3	ND
GRY_41	1:120384192	HT	NBPF7	stop_gained	NM_001047980.2:c.370C>T	p.Arg124Ter	3	ND
GRY_41	19:53762175	HT	VN1R2	frameshift	NM_173856.2:c.548dupT	p.Leu183Phefs*12	1	ND

GRY_41	19:501702	HT	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_41	17:263288	HT	C17orf97	frameshift	NM_001013672.5:c.654_682del	p.Asp218Glufs*43	2	ND
GRY_41	16:15457522	HT	NPIPA5	frameshift	NM_001277325.2:c.989_1047del	p.Thr330Lysfs*7	8	ND
GRY_41	12:40928539	HT	MUC19	stop_gained	NM_173600.2:c.21775A>T	p.Lys7259Ter	123	ND
GRY_41	12:40876263	HT	MUC19	stop_gained	NM_173600.2:c.9509C>A	p.Ser3170Ter	56	ND
GRY_41	12:11214635	HT	PRH1,PRH1- TAS2R14,TAS 2R46	frameshift	NM_001291315.2:c.-133-14848_-133- 14847insGG,NM_001316893.2:c.-133- 14848_-133- 14847insGG,NM_176887.2:c.259_260i nsGG	p.Val87Glyfs*5	2,2,1	ND
GRY_41	11:125818408	HT	VSIG10L2	stop_gained	NM_001365077.1:c.642G>A	p.Trp214Ter	3	ND
GRY_41	9:136253241	HT	STKLD1	frameshift	NM_153710.5:c.306_307delGT	p.Tyr103Profs*9	5	ND
GRY_41	8:94772207	HT	TMEM67	frameshift	NM_153704.6:c.406+1404_406+1407d upAATT	ND	3	ND
GRY_41	8:94772149	HT	TMEM67	frameshift	NM_153704.6:c.406+1345dupT	ND	3	ND
GRY_41	6:126278207	HT	HINT3	frameshift	NM_138571.5:c.84delA	p.Val29Trpfs*43	1	ND
GRY_41	6:33037457	HT	HLA-DPA1	frameshift	NM_033554.3:c.307_308insT	p.Thr103Ilefs*17	2	ND
GRY_41	6:33037456	HT	HLA-DPA1	frameshift	NM_033554.3:c.308_309insG	p.Ile105Aspfs*15	2	ND
GRY_41	6:33037452	HT	HLA-DPA1	frameshift	NM_033554.3:c.311_312delTG	p.Leu104Tyrfs*15	2	ND
GRY_41	6:32551955	HT	HLA-DRB1	frameshift	NM_002124.4:c.301delC	p.Arg101Glyfs*28	2	ND
GRY_41	4:107156436	HT	TBCK	frameshift	NM_001163435.3:c.1439delT	p.Leu480Argfs*53	15	ND
GRY_41	4:88537078	HMZ	DSPP	frameshift	NM_014208.3:c.3264delT	p.Ser1088Argfs	5	ND
GRY_41	3:138244381	HT	CEP70	frameshift	NM_024491.4:c.895delT	p.Tyr299Ilefs*5	11	ND
GRY_41	3:100378552	HT	ADGRG7	frameshift	NM_032787.3:c.1845dupA	p.Pro616Thrfs*40	14	ND
GRY_41	2:110301828	HT	SEPTIN10	frameshift	NM_144710.5:c.*58dupA	ND	11	ND
GRY_41	1:201356001	HT	LAD1	frameshift	NM_005558.4:c.488delG	p.Gly163Alafs*56	3	ND
GRY_68	11:4673934	HT	OR51E1	frameshift	NM_152430.4:c.179_180dupAG	p.Pro61Serfs*99	2	ND
GRY_68	9:34372550	HT	MYORG	frameshift	NM_020702.5:c.373_392del	p.Cys125Alafs*110	2	ND
GRY_68	9:34372548	HT	MYORG	frameshift	NM_020702.5:c.394delC	p.Leu132Trpfs*60	2	ND
GRY_68	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND
GRY_68	11:60264978	HT	MS4A12	frameshift	NM_017716.3:c.187delA	p.Ser63Valfs*10	2	ND
GRY_68	7:64169019	HT	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_68	19:6430131	HT	SLC25A41	frameshift	NM_173637.4:c.405dupG	p.Leu136Alafs*27	3	ND

GRY_68	19:501702	HT	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_68	16:84189328	HT	DNAAF1	frameshift	NM_178452.6:c.715delA	p.Ser239Alafs*12	5	ND
GRY_68	12:57390114	HT	GPR182	frameshift	NM_007264.4:c.1126dupC	p.His376Profs*3	2	ND
GRY_68	11:76895732	HT	MYO7A	frameshift	NM_000260.4:c.3476delG	p.Gly1159Alafs*21	27	ND
GRY_332	17:3374973	HT	SPATA22	splice_donor	NM_001170698.2:c.-74+2_-74+3delTG	ND	1	ND
GRY_332	15:68486361	HT	CALML4	frameshift	NM_033429.3:c.454delG	p.Asp152Thrfs	5	ND
GRY_332	7:64169019	HMZ	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_332	7:45123859	HT	NACAD	frameshift	NM_001146334.2:c.1920delC	p.Val641Leufs*2	2	ND
GRY_332	6:32549580	HT	HLA-DRB1	frameshift	NM_002124.4:c.406delC	p.Gln136Serfs*13	3	ND
GRY_332	5:41064591	HT	MROH2B	frameshift	NM_173489.5:c.436_443delAGGATGA A	p.Arg146Glyfs*8	5	ND
GRY_332	22:19420114	HT	MRPL40	frameshift	NM_003776.4:c.38_42dupTGCGC	p.Pro15Cysfs*33	1	ND
GRY_332	20:2397973	HT	TGM6	frameshift	NM_198994.3:c.1431_1432ins(17)	p.Gly478Argfs*31	10	ND
GRY_332	20:2397971	HT	TGM6	frameshift	NM_198994.3:c.1429_1430insTCTCT	p.Gly477Valfs*28	10	ND
GRY_332	19:49560921	HT	CGB7	splice_donor	NM_001385261.1:c.-1551delG	ND	1	ND
GRY_332	19:36002421	HT	DMKN	frameshift	NM_033317.5:c.810delT	p.Gly271Alafs*106	5	ND
GRY_332	19:17397501	HT	ANKLE1	frameshift	NM_152363.6:c.*141delT		9	ND
GRY_332	19:501702	HT	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_332	19:501702	HT	MADCAM1	frameshift	NM_130760.3:c.700_701insGAGCCTCC CA	p.Asp234Glyfs*64	4	ND
GRY_332	19:501698	HT	MADCAM1	frameshift	NM_130760.3:c.699_700insAACACCAC CTCCCC	p.Asp234Asnfs*77	4	ND
GRY_332	15:44949428	HT	SPG11	frameshift	NM_025137.4:c.733_734delAT	p.Met245Valfs*2	4	ND
GRY_332	14:20482432	HT	OR4K14	frameshift	NM_001004712.2:c.920_921delTG	p.Val307Aspfs	2	ND
GRY_332	12:100631797	HT	DEPDC4	stop_lost	NM_001387201.1:c.*126T>C	ND	11	ND
GRY_332	11:12316389	HT	MICALCL	frameshift	NM_032867.4:c.1411delA	p.Thr471Glnfs*16	3	ND
GRY_332	7:100550860	HT	MUC3A	frameshift	NM_005960.2:c.1432_1433insGT	p.Ser478Cysfs*154	5	ND
GRY_332	7:100550857	HT	MUC3A	frameshift	NM_005960.2:c.1429delA	p.Thr477Profs*154	5	ND
GRY_332	7:100550856	HT	MUC3A	frameshift	NM_005960.2:c.1429_1430delAC	p.Thr477Leufs*66	5	ND
GRY_332	6:32549585	HT	HLA-DRB1	frameshift	NM_002124.4:c.401dupA	p.Thr135Aspfs*23	3	ND
GRY_332	5:96139341	HT	ERAP1	frameshift	NM_016442.5:c.288_289delTA	p.His96Glnfs*8	2	ND
GRY_365	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND

GRY_365	19:17397497	HT	ANKLE1	frameshift	NM_152363.6:c.*136_*140delGTGTT	ND	9	ND
GRY_365	7:64169019	HMZ	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_365	21:36042482	HT	CLIC6	frameshift	NM_053277.3:c.796delG	p.Ala266Argfs*5	1	ND
GRY_365	19:49560921	HT	CGB7	splice_donor	NM_001385261.1:c.-1551delG	ND	1	ND
GRY_365	19:41249997	HT	C19orf54	frameshift	NM_198476.5:c.622dupG	p.Val208Glyfs*17	4	ND
GRY_365	19:17397501	HT	ANKLE1	frameshift	NM_152363.6:c.*141delT	ND	9	ND
GRY_365	17:33879811	HT	SLFN14	frameshift	NM_001129820.2:c.1842dupT	p.His615Serfs*12	5	ND
GRY_365	15:57744359	HT	CGNL1	frameshift	NM_032866.5:c.1936_1937delAG	p.Arg646Aspfs*30	6	ND
GRY_365	15:50926666	HT	TRPM7	frameshift	NM_017672.6:c.920delG	p.Ser307Thrfs*6	8	ND
GRY_365	14:35592463	HT	PRORP	frameshift	NM_014672.4:c.16_19delTTTG	p.Phe6Valfs*16	2	ND
GRY_365	13:79918807	HT	RBM26	frameshift	NM_001366735.2:c.2181dupA	p.Gln728Thrfs*23	15	ND
GRY_365	2:242256951	HT	SEPTIN2	frameshift	NM_004404.5:c.-18+1569_-18+1570delTT	ND	1	ND
GRY_365	2:207527978	HT	DYTN	frameshift	NM_001093730.1:c.1281_1282dupGG	p.Glu428Glyfs*61	11	ND
GRY_365	1:241846884	HT	WDR64	frameshift	NM_001367482.1:c.724dupG	p.Asp242Glyfs*2	7	ND
GRY_365	1:91781486	HT	HFM1	frameshift	NM_001017975.6:c.3025_3026delGT	p.Val1009Tyrfs*6	28	ND
GRY_383	19:54445595	HT	CACNG7	frameshift	NM_031896.5:c.*53delG	ND	6	ND
GRY_383	19:49793867	HT	SLC6A16	frameshift	NM_014037.3:c.1936delA	p.Ser646Alafs*17	11	ND
GRY_383	19:1047508	HT	ABCA7	frameshift	NM_019112.4:c.2126_2132delAGCAGGG	p.Glu709Alafs*86	16	ND
GRY_383	15:90320134	HMZ	MESP2	frameshift	NM_001039958.2:c.546_558delAGGGCAGGGGCAG	p.Gln184Argfs	1	ND
GRY_383	14:20710968	HT	OR11H4	splice_acceptor	NM_001004479.2:c.-11-2A>G	ND	2	ND
GRY_383	11:4945357	HT	MMP26,OR51G1	frameshift	NM_021801.5:c.-144-63939_-144-63938delITG,NM_001005237.1:c.212_213delCA	p.Thr71Argfs*19	3,1	ND
GRY_383	7:99270539	HT	CYP3A5	splice_acceptor	NM_000777.5:c.219-237G>A	ND	4	ND
GRY_383	7:64169019	HMZ	ZNF107	frameshift	NM_001282359.2:c.2544_2545insAA	p.Tyr849Asnfs	4	ND
GRY_383	7:45123859	HT	NACAD	frameshift	NM_001146334.2:c.1920delC	p.Val641Leufs*2	2	ND
GRY_383	6:31324024	HT	HLA-B	frameshift	NM_005514.8:c.539delG	p.Arg181Glufs*33	3	ND
GRY_383	5:140811395	HT	PCDHGA1	frameshift	NM_018912.3:c.2422-62977delC	p.Glu358Lysfs*8	2,1	ND
GRY_383	22:20459343	HT	RIMBP3	frameshift	NM_015672.2:c.1958_1959dupCC	p.Gly654Profs*3	1	ND
GRY_383	21:36042482	HT	CLIC6	frameshift	NM_053277.3:c.796delG	p.Ala266Argfs*5	1	ND
GRY_383	19:49560921	HT	CGB7	splice_donor	NM_001385261.1:c.-1551delG	ND	1	ND

GRY_383	18:28588381	HT	<i>DSC3</i>	<i>frameshift</i>	NM_001941.5:c.1374dupT	p.Thr459Tyrfs*9	10	ND
GRY_383	13:42179396	HT	<i>VWA8</i>	<i>frameshift</i>	NM_015058.2:c.4894delA	p.Met1632Ter	40	ND
GRY_383	8:144942713	HT	<i>EPPK1</i>	<i>frameshift</i>	NM_031308.4:c.4709dupG	p.Asn1571Glnfs*201	2	ND
GRY_383	7:100550857	HT	<i>MUC3A</i>	<i>frameshift</i>	NM_005960.2:c.1429delA	p.Thr477Profs*154	5	ND
GRY_383	7:100550856	HT	<i>MUC3A</i>	<i>frameshift</i>	NM_005960.2:c.1429_1430delAC	p.Thr477Leufs*66	5	ND
GRY_383	6:65301687	HT	<i>EYS</i>	<i>frameshift</i>	NM_001142800.2:c.4073delC	p.Pro1358Glnfs*23	26	ND
GRY_383	5:109961125	HT	<i>TMEM232</i>	<i>frameshift</i>	NM_001039763.4:c.611delT	p.Phe204Serfs*20	7	ND
GRY_383	4:76539580	HT	<i>CDKL2</i>	<i>frameshift</i>	NM_001330724.2:c.222dupA	p.Arg75Thrfs*7	3	ND
GRY_383	3:48422340	HT	<i>FBXW12</i>	<i>frameshift</i>	NM_207102.2:c.926_929delCAGA	p.Thr309Asnfs*7	8	ND
GRY_383	2:206641240	HT	<i>NRP2</i>	<i>frameshift</i>	NM_201266.2:c.2440+9699_2440+9702delGCAC	ND	15	ND
GRY_383	2:179397982	HT	<i>TTN</i>	<i>frameshift</i>	NM_001267550.2:c.103360delG	p.Glu34454Asnfs*3	358	ND
GRY_383	1:159035729	HT	<i>AIM2</i>	<i>frameshift</i>	NM_004833.3:c.787delA	p.Ile263Leufs*2	4	ND
GRY_383	1:89449509	HT	<i>KYAT3,RBMXL1</i>	<i>initiator_codon</i>	NM_001008661.3:c.99+4426dupA,NM_001162536.3:c.1dupA	p.Met1Asnfs*3	2,3	ND
GRY_388	19:54445595	HMZ	<i>CACNG7</i>	<i>frameshift</i>	NM_031896.5:c.*53delG	ND	6	ND
GRY_388	15:90320134	HMZ	<i>MESP2</i>	<i>frameshift</i>	NM_001039958.2:c.546_558delAGGGCAGGGGCAG	p.Gln184Argfs	1	ND
GRY_388	10:122278382	HT	<i>PLPP4</i>	<i>frameshift</i>	NM_001030059.3:c.296delA	p.Asn99Thrfs*5	4	ND
GRY_388	7:45123805	HT	<i>NACAD</i>	<i>frameshift</i>	NM_001146334.2:c.1974delG	p.Lys658Asnfs*25	2	ND
GRY_388	1:28785730	HT	<i>PHACTR4</i>	<i>frameshift</i>	NM_001048183.3:c.160delA	p.Ser54Valfs*12	3	ND
GRY_388	1:26608877	HT	<i>UBXN11</i>	<i>frameshift</i>	NM_183008.2:c.1470_1476delTCCCGGC	p.Gly492Valfs	16	ND
GRY_388	21:36042482	HT	<i>CLIC6</i>	<i>frameshift</i>	NM_053277.3:c.796delG	p.Ala266Argfs*5	1	ND
GRY_388	19:501702	HT	<i>MADCAM1</i>	<i>frameshift</i>	NM_130760.3:c.700_701insGAGCCTCCCA	p.Asp234Glyfs*64	4	ND
GRY_388	17:67149501	HT	<i>ABCA10</i>	<i>frameshift</i>	NM_080282.4:c.4081delG	p.Val1361Cysfs*3	35	ND
GRY_388	17:27899897	HT	<i>TP53I13</i>	<i>frameshift</i>	NM_138349.4:c.1139dupC	p.Pro381Alafs	7	ND
GRY_388	13:95696016	HT	<i>ABCC4</i>	<i>frameshift</i>	NM_005845.5:c.3655delA	p.Ile1219Serfs*11	29	ND
GRY_388	13:21861622	HT	<i>LOC101928764</i>	<i>frameshift</i>	NM_001365454.2:c.489delA	p.Gln164Argfs*7	5	ND
GRY_388	12:40884665	HT	<i>MUC19</i>	<i>frameshift</i>	NM_173600.2:c.17910dupA	p.Pro5971Thrfs*46	57	ND
GRY_388	12:40882388	HT	<i>MUC19</i>	<i>frameshift</i>	NM_173600.2:c.15634delA	p.Ser5212Valfs*7	57	ND
GRY_388	11:125818408	HT	<i>VSIG10L2</i>	<i>stop_gained</i>	NM_001365077.1:c.642G>A	p.Trp214Ter	3	ND

GRY_388	9:107266876	HT	<i>OR13F1</i>	<i>frameshift</i>	NM_001004485.1:c.338_339dupTG	p.Leu114Cysfs*5	1	ND
GRY_388	8:12286231	HT	<i>FAM86B2</i>	<i>frameshift</i>	NM_001137610.3:c.650_653delACAG	p.Asp217Alafs*4	6	ND
GRY_388	8:6668994	HT	<i>XKR5</i>	<i>frameshift</i>	NM_207411.5:c.1786delG	p.Ala596Profs*7	8	ND
GRY_388	7:45123817	HT	<i>NACAD</i>	<i>frameshift</i>	NM_001146334.2:c.1842_1962del	p.Ser615Glnfs*28	2	ND
GRY_388	5:133861426	HT	<i>JADE2</i>	<i>splice_donor</i>	NM_001388185.1:c.-269dupG	ND	1	ND
GRY_388	4:88537078	HMZ	<i>DSPP</i>	<i>frameshift</i>	NM_014208.3:c.3264delT	p.Ser1088Argfs	5	ND
GRY_388	4:9785081	HT	<i>DRD5</i>	<i>frameshift</i>	NM_000798.5:c.1429delC	p.His477Ilefs	1	ND
GRY_388	3:33055705	HT	<i>GLB1</i>	<i>frameshift</i>	NM_000404.4:c.1577dupG	p.Trp527Leufs*5	15	ND

Fonte: Autoria própria.

9.3 INFORMAÇÕES GERAIS DAS 76 VARIANTES (75 GENES) ENCONTRADAS NO GRY_383 E SELECIONADAS NO VARSEQ.

ID da amostra	Chr:Pos	Gene	Efeito	HGVS c.	p.	Éxon	REVEL
GRY_383	1:14925585	<i>KAZN</i>	LoF	NM_201628.3:c.92delC	p.Thr31Argfs*10	1	ND
	1:247615262	<i>OR2B11</i>	LoF	NM_001004492.1:c.23delT	p.Phe8Serfs*2	1	ND
	2:228476140	<i>C2orf83</i>	LoF	NM_020161.4:c.423G>A	p.Trp141Ter	3	ND
	3:49210205	<i>KLHDC8B</i>	LoF	NM_173546.3:c.3G>T	p.Met1Ile	2	0,628
	3:192980889	<i>PLAAT1</i>	LoF	NM_020386.5:c.275dupC	p.Leu93Serfs*21	3	ND
	3:193081962	<i>ATP13A5</i>	LoF	NM_198505.4:c.171delG	p.Trp57Ter	2	ND
	3:194061907	<i>CPN2</i>	LoF	NM_001080513.4:c.1525C>T	p.Gln509Ter	2	ND
	4:68939691	<i>TMPRSS11F</i>	LoF	NM_207407.2:c.319C>T	p.Arg107Ter	4	ND
	5:128301891	<i>SLC27A6</i>	LoF	NM_001017372.3:c.61_65delAAACT	p.Lys21Profs*8	1	ND
	5:150947028	<i>FAT2</i>	LoF	NM_001447.3:c.1465dupG	p.Glu489Glyfs*20	2	ND
	7:72727111	<i>TRIM50</i>	LoF	NM_178125.3:c.1270G>T	p.Glu424Ter	7	ND
	9:70918647	<i>FOXD4L3</i>	LoF	NM_199135.4:c.780C>A	p.Tyr260Ter	1	ND
	9:98231381	<i>PTCH1</i>	LoF	NM_000264.5:c.1899_1902delACAC	p.His634Thrfs*58	14	ND
	9:140110408	<i>NDOR1</i>	LoF	NM_014434.4:c.1493_1505delAGTGGCAGGAGCT	p.Glu498Glyfs*7	12	ND
	10:135084006	<i>ADAM8</i>	LoF	NM_001109.5:c.1642_1643delAT	p.Met548Valfs*35	16	ND

12:11214145	<i>PRH1</i>	LoF	NM_001291315.2:c.-133-14358G>A,NM_001316893.2:c.-133-14358G>A,NM_176887.2:c.749G>A	p.Trp250Ter	2,2,1	ND
13:46288017	<i>CBY2</i>	LoF	NM_152719.3:c.857C>A	p.Ser286Ter	3	ND
13:52951631	<i>THSD1</i>	LoF	NM_018676.4:c.2470_2474delCAGAG	p.Gln824Aspfs*12	5	ND
14:24470138	<i>DHRS4L2</i>	LoF	NM_198083.4:c.475C>T	p.Arg159Ter	4	ND
14:24470691	<i>DHRS4L2</i>	LoF	NM_198083.4:c.629_630insA	p.His211Alafs	6	ND
16:31448156	<i>ZNF843</i>	LoF	NM_001136509.3:c.15delC	p.Phe6Leufs*3	2	ND
17:39458574	<i>KRTAP29-1</i>	LoF	NM_001257309.1:c.518_530delCAACCATCTGTAC	p.Pro173Glnfs*109	1	ND
19:1877675	<i>ABHD17A</i>	LoF	NM_001130111.2:c.529_539delTACGGCATCAG	p.Tyr177Profs*56	4	ND
19:2255311	<i>JSRP1</i>	LoF	NM_144616.4:c.3G>C	p.Met1Ile	2	0,091
19:6897464	<i>ADGRE1</i>	LoF	NM_001974.5:c.420C>G	p.Ser140Arg	5	0,103
19:9236699	<i>OR7G3</i>	LoF	NM_001001958.1:c.924_928dupACCAT	p.Ser310Tyrfs	1	ND
19:52220321	<i>HAS1</i>	LoF	NM_001297436.2:c.825G>A	p.Trp275Ter	3	ND
19:52887146	<i>ZNF880</i>	LoF	NM_001145434.2:c.317delA	p.Asn106Ilefs*5	4	ND
20:2841438	<i>VPS16</i>	LoF	NM_022575.4:c.559C>T	p.Arg187Ter	6	ND
20:48429462	<i>SLC9A8</i>	LoF	NM_015266.3:c.3G>T	p.Met1Ile	1	0,317
21:31744127	<i>KRTAP13-2</i>	LoF	NM_181621.4:c.405T>A	p.Cys135Ter	1	ND
21:46011583	<i>KRTAP10-6</i>	LoF	NM_198688.3:c.783C>A,NM_144991.3:c.83-23694C>A	p.Cys261Ter,	1,2	ND
1:6524501	<i>TNFRSF25</i>	Missense	NM_003790.3:c.476A>G	p.Asp159Gly	5	0,504
1:21928226	<i>RAP1GAP</i>	Missense	NM_002885.4:c.1603G>A	p.Glu535Lys	20	0,533
2:196728886	<i>DNAH7</i>	Missense	NM_018897.3:c.7493T>C	p.Ile2498Thr	41	0,57
2:227915847	<i>COL4A4</i>	Missense	NM_000092.5:c.2996G>A	p.Gly999Glu	33	0,773
3:62355874	<i>FEZF2</i>	Missense	NM_018008.4:c.1264T>G	p.Cys422Gly	5	0,938
3:128622922	<i>ACAD9</i>	Missense	NM_014049.5:c.976G>A	p.Ala326Thr	10	0,874
4:673778	<i>MYL5</i>	Missense	NM_002477.1:c.263T>C	p.Phe88Ser	4	0,713
4:8602934	<i>CPZ</i>	Missense	NM_003652.4:c.173C>T	p.Ser58Leu	2	0,569
5:7826962	<i>ADCY2</i>	Missense	NM_020546.3:c.3254C>T	p.Ser1085Phe	25	0,506
5:140573461	<i>PCDHB10</i>	Missense	NM_018930.4:c.1336G>A	p.Asp446Asn	1	0,694
5:140626737	<i>PCDHB15</i>	Missense	NM_018935.4:c.1591G>A	p.Val531Met	1	0,6
5:168199809	<i>SLIT3</i>	Missense	NM_003062.4:c.1436A>C	p.Lys479Thr	14	0,724
6:31737819	<i>VWA7</i>	Missense	NM_025258.3:c.1159G>C	p.Gly387Arg	8	0,859
6:51923291	<i>PKHD1</i>	Missense	NM_138694.4:c.1342G>C	p.Gly448Arg	16	0,514
6:117130544	<i>GPRC6A</i>	Missense	NM_148963.4:c.431T>G	p.Ile144Arg	2	0,694
6:155561796	<i>TIAM2</i>	Missense	NM_012454.4:c.3301C>T	p.Arg1101Cys	16	0,591

8:57026229	<i>MOS</i>	Missense	NM_005372.1:c.313G>T	p.Ala105Ser	1	0,531
8:105361730	<i>DCSTAMP</i>	Missense	NM_030788.4:c.950T>G	p.Leu317Arg	2	0,511
8:144945191	<i>EPPK1</i>	Missense	NM_031308.4:c.2231A>G	p.Tyr744Cys	2	0,611
10:19569100	<i>MALRD1</i>	Missense	NM_001142308.3:c.3204G>T	p.Gln1068His	20	0,505
10:70949024	<i>SUPV3L1</i>	Missense	NM_003171.5:c.574T>C	p.Tyr192His	5	0,55
11:71819875	<i>LRTOMT</i>	Missense	NM_001330321.2:c.319-1224A>C,NM_001145308.5:c.780T>G	p.Cys260Trp	6,7	0,554
12:52981442	<i>KRT72</i>	Missense	NM_080747.3:c.1283G>T	p.Arg428Leu	7	0,619
12:70091452	<i>BEST3</i>	Missense	NM_032735.3:c.127T>C	p.Tyr43His	2	0,917
12:81101790	<i>MYF6</i>	Missense	NM_002469.3:c.292G>A	p.Ala98Thr	1	0,872
13:99083321	<i>FARP1</i>	Missense	NM_005766.4:c.1930C>T	p.His644Tyr	18	0,535
14:20876282	<i>TEP1</i>	Missense	NM_007110.5:c.317C>T	p.Ser106Phe	2	0,535
14:24976574	<i>CMA1</i>	Missense	NM_001836.5:c.197A>G	p.His66Arg	2	0,627
14:94935901	<i>SERPINA9</i>	Missense	NM_175739.4:c.277C>G	p.Gln93Glu	2	0,536
15:40707168	<i>IVD</i>	Missense	NM_002225.5:c.865G>C	p.Gly289Arg	8	0,888
15:43010290	<i>STARD9</i>	Missense	NM_020759.3:c.13675G>C	p.Val4559Leu	30	0,53
16:30992394	<i>SETD1A</i>	Missense	NM_014712.3:c.4705A>C	p.Lys1569Gln	17	0,584
16:84256619	<i>KCNG4</i>	Missense	NM_172347.3:c.764G>A	p.Cys255Tyr	3	0,574
17:4348389	<i>SPNS3</i>	Missense	NM_182538.5:c.328C>T	p.Arg110Cys	3	0,607
17:7578479	<i>TP53</i>	Missense	NM_000546.6:c.451C>T	p.Pro151Ser	5	0,898
17:10348558	<i>MYH4</i>	Missense	NM_017533.2:c.5291C>T	p.Thr1764Ile	36	0,596
17:17069199	<i>MPRIP</i>	Missense	NM_001364716.4:c.4294C>A	p.Gln1432Lys	16	0,638;0,14
19:1487004	<i>PCSK4</i>	Missense	NM_017573.4:c.916G>A	p.Asp306Asn	8	0,708
19:8168545	<i>FBN3</i>	Missense	NM_032447.5:c.4840G>A	p.Gly1614Ser	39	0,552
19:39233142	<i>CAPN12</i>	Missense	NM_144691.4:c.334G>A	p.Ala112Thr	3	0,775
19:54314164	<i>NLRP12</i>	Missense	NM_144687.4:c.749T>A	p.Ile250Asn	3	0,756
20:25259006	<i>PYGB</i>	Missense	NM_002862.4:c.907G>T	p.Ala303Ser	8	0,762
20:39974594	<i>LPIN3</i>	Missense	NM_022896.3:c.151C>T	p.Arg51Cys	2	0,898
22:46614274	<i>PPARA</i>	Missense	NM_005036.6:c.484C>G	p.Leu162Val	6	0,668

Fonte: Autoria própria.

Legenda: ND – não descrito; LoF – *loss of function*.

9.4 ARTIGO DOS ACHADOS DO SEQUENCIAMENTO DE PAINEL: EM PREPARO PARA SUBMISSÃO

1) Prevalence of germline pathogenic variants in cancer predisposing genes in young patients across sarcoma subtypes

Nathália de Angelis de Carvalho¹, Karina Miranda Santiago¹, Joyce Maria Lisboa Maia², Felipe D'Almeida Costa³, Maria Nirvana Da Cruz Formiga⁴, Diogo Cordeiro De Queiroz Soares⁴, Daniele Paixão Pereira⁴, Celso Abdon Lopes de Mello², Cecilia Maria Lima Da Costa⁵, José Claudio Casali da Rocha⁴, Barbara Rivera⁶, Dirce Maria Carraro^{1,7}, Giovana Tardin Torrezan^{1,7*}

1. Clinical and Functional Genomics Group, International Research Center/CIPE, A.C.Camargo Cancer Center, Sao Paulo 01508-010, Brazil;
2. Clinical Oncology Department, A.C.Camargo Cancer Center, Sao Paulo 01509-900, Brazil;
3. Anatomic Pathology Department, A.C.Camargo Cancer Center, Sao Paulo 01509-900, Brazil;
4. Oncogenetics Department, A.C.Camargo Cancer Center, Sao Paulo 01509-900, Brazil;
5. Pediatric Oncology Department, A.C.Camargo Cancer Center, Sao Paulo 01509-900, Brazil;
6. Rare Tumors Laboratory, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, 08908, Spain;
7. National Institute of Science and Technology in Oncogenomics and Therapeutic Innovation, Sao Paulo, Brazil.

*Corresponding author

Giovana Tardin Torrezan, Clinical and Functional Genomics Group, International Research Center/CIPE, A.C.Camargo Cancer Center, Sao Paulo, Brazil.
giovana.torrezan@accamargo.org.br

Abstract

Sarcomas are a rare and diverse group of cancers occurring mainly in children and young individuals. While some gene-disease associations are known for cancer predisposition syndromes and specific sarcoma subtypes, underlying genetic origin for most cases remains unclear. Our goal was to define the prevalence of germline pathogenic variants (GPVs) in known cancer predisposing genes in children, adolescents, and young adults (AYA) with several types of sarcomas and to compare the clinical and molecular characteristics of patients with and without germline pathogenic variants. We evaluated 177 young patients (<40 years old) diagnosed with soft tissue and bone sarcomas for the presence of GPVs in a hundred cancer predisposing genes. Evaluation of loss of heterozygosity (LOH) in tumor tissue was performed in cases with pathogenic variants. GPVs were detected in 38/177 patients (21.5%), with a positive rate of 15.8% in children

and 21.6% in AYA. These variants were found in genes previously associated with the risk of developing sarcomas (*TP53*, *RBI*, *NF1*, *CHEK2*, *EXT1* and *EXT2*), but also in genes where that risk is still unknown (*ERCC2/3*, *TSC2*, *RAD50*, *FANCM*, and others) or is emerging (*PALB2*, *BRCA2*). GPVs carriers were more likely to present more than one primary tumor than non-carriers (21.1% X 6.5%; $p=0.012$). Of 25 tumors evaluated for LOH, loss of the wildtype allele was detected in nine. Our findings reveal that 1 in each 5 patients with sarcomas under 40 years old presented a GPV, even in patients without family history of cancer. GPVs rates varied from 0% to 33% across sarcoma subtypes and exhibited specific gene-subtypes associations. This pinpoints the urgency of implementing appropriated genetic screening strategies for these individuals and their families.

Introduction

The group of rare tumors of mesenchymal origin called sarcoma mainly affects the younger population, comprising 20% of tumors in children, 10% in adolescents and young adults (1) and only 1% of malignant tumors in adults (2). Sarcomas comprehend more than 100 histological subtypes with a marked heterogeneity (3) and usually an uncertain etiology. Regarding risk factors, it is known that prolonged exposure to ionizing radiation increases the risk of sarcomas, specially bone tumors (4,5). Among inherited factors, individuals with specific Hereditary Cancer Predisposition Syndromes (HCPS) harboring germline pathogenic variants (GPVs) in tumor suppressor genes seem to have an increased risk to develop sarcomas when exposed to radiation (6–8). Furthermore, some specific HCPS are known to be associated per se with a higher risk in developing sarcomas, such as Li-Fraumeni syndrome, Hereditary Retinoblastoma and Neurofibromatosis Type I (9).

The recent expansion of genetic screening and genomic characterization of tumors have provided evidence of a high frequency of GPVs in sarcoma patients, ranging from 13.6 to 28%, even in unselected cohorts and in the absence of family history of cancer (5,10). Also, recent studies focused on reviewing tumor syndromes or searching for germline variants related to specific histology of sarcomas, such as osteosarcomas and rhabdomyosarcomas, have pinpointed the relevance of GPVs in cancer predisposing genes in sarcoma etiology (11–14). Nevertheless, given the rarity of both sarcomas and

most HCPS, it is challenging to establish gene-disease associations and undescribed associations may exist, especially in populations underrepresented in genomic studies.

Within this context, we aimed to describe the prevalence of GPVs associated with sarcoma development in young Brazilian patients. We identified that 1 in each 5 patients with sarcomas under 40 years old presented a GPVs, with variable GPVs rates across sarcoma subtypes and specific gene-subtypes associations. Moreover, we performed molecular analysis of tumors and analysis with a matched control population to build up more evidence of gene-disease associations.

Methods

Patients' cohort

In this study, 177 patients with sarcomas from A.C.Camargo Cancer Center (ACCCC) were included considering the following criteria: 1) patients diagnosed with sarcomas under 40 years old; 2) and patients without known genetic syndromes associated with dysmorphisms and malformations. From 2018 to 2021, 36 patients were included prospectively by the Oncogenetics department, the Pediatric Oncology department, or the Clinical Oncology department. Additionally, we include 141 retrospective patients from blood stored in the institutional biobank, prioritizing patients with osteosarcomas and chondrosarcomas. All study participants provided written informed consent, either specific for the study or for Biobank collection. This study was performed in accordance with the Helsinki Declaration and was approved by the ACCCC Institutional Review Board (CAAE 90036518.6.0000.5432). Genetic testing results were returned to the patients in a post-test counseling.

Germline DNA extraction

Genomic DNA was extracted from the peripheral blood leukocyte at the ACCCC Biobank using the QiaSymphony equipment (Qiagen, NRW, Germany). For the prospective cases, the DNA was obtained from saliva following the purification protocol of the PrepIT-L2P kit (PT-L2P) (DNA Genotek Inc., ON, Canada). Genomic DNA quality and quantity were evaluated using the NanoDrop® ND-1000 spectrophotometer (ThermoFisher Scientific, MA, United States) and Qubit 4 Fluorometer (ThermoFisher Scientific), following the manufacturer's protocol.

Germline sequencing analyses

TP53 and XAF1 specific variants. Germline DNA samples were evaluated using either Sanger sequencing or NGS amplicon sequencing to identify the *TP53*:p.Arg337His variant, as previously described (15,16). For *TP53*:p.Arg337His positive patients, we evaluated the *XAF1*:p.Glu134Ter variant with the same NGS amplicon approach.

Next-Generation Sequencing of multigene panel. 110 patients negative for *TP53*:p.Arg337His were evaluated by NGS sequencing using a customized panel of 113 or 126 cancer predisposing genes (Supplementary table 1). Genes were selected based on previous reports from commercial panels of clinical laboratories (17) and through the panelapp database (18). Briefly, library construction and gene capture were performed with the QIAseq FX DNA Library Kit (Qiagen) or the Lotus DNA Library Prep kit (IDT Technologies) and custom biotinylated probes (xGen Gene Capture Pools - IDT Technologies). Sequencing was performed in the NextSeq 500 platform (Illumina), aiming to obtain a minimum coverage of 30X in 99-100% of the target bases.

Pooled samples screening with multigene panel. For screening GPVs in a larger number of individuals, for 67 patients we applied a two-dimensional DNA pooling strategy similar to the method previously described by Zuzarte et al (19). In this method, the DNAs from samples were individually distributed in 96-well plates and equimolar pooling of the DNAs was carried out in two dimensions, considering the rows and columns of a 96-well plate, generating 8 pools from lines and 12 pools from columns for each plate. The pooled samples were subsequently sequenced using the 113/126 gene panels, as described for the individual samples. This two-dimension strategy allows to identify exactly the individual carrying rare variants, since when a variant is reported in exactly one row pool and one column pool, only one sample can be represented in that condition. A variant was considered to be present when the variant allele was detected in >0.5% of the reads, considering a minimum coverage depth of 100X, in at least one column pool and one row pool.

Variant selection and interpretation. NGS reads were processed using the Isaac Enrichment v2.1 tool (Illumina) for alignment and variant calling. Variants were annotated and filtered using the VarSeq software (Golden Helix). Specific filters were applied to select rare germline variants (Quality ≥ 30 , base coverage $\geq 30X$, variant allele frequency ≥ 0.25 , minor allele frequency [MAF] ≤ 0.01 or absent in population databases [GnomAD, ESP, 1000genomes and ABraOM], coding variants or intronic variants up to

10 base pairs of exon limits). Filtered variants were evaluated for their classification in clinical databases and automated classification tools (ClinVar, VarSome and Franklin Genoox) and were classified according to the American College of Medical Genetics and Genomics (ACMG) recommendations (20). All identified GPVs were confirmed using amplicon sequencing.

Loss of heterozygosity analyses (LOH)

For loss of heterozygosity analysis (LOH), tumor tissues (frozen or formalin fixed paraffin embedded) with a minimum of 40% of tumor cells were selected. Genomic DNA was extracted at the ACCCC Biobank, according to their standard procedures. PCR and amplicon library construction were performed following the Ion Plus Fragment Library kit (Thermo Fisher Scientific), and sequencing was performed in Ion Proton or S5 platforms (Thermo Fisher Scientific). Allele frequency of the variants was visually verified with IGV software (21) and there were considered allelic fractions of $\geq 65\%$ (loss of wildtype allele) (22).

Transcript analysis

We performed gene specific transcript analysis for one case with an intronic likely pathogenic variant in *EXT1* gene. Transcript analysis was performed as previously described by our group (23). Briefly, RNA was extracted from leucocytes for the patient and control samples at the ACCCC Biobank. cDNA was amplified with primers located two exons upstream and two exons downstream the variant (exons 6 and 10). Amplicons were sequenced in Ion Proton (Thermo Fisher Scientific), reads were mapped using CLC Genomics workbench (Qiagen) and STAR, and sashimi plots was generated in IGV software.

Controls cohort

We compared the frequency of GPVs identified in our patients' cohort to the population frequency of GPVs in the same genes using a publicly available dataset of Brazilian individuals (ABraOM database). The ABraOM SABE-WGS-1171 dataset contains whole-genome sequencing from 1,171 unrelated individuals obtained from a census-based sample of elderly individuals from São Paulo, Brazil (24). We evaluated the

19 genes in which we have detected GPVs and classified the ABraOM variants following the same criteria described for the patients.

Statistical analysis

Descriptive statistics were used to summarize the results. Comparison of clinical characteristics between patients with or without GPVs were performed using Pearson Chi-square or Fisher exact test for binary variables, and Mann-Whitney for continuous variables. Two-tailed $p < 0.05$ were deemed significant. For patients and control cohort comparison, Fisher's exact test was used.

Results

Germline analysis and clinical features

Over the study period, we included 177 sarcoma patients (36 prospective and 141 retrospective patients). The mean age of the cohort was 26 years (0-40, except for one patient with sarcoma developed at 54 years). Patients were divided in two groups according to the age of onset: 0 to 14 years (Children – C) and 15 to 40 years (Adolescents and Young Adults – AYA), comprising 20.3% and 79.7% of the patients, respectively (Table 1). Almost 10% of the patients had more than one primary tumor, and 40.7% had a positive family history of cancer reported in medical records.

We compared the clinical features between C and AYA groups, finding no statistical difference (Supplementary table 2). Regarding sarcoma histological subtypes, we observed known expected differences between the age groups, with chondrosarcomas, liposarcomas, synovial sarcomas and leiomyosarcomas being more common in AYAs, while rhabdomyosarcoma was more frequent in children (Supplementary figure 1). It is important to note that we enriched the selection of individuals with osteosarcomas and chondrosarcomas from our Biobank, consequently increasing the frequency of both tumors in our cohort, compared to other subtypes.

Table 1. Clinical features of 177 young patients with sarcomas with and without GPVs.

Features	Total Patients N (%)	With GPV N (%)	VUS/WT N (%)	P-value
Total individuals	177 (100%)	38 (21.5%)	139 (78.5%)	
Sex				.32
F	97 (54.8%)	24 (63.2%)	73 (52.5%)	
M	80 (45.2%)	14 (36.8%)	66 (47.5%)	
Mean age (range)	26 (0 – 54*)	27 (6 – 54*)	26 (0 – 41)	.94
Age group				.58
C (0-14 years)	36 (20.3%)	6 (15.8%)	30 (21.6%)	
AYA (15-40* years)	141 (79.7%)	32 (84.2%)	109 (78.4%)	
Other primary tumors				.012
Positive	17 (9.6%)	8 (21.1%)	9 (6.5%)	
Negative	160 (90.3%)	30 (78.9%)	130 (93.5%)	
Family history of cancer				.64
Positive	72 (40.7%)	15 (39.5%)	57 (41.0%)	
Negative	28 (15.8%)	4 (10.5%)	24 (17.3%)	
No information	77 (43.5%)	19 (50.0%)	58 (41.7%)	

Notes: C - Children; AYA – Adolescents and Young Adults.

* a female patient with sarcoma diagnosis (chondrosarcoma) at 54 years old was included. This case was prospectively included considering her family history of cancer, in which her father had a diagnosis of a bone tumor at 25 years old and was not available for germline material collection (deceased).

As for the germline and molecular analyses, figure 1 depicts the general workflow of the study. First, we screened the cohort for the *TP53*:p.Arg337His variant, which is considered as a founder variant frequent in the Brazilian population. Five patients (3% – 5/177) were detected with this GPV. In these patients, we evaluated the *XAF1*:p.Glu134Ter variant, as recently an haplotype harboring the *TP53*:p.Arg337His and *XAF1*:p.Glu134Ter variants was described to increase the risk of sarcomas and other primary tumors in p.Arg337His carriers (25). From the five *TP53*:p.Arg337His patients, four were positive for *XAF1*:p.Glu134Ter, two had personal and familial history of cancer that fulfilled the 2015 Chompret clinical criteria and none fulfilled the classical Li-Fraumeni criteria (table 2).

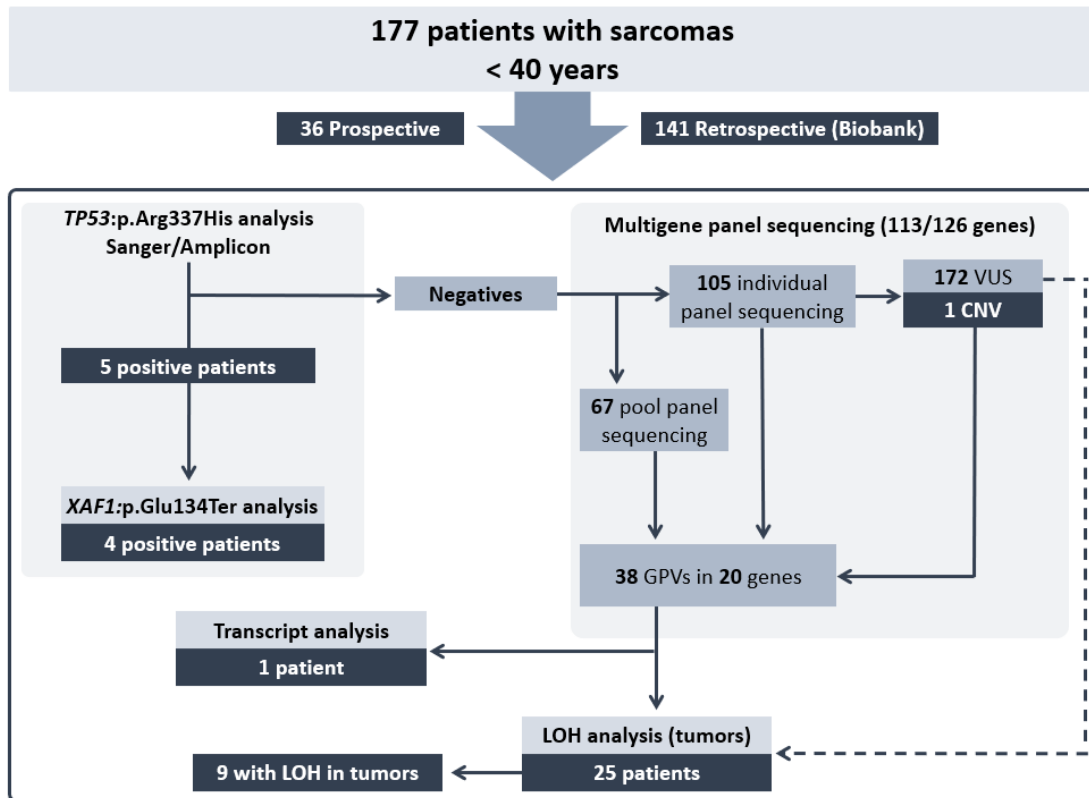


Figure 1. Overview of the study experimental design, methods, and main findings. VUS – variant of uncertain clinical significance; CNV – copy number variants; GPV – germline pathogenic variants; LOH - loss of heterozygosity.

Table 2. Clinical and molecular features of five patients harboring *TP53*:p.Arg337His.

Sample ID	Age at diagnosis, sex	Sarcoma histology	Other tumors	Family history of cancer (Relative, Age)	<i>XAF1</i> :p.Glu134Ter	Clinical criteria
GRY_7	30, M	LMS	No	Gastric cancer (Grandfather, 93)	Present	None
GRY_22	29, F	LMS	No	Breast cancer (Paternal aunt, 39); CNS tumor (paternal aunt, 35); gastric cancer (maternal grandfather, NA)	Present	None
GRY_23	33, M	LMS	No	Prostate cancer and sarcoma (Father, 59, 62); laryngeal neoplasia (Maternal grandmother, 72 smoker); glioblastoma (Paternal cousin, 35); Hodgkin lymphoma (Paternal cousin, 22); colorectal cancer (Paternal uncle, 50 years); Leukemia (Paternal uncle, 5); colorectal cancer (Paternal grandfather, 60).	Absent	Chompret (2015)
GRY_327	33, F	LPS	No	Breast cancer and thyroid carcinoma (Mother, 44, 46); lung cancer (Maternal grandmother, 70); breast cancer (Maternal great-aunt, 63); colorectal cancer (Maternal great-aunt, 50); uterine cancer (Maternal great-aunt, 50 years); CNS tumor (Maternal cousin, 5).	Present	Chompret (2015)
GRY_369	23, M	OS	No	Lung cancer (Maternal aunt, NA).	Present	None

Notes: LMS – leiomyosarcoma; LPS – liposarcoma; OS – osteosarcoma; CNS - central nervous system; NA – not available.

The remaining 172 patients underwent our NGS panel analysis (105 in an individual panel sequencing and 67 in a pool panel sequencing) and 33 (18.6%) patients presented germline pathogenic or likely pathogenic (P/LP) variants. Overall, considering the five *TP53*:p.Arg337His carriers, 21.5% (38/177) patients had GPVs (Supplementary table 3). We compared clinical data from GPVs carriers and non-carriers and identified that GPVs carriers were more likely to present more than one primary tumor ($p=0.012$) (Table 1). The most frequently mutated genes were *TP53* (3.9%), *CHEK2* (2.3%), *BRCA2* (2.3%), and *NF1*, *RBI*, *MUTYH*, *SLX4* (1.7% each) (Figure 2). Besides the *TP53*-p.Arg337His variant, other 3 GPVs were detected in more than one patient: 3 patients had *MUTYH*-p.Gly396Asp; 2 patients had *RBI*-p.Arg251Ter and 2 patients had *MITF*-p.Glu318Lys.

We classified the evaluated cancer predisposing genes according to their evidence of being related to sarcoma predisposition, as described at The Gene Curation Coalition(26) and literature review (Supplementary table 1). We identified variants in genes previously described to increase the risk for sarcomas' development, such as *TP53*, *CHEK2*, *NF1*, *RBI*, *EXT1* and *EXT2*, but also in genes in which this association is emerging, as *BRCA2*, *ERCC2* and *ERCC4* or is unknown, as *MUTYH*, *SLX4*, *MITF*, *AKT1*, *ERCC3*, *FANCM*, *RAD50*, *PALB2* and *TSC2* (Figure 2). Considering only the definitive and emerging genes, 14.7% (26/177) of patients had GPVs. Most genes with unknown relation to sarcomas represent genes associated with recessive cancer predisposing syndromes. Of all patients, three presented P/LP alterations in more than one CPG: one patient with liposarcoma had GPV in *PALB2* and *MITF*, one patient with alveolar sarcoma had alterations in *MUTYH* and *MITF* and one patient with osteosarcoma had GPVs in *TP53* and *SLX4* (Figure 2 and supplementary table 3).

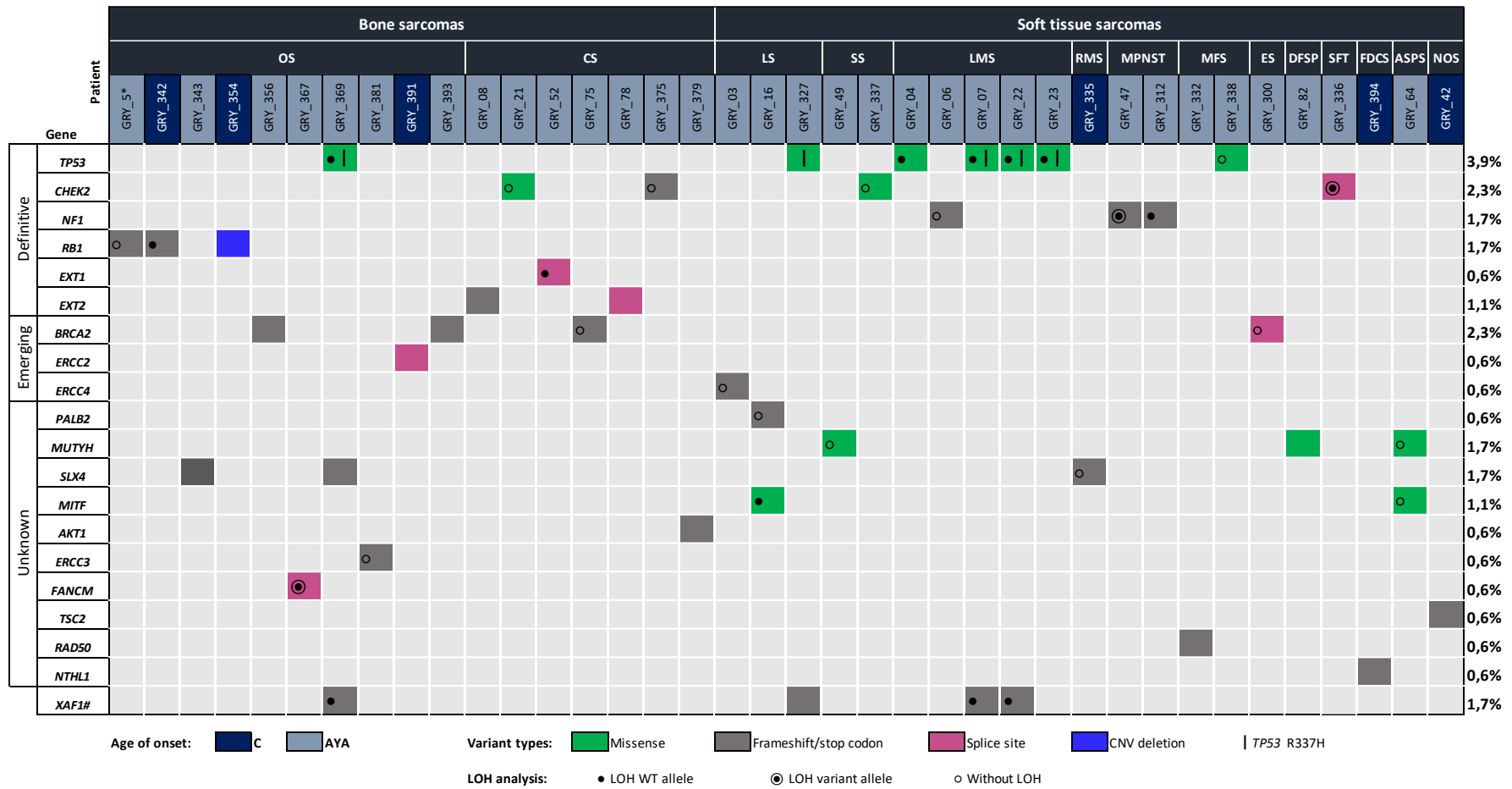
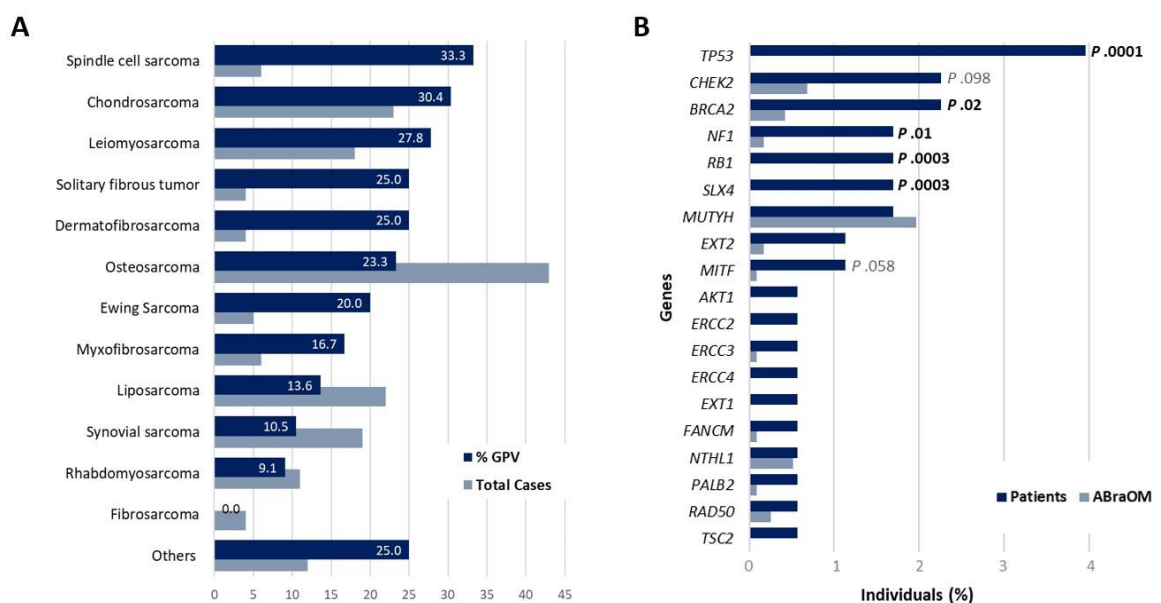


Figure 2. Overview of GPVs detected in 177 sarcoma patients. Sarcomas were divided by histological subtype, between bone tumors and soft tissue sarcomas. On the left, the genes included in the customized panels are represented and divided according to their association to sarcoma predisposition. # *XAF1* is not included in our panel, but alterations in this gene were investigated in patients with *TP53*:p.Arg337His. On the right, the percentages represent the number of patients detected with GPV in the indicated gene and, at the bottom, the types of alterations are described. The results of tumor LOH analyses are also depicted. Patients are colored according to age group. Notes: *patient diagnosed with osteosarcoma, leiomyosarcoma and liposarcoma. OS - osteosarcoma; LS - liposarcoma; CS - chondrosarcoma; SS - synovial sarcoma; LMS - leiomyosarcoma; MPNST - malignant peripheral nerve sheath tumors; MFS - myxofibrosarcoma; ES - Ewing sarcoma; DFSP - dermatofibrosarcoma protuberans; SFT - solitary fibrous tumor; FDCS - follicular dendritic cell sarcoma; ASPS - alveolar soft part sarcoma; sarcoma NOS - not otherwise specified.

Next, we compared the GPV detection rate across sarcoma subtypes and revealed that certain subtypes had higher frequencies of GPVs, as spindle cell sarcoma (33.3%), chondrosarcoma (30.4%) and leiomyosarcoma (27.8%) (Figure 3A). The subtypes with a lower detection of GPVs were fibrosarcoma (0%), rhabdomyosarcoma (9.1%) and synovial sarcoma (10.5%). Likewise, we also observed some expected gene-subtype associations. For instance, we have detected *EXT1* and *EXT2* GPVs only in chondrosarcomas and *RB1* GPVs just in patients with osteosarcomas (1 patient was diagnosed with osteosarcoma, leiomyosarcoma and liposarcoma). Also, there was an enrichment of *TP53* alterations in leiomyosarcomas and of *NF1* in MPNST.

Regarding variants of uncertain clinical significance (VUS), 81.9% patients (86 of 105) were detected with those variants on individual panel sequencing. Of 170 unique variants, 21 were detected in autosomal dominant genes and were considered VUS of interested due to *in silico* prediction of pathogenicity (Revel >0.6) or to Franklin Genoox classification as P/LP (Supplementary Table 4). We selected **XX**¹ prioritized VUS to perform LOH analysis in the tumors attempting for a correlation between the tumor and the alteration.



¹ Análise em andamento.

Figure 3. A: Rates of germline pathogenic variants according to sarcomas subtypes. **B.** Frequency of germline pathogenic variants in genes found in this cohort (177 sarcoma patients) compared with a cohort of healthy elder individuals (1,171 Brazilian individuals from ABraOM database). Spindle cell sarcomas group includes two malignant peripheral nerve sheath tumors. P-values are shown for Fisher exact test, with significant differences ($p < 0.05$) displayed in bold.

Cases with expected non-cancer phenotypes

Some of the genes where GPVs were identified lead to non-cancer phenotypes that represent diagnostic criteria for the associated syndromes. For example, *NF1* carriers usually present multiple cutaneous neurofibromas, multiple café au lait macules, intertriginous freckling, and other features; *TSC2* patients can present multiple abnormalities of the skin, brain, kidney, heart, and lungs; *EXT1/2* patients are expected present multiple osteochondromas (exostoses), which are benign cartilage-capped bone tumors. We identified 7 patients with GPVs in these genes and investigated their clinical phenotypes (Supplementary table 3).

From the three *NF1* carriers identified here, only one had a prior diagnosis of Neurofibromatosis 1 syndrome (GRY_47), with multiple neurofibromas detected since 11 years old. The other two patients were followed in our institution for at least four years, and no other clinical signs of NF1 was described in medical records (although none were seen by a geneticist during this period, and both of them there was loss of follow-up in 2011). One patient with a *TSC2* GPV had a late diagnosis of tuberous sclerosis (TS) (GRY_42). This patient developed three primary tumors: a sarcoma (not otherwise specified) at age 13; a rectal adenocarcinoma at age 29, and Xp11 translocation renal cell carcinoma at age 38. Also, he had West syndrome in his infancy and colonic polyposis at age 35. The clinical diagnosis of TS was only made at 37 years old, with a dermatological evaluation of facial skin lesions.

Regarding the *EXT1/2* genes, three patients with chondrosarcoma presented GPVs in these genes and none of them was prior diagnosed with Hereditary multiple osteochondromas (HMO), previously called hereditary multiple exostoses. One case

(GRY_78) reported removing one osteochondroma at age 10, and one patient (GRY_52) was diagnosed with multiple osteochondromas in lower limbs in imaging exams requested after the results of the genetic testing showed a likely pathogenic variant in *EXT1*. This patient variant was located at the splice donor site of intron 8 (c.1722+1G>C) and, to confirm the effect of this variant on splicing, we performed sequencing of *EXT1* transcripts from blood. This analysis evidenced the complete skipping of exon 8, leading to an in frame deletion of 30 amino acids (c.1633_1722del; p.Met546_Val575del) from the Glycosyl transferase domain (Supplementary figure 2). Based on this finding and the specific phenotype of the patient, the variant was reclassified as pathogenic.

Loss of Heterozygosity (LOH)

To investigate somatic inactivation by deletion of the wildtype allele of GPVs carriers as a possible second hit, we assessed LOH in fresh frozen or FFPE tumors of 25 patients. We observed evidence of LOH of the wildtype allele in 36% of the cases, six patients with GPVs in genes related to sarcomas, as *TP53* (4/4), *NF1* (1/3), *EXT1* (1/1) and *RBI* (1/2), and in *MITF* (1/1), a gene not associated to sarcomas (Figure 2 and Supplementary table S5). Moreover, we searched for LOH in *XAF1* as well, in three tumors with *TP53-XAF1* haplotype found LOH in all of them. For 3 genes, *NF1*, *CHEK2* and *FANCM*, we observed LOH of the variant allele.

Frequency of GPVs compared with healthy individuals

We compared the frequency of P/LP variants identified in our patient's cohort to the frequency of these alterations in a control Brazilian populational cohort obtained from ABraOM database (Figure 3B and Supplementary table S6). The frequency of GPVs detected in the sarcoma cohort was significantly higher in *TP53*, *RBI* and *NF1* (genes definitive associated to sarcomas), and for two genes without established associations to these tumors, such as *BRCA2* and *SLX4*. Moreover, marginal p-values were observed for *CHEK2* and *MITF*.

Discussion

Our study evidenced that 21.5% of unselected young Brazilian patients with sarcomas harbor GPVs in cancer predisposing genes. We found GPVs in genes already known to be related to sarcomas, but also in genes with an unknown association to sarcomas, such as *PALB2*, *MUTYH*, *SLX4*, *MITF*, *AKT1*, *ERCC3*, *FANCM*, *RAD50* and *TSC2*. Considering only the genes with a well-established or emerging association to sarcoma risk, 14.7% of patients had GPVs. Moreover, an additional 11.3% of the patients had VUSs considered damaging by *in silico* prediction tools in an autosomal-dominant cancer predisposing gene. In our population matched case-control comparison, the enrichment of known sarcoma predisposing genes was confirmed for *TP53*, *NF1* and *RBI*, as well as and for *BRCA2*, a gene with emerging evidence for sarcoma risks (5,27), and *SLX4*.

Our findings corroborate with other recent studies, which have found an average of 20% of GPVs in sarcoma patients (5,14,28). In their seminal study, Ballinger et al. (2016) observed 19% of the 1,162 unselected patients with sarcomas (3-93 years old) harbored GPVs in 72 CPGs and almost 50% harbored pathogenic and putative (*in silico* predicted) pathogenic variants. Moreover, compared to a control population, there was an excess burden of putative GPVs in known sarcoma genes, such as *TP53*, but also in *ATM*, *ATR*, *BRCA2*, and *ERCC2* (5). In another multigene panel study evaluating 52 CPGs, 13.6% of 66 patients with sporadic sarcomas under 50 years old carried a GPV, also in genes related and not known to be related to sarcomas (*ATM*, *BRCA2*, *ERCC4*, *FANCC*, *FANCE*, *FANCI*, *MSH6*, *POLE*, *SDHA* and *TP53*) (28). Mirabello et al. (2020) performed exome sequencing of 1,244 individuals (2-80 years old) with osteosarcomas, finding 28% of them with GPVs. In this study, they also performed a variant burden comparison with a control cohort and could identify novel candidate genes, such as *CDKN2A*, *MEN1*, *VHL*, *POT1*, *APC* and *ATRX* (14). Lastly, in a very recent case-control study, Gillani et al (2022) evaluated 141 CPGs in whole genome and exome germline data from 1,147 pediatric

sarcoma patients (enriched for Ewing sarcoma, osteosarcoma, and rhabdomyosarcoma) and more than 10,000 controls. By testing both pan-sarcoma and subtype specific associations, they confirmed the role of genes previously implicated in sarcoma pathogenesis (*TP53*, *DICER1*, *RBI*, *RECQL4*, *NF1*) and also found evidence of association for genes without substantial prior evidence of sarcoma risk (*FANCC*, *FANCA*, *PTPN11*, *RECQL*, *MUTYH*, *BRCA2*, *SDHD*) (27). In addition, other case reports also described CPGs not previously related to sarcomas and other tumors, such as *POT1* (29) and *BAP1* (30).

Although our study is underpowered to confirm differences in genetic predisposition between sarcoma subtypes, we observed qualitative gene-subtypes associations and a range of GPVs detection rates, varying from 0% to 33% across sarcoma subtypes. For instance, well-known gene-subtypes associations were observed in our cohort, as all three *RBI* carriers presented osteosarcomas (one with metachronous leiomyosarcoma and liposarcoma), all 3 *EXT1/2* carriers developed chondrosarcomas and 2 of 3 patients with *NF1* GPV were diagnosed with MPNST. Additionally, *TP53*:p.(Arg337His) were more detected with leiomyosarcomas (60% - 3/5), confirming previous descriptions in the Brazilian population (31) and the distinct behavior of this variant compared to other *TP53* variants, since leiomyosarcoma corresponds to 9.1% of the sarcomas in other *TP53* GPVs carriers (32).

The *TP53*:p.Arg337His variant is a known Li-Fraumeni causing variant associated with a varied risk and penetrance spectrum for different tumors, usually differing its phenotype from the classic LFS (33). It is considered a founder variant in Brazil, as its frequency in this population is higher (MAF = 0.001) (33) compared with worldwide population (MAF gnomAD_v2 1.1 = 0.000012) (34,35). A previous Brazilian study investigated the frequency of *TP53*:p.Arg337His in 502 unselected patients with sarcomas (1-91 years old, mean age 40 years) and observed 8% detection rate, with most carriers being affected by leiomyosarcomas (52.5%) and with ages of onset above 40 years old (82.5%) (36). In our study, the detection rate of this variant was 3% (5/177),

likely due to the earlier ages of onset from our cohort. Moreover, the phenotypes of the probands and their relatives are representative of the variable penetrance and tumor spectrum of this variant, since only three two out of five probands fulfilled the 2015 Chompret criteria and none fulfilled the classical Li-Fraumeni criteria. Recently, a modifier variant, *XAF1*:p.Glu134Ter, occurring as an haplotype with the *TP53*:p.Arg337His was described to increase the risk for sarcoma and subsequent malignancies (25). In that study, the p.Glu134Ter variant was detected in 79% of 203 patients harboring *TP53*:p.Arg337His (25), and in our study, 80% (4/5) of the patients were detected to harbor this compound mutant haplotype.

We identified that 1.7% of our cohort had two GPVs in distinct genes. This phenomenon was recently described in literature as Multilocus Inherited Neoplasia Alleles Syndrome (MINAS) (37,38). There are few studies describing MINAS in sarcoma patients, as revised by McGuigan et al (2022), including cases with GPVs in two autosomal dominant and highly penetrant genes, such as *TP53* with *PTEN* and *NF1* with *BRCA2* (38). In our cases, only one of the three individuals detected with MINAS had one GPV in a highly dominant penetrant gene associated to sarcoma (*TP53*) combined with a monoallelic GPV in a gene of unknown sarcoma risk (and of recessive inheritance – *SLX4*). The other two patients harbored both GPVs in genes not previously related to these tumors (*PALB2* with *MITF* and *MUTYH* with *MITF*). Striking, two MINAS individuals had a GPV in *MITF*, a transcription factor related to cell cycle regulation that plays an important role in melanocytes' homeostasis, and both patients had the same variant (p.Glu318Lys - MAF ABraOM = 0.000427). This rare, moderated penetrance variant is known to be associated to an increased risk of renal cell carcinoma and melanoma (39,40), and was also recently associated with carcinosarcoma risk (41). Moreover, a study from a Brazilian cohort from a Hereditary Cancer Registry described a prevalence of 0.9% for this variant (10/1056 patients), with breast cancer being the most common cancer in probands and their relatives and no description of sarcomas (42).

Besides the occurrence of monogenic high-risk variants leading to sarcoma susceptibility, recent studies have proposed a combined effect of multiple less penetrant variants contributing to sarcoma risks through a polygenic risk model (5). Moreover, polygenic effects and monoallelic germline variants in DNA damage response (DDR) genes were suggested to predispose for young-onset translocation-associated sarcoma subtypes (5), as was later shown to be the case for *FANCC* variants in Ewing sarcomas (27). Our patients with multiple GPVs and patients with monoallelic GPVs in DDR genes usually related with recessive cancer predisposing syndromes (such as *FANCA* and *ERCC* family genes) could represent examples of these polygenic effects. Confirming these smaller effects associations will require large cohorts of patients and controls and proposing specific clinical management for such cases would be even more challenging.

There are several molecular strategies implemented to understand the role of distinct GPVs and genes in the carcinogenesis of a specific tumor. Somatic LOH analysis consider the two-hit model that, in general, occurs in tumor suppressor genes (TSG) inactivation, where the wild-type allele can be lost in tumor tissue due to large deletions. For sarcomas, it is expected to note LOH in TSG where its relation is well-defined, such as *NF1*, *RB1* and *TP53*, what we could observe for several cases in our study. Moreover, some genes can present more complex mechanisms of tumor initiation, as it was described for *EXT1/2* genes in chondrosarcomas, in which a mosaic loss of these proteins is necessary to drive tumorigenesis (43,44). From our three *EXT1/2* cases, LOH was only evaluated in one case, which was positive for wild-type allele loss. Moreover, regarding genes with unknown association to sarcomas, a liposarcoma patient with *MITF* and *PALB2* GPVs had LOH of the *MITF* variant. This case illustrates that LOH data should be interpreted with caveats, especially in genomic unstable tumors like some sarcomas, since *MITF* is considered to influence oncogenesis through a gain of function activity and a second hit is not expected.

We acknowledge that our study has limitations. First, the methodology we used was panel sequencing and, with that, causing variants in non-coding regions of CPGs and

in other genes that could be involved in sarcoma tumorigenesis were not assessed. Second, for some samples we performed panel sequencing using a pool strategy, consequently precluding the detection of CNVs. Third, our results showed several interesting VUSs and most of them were not explored, even though the patients had a strong family history of cancer. Still, in comparison to previous studies, we performed NGS using large multigene panels, included individuals across sarcoma subtypes and with an earlier onset (both pediatric and young adults), from a genetically admixed and underrepresented population, and performed tumor LOH analysis for pathogenic variants, rendering insights into sarcoma genetic predisposition.

Overall, our findings reveal that 1 in each 5 patients with sarcomas under 40 years old presented a germline pathogenic variant in a cancer predisposing gene, including in patients without family history of cancer and without non-cancer clinical signs expected for the underlying syndrome. Considering the current recommendations for genetic screening in sarcoma patients, a substantial number of patients with GPVs would not be detected (45–47). This pinpoints the urgency of implementing appropriated genetic screening strategies for these individuals and their families. GPVs rates varied across sarcoma subtypes and exhibited some specific gene-subtypes associations. Given the great heterogeneity among sarcomas subtypes, it is likely that subtype and even population specific guidelines would be more appropriated. Nevertheless, until larger and powered studies can identify optimal testing strategies, it would not be unreasonable, considering the latest findings, to recommend genetic consultation and testing for all patients with sarcoma.

ACKNOWLEDGMENTS

We acknowledge the patients who participated on the study, and the A.C.Camargo professionals and biobank for sample processing.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

FINANCIAL SUPPORT: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2014/50943-1, 2018/06269-5 and 2018/17027-2); Conselho Nacional de Desenvolvimento Científico e Tecnológico (465682/2014-6 and 426835/2018-2) Coordination for the Improvement of Higher Education Personnel (CAPES - 88887.136405/2017-00, 88882.366031/2019-01 and 88887.509818/2020-00).

AUTHORS CONTRIBUTION:

Conceptualization, D.M.C and G.T.T.; Methodology and experimental analysis, N.A.C. and K.M.S.; Pathological review and analysis, F.D.C.; Formal analysis, N.A.C, K.M.S, B.R.P, G.T.T; Clinical data curation, J.M.L., M.N.C.F., D.C.S.Q., D.P.P., C.M.L.C., J.C.C.R., C.A.L.M.; Writing—original draft preparation, N.A.C.; Writing—review and editing, K.M.S., F.D.C., J.M.L., M.N.C.F., D.C.S.Q., D.P.P., C.M.L.C., J.C.C.R., C.A.L.M, B.R.P., D.M.C. and G.T.T.; Supervision, G.T.T.; Funding acquisition and resources, D.M.C. and G.T.T. All authors have read and agreed to the published version of the manuscript.

6. References

1. Thomas DM, Ballinger ML. Etiologic, environmental and inherited risk factors in sarcomas. *J Surg Oncol.* 2015 Apr;111(5):490–5.
2. Farid M, Ngeow J. Sarcomas Associated With Genetic Cancer Predisposition Syndromes: A Review. *The Oncologist.* 2016 Aug;21(8):1002–13.
3. Choi JH, Ro JY. The 2020 WHO Classification of Tumors of Soft Tissue: Selected Changes and New Entities. *Adv Anat Pathol.* 2021 Jan;28(1):44–58.
4. Wakeford R. Radiation in the workplace-a review of studies of the risks of occupational exposure to ionising radiation. *J Radiol Prot Off J Soc Radiol Prot.* 2009 Jun;29(2A):A61-79.

5. Ballinger ML, Goode DL, Ray-Coquard I, James PA, Mitchell G, Niedermayr E, et al. Monogenic and polygenic determinants of sarcoma risk: an international genetic study. *Lancet Oncol.* 2016 Sep 1;17(9):1261–71.
6. Cohen RJ, Curtis RE, Inskip PD, Fraumeni JF. The risk of developing second cancers among survivors of childhood soft tissue sarcoma. *Cancer.* 2005 Jun 1;103(11):2391–6.
7. Kleinerman RA, Tucker MA, Abramson DH, Seddon JM, Tarone RE, Fraumeni JF. Risk of soft tissue sarcomas by individual subtype in survivors of hereditary retinoblastoma. *J Natl Cancer Inst.* 2007 Jan 3;99(1):24–31.
8. Kleinerman RA. Radiation-sensitive genetically susceptible pediatric subpopulations. *Pediatr Radiol.* 2009 Feb;39 Suppl 1:S27-31.
9. Benjamin RS, Futreal A. Are sarcomas hereditary? *Lancet Oncol.* 2016 Sep;17(9):1179–81.
10. Huang KL, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell.* 2018 Apr 5;173(2):355-370.e14.
11. Hameed M, Mandelker D. Tumor Syndromes Predisposing to Osteosarcoma. *Adv Anat Pathol.* 2018 Jul;25(4):217–22.
12. Schneider KW, Cost NG, Schultz KAP, Svihovec S, Suttman A. Germline predisposition to genitourinary rhabdomyosarcoma. *Transl Androl Urol.* 2020 Oct;9(5):2430–40.
13. Kim J, Light N, Subasri V, Young EL, Wegman-Ostrosky T, Barkauskas DA, et al. Pathogenic Germline Variants in Cancer Susceptibility Genes in Children and Young Adults With Rhabdomyosarcoma. *JCO Precis Oncol.* 2021 Jan 11;5:PO.20.00218.
14. Mirabello L, Zhu B, Koster R, Karlins E, Dean M, Yeager M, et al. Frequency of Pathogenic Germline Variants in Cancer-Susceptibility Genes in Patients With Osteosarcoma. *JAMA Oncol.* 2020 May;6(5):724–34.
15. Torrezan GT, da Silva FCC, Santos ÉMM, Krepischi ACV, Achatz MIW, Junior SA, et al. Mutational spectrum of the APC and MUTYH genes and genotype–phenotype correlations in Brazilian FAP, AFAP, and MAP patients. *Orphanet J Rare Dis.* 2013 Dec;8(1):54.
16. Barbosa MVR, Cordeiro de Lima VC, Formiga MN, Andrade de Paula CA, Torrezan GT, Carraro DM. High Prevalence of EGFR Mutations in Lung Adenocarcinomas From Brazilian Patients Harboring the TP53 p.R337H Variant. *Clin Lung Cancer.* 2020 Mar;21(2):e37–44.
17. Rahman N. Realizing the promise of cancer predisposition genes. *Nature.* 2014 Jan 16;505(7483):302–8.
18. Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet.* 2019 Nov;51(11):1560–5.

19. Zuzarte PC, Denroche RE, Fehringer G, Katzov-Eckert H, Hung RJ, McPherson JD. A two-dimensional pooling strategy for rare variant detection on next-generation sequencing platforms. *PLoS One*. 2014;9(4):e93455.
20. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med Off J Am Coll Med Genet*. 2015 May;17(5):405–24.
21. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant Review with the Integrative Genomics Viewer. *Cancer Res*. 2017 Nov 1;77(21):e31–4.
22. Grasel RS, Felicio PS, de Paula AE, Campacci N, Garcia FA de O, de Andrade ES, et al. Using Co-segregation and Loss of Heterozygosity Analysis to Define the Pathogenicity of Unclassified Variants in Hereditary Breast Cancer Patients. *Front Oncol*. 2020 Oct 2;10:571330.
23. de Angelis de Carvalho N, Niitsuma BN, Kozak VN, Costa FD, de Macedo MP, Kupper BEC, et al. Clinical and Molecular Assessment of Patients with Lynch Syndrome and Sarcomas Underpinning the Association with MSH2 Germline Pathogenic Variants. *Cancers*. 2020 Jul 9;12(7):E1848.
24. Naslavsky MS, Scliar MO, Yamamoto GL, Wang JYT, Zverinova S, Karp T, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. *Nat Commun*. 2022 Dec;13(1):1004.
25. Pinto EM, Figueiredo BC, Chen W, Galvao HCR, Formiga MN, Fragoso MCBV, et al. XAF1 as a modifier of p53 function and cancer susceptibility. *Sci Adv*. 2020 Jun;6(26):eaba3231.
26. DiStefano MT, Goehringer S, Babb L, Alkuraya FS, Amberger J, Amin M, et al. The Gene Curation Coalition: A global effort to harmonize gene–disease evidence resources. *Genet Med*. 2022 Aug 1;24(8):1732–42.
27. Gillani R, Camp SY, Han S, Jones JK, Chu H, O’Brien S, et al. Germline predisposition to pediatric Ewing sarcoma is characterized by inherited pathogenic variants in DNA damage repair genes. *Am J Hum Genet*. 2022 Jun;109(6):1026–37.
28. Chan SH, Lim WK, Ishak NDB, Li ST, Goh WL, Tan GS, et al. Germline Mutations in Cancer Predisposition Genes are Frequent in Sporadic Sarcomas. *Sci Rep*. 2017 Sep 6;7(1):10660.
29. Calvete O, Garcia-Pavia P, Domínguez F, Bougeard G, Kunze K, Braeuninger A, et al. The wide spectrum of POT1 gene variants correlates with multiple cancer types. *Eur J Hum Genet EJHG*. 2017 Nov;25(11):1278–81.
30. Kaszuba MC, Pulido JS, Folpe AL, Pichurin PN, Goodenberger ML, Spinner RJ. Malignant Peripheral Nerve Sheath Tumor in a Patient With BAP1 Tumor Predisposition Syndrome. *World Neurosurg*. 2018 Jan;109:362–4.
31. Volc SM, Ramos CRN, Galvão H de CR, Felicio PS, Coelho AS, Berardineli GN, et al. The Brazilian TP53 mutation (R337H) and sarcomas. *PLoS One*. 2020;15(1):e0227260.

32. Ognjanovic S, Olivier M, Bergemann TL, Hainaut P. Sarcomas in TP53 germline mutation carriers: A review of the IARC TP53 database. *Cancer*. 2012 Mar 1;118(5):1387–96.
33. Custódio G, Parise GA, Kiesel Filho N, Komechen H, Sabbaga CC, Rosati R, et al. Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. *J Clin Oncol Off J Am Soc Clin Oncol*. 2013 Jul 10;31(20):2619–26.
34. Achatz MI, Zambetti GP. The Inherited p53 Mutation in the Brazilian Population. *Cold Spring Harb Perspect Med*. 2016 Dec;6(12):a026195.
35. Formiga MN da C, de Andrade KC, Kowalski LP, Achatz MI. Frequency of Thyroid Carcinoma in Brazilian TP53 p.R337H Carriers With Li Fraumeni Syndrome. *JAMA Oncol*. 2017 Oct 1;3(10):1400–2.
36. Couto PP, Bastos-Rodrigues L, Schayek H, Melo FM, Lisboa RGC, Miranda DM, et al. Spectrum of germline mutations in smokers and non-smokers in Brazilian non-small-cell lung cancer (NSCLC) patients. *Carcinogenesis*. 2017 Oct 26;38(11):1112–8.
37. Whitworth J, Skytte AB, Sunde L, Lim DH, Arends MJ, Happerfield L, et al. Multilocus Inherited Neoplasia Alleles Syndrome: A Case Series and Review. *JAMA Oncol*. 2016 Mar 1;2(3):373.
38. McGuigan A, Whitworth J, Andreou A, Hearn T, Tischkowitz M, Maher ER. Multilocus Inherited Neoplasia Allele Syndrome (MINAS): an update. *Eur J Hum Genet*. 2022 Mar;30(3):265–70.
39. Bertolotto C, Lesueur F, Giuliano S, Strub T, de Lichy M, Bille K, et al. A SUMOylation-defective MITF germline mutation predisposes to melanoma and renal carcinoma. *Nature*. 2011 Dec 1;480(7375):94–8.
40. Paillerets BB de, Lesueur F, Bertolotto C. A germline oncogenic MITF mutation and tumor susceptibility. *Eur J Cell Biol*. 2014 Jan;93(1–2):71–5.
41. Guhan SM, Artomov M, McCormick S, Njauw CN, Stratigos AJ, Shannon K, et al. Cancer risks associated with the germline MITF(E318K) variant. *Sci Rep*. 2020 Dec;10(1):17051.
42. Oliveira LJC, Gongora ABL, Lima FAS, Canedo FSNA, Quirino CV, Pisani JP, et al. Expanding the phenotype of E318K (c.952G > A) MITF germline mutation carriers: case series and review of the literature. *Hered Cancer Clin Pract*. 2021 Jul 21;19:32.
43. Reijnders CMA, Waaijer CJF, Hamilton A, Buddingh EP, Dijkstra SPD, Ham J, et al. No Haploinsufficiency but Loss of Heterozygosity for EXT in Multiple Osteochondromas. *Am J Pathol*. 2010 Oct;177(4):1946–57.
44. de Andrea CE, Zhu JF, Jin H, Bovée JV, Jones KB. Cell cycle deregulation and mosaic loss of *Ext1* drive peripheral chondrosarcomagenesis in the mouse and reveal an intrinsic cilia deficiency: Polarity in peripheral chondrosarcomagenesis. *J Pathol*. 2015 Jun;236(2):210–8.

45. Gronchi A, Miah AB, Dei Tos AP, Abecassis N, Bajpai J, Bauer S, et al. Soft tissue and visceral sarcomas: ESMO-EURACAN-GENTURIS Clinical Practice Guidelines for diagnosis, treatment and follow-up☆. *Ann Oncol Off J Eur Soc Med Oncol*. 2021 Nov;32(11):1348–65.
46. National Comprehensive Cancer Network. Bone Cancer (Version 1.2023 [Internet]. [cited 2022 Sep 25]. Available from:
https://www.nccn.org/professionals/physician_gls/pdf/bone.pdf
47. National Comprehensive Cancer Network. Soft tissue Sarcoma (Version 2.2022). [Internet]. [cited 2022 Sep 25]. Available from:
https://www.nccn.org/professionals/physician_gls/pdf/sarcoma.pdf