



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**FACULDADE DE MEDICINA**  
**DEPARTAMENTO DE MEDICINA CLÍNICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS MÉDICAS**

**ROBERTA TAIANE GERMANO DE OLIVEIRA**

**RASTREAMENTO DE VARIANTES GENÉTICAS EM DNA POLIMERASES COM  
ATIVIDADE TRANSLESÃO NA NEOPLASIA MIELODISPLÁSICA**

**FORTALEZA**

**2023**

ROBERTA TAIANE GERMANO DE OLIVEIRA

RASTREAMENTO DE VARIANTES GENÉTICAS EM DNA POLIMERASES COM  
ATIVIDADE TRANSLEÇÃO NA NEOPLASIA MIELODISPLÁSICA

Tese apresentada ao Programa de Pós-Graduação em Ciências Médicas, do Departamento de Medicina Clínica, da Universidade Federal do Ceará, como parte dos requisitos para obtenção do título de Doutora em Ciências Médicas. Área de concentração: Doenças Crônicas Degenerativas

Orientador: Prof. Dr. Ronald Feitosa Pinheiro

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

D32r de Oliveira, Roberta Taiane Germano.  
Rastreamento de variantes genéticas em DNA polimerases com atividade translesão na neoplasia mielodisplásica / Roberta Taiane Germano de Oliveira. – 2023.  
141 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Faculdade de Medicina, Programa de Pós-Graduação em Ciências Médicas, Fortaleza, 2023.

Orientação: Prof. Dr. Ronald Feitosa Pinheiro.

Coorientação: Prof. Dr. Howard Lopes Ribeiro Junior.

1. Síndromes Mielodisplásicas. 2. Sequenciamento de Nucleotídeos em Larga Escala. 3. DNA Polimerase. 4. Mutação. I. Título.

CDD 610

---

ROBERTA TAIANE GERMANO DE OLIVEIRA

RASTREAMENTO DE VARIANTES GENÉTICAS EM DNA POLIMERASES COM  
ATIVIDADE TRANSLEÇÃO NA NEOPLASIA MIELODISPLÁSICA

Tese apresentada ao Programa de Pós-Graduação em Ciências Médicas, do Departamento de Medicina Clínica, da Universidade Federal do Ceará, como parte dos requisitos para obtenção do título de Doutora em Ciências Médicas. Área de concentração: Doenças Crônicas Degenerativas

Orientador: Prof. Dr. Ronald Feitosa Pinheiro

Aprovada em: 23/05/2023

BANCA EXAMINADORA

Ronald Feitosa Pinheiro (Orientador)  
Universidade Federal do Ceará (UFC)

Sílvia Maria Meira Magalhães  
Universidade Federal do Ceará (UFC)

Sara Teresinha Olalla Saad  
Universidade Estadual de Campinas (UNICAMP)

Luis Felipe Pinto  
Instituto Nacional de Câncer (INCA)

Antônio Roberto Lucena de Araújo  
Universidade Federal de Pernambuco (UFPE)

FORTALEZA

2023

*Dedico este trabalho à Ariadna, amiga de infância e irmã de coração, vítima da COVID-19.*

## AGRADECIMENTOS

Há 13 anos o sonho de ingressar na Universidade pública estava se realizando. Com muito esforço, dedicação e o suporte da minha família, consegui alcançar um dos meus maiores objetivos e agora me vejo finalizando mais uma etapa. Os últimos três anos foram especialmente difíceis. Uma pandemia que tirou tantas pessoas de nossa vida, por conta de um Governo criminoso e cruel, mas que não conseguiu tirar de mim e da minha família a realização deste sonho. Chego ao fim desta jornada sendo muito grata. Grata à minha família, que abdicou de tanta coisa para me ver realizando esse sonho, que não é só meu, é nosso. É como se diz “sonho que se sonha só é apenas um sonho. Sonho que se sonha junto, torna-se realidade. E a realização desse sonho só pode se tornar possível graças à incalculável ajuda de muitas pessoas.

A meus avós, **Júlio e Antônia**, por serem meu pai-vovô e mãe-vovó. Por terem me criado de uma forma simples, mas singela. À minha avó por ter tido mão de ferro nas horas em que era preciso, por me deixar na escola todos os dias, por catar manga com a gente na reitoria da UFC. Lugar que eu olhava e não acreditava que um dia estaria lá. Hoje estou por causa de você, Dona Antônia. Não só como aluna e possivelmente Doutora rsrs, mas como funcionária. Ao meu avô, por ser meu tudo. Por trazer à minha memória até hoje os dias que ele fazia nosso café da manhã antes de ir à escola. Pelo esforço que sempre fez para continuar trabalhando e nos trazer o pão de cada dia, assim como o dinheiro das minhas passagens para ir à faculdade durante o cursinho e durante minha graduação... Tudo que sou hoje, devo a vocês.

Com um grande carinho, agradeço às minhas irmãs, **Natalia e Taline**. Por passarem pelas dificuldades da vida de mãos dadas comigo, por serem para mim exemplos de mulheres fortes, cada uma ao seu modo. Obrigada, Taline por me salvar várias vezes durante o recreio na escola, se não fosse você talvez eu não estivesse viva para chegar até aqui hahaahha. Obrigada Natalia por estar sempre disponível a nos ajudar nas emergências do dia a dia, você nasceu com o dom da enfermagem e desejo muito que consiga alcançar esse sonho. Eu as amo imensamente.

Ao meu tio, **Pe. Julio**, por ter me levado para conhecer a biblioteca dos padres em sua casa e ter acendido em mim a chama do estudo. Por ter sido para mim a maior inspiração da pessoa que eu quero me tornar. Um grande exemplo de homem, militante e servo de Deus.

À minha mãe, **Jacira**, por ser para mim o maior exemplo de uma mulher guerreira e batalhadora. Pelo seu apoio incondicional durante minha trajetória de vida, sendo diversas vezes meu alicerce e por sempre ter se esforçado para me educar, me ensinando as tarefinhas de casa

todos os dias, mesmo sem ter terminado o ensino fundamental... Por ter abdicado de tanto por nós três. Por todo amor que tem por mim. Obrigada mãe por seus ensinamentos e esforço sem medida. Eu a amo incondicionalmente. Sem você nada disso seria possível.

Aos meus amigos de infância, **Romichele, Nayara, Ariadna, Ronaldo e Maxwell**, por estarem presentes durante grande parte dos passos dados para a realização deste sonho e por sempre me apoiarem e darem forças para dar mais um passo. Em especial à **Romichele** e a **Nayara**, pela paciência sem limites comigo, por garantirem todo o suporte que eu precisava nos maiores momentos de dificuldade, por sempre me encorajarem a ser melhor do que eu posso ser e simplesmente por fazerem parte da minha vida. Vocês são a família que Deus me permitiu escolher. Ariadna, tenho certeza de que você estaria muito orgulhosa neste dia. E mais certeza ainda que estaríamos bêbadas no final de semana seguinte. Você faz falta. Te amo.

Ao **Lucas Bayde**, com todo carinho, por ser um grande companheiro. Obrigada pelos cafés da manhã de domingo discutindo política ou esporte, obrigada pelas comidinhas preparadas com tanto carinho, obrigada por todo o suporte que tem me proporcionado e por conseguir lidar com minhas chatices e reclamações do dia a dia, rrsrs. O seu apoio foi um dos grandes catalisadores nessa jornada.

À **Dona Layre, Sr. Asdrubal, Aarão e Camila**, a minha segunda família. Vocês não têm ideia de como o carinho e o suporte que vocês me deram todo esse tempo fez essa jornada um pouco mais leve. Agradeço à vida pela oportunidade de ter conhecido pessoas tão bondosas.

Aos meus amigos da UNILAB, **Tatyane, Sarah, Lívia, Davino, Camila, Julie, Henrique, Fernanda, Débora e Simone**, meu muito obrigada pelo companheirismo e ajuda no dia a dia durante os anos de UNILAB. Foi incrível trabalhar com vocês e, apesar da ausência, tenho um enorme carinho por cada um.

Aos estudantes do **Laboratório de Citogenômica do Câncer**, muito obrigada pelo companheirismo do dia a dia e por estarem presentes nessa jornada auxiliando sempre que eu preciso. Apesar de darem um certo trabalho, acreditem, vocês são um suporte e tanto.

Ao meu grupinho, **Dani, Gabriel e Wesley**, que estão disponíveis para absolutamente tudo que eu preciso. Obrigada pelo apoio sem limites durante todos esses anos de Graduação, Mestrado e Doutorado. Pelas madrugadas de experimentos e por aguentarem meus desesperos e angústias. Dani, estou muito orgulhosa da pesquisadora que você se tornou. Você é uma grande inspiração.

A meu orientador, **Ronald Feitosa**, por acreditar no meu potencial e ter me dado a oportunidade de trabalhar em seu laboratório, sendo um grande orientador acadêmico e um médico sem igual. Muito obrigada por ter me repassado e continuar repassando tantos ensinamentos, tanto academicamente quanto pessoalmente. Sua vontade de fazer pesquisa séria, honesta e de qualidade (não seja uma pessoa medíocre, hahaha), atrelada à vontade de melhorar a vida dos pacientes com SMD me mostrou que é esse caminho que quero seguir. Infelizmente o seu maior defeito é não ser torcedor do voção.

Com um carinho especial, ao meu amigo **Howard**. Pelos conselhos, incentivo e orientações durante a graduação e o Mestrado. Por sempre me expor meus erros e me incentivar a melhorar e por acreditar no meu potencial. Muito obrigada pelos conselhos e orientações até aqui, você tem grande influência na minha vida acadêmica.

Um agradecimento muito especial **aos pacientes diagnosticados com SMD**, que me ensinaram muito sobre a vida durante o período que fiquei no ambulatório e que contribuem para o avanço do conhecimento da doença, nos permitindo realizarmos nossas pesquisas através da doação de suas amostras.

**Aos participantes da banca**, por terem aceitado avaliar este trabalho e disponibilizado um pouco do seu tempo, engrandecendo o mesmo com suas considerações.

À **Ivone**, pela enorme paciência durante esses dois anos e por estar sempre disposta a resolver os problemas da melhor forma possível.

Ao **CNPq** e à **CAPES** pelo apoio financeiro a essa pesquisa.

À **Universidade Federal do Ceará**, pela oportunidade de fazer o Doutorado e por todo apoio financeiro concedido durante esses anos

E, finalmente, ao nosso **Presidente Luiz Inácio Lula da Silva**, que, através dos programas de incentivo a famílias de baixa renda, garantiu nosso prato na mesa para termos forças e lutarmos pelo nosso lugar na Universidade Pública.

*“Sua determinação dita seu destino”*

***Livro Anne With An E***

## RESUMO

A Neoplasias Mielodisplásica (SMD) é o câncer de medula óssea mais comum em idosos acima de 65 anos, caracterizada por citopenias, displasias, hematopoese ineficaz e risco de progressão para leucemia mieloide aguda. Alterações citogenéticas estão presentes em 40-60% dos casos e 94% dos pacientes apresentam pelo menos uma mutação oncogênica. Acredita-se que estas alterações sejam decorrentes de lesões no DNA ocasionados por processos oxidativos endógenos ou por exposição ambiental. A tolerância a danos no DNA não passíveis de reparo é realizada por DNA polimerases de síntese translesão (TLS), que transpassam ou corrigem as lesões, porém, podem gerar mutações pontuais, sendo importante seu estudo nas doenças oncológicas. O objetivo deste estudo foi realizar um *screening* genético para identificar, e classificar variantes em genes de DNA polimerases com atividade TLS (*REV3L*, *POLQ*, *REVI*, *POLI*, *POLH*, *POLK*, *POLL* e *POLN*) em pacientes portadores de SMD, relacionando os achados com alterações cromossômicas e subtipos da doença de menor e maior risco de evolução para LMA. Foram sequenciadas, por NGS-Illumina, amostras de medula óssea de 50 pacientes diagnosticados com SMD, 17 classificados como SMD com excesso de blastos (SMD-EB) e 13 pacientes classificados como SMD com sideroblastos em anel (SMD-SA). O estudo citogenético foi realizado para caracterização prognóstica dos pacientes. Para análise das variantes, foram utilizadas cinco ferramentas de predição de patogenicidade *in silico* (*Ensembl VEP*, *Mutation Taster*, *Sift*, *Mutpred* e *Provean*) e, para análise de oncogenicidade, foram utilizados padrões para classificação de oncogenicidade de variantes somáticas (SOP), recomendados pelo *Clinical Genome Resource* (ClinGen), *Cancer Genomics Consortium* (CGC) e *Variant Interpretation for Cancer Consortium* (VICC). Foram selecionadas para classificação de oncogenicidade neste estudo, variantes com perda de função (LoF) e variantes registradas no COSMIC Database. Para o gene *POLQ*, foram identificadas 194 variantes, 85 variantes *missense*, 38 variantes sinônimas e 17 variantes do tipo perda de função (*Lof - frameshift*, *inframe*, *stop gained*, *stop loss*, *splice donor* e *splice acceptor*). Das variantes analisadas, 10 foram classificadas como VUS e 1 como Provavelmente Benigna. Para o gene *REV3L*, foram identificadas 284 variantes (85 *missense*, 38 sinônimas e 17 com perda de função). Das variantes analisadas, 18 foram classificadas como Potencialmente Oncogênicas, 7 VUS e 3 Benignas. Para o gene *POLI* foram identificadas 58 variantes (22 *missense*, 7 sinônimas e 17 com perda de função). Das variantes analisadas, 3 foram classificadas como Potencialmente Oncogênicas, 1 VUS e 2 como Benignas ou Provavelmente Benignas. Para o gene *REVI* foram identificadas 108 variantes (39 *missense*, 12 sinônimas e 8 com perda de função). Das variantes analisadas, 2 foram classificadas como Potencialmente Oncogênicas e 4 como VUS. Para o gene *POLH* foram identificadas 22 variantes (18 *missense* e 4 sinônimas). Não identificamos variantes oncogênicas no gene *POLH*. Pacientes do subtipo SMD-EB tiveram alta incidência de variantes LoF nos genes *POLQ* e *REV3L*, comparado a pacientes do subtipo SMD-SA, representando perda do mecanismo de síntese translesão nestes pacientes. Pacientes do subtipo SMD-SA com cariótipo normal apresentaram menor frequência de variantes LoF e maior frequência de variantes *missense* para os genes *REVI* e *POLI*, bem como altas taxas de variantes não-LoF, o que pode representar um biomarcador de instabilidade genômica na SMD. Um paciente com cariótipo complexo apresentou acúmulo de mutações LoF nos genes *POLQ* e *REV3L*, o que sugere que a grande instabilidade genética observada nos pacientes com múltiplas alterações cromossômicas possa estar presente devido à ausência do mecanismo TLS. Estes resultados reforçam o impacto de mutações no processo fisiopatológico da SMD, em especial acometendo genes não incluídos em escores de risco prognóstico.

**Palavras-chave:** Síndromes Mielodisplásicas. Sequenciamento de Nucleotídeos em Larga Escala. DNA Polimerase. Mutação.

## ABSTRACT

Myelodysplastic Neoplasms (MDS) is the most common bone marrow cancer in the elderly over 65 years, characterized by cytopenias, cell dysplasias, ineffective hematopoiesis, and risk of progression to acute myeloid leukemia. Cytogenetic alterations are present in 40-60% of cases and 94% of patients have at least one oncogenic mutation. These changes are believed to be due to DNA damage caused by endogenous oxidative processes or environmental exposure. Tolerance to non-repairable DNA damage is performed by translesion synthesis DNA polymerases (TLS), which pierce or correct the lesions; however, they can generate point mutations, and their study is important in oncological diseases. The aim of this study was to carry out a genetic screening to identify, characterize and classify variants in DNA polymerase genes with TLS activity (*REV3L*, *POLQ*, *REV1*, *POLI*, *POLH*, *POLK*, *POLL*, and *POLN*) in patients with MDS, relating the findings with chromosomal alterations and disease subtypes of lower and higher risk of progression to AML. Bone marrow samples from 50 patients diagnosed with MDS were sequenced by NGS-Illumina, with 17 classified as MDS with excess blasts (SMD-EB) and 13 classified as MDS with ring sideroblasts (SMD-SA). The cytogenetic study was performed for the prognostic characterization of the patients. For the analysis of the variants obtained, five *in silico* pathogenicity prediction tools were used (Ensembl VEP, Mutation Taster, Sift, Mutpred, and Provean) and, for the oncogenicity analysis of these variants, standards for classification of oncogenicity of somatic variants (SOP), recommended by the Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC) and Variant Interpretation for Cancer Consortium (VICC). This study selected Loss-of-function (LoF) variants registered in the COSMIC Database for oncogenicity classification. For the *POLQ* gene, 194 variants, 85 missense variants, 38 synonymous variants, and 17 loss-of-function variants (LoF - frameshift, inframe, stop gained, stop loss, splice donor, and splice acceptor) were identified. Of the analyzed variants, 10 were classified as VUS and 1 as Probably Benign. For the *REV3L* gene, 284 variants were identified (85 missense, 38 synonyms, and 17 with loss of function). Of the analyzed variants, 18 were classified as Potentially Oncogenic, 7 as VUS, and 3 as Benign. For the *POLI* gene, 58 variants were identified (22 missense, 7 synonyms, and 17 with loss of function). Of the analyzed variants, 3 were classified as Potentially Oncogenic, 1 as VUS, and 2 as Benign or Probably Benign. For the *REV1* gene, 108 variants were identified (39 missense, 12 synonyms, and 8 with loss of function). Of the analyzed variants, 2 were classified as Potentially Oncogenic and 4 as VUS. For the *POLH* gene, 22 variants were identified (18 missense and 4 synonyms). We did not identify oncogenic variants in the gene *POLH*. Patients of the SMD-EB subtype had a high incidence of LoF variants in the *POLQ* and *REV3L* genes, compared to patients of the SMD-SA subtype, representing loss of the translesion synthesis mechanism in these patients. Patients of the SMD-SA subtype with normal karyotype had less frequency of LoF variants and higher frequency of missense variants for the *REV1* and *POLI* genes, as well as high rates of non-LoF variants, which may represent a biomarker of genomic instability in MDS. A patient with a complex karyotype showed an accumulation of LoF mutations in the *POLQ* and *REV3L* genes, which suggests that the great genetic instability observed in patients with multiple chromosomal alterations may be present due to the absence of the TLS mechanism. These results reinforce the impact of genetic mutations on the pathophysiological process of MDS, especially affecting genes not included in current prognostic risk scores.

**Keywords:** Myelodysplastic Syndrome. NGS. DNA Polymerase. Mutation

## LISTA DE FIGURAS

<b>Figura 1:</b> Representação esquemática da patogênese e evolução da Neoplasia Mielodisplásica .....	19
<b>Figura 2:</b> Categorias de risco e escores prognósticos segundo o IPSS-R .....	25
<b>Figura 3:</b> Representação esquemática dos Mecanismos de Reparo e Tolerância a danos no DNA.....	29
<b>Figura 4:</b> Representação esquemática da Ativação do Mecanismo Translesão .....	31
<b>Figura 5:</b> Representação esquemática do Mecanismo de tolerância a danos no DNA através da síntese translesão e sua implicação na carcinogênese.....	32
<b>Figura 6:</b> Representação esquemática da Metodologia de Citogenética por Bandeamento-G .....	43
<b>Figura 7:</b> Delineamento experimental do estudo .....	53
<b>Figura 8:</b> Esquematização da metodologia utilizada para análise de variantes realizada neste estudo	54
<b>Figura 9:</b> Representação dos resultados citogenéticos encontrados <i>nos pacientes com SMD incluídos no estudo</i> .....	57

## LISTA DE TABELAS

<b>Tabela 1:</b> Classificação clínica da SMD segundo os critérios da OMS 2016 .....	24
<b>Tabela 2:</b> Regiões sequenciadas dos genes REV3L, POLQ, POLH, REV1, POLI, POLL, POLM, POLN e POLK e o número de CDS de cada gene .....	44
<b>Tabela 3:</b> Principais classificações de variantes de acordo com o Ensembl VEP .....	49
<b>Tabela 4:</b> Ferramentas de análise utilizadas neste estudo .....	52
<b>Tabela 5:</b> Descrição clínica dos pacientes com Neoplasia Mielodisplásica incluídos no estudo .....	56
<b>Tabela 6:</b> Análise in sílico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene POLQ .....	62
<b>Tabela 7:</b> Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene POLQ .....	65
<b>Tabela 8:</b> Análise in sílico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene REV3L .....	71
<b>Tabela 9:</b> Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene REV3L .....	76
<b>Tabela 10:</b> Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene POLH .....	81
<b>Tabela 11:</b> Análise in sílico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene POLI .....	84
<b>Tabela 12:</b> Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene POLI .....	87
<b>Tabela 13:</b> Análise in sílico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene REV1 .....	91
<b>Tabela 14:</b> Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene REV1 .....	93

## LISTA DE ABREVIATURAS E SIGLAS

AA	<i>Aminoacids</i>
ACMG	<i>American College of Medical Genetics and Genomics</i>
AMP	<i>Association for Molecular Pathology</i>
ANC	<i>Contagem absoluta de neutrófilos</i>
ATM	<i>Ataxia telangiectasia mutada</i>
BER	<i>Base excision repair</i>
BRCT	<i>BRCA1 C terminal</i>
CGC	<i>Cancer Genomics Consortium</i>
CPH	<i>Células progenitoras hematopoiéticas</i>
CLINGEN	<i>Clinical Genome Resource</i>
COSMIC	<i>Catalogue of somatic mutations in cancer</i>
DDR	<i>DNA damage response</i>
DDT	<i>DNA damage tolerance</i>
dbSNP	<i>Database Single-nucleotide polymorphism</i>
dNTP	<i>Desoxirribonucleotídeo trifosfato</i>
FAB	<i>Franco-Americano-Britânica</i>
FISH	<i>Hibridização in situ por fluorescência</i>
FRA6F	<i>Sítio frágil localizado no cromossomo 6</i>
gnmoAD	<i>The Genome Aggregation Database</i>
GRCh37/ hg19	<i>Genome Reference Consortium build 37/hg19</i>
Hb	<i>Hemoglobina</i>
HR	<i>Recombinação homóloga</i>
IPSS-R	<i>Revised-Índice prognostic score system</i>
KCl	<i>Cloreto de potássio</i>
LMA	<i>Leucemia mieloide aguda</i>
LOF	<i>Loss of function</i>
MMR	<i>Reparo de erros de emparelhamento de bases</i>
MO	<i>Medula óssea</i>
mRNA	<i>RNA mensageiro</i>
MUPRED	<i>Mutation prediction</i>
NCBI	<i>The National Center for Biotechnology Information</i>
NER	<i>Reparo por excisão de nucleotídeos</i>

NGS	<i>Next Generation Sequencing</i>
NHEJ	<i>Junção por extremidades não-homólogas</i>
NR	<i>Neutropenia refratária</i>
OMS	<i>Organização Mundial da Saúde</i>
OVS1	<i>Oncogenic very Strong-1</i>
OS1	<i>Oncogenic Strong-1</i>
OS2	<i>Oncogenic Strong-2</i>
OS3	<i>Oncogenic Strong-3</i>
OM1	<i>Oncogenic Moderate-1</i>
OM2	<i>Oncogenic Moderate-2</i>
OP1	<i>Oncogenic Supporting-1</i>
OP2	<i>Oncogenic Supporting-1</i>
OP3	<i>Oncogenic Supporting-3</i>
OP4	<i>Oncogenic Supporting-4</i>
OM3	<i>Oncogenic Moderate-3</i>
OM4	<i>Oncogenic Moderate-4</i>
PCNA	<i>Antígeno nuclear de proliferação celular</i>
PCR	<i>Reação em cadeia da polimerase</i>
POLH	<i>DNA polymerase eta</i>
POLI	<i>DNA polymerase iota</i>
POLK	<i>DNA polymerase kappa</i>
POLL	<i>DNA polymerase lambda</i>
POLN	<i>DNA polymerase nu</i>
POLQ	<i>DNA polymerase theta</i>
POLZ	<i>DNA polymerase zeta</i>
PROVEAN	<i>Protein Variation Effect Analyze</i>
RAD18	<i>E3 ubiquitin-protein ligase</i>
REF	<i>Referência</i>
REV1	<i>DNA directed polymerase</i>
REV3L	<i>DNA directed polymerase zeta catalytic subunit</i>
RPMI	<i>Roswell Park Memorial Institute medium</i>
SBVS1	<i>Somatic Benign Very Strong-1</i>
SBS1	<i>Somatic Benign Strong-1</i>

SBS2	<i>Somatic Benign Strong-2</i>
SBP1	<i>Somatic Benign Supporting-1</i>
SBP2	<i>Somatic Benign Supporting-2</i>
SEER	<i>Surveillance, epidemiology and end Results</i>
SF3B1	<i>Splicing factor 3b subunit 1</i>
SIFT	<i>Sorting intolerant from tolerant</i>
SMD	<i>Neoplasia mielodisplásica</i>
SMD-EB1	<i>SMD com excesso de blastos 1</i>
SMD-EB2	<i>SMD com excesso de blastos 2</i>
SMD-SA	<i>SMD com sideroblastos em anel</i>
SMD – SA- DU	<i>SMD com sideroblastos em anel e displasia em uma única linhagem</i>
SMD- SA- DM	<i>SMD com sideroblastos em anel e displasia em múltiplas linhagens</i>
SNP	<i>Single nucleotide polymorphism</i>
SNV	<i>Single nucleotide variant</i>
SOP	<i>Standards operation procedure</i>
TdT	<i>Transferase deoxinucleotídica terminal</i>
TLS	<i>Translesão</i>
TP53	<i>Tumor protein 53</i>
ub-PCNA	<i>Antígeno nuclear de proliferação celular ubiquitinado</i>
UBZ	<i>Domínio de ligação à ubiquitina</i>
VAF	<i>Variant Allele Frquency</i>
VEP	<i>Variant Effect Predictor</i>
VICC	<i>Variant Interpretation for Cancer Consortium</i>

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>16</b>
1.1. NEOPLASIA MIELODISPLÁSICA .....	16
1.1.1. Aspectos gerais e epidemiologia.....	16
1.1.2. Etiologia.....	16
1.1.3. Mecanismos de Patogênese da Neoplasia Mielodisplásica .....	17
1.1.4. O papel das alterações cromossômicas na patogênese da SMD .....	20
1.1.5. Classificação Clínica da Neoplasia Mielodisplásica.....	22
1.1.6. Estratificação Prognóstica da Neoplasia Mielodisplásica .....	25
1.2. LESÕES AO DNA E MECANISMOS DE RESPOSTA AOS DANOS.....	26
1.3. MECANISMOS DE REPARO DO DNA .....	26
1.4. MECANISMOS DE TOLERÂNCIA A DANOS NO DNA: SÍNTESE TRANSLESÃO.....	29
1.4.1. A baixa fidelidade do mecanismo TLS: quando o relacionamento com o DNA pode falhar .....	33
1.4.2. Famílias de DNA Polimerases .....	33
1.4.2.1. Família A: DNA Polimerases TLS $\theta$ (POLQ) e $\nu$ (POLN).....	33
1.4.2.2. Família B: DNA Polimerase TLS $\zeta$ (POLZ) – REV3L.....	35
1.4.2.3. Família Y: DNA Polimerases TLS $\eta$ (POLH), $\kappa$ (POLK), $\iota$ (POLI) e REV1 .....	36
<b>2. OBJETIVOS</b> .....	<b>41</b>
2.1. OBJETIVO GERAL.....	41
2.2. OBJETIVOS ESPECÍFICOS.....	41
<b>3. PACIENTES, MATERIAIS E MÉTODOS</b> .....	<b>42</b>
3.1. ASPECTOS ÉTICOS DA PESQUISA .....	42
3.2. OBTENÇÃO DAS AMOSTRAS DE CÉLULAS DE MEDULA ÓSSEA .....	42
3.3. CARIÓTIPO POR BANDEAMENTO-G .....	43
3.4. SEQUENCIAMENTO-ALVO DOS GENES <i>REV3L</i> , <i>POLQ</i> , <i>POLH</i> , <i>REV1</i> , <i>POLI</i> , <i>POLL</i> , <i>POLM</i> , <i>POLN</i> , <i>POLK</i> .....	44
3.4.1. Extração de DNA Genômico.....	44
3.4.2. Customização do Painel Genético .....	44
3.4.3. Preparo da Biblioteca.....	44
3.4.4. Análise das variantes obtidas nos genes <i>POLQ</i> , <i>REV3L</i> , <i>POLH</i> , <i>REV1</i> , <i>POLI</i> , <i>POLL</i> , <i>POLM</i> , <i>POLN</i> , <i>POLK</i> .....	46
3.4.4.1. Análise Primária: Qualidade da corrida, filtragem e alinhamento com genoma de referência.....	46
3.4.4.2. Análise Secundária: Chamada de variantes.....	47
3.4.4.3. Análise terciária: Classificação de Patogenicidade e Oncogenicidade .....	47
<b>4. RESULTADOS</b> .....	<b>55</b>
4.1. CARACTERIZAÇÃO DOS PACIENTES .....	55
4.2. CLASSIFICAÇÃO E CARACTERIZAÇÃO DAS VARIANTES.....	58
4.2.1. Caracterização, análise <i>in silico</i> da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene <i>POLQ</i> .....	58
4.2.1.1. Variantes com Potencial Perda de Função (LoF).....	59
4.2.1.2. Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP).....	64
4.2.2. Caracterização, análise <i>in silico</i> da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene <i>REV3L</i> .....	69
4.2.2.1. Variantes com Potencial Perda de Função (LoF).....	69

4.2.2.2.	Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)	74
4.2.3.	<i>Caracterização, análise in silico da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene POLH</i> .....	79
4.2.3.1.	Variantes com Potencial Perda de Função (LoF).....	79
4.2.3.2.	Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)	79
4.2.4.	<i>Caracterização, análise in silico da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene POLI</i> .....	82
4.2.4.1.	Variantes com Potencial Perda de Função (LoF).....	82
4.2.4.2.	Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)	86
4.2.5.	<i>Caracterização, análise in silico da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene REVI</i> .....	89
4.2.5.1.	Variantes com Potencial Perda de Função (LoF).....	89
4.2.5.2.	Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)	92
<b>5.</b>	<b>DISCUSSÃO</b> .....	<b>95</b>
<b>6.</b>	<b>CONCLUSÕES</b> .....	<b>106</b>
<b>6.</b>	<b>LIMITAÇÕES E PERSPECTIVAS FUTURAS</b> .....	<b>107</b>
<b>7.</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>108</b>

## 1. INTRODUÇÃO

### 1.1. Neoplasia Mielodisplásica

#### 1.1.1. Aspectos gerais e epidemiologia

A Neoplasia Mielodisplásica (SMD) é um câncer da medula óssea caracterizado por distúrbios de diferenciação e maturação celular que tem como consequência displasias em uma ou mais linhagens celulares, citopenia(s) no sangue periférico, medula óssea (MO) hipercelular e risco de evolução para Leucemia Mieloide Aguda (LMA) em 30% a 40% dos casos (ARBER et al., 2016; HASSERJIAN, 2019).

A fisiopatologia da doença compreende múltiplos estágios e pode acometer diversas vias de sinalização, sendo a grande heterogeneidade da SMD o que a torna uma das doenças mais desafiadoras da oncohematologia. Seu curso clínico pode ser dividido em duas fases distintas, alto risco e baixo risco de progressão para LMA, tendo como principais critérios o número de linhagens celulares displásicas, a porcentagem de blastos na medula óssea, que aumenta conforme a evolução da doença, e a presença de alterações moleculares específicas (VARDIMAN et al, 2008; BEJAR, 2014; ALAGGIO et al, 2022).

A SMD é reconhecida como “a doença do idoso” por acometer, prioritariamente, indivíduos com idade superior a 70 anos, sendo mais comum em indivíduos do sexo masculino. Em 2016, a Organização Mundial de Saúde (OMS) publicou a epidemiologia da SMD: em indivíduos acima de 70 anos, a incidência é de 20 casos por 100.000 de habitantes, porém foi estimado que a real incidência nessa faixa de idade seja de 75 casos por 100.000 habitantes, devido à baixa notificação de casos de câncer em geral. Apesar da fisiopatologia desta neoplasia ainda não estar completamente elucidada, sabe-se que agentes citotóxicos que causam danos à medula óssea, como o benzeno, cigarro e venenos agrícolas, são possíveis agentes etiológicos da SMD (CHAMSEDDINE et al., 2016; COSTA et al., 2021; SWERDLOW et al., 2017)

#### 1.1.2. Etiologia

Estudos epidemiológicos demonstram diversos fatores de risco associados ao desenvolvimento da SMD, como a exposição prolongada ao benzeno, tabaco, formol e venenos químicos agrícolas (pesticidas, herbicidas e fertilizantes) bem como a radiação ionizante (SEKERES, 2010; CHAMSEDDINE et al, 2016; BOWEN et al, 2013; ADES et al, 2014). Algumas doenças congênitas, como a anemia de Fanconi, podem, ocasionalmente, levar ao

desenvolvimento da doença (ALTER et al, 2014). No entanto, no que tange à alterações congênitas, embora existam casos raros descritos na literatura de SMD com componente familiar, não foram encontradas correlações significativas da doença como condição hereditária ou congênita, apesar de conhecermos mutações germinativas associadas à SMD com componente familiar (LIEW et al, 2011)

Os fatores etiológicos da SMD se tornaram uma importante linha de pesquisa de nosso grupo, com alguns trabalhos já publicados e outros em andamento. Recentemente, relatamos importantes associações entre a exposição crônica a agrotóxicos e a desregulação na expressão de genes relacionados ao sistema de reparo do DNA, bem como alterações em marcadores epigenéticos associadas aos mecanismos de síntese translesão e identificação de alterações cromossômicas na medula óssea de trabalhadores rurais expostos a agrotóxicos, demonstrando o impacto da exposição a venenos agrícolas nas células da medula óssea (COSTA et al, 2021; CAVALCANTE et al, 2022)

### ***1.1.3. Mecanismos de Patogênese da Neoplasia Mielodisplásica***

O modelo proposto mais aceito para fisiopatologia da SMD é o modelo de múltiplos passos que engloba alterações funcionais nas células tronco hematopoéticas (CTH) que podem ser causadas pelo envelhecimento e a senescência celular (por exemplo, devido ao encurtamento dos telômeros), pela exposição a agentes genotóxicos, como os venenos agrícolas, ou por alguma condição inflamatória no microambiente medular. Essas alterações funcionais são decorrentes de modificações moleculares, genéticas e epigenéticas e, como consequência, ocasionam imunidade inata desregulada, hematopoese ineficaz e alterações no processo de apoptose, mecanismos de reparo do DNA e maquinaria do ciclo celular, sendo tais fenômenos gerados de modo independente ou inter-relacionados (Figura 1) (GARCIA-MANERO et al, 2016; ADES et al, 2014; CHAMSEDDINE et al, 2016).

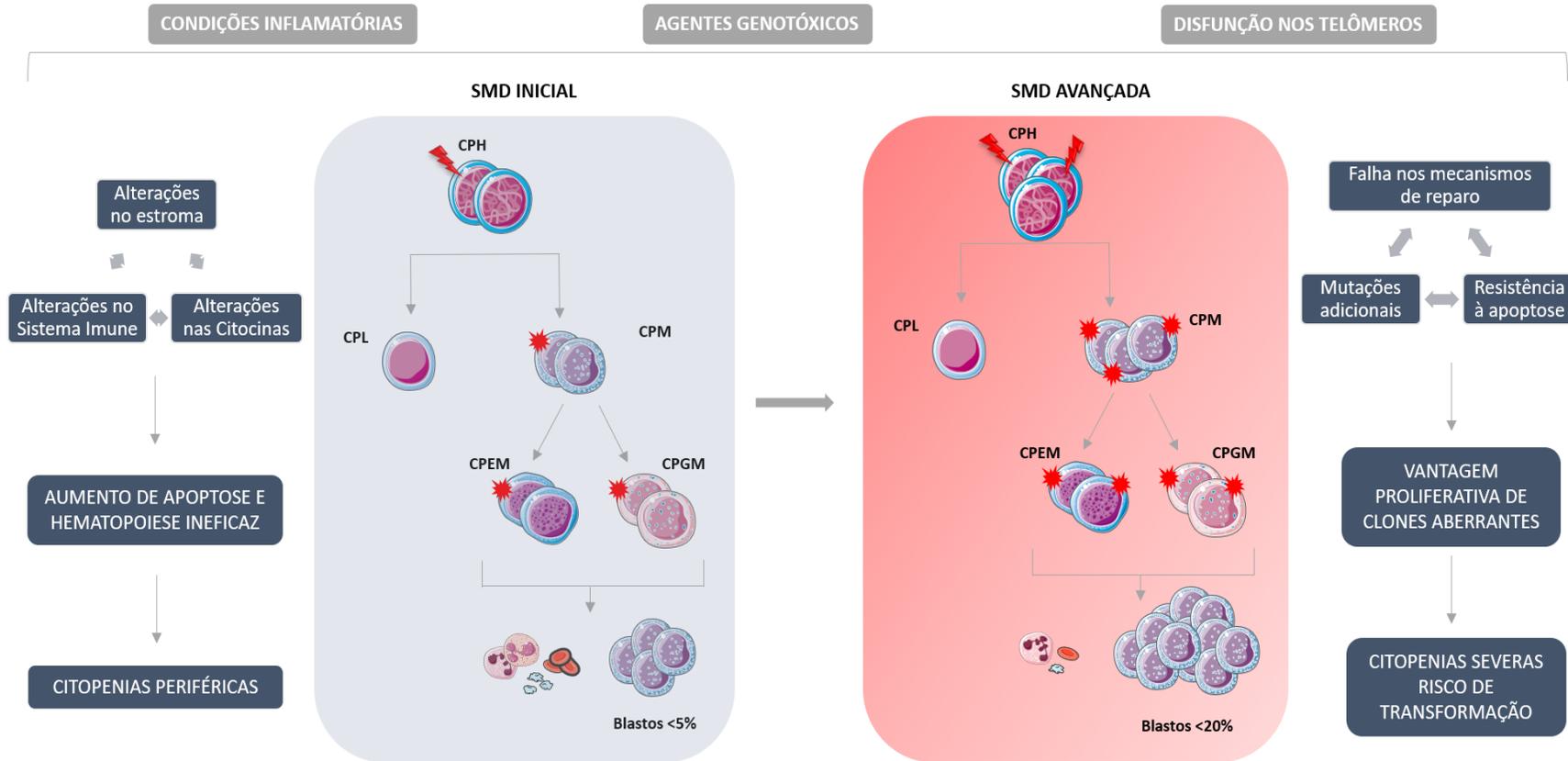
No modelo de múltiplos passos para a patogênese da SMD, inicialmente, vê-se que condições inflamatórias do microambiente medular decorrente da senescência provocariam alterações no estroma, citocinas e sistema imune (NAEIM; RAO; GRODY, 2008). Associados a essas alterações, agentes genotóxicos causariam lesão ao DNA das células tronco hematopoéticas (CTH) levando ao aparecimento de um clone anormal precursor de células hematopoéticas disfuncionais e morfologicamente displásicas (NAEIM; RAO; GRODY, 2008). O acúmulo dessas lesões em mecanismos de controle do ciclo celular, transcrição de

genes supressores tumorais, reparo do DNA ou mecanismos epigenéticos podem levar à dominância das CTH anormais mutadas (Figura 1).

A aquisição de uma ou múltiplas mutações genéticas nas CTH podem causar falhas na diferenciação, enquanto defeitos subsequentes, possivelmente em oncogenes, causam o bloqueio da diferenciação, estimulam a proliferação e afetam a susceptibilidade à apoptose de progenitores eritróides e mielóides, causando a expansão do clone anormal. Esse quadro configura a hematopoese ineficaz e o fenótipo clínico característico da doença, que leva a citopenias no sangue periférico apesar da medula óssea hipercelular (OLNEY; LE BEAU, 2002). Quando a SMD evolui para um estágio avançado ocorre a diminuição da apoptose, promovida por falhas nos mecanismos de reparo, pelo bloqueio da diferenciação e o acúmulo de mutações genéticas subclonais que garantem vantagem proliferativa significativa à população mutante sobre o tipo selvagem, o que pode levar à evolução da doença para LMA (Figura 1).

Na presença de um mecanismo de reparo totalmente operante que responda e conserte as lesões ao DNA causadas por agentes genotóxicos ou por processos endógenos, são poucas as chances de que as mutações que resultam em diferenciação anormal ou levam ao aumento da proliferação, ocorram dentro da mesma célula. No entanto, essa probabilidade e, portanto, o risco de desenvolver SMD ou LMA, pode aumentar drasticamente quando a maquinaria de reparo é deficiente. Alternativamente, o uso inadequado de um mecanismo de reparo de DNA ou a presença de um mecanismo de tolerância a danos de baixa fidelidade, pior ainda, a total ausência dos dois mecanismos de resposta à danos, também pode aumentar o risco de acumular múltiplas mutações no DNA na mesma célula (ZHOU et al, 2015), levando ao fenótipo da SMD.

**Figura 1:** Representação esquemática da patogênese e evolução da Neoplasia Mielodisplásica



**Legenda:** Danos no DNA e alterações epigenéticas são causados pela exposição prolongada a condições inflamatórias, a agentes genotóxicos e/ou disfunção dos telômeros, causada pelo envelhecimento e senescência celular. Tais lesões e alterações tem como consequência mudanças funcionais nas células tronco hematopoéticas (CTH), o que pode levar ao desenvolvimento de SMD. Na SMD inicial, há um aumento da apoptose e hematopoiese ineficaz, causadas por alterações no microambiente medular, na secreção de citocinas e no sistema imune inato. Essas mudanças desencadeiam a supressão das células saudáveis, levando à presença de citopenias periféricas. Na SMD avançada, ocorrem falhas nos mecanismos de reparo, mutações adicionais na CTH e, além disso, o clone displásico adquire resistência à apoptose e vantagem proliferativa, o que traz risco aumentado de transformação para LMA. **Abreviações:** CPH: Célula progenitora Hematopoética; CPL: célula precursora linfóide; CPM: célula precursora mieloide; CPEM: Célula Precursora de Eritrócitos e Megacariócitos; CPGM: Célula Precursora de Granulócitos e Monócitos. **Fonte:** Adaptado de Elias et al (2013) e Chamseddine et al (2016).

#### **1.1.4. O papel das alterações cromossômicas na patogênese da SMD**

A citogenética é a mais importante variável na determinação do prognóstico da SMD (BERNASCONI et al., 2013). Aproximadamente 40-60% dos pacientes com SMD *de novo* e até 80% dos pacientes com SMD secundária, em que o desenvolvimento da doença ocorre após um período prévio de exposição à radio e/ou quimioterápicos, apresentam alterações cromossômicas (CHAMSEDDINE et al., 2016; GREENBERG et al., 1997, 2012; SWERDLOW et al., 2016)

Ao contrário de outras neoplasias hematológicas caracterizadas por alterações cromossômicas equilibradas, tais como translocações recíprocas e inversões ocorridas em LMA, as quais resultam em mutações dominantes e ativação de oncogenes, a SMD está geralmente associada a alterações cromossômicas não equilibradas (FEARON et al., 2002). Muitas das alterações cromossômicas recorrentes na SMD levam à perda de material genético e conseqüente inativação de genes supressores tumorais (FEARON et al., 2002), sendo as alterações cromossômicas mais frequentes, a deleção dos cromossomos 5 e 7, a trissomia do cromossomo 8 e a nulissomia do cromossomo Y. (OLNEY; LE BEAU, 2002).

A maioria dessas anormalidades cromossômicas representam eventos genéticos secundários resultantes da instabilidade genômica intrínseca à doença ao invés de serem parte da patogênese primária, com exceção da del(5q), a qual apresenta relação direta com a base fisiopatológica de um subgrupo, a SMD com del(5q) isolada. A fisiopatologia do fenótipo da SMD com del(5q) isolada é associada à haploinsuficiência dos genes *RPS14* e *CSNK1A1*, localizados na região deletada do cromossomo 5, provavelmente causada pela ação da proteína p53. Os pacientes com del(5q) isolada possuem uma média de sobrevida alta com menor risco de evolução para LMA, além de responderem ao tratamento com lenalidomida, o qual diminui o clone aberrante, porém, a presença de mutações no gene *TP53* aumenta a resistência ao tratamento (JAN; SPERLING; EBERT, 2021; SHAHRABI et al., 2016; SWERDLOW et al., 2017b).

#### **1.1.5 O papel das mutações na patogênese da SMD**

No final dos anos 80, Hirai e colaboradores (1988) publicaram o primeiro relato de uma mutação somática em pacientes com Neoplasia Mielodisplásica. Foi identificada uma mutação pontual no códon 13 responsável pela ativação do oncogene *NRAS* em três pacientes com SMD. Partindo da observação de que os três pacientes evoluíram para LMA após um ano de seguimento, os pesquisadores sugeriram a participação do gene *NRAS* na progressão e

evolução da doença. Daí em diante, diversos estudos passaram a realizar pesquisas envolvendo mutações nos oncogenes da família RAS em pacientes com Síndrome SMD. Estas pesquisas deram o pontapé inicial nos estudos envolvendo mutações nesta doença a como fator primordial na sua patogênese. A descoberta de mutações nos genes da família RAS foram seguidas, nos 10 anos subsequentes, por evidências de mutações nos genes *FMS* e *TP53* e associações dessas mutações com a sobrevida dos pacientes. Iniciavam-se, assim, as associações entre mutações no DNA de pacientes com SMD com sua etiologia, patogênese, diagnóstico e prognóstico (HIRAI, 1988)

O advento do Sequenciamento de Nova Geração (do inglês, *Next Generation Sequencing - NGS*) revolucionou o manejo dos pacientes com SMD e proporcionou novos conhecimentos acerca das bases genéticas e moleculares de sua patogênese, permitindo a descoberta das principais mutações que estariam envolvidas com o desenvolvimento e progressão da doença. No início dos anos 2000, três estudos foram marcantes nas pesquisas com alterações moleculares na SMD, com a identificação de mais de 40 mutações ocorrendo em genes de diferentes vias de sinalização identificadas de forma recorrente na SMD. Bejar e colaboradores (2011) avaliaram um painel de 111 genes em 439 pacientes e teve como resultado a detecção de mutações em 18 genes, com 51% dos pacientes apresentando pelo menos uma mutação pontual. No ano de 2013, Papaemmanuil e colaboradores avaliaram um painel também de 111 genes em 738 pacientes, identificando 2260 SNVs (variante de nucleotídeo único, do inglês - *single nucleotide variant*) em 43 genes. Um ano após a avaliação feita por Papaemmanuil (2013), Haferlach e colaboradores (2014) desenvolveram um painel com 104 genes, avaliado em 944 pacientes e identificaram 2764 variantes, dentre SNV's e indels (inserções e deleções) em 96 dos 104 genes. Estas pesquisas de *screening* genético permitiram as primeiras identificações de mutações que são, hoje, biomarcadores na doença, e possuem grande relevância clínica, impactando diretamente o prognóstico e a sobrevida dos pacientes (BEJAR et al, 2011; PAPAEMMANUIL et al, 2013; HAFERLACH et al, 2014).

Mais de 20 anos após os primeiros estudos com mutações em SMD, nosso grupo de pesquisa teve o prazer de participar da construção de um sistema de prognóstico que soma as mutações oncogênicas à estratificação de risco dos pacientes. Elsa e colaboradores (2022) publicaram o *International Prognostic Scoring System-Molecular (IPSS-M)*, um método prognóstico validado que incorpora mutações somáticas aos dados clínicos e citogenéticos para melhorar a estratificação de risco dos pacientes com SMD, com o desenvolvimento, inclusive, de uma calculadora de acesso aberto para uso na prática clínica. Até então, a estratificação de risco e a tomada de decisão terapêutica na SMD eram baseadas no *International Prognostic*

*Scoring System –Revised* (IPSS-R), que considera parâmetros hematológicos e alterações citogenéticas (ELSA et al, 2022)

Resumidamente, o IPSS-M incorpora as seguintes informações moleculares à patogênese e estratificação de risco da SMD:

- Em 94% dos casos há, pelo menos, uma mutação oncogênica;
- Cada paciente apresenta, em média, a presença de quatro mutações oncogênicas;
- 121 genes se encontram recorrentemente mutados;
- Existem 9.254 mutações oncogênicas na SMD;
- A severidade da SMD se relaciona proporcionalmente com o número de mutações oncogênicas;
- A perda de heterozigose é um fenômeno frequente no genoma de pacientes com SMD;
- Pacientes com duas ou mais mutações no *TP53* (fenômeno multi-hit) tem um prognóstico sombrio, com menor sobrevida comparado aos pacientes com apenas uma mutação.
- Comparado com o IPSS-R, o IPSS-M melhorou a discriminação prognóstica em todos os desfechos clínicos e reestratificou 46% dos pacientes;

O IPSS-M se torna hoje ferramenta fundamental para diagnóstico, prognóstico, desenvolvimento e validação de estratégias de tratamento para pacientes com SMD. Um estudo de validação do IPSS-M publicado por Sauta e colaboradores (2023) demonstrou que houve mudança no grupo de risco em 46% dos pacientes avaliados, comparado com a estratificação baseada apenas no IPSS-R (SAUTA et al, 2023). Dessa forma, os dados moleculares mostram o grande benefício de direcionar a pesquisa de mutações usando NGS para prognóstico e estratificação de risco do paciente com SMD.

#### ***1.1.5. Classificação Clínica da Neoplasia Mielodisplásica***

O termo “Síndrome Mielodisplásica” foi formalmente introduzido em 1982 em uma revisão da classificação Franco-Americano-Britânica (FAB) publicada em 1976 (BENNET et al., 1976). Desde então, diversos sistemas de classificação têm sido propostos com o intuito de auxiliar no manejo do paciente, decisão terapêutica, estimar sua sobrevida e risco de transformação para LMA.

No intuito de incluir informações emergentes e aperfeiçoar o valor prognóstico da doença, em 2001, 2008, 2016 e 2022 a OMS vem aperfeiçoando a classificação inicial da doença, sempre introduzindo características biológicas e genéticas para definir o significado clínico de cada entidade patológica que compõe a SMD: clínica, morfologia, citogenética e alterações moleculares (VARDIMAN et al, 2008, SWERDLOW et al, 2016; KHOURY et al, 2022).

No ano de 2016, foi dado o primeiro passo na nova era de classificação da doença: a era molecular. A OMS revisou novamente sua classificação e encorajou a pesquisa de mutação no gene *SF3B1*, por apresentar forte relação com a porcentagem de sideroblastos em anel na MO, definindo a categoria SMD-SA. Esta foi a única mutação adicionada à classificação da OMS de 2016, a qual é definida pela presença de citopenias ao diagnóstico, quantidade de linhagens mielóides atingidas por displasias, porcentagem de blastos na MO e/ou no sangue periférico e presença de sideroblastos em anel. Apenas uma alteração citogenética é definidora de subtipo, a SMD com del(5q) isolada (ARBER et al, 2016).

O risco de evolução para LMA varia de acordo com a classificação do paciente, com subtipos apresentando baixo risco de evolução, um curso clínico que está associado a maior sobrevida, enquanto outros subtipos tem características clínicas mais graves, como o aumento de blastos e apresentam maior risco de evolução, com menor sobrevida. Baseado nos dados de sobrevida e evolução para LMA, os subgrupos da SMD classificados pela OMS 2016 podem ser divididos em baixo risco (SMD com sideroblastos em anel) e alto risco de evolução para LMA (SMD com excesso de blastos) (ARBER et al., 2016).

A classificação mais atual da OMS, publicada em 2022, deu ênfase às alterações genéticas e trouxe grandes avanços incorporando mutações oncogênicas na estratificação de subtipos da doença. Resumidamente, esta nova classificação dividiu a SMD em dois grandes grupos principais: SMD com anormalidades genéticas definidoras e SMD definida pela morfologia, em que os critérios morfológicos foram padronizados e o número de subtipos reduzido (KHOURY et al, 2022).

Dentro do grupo de SMD com anormalidades genéticas definidoras, destacamos o subtipo com inativação bialélica de *TP53*, definido pela presença de mutações multi-hit neste gene. Adicionalmente, a presença de uma mutação no gene *SF3B1* e baixa porcentagem de blastos na MO substitui a entidade anterior de SMD com sideroblastos em anel (MDS-SA) e é incorporado como um novo subtipo, SMD-SF3B1.

No entanto, a incorporação das pesquisas moleculares à prática científica ou à prática clínica ainda é uma realidade distante para países da América Latina, como o Brasil,

dificultando o uso da classificação da OMS de 2022 nestes pacientes. Dessa forma, no estudo em questão, utilizamos a classificação da OMS de 2016 por não ser factível a realização de testes genéticos adicionais nestes pacientes. Os critérios de classificação da OMS de 2016 podem ser consultados na tabela abaixo:

**Tabela 1:** Classificação clínica da SMD segundo os critérios da OMS 2016

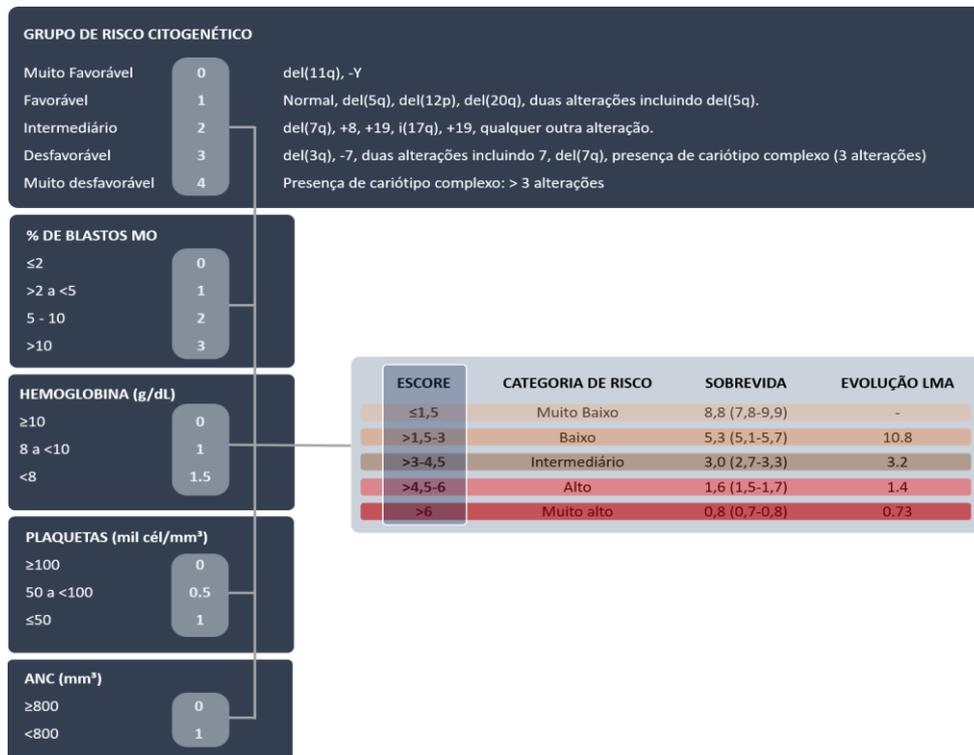
SUBTIPO DE SMD	SANGUE PERIFÉRICO	MEDULA ÓSSEA
(SMD-DU)	$\leq 1\%$ de Blastos; Uni ou Bicitopenia	Displasia em 1 linhagem; $5\%^*$ de Sideroblastos em anel; $\leq 5\%$ de blastos
(SMD-DM)	Citopenia em 1 ou mais linhagens; $\leq 1\%$ de blastos	Displasia em 2 ou 3 linhagens; $5\%$ de Sideroblastos em anel*; $\leq 5\%$ de blastos
<b>SMD COM SIDEROBLASTOS EM ANEL (SMD-SA)</b>		
(SMD-SA-DU)	$\leq 1\%$ de Blastos; Uni ou Bicitopenia	Displasia em 1 linhagem; $\geq 15\%$ ou $\geq 5\%^*$ de Sideroblastos em anel; $\leq 5\%$ de blastos
(SMD-SA-DM)	Citopenia em 1 ou mais linhagens; $\leq 1\%$ de blastos	Displasia em 2 ou 3 linhagens; $\geq 15\%$ ou $\geq 5\%^*$ de Sideroblastos em anel; $\leq 5\%$ de blastos
<b>SMD com Excesso de Blastos (SMD-EB)</b>		
(SMD-EB1)	Citopenia em 1 ou mais linhagens; 2 a 4% de blastos	Displasia em 0,1, 2 ou 3 linhagens; 5 a 9% de blastos
(SMD-EB2)	Citopenia em 1 ou mais linhagens; 5 a 19% de blastos	Displasia em 0,1, 2 ou 3 linhagens; 10 a 19% de blastos
<b>SMD inclassificável (SMD-i)</b>		
SMD-i com $1\%$ de blastos	Citopenia em 1 ou mais linhagens; $\leq 1\%'$ de blastos	Displasia em 1 ou mais linhagens; $\leq 5\%$ de blastos
SMD-i com displasia em única linhagem e pancitopenia	Citopenia em 3 linhagens; $\leq 1\%$ de blastos	Displasia em 1 linhagem; $\leq 5\%$ de blastos
SMD-i baseada em alterações citogenéticas	Citopenia em 1 ou mais linhagens; $\leq 1\%$ de blastos	Ausência de displasias; $\leq 15\%$ de Sideroblastos em anel; $\leq 5\%$ de blastos
SMD com del(5q) isolada	$\leq 1\%$ de Blastos; Uni ou Bicitopenia	Displasia em 1 ou mais linhagens; $\leq 5\%$ de blastos; del(5q) isolada ou com 1 alteração adicional, exceto del(7q)-7

**Legenda:** SMD-DU (SMD com displasia unilinhagem), SMD-DM (SMD com displasia multilinhagem), SMD-AS (SMD com sideroblastos em anel), SMD-SA-DU (SMD com sideroblastos em anel e displasia unilinhagem), SMD-SA-DM (SMD com sideroblastos em anel e displasia multilinhagem), SMD-EB (SMD com excesso de blastos).

### 1.1.6. Estratificação Prognóstica da Neoplasia Mielodisplásica

Por ser uma neoplasia de natureza extremamente heterogênea, se tornou necessário o desenvolvimento de ferramentas que possibilitem uma avaliação prognóstica mais acurada com relação a sobrevida e evolução para LMA, a fim de auxiliar nas decisões clínicas e terapêuticas. Em 1997, Greenberg e colaboradores criaram o Índice Internacional de Escore Prognóstico (IPSS) para SMD, um tipo de estratificação prognóstica que permite prever o risco de transformação leucêmica assim como a mediana de sobrevida esperada para cada paciente (GREENBERG et al., 1997). Este sistema de predição de prognóstico teve a colaboração da Prof. Dra. Sílvia Magalhães, coordenadora do nosso grupo de pesquisa. Em 2012 foi publicada a revisão do IPSS (IPSS-R), fazendo um refinamento nos parâmetros, sendo ainda hoje o principal sistema prognóstico utilizado na SMD. As variáveis utilizadas no IPSS-R são: alterações citogenéticas, porcentagem de blastos na medula óssea, valores de hemoglobina, plaquetas e contagem absoluta de neutrófilos (ANC) e após análise multivariada com tempo de sobrevida e transformação para LMA, foram criadas 5 categorias de risco: muito baixo, baixo, intermediário, alto e muito alto risco (Figura 2) (GREENBERG et al., 2012). A figura 2 representa os grupos de risco definidos pelo IPSS-R.

**Figura 2:** Categorias de risco e escores prognósticos segundo o IPSS-R



Fonte: Adaptado de Bejar (2014).

## 1.2. Lesões ao DNA e Mecanismos de resposta aos danos

Estima-se que mais de 10.000 lesões de fita simples no DNA sejam produzidas diariamente (HOEIJMAKERS, 2009). Essas lesões podem ser causadas por moléculas reativas endogenamente produzidas ou por agentes ambientais, como luz ultravioleta, radiação ionizante, metais pesados, quimioterápicos, formol e venenos agrícolas (LINDAHL, 1993; HWA YUN, 2018). Quando não reparadas, essas lesões podem levar a mutações pontuais ou alterações cromossômicas em larga escala que ameaçam a viabilidade celular. Por exemplo, a luz UV e o Raio-X induzem dímeros de pirimidina e quebras de fita simples e dupla na hélice do DNA, causando alterações na replicação e transcrição de genes, falhas reprodutivas, mutações e morte celular (OZTURK e DEMIR, 2011). Os efeitos a longo prazo resultantes dessas falhas e mutações contribuem para o envelhecimento celular, para doenças relacionadas à senescência e neurodegeneração e para a carcinogênese (IYAMA e WILSON, 2013; LANGIE et al., 2015).

Os estudos no campo de respostas a danos no DNA se iniciaram de forma indireta na primeira metade do século XX, nos Estados Unidos da América (EUA) (FRIEDBERG, 2008). O aumento das pesquisas para confecção de armas atômicas trouxe uma maior atenção da comunidade científica para os efeitos da radiação em organismos biológicos. No começo da década de 1940 já se tornava evidente que agentes que levam a alterações mutagênicas (tais como radiação ionizante e ultravioleta) são capazes de interagir e causar danos no material genético das células. Em adição a isso, emergiam evidências que sugeriam que organismos vivos eram capazes de se recuperar dos efeitos letais destes agentes (HOLLAENDER & CURTIS, 1935).

Com o passar dos anos, diversos mecanismos de vigilância, reparo e tolerância aos danos no DNA evoluíram com o intuito de impedir suas consequências deletérias, detectando e sinalizando a presença de lesões e promovendo seu reparo. Existem situações em que as lesões no DNA não podem ser totalmente reparadas e as células são forçadas a tolerar esses danos. Assim, a resposta aos danos no DNA pode resultar em reparo ou em tolerância aos danos e podem diferir quanto as várias classes de lesões, mas todas tem o mesmo objetivo: manutenção da estabilidade genômica (Figura 5) (LINDAHL et al, 1999).

## 1.3. Mecanismos de Reparo do DNA

Para neutralizar lesões no DNA, a célula coloca em funcionamento uma rede coordenada de cascata de sinalização denominada *DNA Damage Response* (DDR). A DDR

detecta a lesão no DNA e transmite um sinal para ativar o controle do ponto de verificação do ciclo celular a fim de retardar a progressão do ciclo celular e permitir que o dano seja reparado pelo recrutamento de mecanismos de reparo específicos para cada tipo de dano (GHOSAL, 2013).

Existem cinco vias principais de reparo de danos que operam em mamíferos, com variações em cada uma delas, dependendo do tipo de lesão: reparo por excisão de nucleotídeos (NER), reparo por excisão de bases (BER), o reparo de erros de emparelhamento de bases (MMR), que são mecanismos de reparo de fita simples e a recombinação homóloga (HR) e junção por extremidades não-homólogas (NHEJ) que figuram entre os mecanismos de reparo de fita dupla de DNA (Figura 3) (IYAMA e WILSON, 2013).

Os mecanismos atuantes no reparo de fita simples por excisão consistem na remoção de uma região danificada por nucleases especializadas e, em seguida, para preencher a lacuna, a síntese de DNA é realizada por uma DNA polimerase que utiliza a fita não-danificada como molde. Estes mecanismos reparam os dímeros timina-timina, o tipo mais comum de dano causado pela luz UV, que interferem tanto na replicação quanto na transcrição do DNA. O reparo por excisão também pode corrigir regiões de DNA contendo bases alteradas pela inserção de grandes grupos químicos, por exemplo, carcinógenos como benzo-pireno. A NER lida com a ampla classe de lesões exógenas que distorcem a hélice, interferem no pareamento de bases e geralmente obstruem a transcrição e a replicação normal. Por sua vez, pequenas alterações químicas de bases, que podem ou não impedir a transcrição e a replicação, são alvos de BER. As lesões reparadas por BER estão principalmente, mas não exclusivamente, relacionadas a danos de origem endógena (KAO et al., 2005; IYAMA e WILSON, 2013). Já o mecanismo MMR atua principalmente no reconhecimento e reparação de erros nas inserções de bases e presença de deleções incorporadas no DNA que podem surgir durante a replicação e na recombinação do DNA (IYAMA e WILSON, 2013).

As lesões de fita dupla são as mais deletérias formas de dano ao DNA, uma vez que as ambas as fitas da molécula são afetadas e podem contribuir para a formação de anormalidades cromossômicas, morte celular e transformação neoplásica (BOHGAKI; HAKEM, 2010; IYAMA e WILSON, 2013). Surgem pela ação de espécies reativas de oxigênio (ROS) ou induzidos por fontes exógenas como terapia antineoplásica, radiação ionizante e produtos químicos genotóxicos (IYAMA e WILSON, 2013).

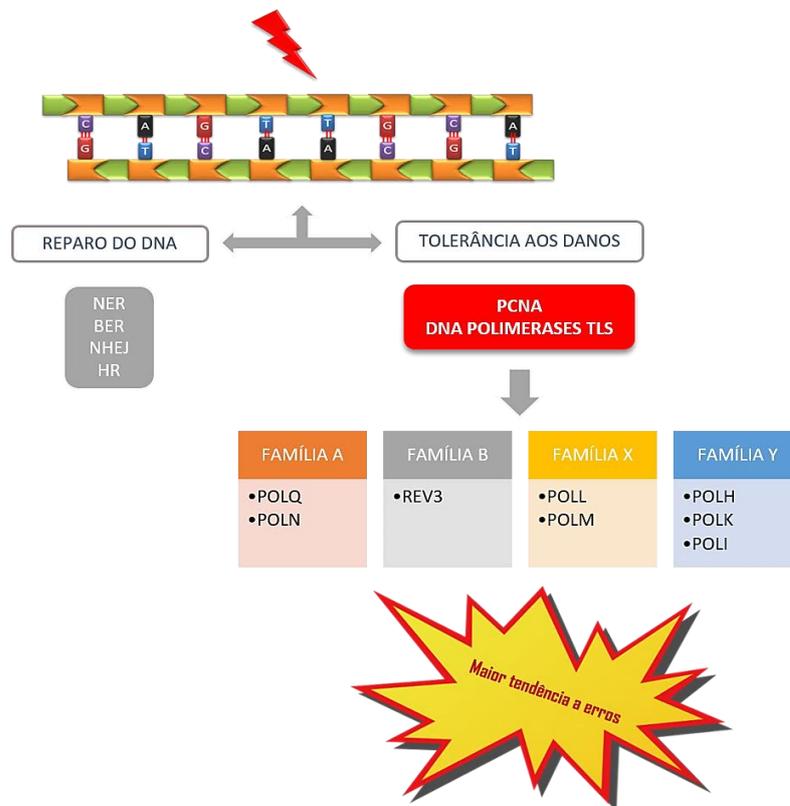
A via HR requer uma cromátide-irmã homóloga para sua execução, ao passo que a via NHEJ pode ocorrer em células que estejam ou não em divisão, não necessitando de uma sequência homóloga de controle, o que faz dela uma via com maiores chances de erro

(OZTURK e DEMIR, 2011). Diversas doenças humanas derivam-se de deficiências em HR e NHEJ e exibem defeitos no desenvolvimento, imunológicos e neurológicos, assim como sensibilidade à radiação, fenótipos de envelhecimento precoce e predisposição ao câncer (IYAMA e WILSON, 2013).

No estudo de Ribeiro e colaboradores (2013), o qual compõe o nosso grupo de pesquisa, foi demonstrado que variações polimórficas do gene *ATM* (*Ataxia Telangiectasia Mutada* – gene com importante atividade de sinalização de lesão no DNA) são associadas ao desenvolvimento de SMD de baixo risco e que polimorfismos dos genes de reparo do DNA são associados à patogênese da Neoplasia Mielodisplásica (RIBEIRO, 2013).

Uma vez reparado o dano, a célula volta a entrar no ciclo celular, o genoma é duplicado de uma maneira livre de erros por DNA polimerases de alta fidelidade e transmitido para as células filhas. Durante a fase S, quando o DNA está sendo replicado, se as DNA polimerases de alta fidelidade encontram lesões volumosas ou distorções na estrutura secundária da molécula de DNA elas param a replicação, uma vez que seus sítios ativos são incapazes de acomodar esse tipo de lesão. Isso resulta na imobilização da forquilha de replicação e desacoplamento das polimerases replicativas. Essa prolongada estagnação pode levar ao colapso da forquilha de replicação e, em última análise, à instabilidade do genoma. Diversas enzimas atuam nesse ponto à fim de reestabilizar a forquilha e reiniciar a replicação. Assim, uma segunda resposta ao DNA danificado, denominado *DNA Damage Tolerance* (DDT) atua para promover a replicação através e além de um modelo alterado, deixando que o dano seja reparado em um momento posterior, reduzindo assim o risco geral de colapso da forquilha de replicação e instabilidade do genoma (figura 3) (POWERS, 2018; GHOSAL, 2013; ANDERSEN, 2008).

**Figura 3:** Representação esquemática dos Mecanismos de Reparo e Tolerância a danos no DNA



**Legenda:** NER: Reparo por excisão de nucleotídeos; BER: Reparo por excisão de bases; NHEJ: junção por extremidades não-homólogas; HR: Recombinação Homóloga; PCNA: Antígeno Nuclear de Proliferação Celular; TLS: Translesão; POLQ: DNA Polimerase theta; POLN: DNA Polimerase nu; POLL: DNA Polimerase lambda; POLI: DNA Polimerase iota; POLH: DNA Polimerase eta; POLK: DNA Polimerase kappa; REV3L: DNA Directed Polymerase Zeta Catalytic Subunit; REV1: DNA Directed Polymerase.

**Fonte:** Elaborada pelo autor

#### 1.4. Mecanismos de tolerância a danos no DNA: Síntese Translesão

As DNA polimerases replicativas são particularmente específicas para pares de bases de DNA normais, mas seu sítio ativo não é capaz de acomodar bases alteradas ou com presença de lesões volumosas. Para resolver este impasse, as células dispõem de um importante mecanismo para promover a passagem através destes danos: a síntese Translesão (Figura 2). Este mecanismo envolve a atividade de DNA polimerases especializadas capazes de sintetizar o DNA através da lesão (LIVNEH, 2010; POWERS e WASHINGTON, 2018). Essas polimerases não replicativas com atividade translesão (TLS) são uma classe de polimerases com baixa processabilidade e baixa fidelidade, mas que possuem grande flexibilidade conformacional, com um sítio catalítico maior que as permitem moldar-se sobre distorções no DNA. Essas enzimas replicam o DNA danificado de uma maneira livre de erros (quando fazem o pareamento correto de bases), de uma maneira propensa a erros (quando fazem o pareamento

incorreto de bases) ou transpassam a lesão sem removê-la e continuam a replicar um curto trecho do DNA evitando o colapso da forquilha de replicação e a geração de instabilidade cromossômica (Figura 4). (CHANG, 2009; FRIEDBERG, 2002; LIVNEH, 2010)

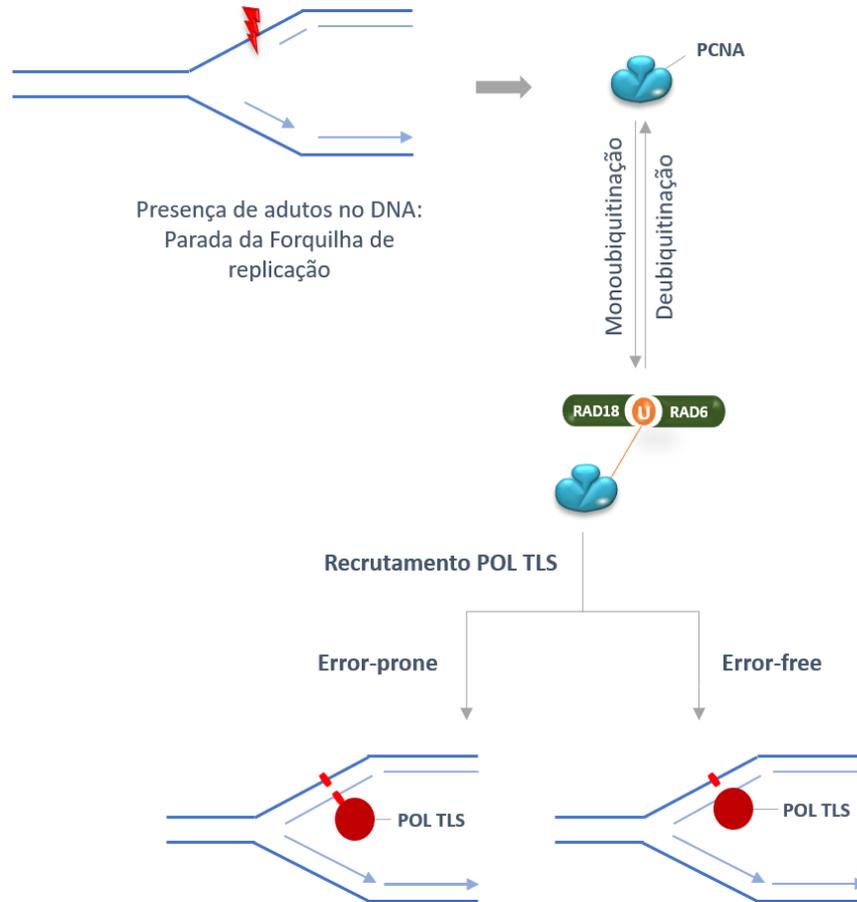
A ativação e regulação do mecanismo TLS é mediada em 3 fases: (1) Acesso ao DNA. Como as polimerases TLS são mutagênicas, seu acesso ao DNA deve ser restrito ao local da lesão. (2) A seleção de uma DNA polimerase específica para sua lesão cognata e (3) Troca ou não da Polimerase TLS por outra no momento da extensão da fita. Essas regulações são intermediadas por modificações do antígeno nuclear de proliferação celular (PCNA) (DAIGAKU, 2010; ZHOU, 2017), que, após a parada na forquilha de replicação, é mono ou poliubiquitinada pelo complexo RAD6-RAD18 na lisina 164 (Figura 4). (DAIGAKU, 2010; KANNOUCHE, 2004; BIENKO, 2005; ANDERSEN, 2008; FOX, 2011; MCINTYRE, 2015; KANAOKA, 2017) (Figuras 4 e 5).

O modelo atual para o processo TLS envolve pelo menos dois eventos de troca de polimerase. Na primeira troca, a DNA polimerase replicativa bloqueada é substituída por uma polimerase TLS capaz de replicar sobre a lesão do DNA. O segmento de fita adjacente à lesão é então estendido pela mesma ou outra polimerase TLS. A etapa de extensão permite que a lesão escape da detecção da atividade de revisão exonucleásica 3' a 5' da DNA polimerase replicativa. Isto é seguido pelo segundo interruptor, que restaura a DNA polimerase replicativa de alta fidelidade ao molde de DNA para retomar a replicação do DNA (ANDERSEN, 2008; ZHAO, 2017) (Figuras 4 e 5).

Uma vez que existe uma grande diversidade química de lesões ao DNA, cada uma dessas polimerases possui uma ou mais lesões que podem contornar com precisão e eficiência. Essas lesões são conhecidas como lesões cognatas das polimerases não replicativas ou polimerases TLS. A precisão e a eficiência do mecanismo TLS depende de qual das polimerases TLS disponíveis é escolhida para contornar o dano ao DNA. Se o dano for uma lesão cognata para a polimerase escolhida, a precisão e a eficiência da TLS serão relativamente altas. Se o dano não for uma lesão cognata da polimerase, a precisão e a eficiência serão muito menores (POWERS e WASHINGTON, 2018).

Dessa forma, além dos mecanismos de reparo do DNA, maquinaria celular dispõe dessas polimerases especializadas em corrigir ou transpassar as lesões, porém, por serem de baixa fidelidade, elas possuem maior tendência à incorporar erros, isto é, tem grande potencial mutagênico, potencial este que depende de alterações na sua ativação, proporcionadas por mudanças de expressão ou presença de mutações (ANDERSEN et al, 2008; BERTOLIN, MANSILLA, GOTTIFREDI, 2015)

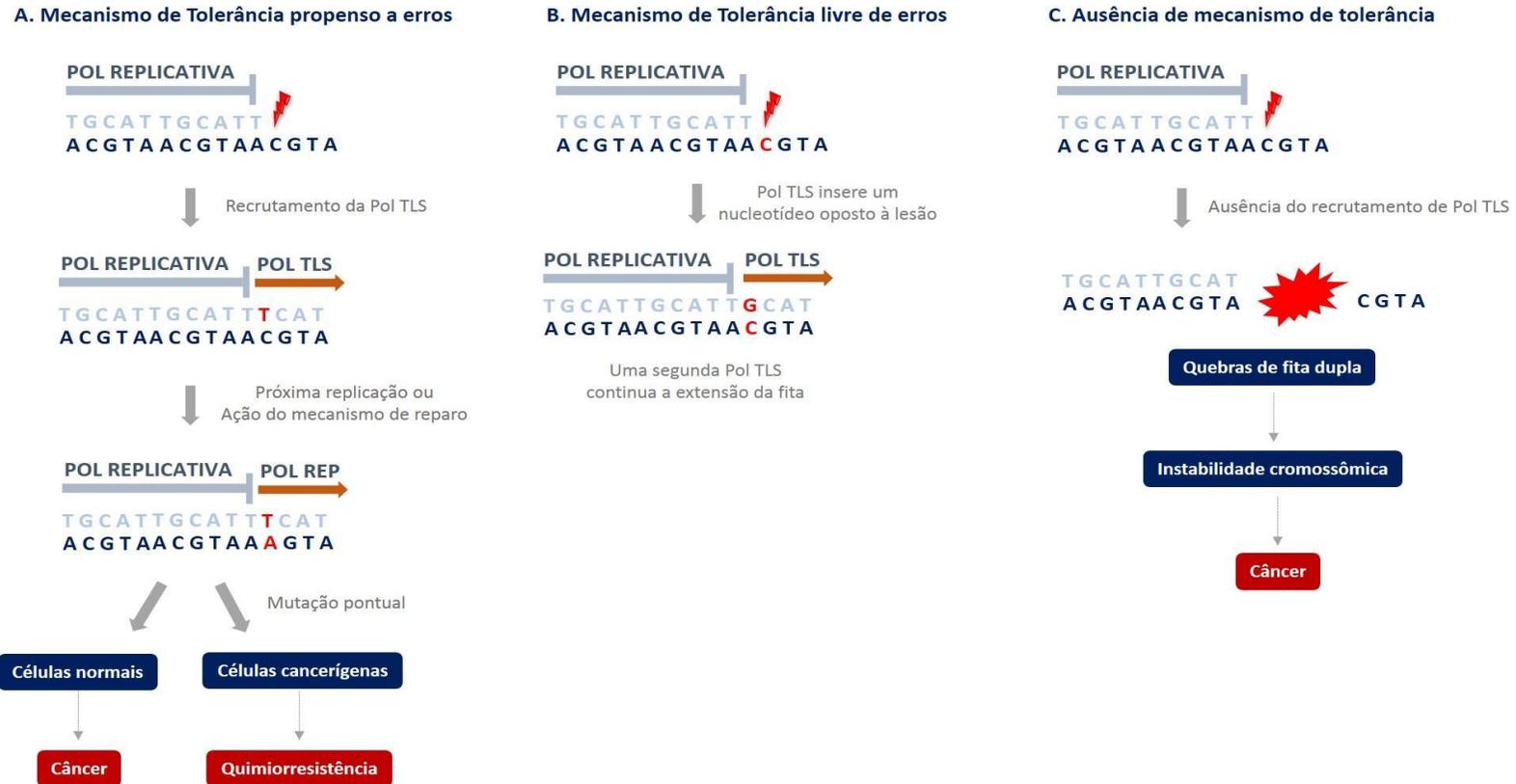
**Figura 4:** Representação esquemática da Ativação do Mecanismo Translesão



**Legenda:** Ao se deparar com lesões no DNA que causam o bloqueio da forquilha de replicação, a DNA Polimerase replicativa não consegue dar continuidade à replicação por não conseguir acomodar lesões volumosas em seu sítio ativo. Dessa forma, ocorre a ativação dos mecanismos de síntese translesão. A ubiquitinação da proteína PCNA pelo complexo RAD18-RAD6 na lisina 164 proporciona a troca da Polimerase replicativa pela DNA Polimerase TLS cognata da lesão presente na fita de DNA. A POL TLS realiza a síntese translesão por meio de um mecanismo que pode ser livre de erros (error-free) ou propenso a erros (error-prone).

**Fonte:** Elaborada pelo autor

**Figura 5:** Representação esquemática do Mecanismo de tolerância a danos no DNA através da síntese translesão e sua implicação na carcinogênese



**Legenda:** A presença de lesões no DNA bloqueia as forquilha de replicação e impede o seguimento da replicação. **A.** Após a parada da forquilha de replicação, uma polimerase TLS é recrutada para o ponto da lesão. No entanto, a POL TLS pode incorporar uma base errada a fim de continuar a replicação. Na próxima replicação, esse nucleotídeo que foi incorporado erroneamente dará origem a uma mutação pontual. Essa mutação pontual pode contribuir para o desenvolvimento de neoplasias ou, em células neoplásicas, garantir resistência à agentes terapêuticos. **B)** No mecanismo TLS livre de erros, a POL TLS incorpora uma base pareada de forma correta e uma mudança para uma segunda DNA polimerase especializada pode ocorrer para estender parte da fita. Após isso, ocorre a troca para uma POL Replicativa, dando continuidade à replicação. **C)** Na ausência completa de uma polimerase TLS, não ocorre desvio translesão, o que acarreta o colapso da forquilha de replicação, levando a quebras na dupla fita e instabilidade cromossômica, que podem contribuir para a carcinogênese.

**Fonte:** Adaptado de Lange et al (2010)

#### **1.4.1. A baixa fidelidade do mecanismo TLS: quando o relacionamento com o DNA pode falhar**

As paradas na forquilha de replicação são resolvidas com maior frequência usando o mecanismo livre de erros da translesão citado anteriormente. No entanto, a replicação pode prosseguir em um processo propenso a falhas. As células arriscam induzir mutações pela incorporação incorreta de bases causadas pelo processo TLS a fim de desbloquear a forquilha de replicação causada pelos adutos de DNA e evitar, assim, a instabilidade cromossômica por quebras de fita dupla e a morte celular. Assim, a solução de tolerância a danos geralmente vem às custas de uma maior taxa de erros de incorporação, uma vez que as DNA polimerases TLS são de baixa fidelidade e isso faz com que sejam propensas a cometer erros. Diante disso, esse processo acaba sendo responsável pela maioria das mutações pontuais induzidas por danos e, portanto, é particularmente relevante para a carcinogênese (CHANG, 2009; FRIEDBERG, 2015), tendo um papel etiológico na maioria dos cânceres induzidos por fatores ambientais (ADAR, 2009; SABINE, 2011; JANSEN, 2015).

#### **1.4.2. Famílias de DNA Polimerases**

A partir dos dados de análise filogenética e caracterização bioquímica publicados, existem, pelo menos, oito enzimas atuantes no mecanismo de síntese TLS em mamíferos, estando agrupadas em quatro famílias: família Y, que inclui Pol  $\eta$  (POLH), Pol  $\iota$  (POLI), Pol  $\kappa$  (POLK) e REV1. Família B, que inclui Pol  $\zeta$  (POLZ), cuja subunidade catalítica é REV3L. Família A, que inclui Pol  $\theta$  (POLQ) e Pol  $\nu$  (POLN) e família X, a qual inclui Pol  $\lambda$  (POLL) (Figura 5). Cada uma das polimerases TLS tem diferentes especificidades de substrato para diferentes tipos de danos no DNA (POWERS, 2018; LANGE, 2011; MAKRIDAKIS, 2012; GHOSAL, 2013).

##### **1.4.2.1. Família A: DNA Polimerases TLS $\theta$ (POLQ) e $\nu$ (POLN)**

Descrita pela primeira vez em 1999 por Sharief e colaboradores, *POLQ* está localizado no braço longo do cromossomo 3 (SHARIEF, 1999; MAGA, 2002; SEKI, 2003; SEKI, 2004) e codifica um polipeptídeo de 2.592 aminoácidos com uma estrutura bioquímica composta pela fusão gênica de um domínio polimerase na região C-terminal, um domínio helicase do tipo *Hel218* na região N-terminal e um terceiro domínio central com função ainda indeterminada. Embora a fusão gênica do tipo helicase-polimerase seja única para os eucariotos

superiores, as proteínas do tipo polimerase-helicase são multifuncionais, já foram previamente identificadas em bactérias e vírus e normalmente atuam no início da replicação, bem como no reparo do DNA e no mecanismo de tolerância a danos. O domínio polimerase permite a atuação de POLQ na síntese translesão, enquanto o domínio helicase aparenta possibilitar sua atuação nos mecanismos de reparo do DNA. Dentro da Família A, essa multifuncionalidade é exclusiva da enzima POLQ (BEAGAN, 2016; MALABY, 2017).

POLQ realiza a síntese translesão através de sítios apurínicos e apirimidínicos (AP), lesões causadas por espécies reativas de oxigênio, como sítios de timina glicol (Tg) e tem a capacidade de catalisar tanto a inserção de nucleotídeos opostos às lesões quanto as etapas de extensão (SEKI, 2004). Por outro lado, POLQ não consegue incorporar bases em sítios opostos a lesões causadas por radiação UV, sendo necessária a coparticipação de POLI para realizar a extensão (SEKI e WOOD, 2008). Dentre as polimerases da família A, a fidelidade durante a incorporação de dNTP's de POLQ é a mais baixa (YOUSEFZADEH, WOOD, 2013).

Camundongos com a mutação missense do tipo *chaos1* (fenótipo atribuído a transição de T>C que causa uma alteração de aminoácidos de serina para prolina no resíduo 1932 de POLQ) se desenvolvem normalmente, mas demonstram frequências elevadas de micronúcleos espontâneos e micronúcleos induzidos por radiação, indicando algum papel desta enzima na manutenção da instabilidade genômica nos eritroblastos (SHIMA, 2003). No entanto, a viabilidade de camundongos com defeitos em *POLQ* é severamente comprometida com a ocorrência de uma mutação adicional no gene *ATM*, demonstrando que estas enzimas tem relações de semiletalidade sintética, mas operam de formas distintas na embriogênese de mamíferos (SHIMA, 2004).

Adicionalmente, Goff e colaboradores (2009) realizaram experimentos com cultura de célula da medula óssea de camundongos do tipo *chaos 1*. Células Polq (-/-) apresentaram um fenótipo mais sensível à radiação gama do que as células Polq (+/+). Adicionalmente, células Polq (-/-) apresentaram moderada sensibilidade à bleomicina em comparação com células Polq (+/+) e não demonstraram hipersensibilidade ao paraquat ou ao peróxido de hidrogênio. Camundongos Polq (-/-) apresentaram mais reticulócitos micronucleados espontâneos e induzidos por radiação do que os camundongos Polq (+/+) e (+/-). A sensibilidade à radiação ionizante e bleomicina de células do estroma da medula óssea defeituosas para POLQ e o aumento de micronúcleos nas hemácias demonstram o papel dessa DNA polimerase na tolerância a danos no DNA que podem levar a quebras na dupla fita (GOFF, 2009).

Ainda sobre alterações em *POLQ* que promovem hipersensibilidade à agentes genotóxicos, um estudo com siRNA (RNA de interferência) identificou que o *knockdown* de

POLQ resulta na radiosensibilização de linhagens celulares tumorais, tendo pouco ou nenhum efeito nas linhagens celulares de tecidos normais. Esses achados levantam a possibilidade de que a inibição de *POLQ* possa ser usada clinicamente para causar radiosensibilização de células neoplásicas (HIGGINS, 2010).

A literatura mostra alterações de expressão do gene *POLQ* em câncer de pulmão, mama, estômago, colorretal e carcinoma de células escamosas bucal e associações dessas alterações com a diminuição da sobrevida nesses pacientes. O estudo de Lemée e colaboradores (2010) relatou níveis de expressão mais elevados em pacientes com tumores de câncer de mama do tipo triplo-negativo, um subtipo mais agressivo da doença, bem como outros pesquisadores demonstraram que altos níveis de expressão de *POLQ* levam ao atraso na progressão da forquilha de replicação e aumentam a quantidade de alterações cromossômicas e de mutações somáticas pontuais em câncer de mama (HIGGINS, 2010; LEMÉE, 2010; KAWAMURA, 2004; ALLERA-MOREAU, 2012; LESSA, 2013; CECCALDI ET AL., 2015). No entanto, são escassos na literatura os relatos de alterações neste gene, sejam de expressão ou de mutações, em neoplasias mieloides, não existindo, até o momento, nenhum estudo de sua funcionalidade na Neoplasia Mielodisplásica.

*POLN* está localizado na região cromossômica 4p16.2, codificando uma proteína de 900 resíduos de aminoácidos com uma massa molecular de 100 kDa. Diferente da enzima *POLQ*, *POLN* possui apenas o domínio DNA polimerase. Estudos *in vitro* demonstraram que a polimerase  $\nu$  é uma enzima altamente propensa a erros, no entanto passa eficientemente sítios de Tg. Geralmente sua atividade enzimática favorece a incorporação de Timina (T) quase tão frequentemente quanto Citosina (C) para o nucleotídeo Guanina (G), resultando em frequentes transições de GC para AT (ARANA, 2007). Quase 50% dos carcinomas de mama exibem mutações no gene *POLN* e sua baixa expressão tem sido correlacionada com a menor eficiência do mecanismo de recombinação homóloga (HR), demonstrando que deve existir correlação entre sua atividade translesão e esse mecanismo (MOLDOVAN et al, 2010).

#### 1.4.2.2. Família B: DNA Polimerase TLS $\zeta$ (*POLZ*) – *REV3L*

*POLZ* é uma holoenzima tetramérica, a qual possui uma subunidade catalítica *REV3L* e uma subunidade estrutural *REV7*, assim como duas subunidades acessórias, *POLD2/p50* e *POLD3/p66*, sendo a primeira enzima com atividade TLS identificada (BARANOVSKIY, 2012). É importante salientar que *POLZ* não possui atividade exonucleásica de revisão, 3'-5'. Sua subunidade catalítica *REV3L* localiza-se na região

cromossômica 6p21, sendo essa a posição do sítio frágil FRA6F, que se encontra deletada em diversos tipos de leucemias e tumores sólidos (MORELLI et al, 2002). Ademais de ser a polimerase mais propensa a causar erros devido a sua alta atividade de ultrapassar as lesões, POLZ e sua subunidade catalítica REV3L são de extrema importância na defesa contra agentes genotóxicos (MAKAROVA, 2015; SUZUKIA, 2016).

Em modelos murinos, camundongos deficientes de *REV3L* morrem ainda na embriogênese, que é acompanhada por apoptose de células dentro do embrião, demonstrando que a presença de mutações nessa polimerase, ainda no período embrionário, é letal para camundongos (O-WANG et al, 2002). Além disso, estudos demonstram aumento da expressão de *REV3L* após indução de danos no DNA de um modo dependente da proteína p53 e após tratamento das células com cisplatina (KRIEG, ET AL 2006; WU ET AL, 2004). Foi demonstrado que o *knockout* de *REV3L* concomitantemente com a deficiência de *TP53* em fibroblastos leva ao aumento da instabilidade cromossômica (WITTSCHIEBEN ET AL, 2010).

No tecido hematopoiético, camundongos com supressão do gene *REV3L* desenvolveram linfoma de células T, assim como também houve o aparecimento de tumores mamários. Foram identificadas também alterações pré-neoplásicas em tecidos glandulares adjacentes aos tumores mamários em camundongos com silenciamento do gene *REV3L*, o que sugere que sua deleção pode ocasionar a tumorigênese (WITTSCHIEBEN ET AL, 2010). Em humanos, a baixa fidelidade da *POLZ* tem como consequência a instabilidade cromossômica (LANGE ET AL., 2011; WITTSCHIEBEN ET AL., 2010).

#### 1.4.2.3. Família Y: DNA Polimerases TLS $\eta$ (*POLH*), $k$ (*POLK*), $\iota$ (*POLI*) e *REV1*

As DNA polimerases da família Y são extremamente conservados e apresentam particularidades, as quais podem estar presentes em apenas uma delas ou que pode ser compartilhada por mais de uma. A enzima *REV1*, por exemplo, possui uma extensão na porção N-terminal na qual se encontra o domínio BRCT. Este domínio é importante para a regulação da síntese translesão, já que ele se liga a proteínas fosforiladas por ATR e ATM, essenciais para a resposta a estresse durante a replicação celular (MANKE ET AL., 2003). Outras duas regiões conservadas de grande importância são os domínios de ligação ao PCNA e os motivos estruturais dedos de zinco, conhecidos por terem participação no acoplamento de uma proteína ao DNA.

REV1 é composta por 1251 aminoácidos, seu gene codificador está localizado no cromossomo 2 (2q11.2) e a proteína tem um peso molecular de 138kDa. Ela possui dois domínios de ligação à Ubiquitina, UBM1 e UBM2, um domínio C-terminal de ~100 aminoácidos e um domínio BRCT. Como a levedura Rev1, o REV1 humano possui uma atividade de dCMP transferase que, preferencialmente, insere uma Citosina oposta a uma Guanina. No entanto, o REV1 humano pode inserir uma citosina oposta a um sítio AP ou a uma uracila. (LIN, 1999). Apesar da atividade catalítica de REV1, tem sido demonstrando que sua função mais importante na síntese translesão é a reguladora e não a catalítica (LAWRENCE, 2004; SASATANI, 2017). Sua função reguladora é demonstrada em um modelo onde REV1 tem a função principal de alocar as polimerases TLS em regiões com danos no DNA e modular a mudança da polimerase no local da lesão. Assim, pensa-se que REV1 possa desempenhar um papel de recrutadora e reguladora, assim como PCNA (GUO, 2003).

Estudo demonstram que a diminuição dos níveis de expressão de *REV1* em modelos *in vitro* resulta em uma menor taxa de indução de mutações, indicando o potencial dessa via como alvo terapêutico para abrandar mutações genéticas (BOUDSOCQ et al., 2002). Mutações pontuais identificadas no gene *REV1* foram associadas a um risco aumentado de câncer cervical (HE et al., 2008), o que denota a necessidade de realizar mais investigações acerca do papel de REV1 na etiologia e na progressão do câncer.

A proteína POLH é composta por 713 resíduos de aminoácidos e seu gene está localizado no cromossomo 6. Estruturalmente, ela consiste em uma região de interação com REV1 (RIR), um domínio de ligação à ubiquitina (UBZ), um domínio de localização nuclear e regiões peptídicas de interação com PCNA (PIP) (KNOBEL E MARTI, 2011; UCHIYAMA, TERUNUMA, HANAOKA, 2015).

Até o momento, POLH é a única polimerase descrita cuja deficiência está associada a uma doença humana, o Xeroderma Pigmentosum Variante (XP-V). Em 1999, após a clonagem do cDNA do gene XPV/POLH, identificou-se seu produto gênico como sendo POLH. Esses pacientes representam aproximadamente 20% dos pacientes diagnosticados com XP e exibem um fenótipo clínico mais favorável do que os pacientes com XP clássico. Além disso, pacientes que apresentam mutações no gene POLH/XPV exibem maior incidência de tumores de pele (CORDONNIER, 1999; JOHNSON, 1999a; MASUTANI, 1999a).

Essa associação é evidenciada, também, pelo fato de que POLH é especializada na síntese translesão de danos causados pela luz UV, adicionando, na maioria dos casos, A-A (adeninas) opostas a fotoprodutos de dímeros de pirimidina de ciclobutano (CPDs) formados por T-T. Na ausência de POLH, outra polimerase TLS pode ser recrutada para fazer o bypass

(síntese translesão), dos CPDs mas de forma menos eficiente e menos precisa. Isso resulta no aumento da taxa de mutagênese induzida por UV e carcinogênese em células XP-V. (BIERTUMPFEL, 2010; MASUTANI, 2000).

Além da síntese translesão de CPDs, POLH também está envolvida no bypass dos 6,4 PPs (do inglês, 6-4 photoproducts). A diferença é que enquanto o primeiro mecanismo necessita apenas de uma polimerase, o segundo necessita da participação de mais de duas polimerases TLS e tem maior potencial mutagênico: POLH ou POLI incorporam o nucleotídeo oposto à lesão e POLZ (REV3L) realiza a extensão da fita. Sua participação também é evidenciada em outras lesões não relacionadas a fotolesão, como 8-oxoG, O<sup>6</sup>-MeG, timina glicol (Tg), adutos de cisplatina e sítios AP (CHAI, 2010; HARACKSA, 2000a; HARACKSA 2000b; KUSUMOTO, 2002)

A proteína POLK é composta por 870 resíduos de aminoácidos e seu gene está localizado na região cromossômica 5q13. POLK se distribui uniformemente pelo núcleo de células não danificadas, porém, pode se concentrar em focos de replicação quando essas células sofrem algum dano genotóxico (BERGOGLIO ET AL., 2002), característica essa compartilhada pelas demais DNA polimerases que realizam síntese translesão. A enzima POLK participa ativamente da etapa de extensão da fita após o pareamento de bases. Seu diferencial das outras DNA polimerases da família Y é a especificidade de incorporação, sendo a enzima mais fiel dentre as DNA polimerases, ela incorpora nucleotídeos erroneamente com uma frequência de aproximadamente  $10^{-3}$  a  $10^{-4}$  (JOHNSON ET AL., 2000), frequência esta menor que a da Pol $\eta$ , Pol $\iota$ , Rev1. Entretanto, a DNA polimerase kappa tem uma alta capacidade de estender extremidades 3'OH que estão pareadas incorretamente (mismatches), fazendo isso a uma frequência de  $10^{-1}$  a  $10^{-2}$  (WASHINGTON ET AL., 2002).

Uma das características mais marcantes de POLK é a sua tendência a gerar erros de mudança de fase de leitura, as chamadas mutações frameshifts, principalmente por deleção de um nucleotídeo. Essas frameshifts são geradas quando POLK desloca – para fora da hélice do DNA – o nucleotídeo da fita molde que está pareado errado, fazendo com que o nucleotídeo do iniciador fique corretamente pareado com a próxima base da fita molde (WOLFLE, ET AL., 2003). Uma vez que cada códon determina um aminoácido, a ocorrência de deleções, duplicações ou inserções pode acarretar modificações no tripleto de bases que forma cada códon, tendo como consequência mudanças na codificação dos aminoácidos e havendo ainda a possibilidade de se formar um códon de parada prematuro.

Estudos *in vitro*, avaliando os tipos de lesões nas quais a DNA polimerase kappa consegue realizar a síntese translesão, demonstraram que esta enzima é capaz de replicar DNA

contendo resíduos de 8-oxoguanina (8-oxoG) induzidos por benzopireno, mas é ineficiente na síntese através de dímeros de timina e TT (6-4) fotoprodutos (ZHANG ET AL., 2000; OGI, 2002).

Alguns pesquisadores propuseram que POLK poderia estar envolvida na replicação de DNA contendo adutos de benzo[a]pireno-dihidrodiol epóxido (BPDE). Esse composto possui grande semelhança química com a molécula de colesterol (da qual ele é um derivado), e a sua ligação à guanina provoca uma distorção espacial na dupla hélice do DNA, impedindo sua replicação normal. O BPDE é um dos compostos químicos presentes nos cigarros, e alguns estudos apontam uma forte associação entre o câncer de pulmão e o aumento da expressão de *POLK*, inclusive induzindo a danos no DNA e correlacionando-a com presença de aneuploidias (O-WANG et al., 2009; BAVOUX ET AL, 2005).

Há relatos da alta frequência de mutações no gene *POLK* em pacientes com câncer de próstata (YADAV et al, 2015). Ensaios *in vivo* mostram que camundongos *knockout* *Polk* -/- apresentam um fenótipo mutante ligado ao envelhecimento e acúmulo de lesões endógenas. Em adição a isto, ensaios *in vitro* mostraram que *POLK* possui a capacidade de replicar DNA contendo resíduos de timina glicol (Tg), um tipo bem conhecido de lesão oxidativa (PAULA L FISCHHABER ET AL., 2002).

A proteína POLI é composta por 715 resíduos de aminoácidos e seu gene está localizado no cromossomo 18. Um estudo com fibroblastos humanos revelou que células com baixa expressão de POLI exibiam hipersensibilidade a danos no DNA induzidos por peróxido de hidrogênio ou menadiona. Adicionalmente linhagens celulares com baixa expressão de *POLI* demonstraram redução da atividade do reparo por excisão de bases (BER). Estes estudos revelam o importante papel de POLI na proteção de danos contra o estresse oxidativo (BAVOUX ET AL, 2005; SALE, J. E.; LEHMAN, A. R.; WOODGATE, R., 2012)

POLI apresenta baixa fidelidade de incorporação de nucleotídeos, e devido a esse padrão incorreto de incorporação observado *in vitro*, sugeriu-se que Pol  $\iota$  seria uma boa candidata a participar do processo de hipermutação somática das imunoglobulinas *in vivo*. De fato, mais recentemente, foi descoberto o seu papel fundamental em eventos de hipermutação somática em genes de imunoglobulinas, onde Pol  $\iota$  e Pol  $\zeta$  são responsáveis pela geração de mutações em tandem (VAISMAN; WOODGATE, 2020; MAUL et al., 2016).

A função celular de POLI ainda não foi totalmente explicada, contudo, estudos demonstram que está relacionada à proteção contra danos oxidativos, de modo que variantes identificadas na sua subunidade catalítica causaram redução em sua atividade e foram capazes de aumentar a susceptibilidade individual de pacientes expostos a agentes genotóxicos. Além

disso, mutações pontuais no gene *POL1* foram associadas ao aumento do risco de adenocarcinoma pulmonar e carcinoma de células escamosas, relevando seu papel na carcinogênese (YEOM et al., 2020; SAKIYAMA et al., 2005).

Não existem relatos na literatura do estudo do mecanismo TLS em pacientes com Neoplasia Mielodisplásica, sendo esta a primeira pesquisa que se propõe a identificar mutações pontuais nas DNA Polimerases com atividade Translesão que possam contribuir com a fisiopatologia da doença.

## **2. OBJETIVOS**

### **2.1. Objetivo Geral**

Realizar o rastreamento, caracterização e anotação de variantes genéticas presentes nas enzimas DNA Polimerases com atividade translesão, *REV3L*, *POLQ*, *POLH*, *REVI*, *POLI*, *POLL*, *POLM*, *POLN* e *POLK* em pacientes com Neoplasia Mielodisplásica.

### **2.2. Objetivos Específicos**

1. Caracterizar os pacientes diagnosticados com Neoplasia Mielodisplásica de acordo com alterações citogenéticas, subtipo da doença e classificação de risco;
2. Pesquisar variantes nas regiões codificadoras (CDS) dos genes de DNA Polimerases com atividade translesão por sequenciamento de última geração (NGS);
3. Caracterizar e realizar a anotação genética das variantes identificadas;
4. Realizar a predição *in silico* de patogenicidade das variantes identificadas;
5. Realizar a classificação de oncogenicidade das variantes identificadas;

### **3. PACIENTES, MATERIAIS E MÉTODOS**

#### **3.1. Aspectos éticos da Pesquisa**

A pesquisa foi submetida e aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal do Ceará/PROPESQ-UFC (número de processo: 96542518.8.0000.5054) através do sistema da Plataforma Brasil, utilizando-se de Termos de Consentimento Livre e Esclarecido (TCLE) apresentados na sessão apêndice. A equipe executora desta pesquisa comprometeu-se a cumprir todas as diretrizes e normas reguladoras descritas na Resolução nº 466 de 12 de dezembro de 2012 do Conselho Nacional de Saúde que aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. A representação esquemática do delineamento experimental deste estudo encontra-se representado na figura 7.

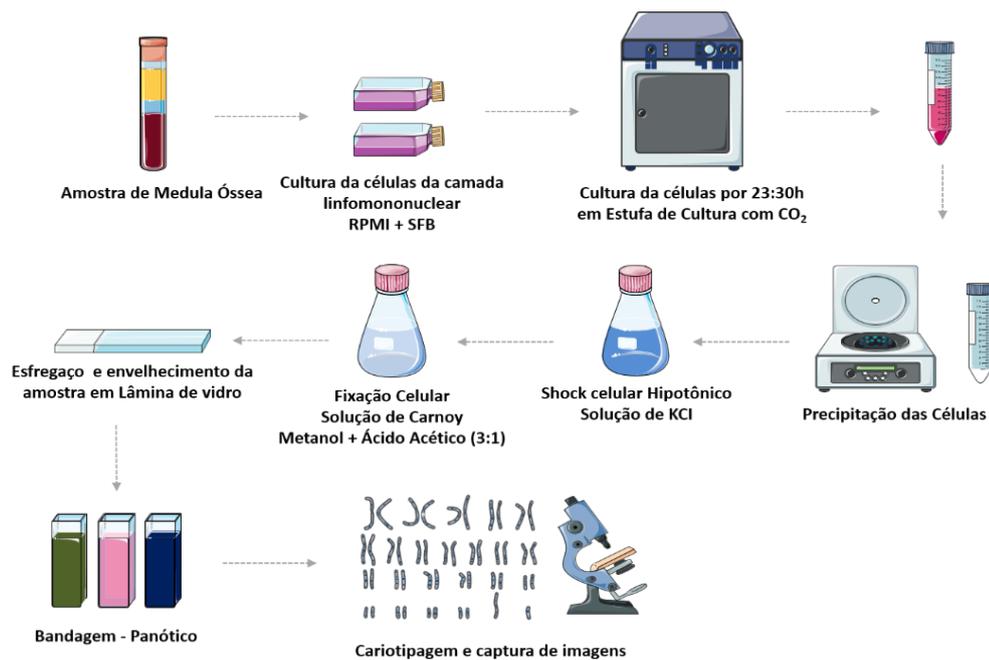
#### **3.2. Obtenção das amostras de células de medula óssea**

Foram coletados, em tubos contendo heparina e EDTA, 3ml de medula óssea de 50 pacientes diagnosticados com SMD para os procedimentos de Citogenética Clássica e 5mL para os procedimentos de extração de DNA a fim de realizar o sequenciamento por NGS. Foram sequenciados 9 genes em 50 amostras de *pool* celular de medula óssea de pacientes diagnosticados com SMD, dos subtipos SMD com Sideroblastos em anel (SMD-AS) e SMD com Excesso de blastos (SMD-EB), subtipos de baixo de alto risco de evolução para LMA. (SWERDLOW et al, 2008), atendidos em um hospital terciário de referência de 2010 a 2018. Para a separação das células do *pool* celular da medula óssea, realizou-se a transferência da amostra para um tubo do tipo Falcon de 50mL onde foi lavada com solução de lise (25mL de solução de cloreto de amônio 0,144M e bicarbonato de amônio 0,01 M). O conteúdo foi agitado lentamente por 3 minutos e centrifugado a 13200 rpm por 10 minutos a 4° C. Em seguida foi desprezada a fase aquosa e acrescentado 250µL de PBS, dependendo do volume do material obtido. Posteriormente foi acrescentado 750uL de Trizol LS Reagent® (Invitrogen, EUA) para cada 250uL de volume de PBS aplicado no pool medular. O material foi homogeneizando com pipeta até dissolução completa. Após este procedimento, o material foi armazenado no freezer a -80°C e, posteriormente, utilizado para os procedimentos de extração de DNA.

### 3.3. Cariótipo por bandeamento-G

A Citogenética foi realizada de acordo com os protocolos já estabelecidos pelo Laboratório de Citogenômica do Câncer da Universidade Federal do Ceará, segundo a técnica descrita por Chauffaille e adaptada por Pinheiro (PINHEIRO et al, 2009). A medula óssea foi colhida em heparina e de forma estéril e foi dividida em dois frascos contendo 7mL de meio RPMI 1640 (pH 7,0), 3mL de soro fetal bovino e 100µL de L-glutamina. Este material foi cultivado por 24 horas em estufa a 37°C. Uma hora antes do término da cultura foram adicionados 50uL de colchicina (Colcemid®) por 30 minutos. Em seguida, o material foi centrifugado e ressuspenso em solução hipotônica de KCl 0,075M e fixado em solução de ácido acético e metanol (proporção 3:1), por 4 vezes. Para confecção das lâminas, o material foi gotejado em lâminas de microscopia óptica e em seguida foram aquecidas por 3 minutos em micro-ondas na potência alta. O bandeamento foi realizado pela técnica de tripsina e as bandas coradas com o Kit Panótico® e Giemsa. Foram analisadas pelo menos 20 metáfases de cada paciente, sendo capturadas em sistema computadorizado com software para cariotipagem: CytoVision® (Figura 6). Os resultados das análises citogenéticas de medula óssea foram estabelecidos de acordo com os critérios do Sistema Internacional de Nomenclatura Citogenética Humana (ISCN) (SIMONS et al., 2013).

**Figura 6:** Representação esquemática da Metodologia de Citogenética por Bandeamento-G



**Fonte:** Elaborada pelo autor

### 3.4. Sequenciamento-alvo dos genes *REV3L*, *POLQ*, *POLH*, *REV1*, *POLI*, *POLL*, *POLM*, *POLN*, *POLK*

#### 3.4.1. Extração de DNA Genômico

O DNA genômico foi extraído previamente a partir das amostras de medula óssea dos pacientes utilizando o Kit de Extração de DNA genômico *PureLink™ Genomic DNA Mini Kit* (Invitrogen) e seguindo as instruções do fabricante. Em seguida foi realizada a avaliação da qualidade e concentração das amostras de DNA genômico provenientes dos 50 pacientes diagnosticados com Neoplasia Mielodisplásica incluídos neste estudo.

#### 3.4.2. Customização do Painel Genético

O painel genético customizado foi desenhado utilizando a ferramenta *online Desing Studio* (Illumina). As regiões-alvo compreenderam todas as regiões CDS de 9 genes codificadores de DNA Polimerases com atividade TLS (*REV3L*, *POLQ*, *POLH*, *REV1*, *POLI*, *POLL*, *POLM*, *POLN*, *POLK*), com uma cobertura total de 98%. Abaixo são apresentadas as regiões que foram avaliadas e o número de CDS de cada gene (Tabela 2).

**Tabela 2:** Regiões sequenciadas dos genes *REV3L*, *POLQ*, *POLH*, *REV1*, *POLI*, *POLL*, *POLM*, *POLN* e *POLK* e o número de CDS de cada gene

GENE	CROMOSSOMO	INÍCIO	FIM	NÚMERO DE CDS's
<i>POLH</i>	chr6	43,543,878	43,588,250	10/10 CDS
<i>REV3L</i>	chr6	111,620,234	111,804,414	32/32 CDS
<i>POLQ</i>	chr3	121,150,273	121,264,853	30/30 CDS
<i>POLI</i>	chr18	51,795,849	51,824,604	10/10 CDS
<i>REV1</i>	chr2	100,016,938	100,106,480	23/23CDS
<i>POLL</i>	chr10	103,338,639	103,348,027	07/07 CDS
<i>POLM</i>	chr7	44,111,847	44,122,129	05/05 CDS
<i>POLN</i>	chr4	2,073,645	2,230,958	24/24 CDS

#### 3.4.3. Preparo da Biblioteca

O preparo das bibliotecas de DNA foi realizado de acordo com o procedimento padrão recomendado pelo fabricante, utilizando dois kits apropriados (*TruSeq Custom Amplicon Assay Kit* e *AmpliSeq Custom*).

A quantidade de DNA necessária (*DNA input*) para cada corrida foi de 10ng com uma relação A260/280 de 1,8-2,0. O tamanho do amplicon foi, em média, de 175 e 275bp. O processo experimental iniciou com o preparo da biblioteca, seguiu com a amplificação dos

*amplicons* adicionados de adaptadores de amostras em um processo de PCR. Logo após, ocorreu a avaliação da qualidade da corrida, alinhamento com genoma de referência (análise primária), chamada de variantes (análise secundária) e caracterização e predição *in silico* (análise terciária).

De uma maneira sucinta, durante o preparo das bibliotecas e sequenciamento, foram realizadas:

- 1. Hibridização do Pool de Oligos.** Os oligos específicos para cada região alvo de interesse de cada DNA polimerase foi hibridizado ao DNA genômico do paciente (10ng) em reações com TSO/OHS2/ACD1 em termociclador pré-aquecido a 95°C com gradual redução da temperatura para 40°C em 80 minutos conforme orientações do fabricante (*TruSeq Custom Amplicon Assay Kit - Illumina*).
- 2. Remoção dos Oligos não hibridizados.** Os oligos não hibridizados foram removidos por processo de filtragem após adição dos reagentes ELM4/SW1/UB1. O processo seguiu com extensão-ligação mediante adição de 45uL de ELM41. Foram adicionados adaptadores de amostras (*index*) para posterior amplificação por PCR. Para esta fase, foram adicionados NaOH, PMM2 e TDPI. Após amplificação por PCR, foi realizada a limpeza do produto (*PCR clean UP*) com o uso de AMPure XP beads após centrifugação e repouso em placa magnética.
- 3. Normalização da Biblioteca.** O processo seguiu com a normalização da biblioteca mediante adição dos tampões/reagentes LNA1/LNB1/LNW1/LNS2. Para gerar os *clusters* e sequenciar, volumes iguais da biblioteca normalizada foram combinados, diluídos e hibridizados em tampão (PAL 6uL+ HT1 594ul), aquecidos e desnaturados para posterior carregamento no MiSeq.
- 4. Sistema Miseq.** O sistema MiSeq integra a clusterização, o sequenciamento propriamente dito e o alinhamento com o genoma de referência para geração dos arquivos BAM, BAI e VCF. Foi utilizado o kit de reagentes do MiSeq composto de lâmina de fluxo (*Flow cell*), frasco PR2 e cartucho do tipo V3 – 600 ciclos. Assim, foram executados, em média, 300 ciclos do tipo *pair-end* (2x 301= 602 ciclos). Foram adicionados 600uL da biblioteca em poço especificado no cartucho de reagentes e a *flow cell* foi higienizada para posterior carregamento. Após início do procedimento, foram acompanhados o Escore

Q30 de todos os ciclos. Após o término da execução, a análise inicial de qualidade e alinhamento foi realizada no sistema Miseq repórter.

#### **3.4.4. Análise das variantes obtidas nos genes *POLQ*, *REV3L*, *POLH*, *REVI*, *POLI*, *POLL*, *POLM*, *POLN*, *POLK***

##### *3.4.4.1. Análise Primária: Qualidade da corrida, filtragem e alinhamento com genoma de referência*

Para o sequenciamento das 30 amostras inicialmente incluídos neste estudo, utilizamos o kit *Truseq custom amplicon* (illumina), com posterior análise primária e secundária realizada pela ferramenta *Truseq Amplicon*, fornecida na plataforma *Basespace* da empresa. No entanto, para o segundo grupo de amostras sequenciadas (20), não foi possível a utilização do mesmo kit, o qual havia sido descontinuado, e optamos pelo uso do novo kit de sequenciamento para painéis customizados, *AmpliSeq* (illumina). Após o sequenciamento das amostras, a análise de dados secundária forneceu uma chamada de variantes incompleta, os quais não consideramos nesta análise por não condizerem com o perfil apresentado pelas amostras iniciais. Como exemplo, a chamada de variante dos pacientes, individualmente, identificou uma quantidade de variantes bem menor do que o que foi apresentado no primeiro grupo de amostras (ex. 111, 146 variantes nos pacientes do primeiro grupo X 36, 42, 21 variantes em pacientes do segundo grupo). Fato que se repetiu para a maioria das amostras sequenciadas neste último grupo. Dessa forma, optamos por seguir com a análise terciária apenas para as amostras do primeiro grupo (30), das quais temos certeza acerca da chamada de variantes.

Após o sequenciamento, seguindo a etapa de pré-processamento, as *reads* geradas foram então processadas para remoção de adaptadores e bases de baixa qualidade (*Phred quality score* <Q30). Para análise de qualidade da corrida foi considerado o índice de score de qualidade Q30, que foi de 93,60% e 60%, nas duas corridas realizadas. Os arquivos obtidos após processamento de qualidade foram exportados no formato FASTQ utilizando a plataforma *Basespace* (Illumina) e os fragmentos de DNA paired-end emparelhados (175 e 275bp) foram demultiplexados em arquivos BAM e BAI individuais, alinhados ao genoma referência (Homo

sapiens UCSC GRCh37/hg19) através do software integrado *Illumina Miseq Reporter Software* 2.4 (Figura 7)

#### 3.4.4.2. Análise Secundária: Chamada de variantes

Para a chamada de variantes (*variant calling*) foram utilizados os softwares IGV ([https://software.broadinstitute.org/.](https://software.broadinstitute.org/)) e NextGENe (<https://softgenetics.com/NextGENe.php>), utilizando parâmetros específicos para garantir a exclusão de variantes errôneas, de modo que somente as variações de nucleotídeo único (SNVs) e indels com  $\geq 2\%$  de frequência do alelo variante (VAF) e mais de 400 reads de profundidade foram consideradas significativas. Após a análise de chamada de variantes foram obtidos os arquivos VCF e extraídos para planilha de Excel, onde as variantes foram organizadas e filtradas para levantamento dos dados. A representação ilustrativa do processo metodológico de análise de variantes está apresentada na Figura 7.

#### 3.4.4.3. Análise terciária: Classificação de Patogenicidade e Oncogenicidade

*Anotação, caracterização e predição in silico para classificação das variantes identificadas nos genes REV3L, POLQ, POLH, REV1 e POLI*

Foram considerados para avaliação e análise terciária neste estudo somente os genes *REV3L, POLQ, POLH, REV1* e *POLI*. Os genes *POLL, POLM, POLN* e *POLK* serão analisados em um estudo posterior. Para a caracterização e classificação das variantes obtidas já conhecidas como **polimorfismos (dbSNP)** ou **com potencial oncogênico**, a partir das amostras de pacientes com Neoplasia Mielodisplásica presentes no estudo, foram utilizados o banco de dados dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>) e a plataforma COSMIC Database (*Catalogue of Somatic Mutations in Cancer* - <https://cancer.sanger.ac.uk/cosmic>). A busca nas plataformas permitiu a coleta dos números de identificação (ID - dbSNP e ID - COSMIC) das variantes já conhecidas e caracterização das mesmas a partir de informações atualmente disponíveis nas duas principais bases de dados sobre SNPs e mutações somáticas potencialmente oncogênicas já catalogadas, para posterior predição de patogenicidade/oncogenicidade.

*Predizendo o impacto da variante na proteína codificada – Uso do Ensembl VEP (Variant Effect Predictor)*

A análise *in silico* teve início com a predição da consequência que a variante teria para o transcrito ou a proteína codificada. Para classificação das variantes de acordo com a mudança de aminoácido, consequência e tamanho do efeito gerado no transcrito ou na proteína, foi utilizada a base de dados **Ensembl VEP** (*Variant Effect Predictor* - <https://www.ensembl.org/Multi/Tools/VEP>). A abordagem de cálculo do Ensembl VEP pode ser aplicada tanto a variantes germinativas como a variantes somáticas. O efeito das variantes pode ser dividido em: **Alto Impacto** (*High Impact*), **Moderado Impacto** (*Moderate Impact*), **Baixo Impacto** (*Low Impact*) e **Modificador** (*Modifier Impact*). Variantes do tipo *Alto Impacto* são aquelas que causam um grande impacto na proteína, causando, provavelmente, o truncamento da mesma ou perda de sua função, conhecidas como LoF (do inglês, *loss of function*). Variantes do tipo *Moderado Impacto* são aquelas que não conseguem causar perda total de função, mas podem diminuir a efetividade da proteína. Variantes de *Baixo Impacto* são aquelas consideradas inofensivas à proteína ou improváveis de alterar o seu comportamento. Variantes classificadas como *Modificadoras* são aquelas em que não se tem evidência do impacto para a proteína. Acontecem geralmente em regiões não codificadoras de genes.

A classificação completa de consequências das variantes para transcrito ou para a proteína podem ser acessadas no seguinte link: *Variant Effect Predictor* - <https://www.ensembl.org/Multi/Tools/VEP>. A tabela abaixo ilustra os principais termos de classificação obtidos nas variantes estudadas:

**Tabela 3:** Principais classificações de variantes de acordo com o *Ensembl VEP*

CONSEQUÊNCIA	DESCRIÇÃO	IMPACTO
<b>splice_acceptor_variant</b>	Altera a região de até duas bases na extremidade 3' do intron	Alto
<b>splice_donor_variant</b>	Altera a região de até duas bases na extremidade 5' do intron	Alto
<b>stop_gained</b>	Pelo menos uma base de um códon é alterada, resultando em um códon de parada prematuro, levando a um transcrito encurtado	Alto
<b>frameshift_variant</b>	Causa uma interrupção do quadro de leitura traducional porque o número de nucleotídeos inseridos ou excluídos não é um múltiplo de três	Alto
<b>stop_lost</b>	Pelo menos uma base do códon de término (stop) é alterada, resultando em um transcrito alongado	Alto
<b>inframe_insertion</b>	Uma variante não sinônima do quadro que insere bases na sequência de codificação	Alto
<b>inframe_deletion</b>	Uma variante inframe não sinônima que exclui bases da sequência de codificação	Moderado
<b>missense_variant</b>	Uma variante que muda uma ou mais bases, resultando em uma sequência de aminoácidos diferente, mas com preservação do comprimento da cadeia polipeptídica	Moderado
<b>splice_region_variant</b>	Uma variante que causa mudança dentro da região do sítio de splicing, seja dentro de 1-3 bases do éxon ou de 3-8 bases do íntron	Baixo
<b>synonymous_variant</b>	Uma variante em que não há alteração no aminoácido codificado	Baixo
<b>5_prime_UTR_variant</b>	Uma variante que ocorre na região 5'UTR, que está diretamente upstream ao códon de iniciação.	Modificadora
<b>3_prime_UTR_variant</b>	Uma variante que ocorre na região 3'UTR, que está diretamente downstream ao códon de iniciação. Esta região geralmente está envolvida em processos regulatórios como estabilidade do RNA, tradução e localização do mRNA.	Modificadora
<b>intron_variant</b>	Uma variante de transcrição que ocorre dentro de um íntron	

### ***Descobrimo o potencial patogênico das variantes: Análise funcional in silico de predição***

Uma análise funcional integrada *in silico* das variantes dos genes **REV3L**, **POLQ**, **POLH**, **REVI** E **POLI** foi conduzida para avaliação da patogenicidade de variantes que já possuíam identificação nas bases de dados dbSNP ou COSMIC Database. As variantes classificadas como LoF - perda de função pelo **Ensembl VEP** foram consideradas como claramente patogênicas e seguiram diretamente para análise de oncogenicidade (SOP-ACMG/AMP). Por se tratar de um estudo de *screening*, variantes desconhecidas e identificadas fora desse contexto não tiveram a análise funcional realizada por esta pesquisa e serão analisadas em um estudo posterior.

A avaliação acerca da patogenicidade das substituições de aminoácidos provocadas pelas variantes anotadas nas regiões codificadores dos genes em questão, foi realizada utilizando três diferentes algoritmos abertos: **SIFT** (*Sorting Intolerant from Tolerant* - <http://sift.bii.a-star.edu.sg/>), o qual prevê se uma substituição de aminoácido afeta a função da proteína com base na homologia de sequência e nas propriedades físicas dos aminoácidos, caracterizando-a como “Prejudicial” ou “Tolerada”; **PROVEAN** (*Protein Variation Effect Analyzer* - <http://provean.jcvi.org/index.php>), software que prevê se uma substituição de aminoácido ou indel tem impacto na função biológica de uma proteína utilizando abordagem de computação rápida para obter scores de alinhamento de pares de sequências provenientes do genoma anotado na plataforma *Ensembl*, resulta em variante “Deléteria” ou “Neutra”, **MutPred** (<http://mutpred.mutdb.org/>), método baseado em Machine Learning que integra dados genéticos e moleculares disponíveis nas bases de dados Human Gene Mutation Database (HGMD), SwissVar e dbSNP, desenvolvido para classificar as substituições de aminoácidos como “Patogênicas” ou “Benignas” em humanos, prevendo seu impacto em mais de 50 propriedades diferentes de proteínas e a ferramenta **Mutation Taster** (<http://www.mutationtaster.org/>), que emprega um classificador de Bayes, alimentado constantemente com dados de polimorfismos do projeto 1000 Genomas (AUTON et al, 2015) na previsão do potencial patogênico de alterações genéticas através do cálculo de probabilidade da mutação ser “Causadora de Doença” ou apenas um “Polimorfismo” sem impacto na doença.

As variantes selecionadas para análise neste estudo foram classificadas como patogênicas se dois ou mais softwares fornecessem previsões semelhantes entre si, adotando-se um score de 0 a 4 de acordo com os resultados obtidos para cada SNV nas ferramentas utilizadas, sendo 4 o valor mais alto e deletério para a previsão de patogenicidade das variantes, de acordo com método descrito por Pousada e colaboradores (POUSADA et al., 2016). Como

apontado anteriormente, as variantes identificadas que levam à perda de função da proteína (LoF – *loss of function*), do tipo mudança de quadro de leitura (*frameshift*), *inframe*, *introdução de códon de parada prematuro*, (*stop gained*), *stop lost*, *splice acceptor* e *splice donor*, foram consideradas como claramente patogênicas, devido às consequências de alto potencial deletério atreladas, as quais comumente comprometem processos transcricionais e traducionais, gerando proteínas truncadas ou não funcionais (POUSADA et al)

***Descobrendo o potencial oncogênico das variantes: classificação utilizando o SOP(ACMG/AMP/VICC)***

Após predição *in silico* de patogenicidade, as variantes selecionadas registradas no COSMIC Database ou fortes candidatas a uma associação causal com o fenótipo da doença (variantes do tipo LoF - perda de função), foram classificadas quanto ao seu grau de oncogenicidade de acordo com o procedimento operacional padrão (SOP) do *Clinical Genome Resource (ClinGen)*, *Cancer Genomics Consortium (CGC)*, and *Variant Interpretation for Cancer Consortium (VICC)* (HORAK et al, 2022).

O SOP - *Standard Operating Procedure*, foi um consenso de classificação para variantes oncogênicas publicado em 2022, após esforços de um grupo de trabalho composto por membros do *ClinGen Somatic Clinical Domain Working Group*, *ClinGen Germline/Somatic Variant Subcommittee*, *Cancer Genomics Consortium (CGC)* e *Variant Interpretation for Cancer Consortium (VICC)*. O grupo adaptou as diretrizes de patogenicidade para linhagem germinativa padronizadas pelo *American College of Medical Genetics and Genomics (ACMG)* e *Association for Molecular Pathology (AMP)* para categorizar de forma sistemática e abrangente as evidências de oncogenicidade para variantes somáticas. As diretrizes do SOP(ACMG/AMP/VICC), como é denominado, definem oncogenicidade como significando a patogenicidade da variante no contexto de uma doença neoplásica. Categorizando especificamente variantes somáticas por seu potencial de conferir vantagens de crescimento e sobrevivência em células tumorais (HORAK et al, 2022; RICHARDS et al, 2015).

O Procedimento Operacional Padrão (SOP) adaptado e desenvolvido para variantes somáticas é constituído por um conjunto de critérios envolvendo três campos principais de avaliação: dados populacionais; dados computacionais e preditivos e dados funcionais. Para cada variante, em cada um destes critérios, atribuiu-se um código com peso de evidência e

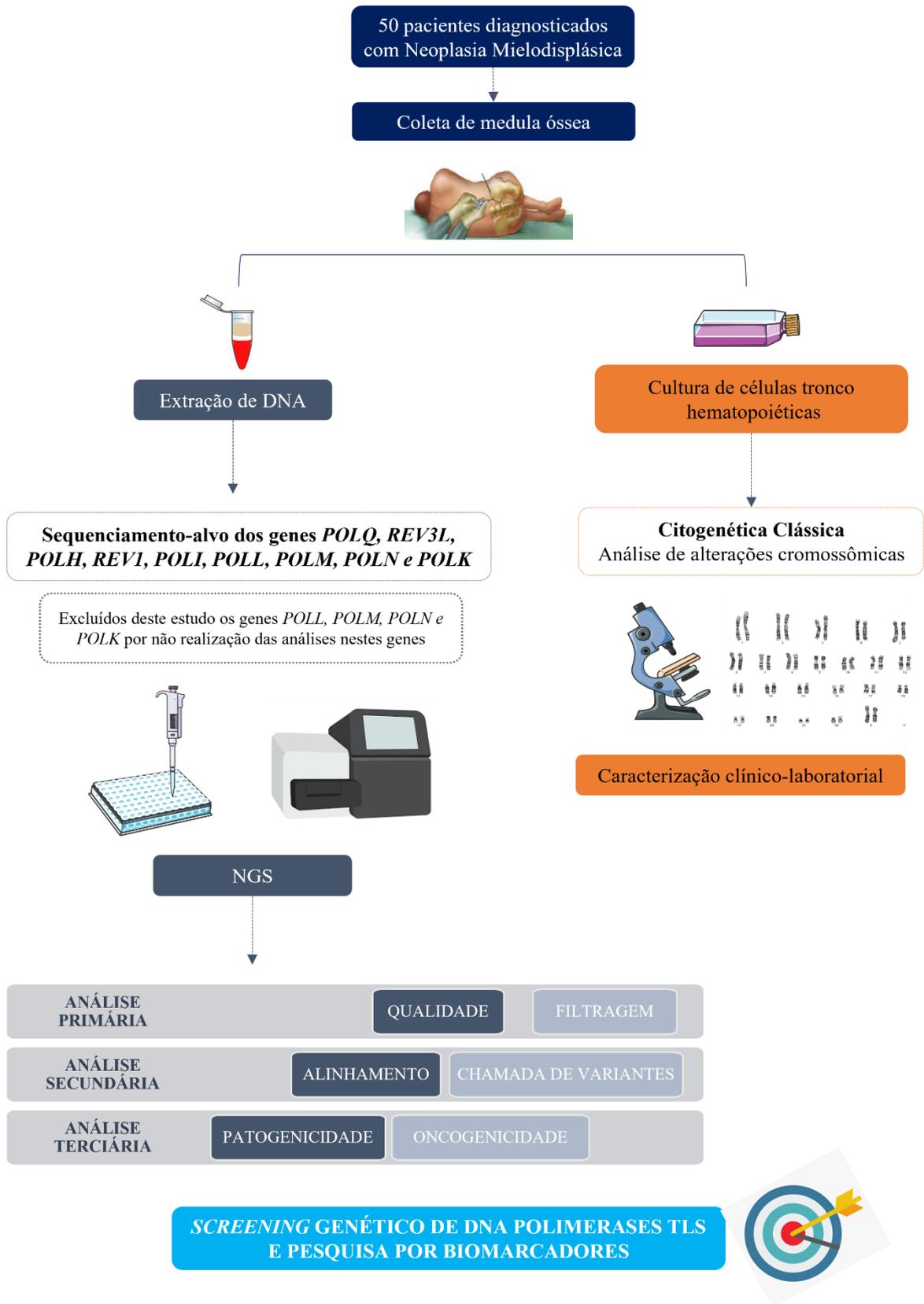
pontuações definidas. Após o somatório de todas as pontuações de evidência, variantes somáticas de nucleotídeo único (SNVs) e pequenas inserções/deleções podem ser atribuídas a uma das cinco categorias: Oncogênica (contribui diretamente para o desenvolvimento da doença), Provavelmente Oncogênica (há uma alta probabilidade - mais de 90% - de que esta variante seja causadora de doença), Variante de Significado Incerto (não há evidências científicas suficientes para apoiar uma classificação mais definitiva desta variante), Provavelmente Benigna (a atual evidência científica é insuficiente para provar de forma conclusiva que esta variante seja benigna) e Benigna (esta variante não causa doença), auxiliando sua interpretação clínica (anexo A). Variantes que não alcançam a pontuação para nenhuma dessas categorias ou possuem classificações conflitantes entre elas e que não há evidência científica suficiente para assim classificá-las, são categorizadas como variantes de significado clínico desconhecido ou incerto, comumente conhecidas como VUS (HORAK et al, 2022). A relação de critérios com suas devidas pontuações e evidências sugeridas encontra-se no anexo A deste trabalho (anexo A). As ferramentas de análise utilizadas neste estudo estão descritas na Tabela 4.

**Tabela 4:** Ferramentas de análise utilizadas neste estudo

FERRAMENTA	DESCRIÇÃO	FUNÇÃO
<b>Ensembl VEP</b>	Banco de dados genético-molecular e análise <i>in silico</i> de patogenicidade	Utilizado para classificação das variantes quanto à consequência para a proteína codificada
<b>dbSNP</b>	Banco de dados genético-molecular de variantes polimórficas	Utilizado para busca de variantes já conhecidas e coleta de número de identificação
<b>COSMIC Database</b>	Banco de dados genético-molecular de variantes com potencial oncogênico	Utilizado para busca de variantes já conhecidas e coleta de número de identificação
<b>Mutation Taster</b>	Avaliação <i>in silico</i> do impacto funcional de variantes alélicas	Utilizado para análise de predição de patogenicidade
<b>SIFT</b>	Avaliação <i>in silico</i> do impacto funcional de variantes alélicas <i>nonsynonymous</i> e <i>missense</i>	Utilizado para análise de predição de patogenicidade
<b>Provean</b>	Avaliação <i>in silico</i> do impacto funcional de variantes alélicas	Utilizado para análise de predição de patogenicidade
<b>MutPred</b>	Avaliação <i>in silico</i> do impacto funcional de variantes alélicas	Utilizado para análise de predição de patogenicidade
<b>SOP (ACMG/AMP/VICC)</b>	Avaliação de oncogenicidade baseada em dados populacionais, dados computacionais e preditivos e dados funcionais	Utilizado para classificação de oncogenicidade

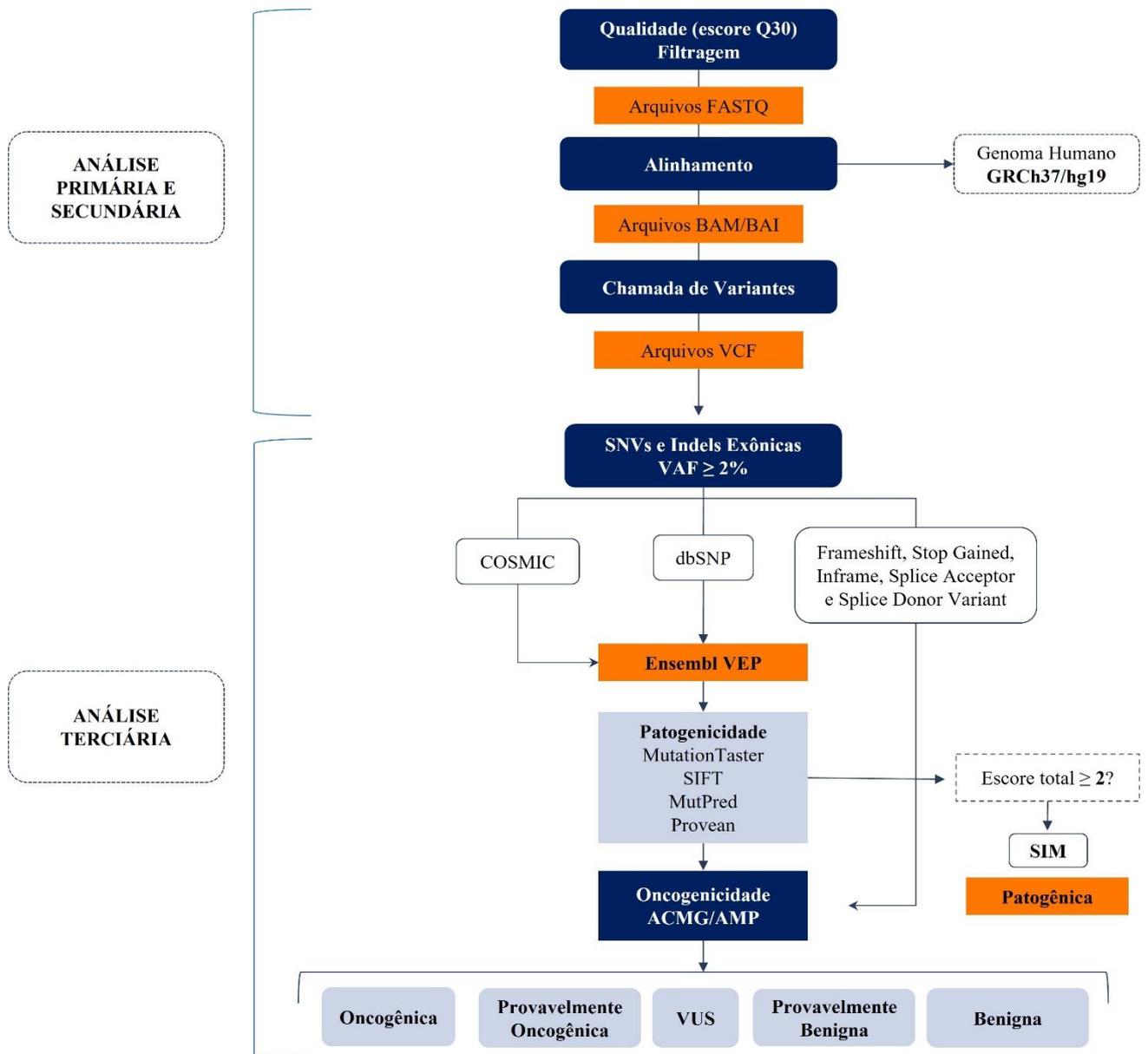
**Legenda:** Ensembl VEP: *Variant Effect Predictor*; dbSNP: *Database Single-nucleotide polymorphism*; COSMIC: *Catalogue of somatic mutations in cancer*; SIFT: *Sorting intolerant from tolerant*; Provean: *Protein Variation Effect Analyzer*; MutPred: *mutation prediction*; SOP(ACMG/AMP/VICC): *standards operation procedure*

**Figura 7:** Delineamento experimental do estudo



Fonte: Elaborada pela autora

**Figura 8:** Esquemática da metodologia utilizada para análise de variantes realizada neste estudo



**Fonte:** Elaborada pela autora

## 4. RESULTADOS

### 4.1. Caracterização dos pacientes

Realizamos o sequenciamento-alvo e análise por abordagem de bioinformática dos genes de DNA Polimerase com atividade TLS em 30 pacientes diagnosticados com Neoplasia Mielodisplásica (Tabela 5). A média de idade dos pacientes incluídos neste estudo foi de 72 anos. Quanto às variáveis sociodemográficas, frente à variável gênero, foram incluídos 13 indivíduos do sexo feminino (43,33%) e 17 do sexo masculino (56,67%) (Tabela 5).

Em relação à classificação clínica baseada nos critérios da OMS de 2016, 17 (56.67%) pacientes foram classificados como SMD com excesso de blastos e 13 (43.33%) como SMD com sideroblastos em anel (ARBER et al., 2016).

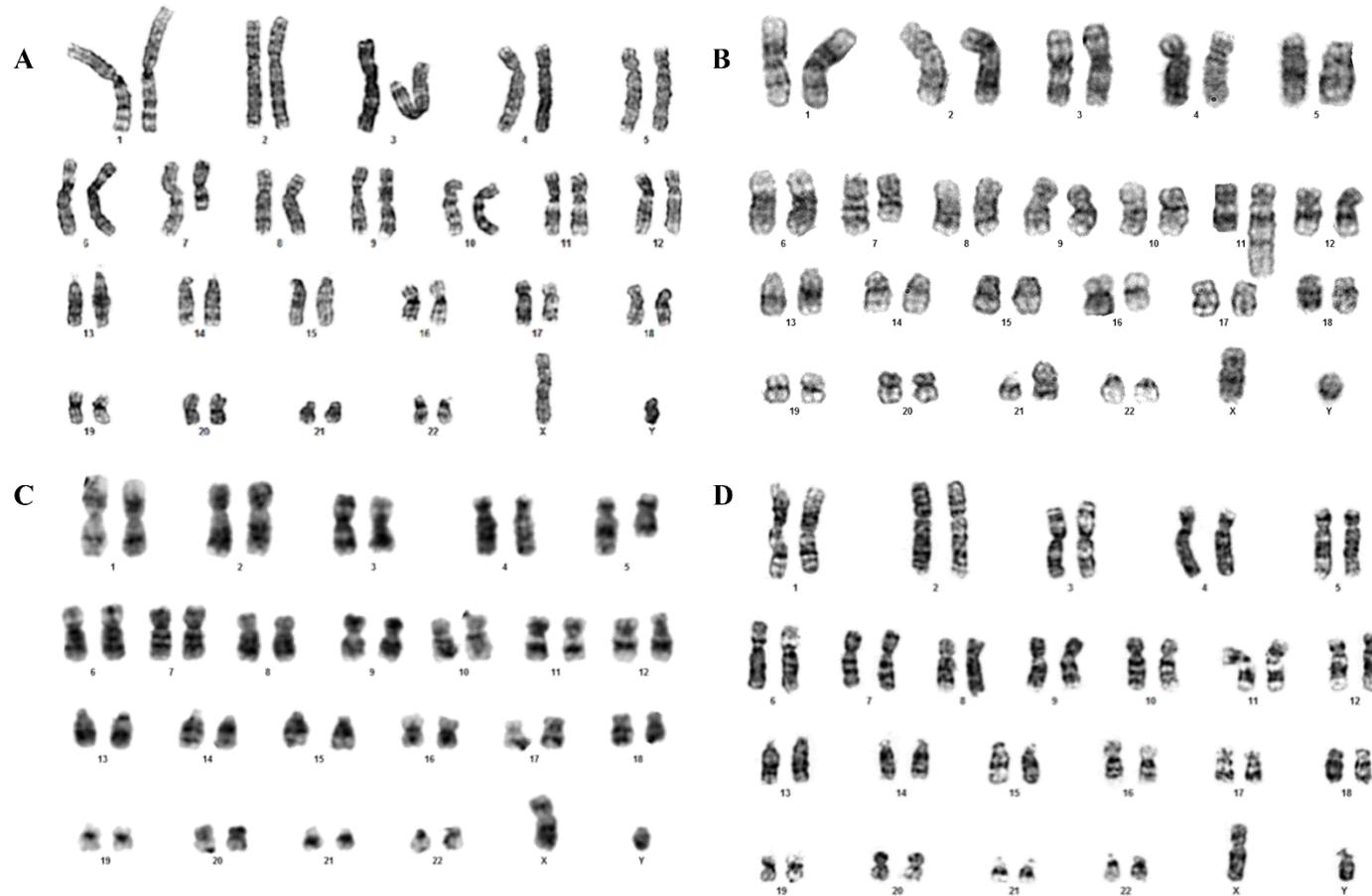
Com relação à frequência das alterações citogenéticas nos pacientes incluídos no estudo, identificou-se que 12 (40%) indivíduos possuíam cariótipo normal, em contraste à 18 (60%) dos pacientes que possuía cariótipo alterado. Quanto ao grupo de risco determinado através IPSS-R, 10 pacientes pertencem ao grupo de risco muito baixo e baixo (33.33%), enquanto 11 pacientes pertencem ao grupo de muito alto e alto risco 36.67% e 4 (13.33%) pacientes tiveram o grupo de risco classificado como intermediário. Em 5 pacientes (16.67%) não foi possível a estratificação prognóstica por ausência de resultado citogenético (tabela 5). A figura 9 exemplifica os resultados da análise citogenética encontrados nesse estudo

Quanto à exposição à produtos tóxicos ao longo da vida, 14 pacientes (46.67%) relataram exposição crônica à produtos tóxicos e 16 (53.33%) pacientes não relataram exposição, sendo o agrotóxico o produto tóxico mais relatado por estes pacientes. A caracterização do perfil clínico-laboratorial dos pacientes incluídos neste estudo, encontra-se na tabela a seguir:

**Tabela 5:** Descrição clínica dos pacientes com Neoplasia Mielodisplásica incluídos no estudo

Paciente	Sexo	Idade	Cariótipo - ISCN 2016	OMS 2016	IPSS-R	Exposição à tóxicos
1	M	73	46,XY[20]	SMD-EB1	Baixo	Não
2	F	84	Ausencia de Metáfase	SMD-SA	-	Não
3	M	63	47,XY,+8[6]/47,XY,del(7)(q32)[3]/46,XY[2]	SMD-EB2	Alto	Agrotóxicos
4	M	73	Ausência de metáfase	SMD-EB2	-	Óleo diesel
5	F	27	46,XX[20]	SMD-SA	Baixo	Agrotóxicos
6	F	72	46,XX[20]	SMD-AS	Baixo	Inseticidas
7	M	91	46,X-Y[4]/46,XY[16]	SMD-SA	Muito Baixo	Não
8	M	69	45,XY,del(3)(p22),del(4)(q21),-5,-6,-7,+add(16)(p13),add(20)(p13),+mar[20]	SMD-SA-DM	Muito Alto	Não
9	M	63	37,X,-2,-3,-9,-11,-12,-15,-16,-18,-Y[8]/46,XY,del(5)(q15q33)[5]/46,XY[6]	SMD-EB1	Muito Alto	Solventes, tintas
10	M	80	45,X,-Y[18]/46,XY[7]	SMD-SA	Muito Baixo	Solventes, tintas
11	F	70	Ausência de Metáfases	SMD-SA	-	Não
12	F	87	Ausência de metáfase	SMD-SA	-	Não
13	F	81	46,XX[5]	SMD-SA	Muito Baixo	Não
14	F	56	46, XX [11]	SMD-SA	Muito Baixo	Querosene
15	M	75	46,XY,del(7)(q31)[3]/45,XY,-5,del(7)(q31)[5]/46,XY,del(5)(q15),del(7)(q31)[8]/46,XY,del(7)(q31),add(11)(q24)[4]/46,XY,del(5)(q15),del(7)(q31),add(11)(q24)/46,XY[5]	SMD-EB1	Muito Alto	Agrotóxicos
16	F	72	Ausência de metáfase	SMD-EB1	-	Não
17	M	73	47,XY,+8[3]/48,XY,+8,+12,dup(1)(q21q44)[5]/46,XY[12]	SMD-EB2	Alto	Não
18	F	77	46,XX[20]	SMD-SA-DM	Baixo	Não
19	M	55	45,XY,-7[15]/46,XY,-7,+mar[5]	SMD-EB2	Muito Alto	Sinteco, Verniz, Baygon
20	M	92	46,XY,del(5)(q32)[13]/46,XY[17]	SMD-SA-DM	Baixo	Agrotóxicos
21	M	58	46,XY[20]	SMD-EB1	Intermediário	Verniz
22	M	79	46,XY[20]	SMD-EB1	Intermediário	Não
23	M	64	46,XY,del(11)(q23)[3]/46,XY[12]	SMDA-EB1	Intermediário	Colagem de fábrica
24	M	84	46,XY[20]	SMD-EB1	Alto	Não
25	F	84	46,XX[20]	SMD-EB1	Alto Risco	Não
26	F	81	46,XX[20]	SMD-EB1	Alto Risco	Não
27	F	82	47,XX,+8[9]/47,XX,+8,del(20)(q12)[5]/46,XX[6]	SMD-SA-DM	Intemediário	Lampião de Gás e Querosene
28	F	71	46,XX[20]	SMD-EB-I	Baixo risco	Não
29	M	76	46,XY,del(5)(q31)[2]/46,XY,del(5)(q31),-7,+8[16]	SMD-EB2	Alto Risco	Tiner Paulista
30	M	55	45,X,-Y[15]/45,X,-Y,del(5)(q32)[3]/46,XY[2]	SMD-EB2	Muito Alto	Não

**Figura 9:** Representação dos resultados citogenéticos encontrados nos pacientes com SMD incluídos no estudo



**A. ID03**, sexo masculino, apresentando cariótipo alterado: 47,XY,+8[6]/47,XY,del(7)(q32)[3]/47,XY,del(7)(q32),+8[7]/46,XY[2]. **B. ID15**, sexo masculino, apresentando cariótipo alterado: 46,XY,del(7)(q31)[3]/45,XY,-5,del(7)(q31)[5]/46,XY,del(5)(q15),del(7)(q31)[8]/46,XY,del(7)(q31),add(11)(q24)[4]/46,XY,del(5)(q15),del(7)(q31),add(11)(q24)/46,XY[5], 46,X-Y[4]/46,XY[16]. **C. ID29**, sexo masculino, apresentando cariótipo alterado: 46,XY,del(5)(q32)[13]/46,XY[17]. **D. ID21**, sexo masculino, apresentando cariótipo normal 46,XY[10]

## 4.2. Classificação e caracterização das variantes

Para este estudo, optamos por realizar, primeiramente, a análise e classificação das variantes caracterizadas como perda de função – LoF pelo *Ensembl Variant Effect Predictor*. São elas: *frameshift*, *inframe*, *stop gained*, *splice acceptor variant*, *splice donor variant*. Posteriormente, realizamos a predição e classificação das variantes depositadas na base de dados COSMIC Database e caracterizamos as variantes polimórficas já depositadas nas bases de dados dbSNP. As variantes desconhecidas **não** classificadas como perda de função (LoF - do inglês *loss of function*) (*missense* e *synonymous*) bem como variantes identificadas em regiões não-codificadoras (íntrons), serão analisadas em um estudo posterior.

O banco de dados completo, incluindo todas as variantes identificadas e anotadas, mesmo as que não foram alvo de análise neste estudo, contendo informações acerca da localização, consequência resultante para a proteína, número de amostras nas quais as variantes foram detectadas, VAF e pontuações para classificação SOP, está disponível para consulta.

Com relação à classificação de oncogenicidade SOP (ACMG/AMP/VICC), a ausência de recorrência das variantes em outros estudos foi o fator que mais contribuiu para a evidência incerta (classificações VUS) de oncogenicidade atribuída à algumas variantes.

### 4.2.1. Caracterização, análise *in silico* da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene *POLQ*

Foram identificadas 194 variantes no gene *POLQ*, das quais 140 foram exônicas, classificadas pelo *Ensembl VEP* de acordo com sua consequência para a proteína: 85 variantes *missense*, 38 variantes *synonymous*, 10 variantes *frameshift*, 3 variantes do tipo *stop gained* e 4 variantes do tipo *Splice Acceptor Variant* e *Splice Donor Variant*.

Dentre as variantes identificadas, após busca em bases de dados, identificamos 36 variantes registradas no banco de dados *Single Nucleotide Polymorphism* (dbSNP) e 03 registradas pelo banco de dados *COSMIC Database*. Sendo assim, nesta coorte, classificamos e analisamos a patogenicidade e oncogenicidade de variantes com alto impacto na proteína (10 variantes *frameshift*, 3 variantes do tipo *stop*

*gained* e 4 variantes do tipo *Splice Acceptor Variant e Splice Donor Variant*) (tabelas 6 e 7) e variantes registrada no COSMIC que seriam potencialmente oncogênicas (tabela 6).

#### 4.2.1.1. Variantes com Potencial Perda de Função (LoF)

Durante a avaliação de patogenicidade, foram detectadas variantes cujas consequências são seriamente prejudiciais para a proteína resultante (variantes potencialmente *loss of function*, perda de função). São elas: 10 variantes *frameshift*, 3 variantes do tipo *stop gained* e 4 variantes do tipo *Splice Acceptor Variant e Splice Donor Variant* (tabela 6). Todas essas variantes são consideradas como **Claramente Patogênicas** pelos softwares de predição devido às consequências de alto potencial deletério, as quais comumente comprometem processos transcricionais e traducionais, gerando proteínas truncadas ou não funcionais.

A **primeira variante** de consequência *frameshift* (p.Glu1678AspfsTer24), é caracterizada pela deleção de uma timina na posição (ch3)121206743, e foi detectada em 08 dos 30 pacientes avaliados (IDs: 15,17,20,23,26,27,28,30). A frequência da variante não foi encontrada no gnomAD e, por este motivo, para a classificação de oncogenicidade utilizando os critérios SOP-ACMG/AMP, recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (variante de significado incerto). Dos 08 pacientes nos quais a variante foi identificada, 06 foram diagnosticados com SMD-EB 1 ou 2 e 04 relataram alta exposição à produtos tóxicos ao longo da vida (agrotóxicos, querosene e verniz) (tabela 6).

A **segunda variante** de consequência *frameshift* (c.5031\_5031delA), é caracterizada pela deleção também de uma timina na posição (ch3)121206746, e foi detectada em 04 dos 30 pacientes avaliados (IDs: 15,17,23,24). A frequência alélica da variante não foi encontrada no gnomAD e, por este motivo, para a classificação de oncogenicidade utilizando os critérios SOP-ACMG/AMP, recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (variante de significado incerto). Os 04 pacientes nos quais a variante foi identificada foram diagnosticados com SMD-EB 1 ou 2 (tabela 6).

A **terceira variante** *frameshift*, POLQ(NM\_199420.4):c.5016del (p.Asn1672LysfsTer4), localizada na região (ch3)121206761, teve como consequência a substituição do aminoácido asparagina por um códon de terminação (Ter) (p.Asn1672LysfsTer). A frequência da variante não foi encontrada no gnomAD e, por este motivo, para a classificação de oncogenicidade utilizando os critérios ACMG/AMP,

recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (variante de significado incerto). A variante foi identificada em 11 dos 30 pacientes (IDs 15,17,18,19,20,21,22,23,24,25,27,28,30) avaliados (tabela 6).

Para a **quarta variante**, POLQ(chr3)121206771ATTC>TTT, identificada na região (chr3)121206771, não foi possível realizar a classificação e predição de patogenicidade e oncogenicidade, pois o software não foi capaz de identificar a região da variante (tabela 6).

A **quinta variante**, POLQ(chr3)121190925CA>C, localizada na região (chr3)121190925, foi identificada em apenas uma amostra analisada (ID 17), no entanto, não foram realizadas análises devido à baixa qualidade da chamada de variante na amostra analisada (phred=30) (tabela 6).

A **sexta variante** do tipo *frameshift* (p.Leu672PhefsTer9) foi identificada em 03 pacientes (IDs 17,20,28), dois diagnosticados com SMD-EB e um com SMD-SA, é caracterizada pela deleção de uma timina na região (chr3)121217460. A frequência da variante também não foi encontrada no gnomAD e, por este motivo, para a classificação de oncogenicidade utilizando os critérios ACMG/AMP, recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (tabela 6).

A **sétima variante** do tipo *frameshift*, POLQ(NM\_199420.4):c.6411del (p.Leu2137PhefsTer5), foi identificada em 06 pacientes (IDs 18,20,21,26,28,30) e é caracterizada pela deleção de uma timina na região (chr3)121192328. Não há frequência da variante no gnomAD e, por este motivo, para a classificação de oncogenicidade utilizando os critérios ACMG/AMP, recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (tabela 6).

A **oitava variante** do tipo *frameshift* (p.Asn1669LysfsTer7) foi identificada em 08 pacientes (IDs 18,19,20,21,24,25,26,28) e é caracterizada pela deleção de uma timina na região (chr3)121206770. Não há frequência da variante no gnomAD e, por este motivo, para a classificação de oncogenicidade utilizando os critérios ACMG/AMP, recebeu a pontuação para a classificação de oncogenicidade utilizando os critérios ACMG/AMP, recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (tabela 6).

A **nona e 10ª variantes** *frameshift*, localizadas nas regiões (chr3)121208472 e (chr3)121190921, respectivamente, foram identificadas em apenas uma amostra analisada (ID 27 e 28), no entanto, não foram realizadas análises devido à baixa qualidade

e baixa cobertura da chamada de variante na amostra analisada (phred=30 e 32/depth=52 e 65) (tabela 6).

Foram detectadas 03 variantes do tipo *stop gained* - POLQ(NM\_199420.4):c.6241C>T (p.Gln2081Ter), POLQ(NM\_199420.4):c.1498G>T (p.Glu500Ter), POLQ(NM\_199420.4):c.6643C>T (p.Gln2215Ter) - todas classificadas como VUS após realização de pontuação seguindo os critérios de oncogenicidade ACMG/AMP. Duas das variantes, localizadas nas regiões (chr3)121195551 e (chr3)121230847, respectivamente, foram identificadas no mesmo paciente, classificado como subtipo SMD-AS (ID12), porém sem avaliação citogenética. Recebe destaque a variante POLQ(NM\_199420.4):c.6643C>T (p.Gln2215Ter), relatada no COSMIC como potencialmente oncogênica. Esta variante foi classificada como VUS seguindo os critérios de oncogenicidade deste trabalho (tabela 6).

As variantes do tipo *Splicing Donor Variant e Splicing Acceptor Variant* são consideradas de alto impacto na função da proteína por acontecerem em regiões de splicing e, por este motivo, também são classificadas como variantes de perda de função.

A variante do tipo *Splicing Donor Variant*, localizada na região (chr3)121260194, teve como consequência a substituição de uma adenina por citosina, em uma região que altera as duas bases na extremidade 3' do íntron, importante para processos de *splicing*. Apesar deste tipo de variante demonstrar ter alto impacto na proteína, provavelmente gerando uma proteína truncada, a variante não foi analisada ou teve sua predição realizada por apresentar baixíssima cobertura (21) e baixa qualidade de leitura (42/100) (tabela 6).

Dentre as três variantes do tipo *Splicing Acceptor Variant*, destacamos a variante POLQ(NM\_199420.4):c.5774-1G>A, descrita como o polimorfismo rs761787398. Ademais de ter apresentado baixa cobertura nas amostras detectadas, o polimorfismo foi detectado em cinco pacientes (IDs 17,21,22,25,26), sua frequência alélica não foi encontrada no banco de dados gnomAD e, utilizando os critérios SOP-ACMG/AMP, recebeu a pontuação OP4 – oncogênica de suporte e foi classificada como VUS (tabela 6)

**Tabela 6:** Análise in silico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene *POLQ*

VARIANTE	PACIENTE ID/VAF	SOP (ACMG/AMP/VICC)	MUTATION TASTER	SIFT	PROVEAN	MUTPRED	ESCORE TOTAL
POLQ(NM_199420):c.5034_5034del (p.Glu1678AspfsTer24)	15(6,0%),17(2,9%),20(4,7%),23(5,6%),23(3,9%),27(3,6%),28(5,3%),30(4,5%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLQ(NM_199420):c.5031_5031delA	15(4,7%),17(2,9%),23(6,1%),24(4,8%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLQ(NM_199420.4):c.5016del (p.Asn1672LysfsTer4)	15(5,1%),17(3,8%),18(3,9%),19(5,5%),20(3,7%),21(4,5%),22(4,8%),23(5,6%),24(4,8%),25(4,2%),27(15,1%),28(4,9%),30(8,7%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLQ(chr3)121206771ATTC>TTT	15(9,0%),17(10%),22(11,1%),23(13,3%),27(6,1%)	-	-	-	-	-	-
POLQ(chr3)121190925CA>C	17(6,8%)	-	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLQ(NM_199420.4):c.2016del (p.Leu672PhefsTer9)	17(10%),20(17,4%),28(31,8%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLQ(NM_199420.4):c.6411del (p.Leu2137PhefsTer5)	18(13,8%),20(10,4%),21(15,4%),26(9,3%),28(8,1%),30(13%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLQ(NM_199420.4):c.5007del (p.Asn1669LysfsTer7)	18(9,8%),19(6,8%),20(7,4%),21(14,1%),24(7,8%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-

	,25(8,8%),26(9,6%),28(7,4%)							
POLQ(chr3)121208472AG>A	27(9,6%)	-	-	-	-	-	-	-
POLQ(chr3)121190921AG>A	28(7,7%)	-	-	-	-	-	-	-
POLQ(chr)121260194A>C	17(19%)	-	-	-	-	-	-	-
POLQ(NM_199420.4):c.5774-1G>A	17(22,2%),21(13,5%),22(11,4%),25(12,5%),26(12,1%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-	
POLQ(NM_199420.4):c.6241C>T (p.Gln2081Ter)	12(6,5%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-	
POLQ(NM_199420.4):c.1498G>T (p.Glu500Ter)	12(6,6%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-	
POLQ(NM_199420.4):c.6643C>T (p.Gln2215Ter)	18(10%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-	
POLQ(NM_199420.4):c.7271A>T (p.Asn2424Ile) COSM1693249	25(10,8%)	Provavelmente Benigna	Polimorfismo	Tolerada	Deletéria	Deletéria	2	

**Legenda.** Variante: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; VUS: variante de significado incerto; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; PROVEAN/SIFT/MutationTaster/MutPred: resultado da análise de patogenicidade *in silico* das variantes de acordo com cada ferramenta. Escore total: número total de ferramentas que concordaram entre si acerca do potencial patogênico da variante. Publicações: número de referências encontradas para determinada variante na literatura. Células contendo “-” referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora.

#### 4.2.1.2. Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)

Seguindo os critérios de análise sugeridos por este trabalho, analisamos as três variantes detectadas neste estudo que já haviam sido registradas na base de dados COSMIC (c.6643C>T - COSM4825906; c.7271A>T - COSM1693249; ch3-121202429). Estão localizadas nas regiões (ch3)121190912, (ch3)121158957 e (ch3)121202429, respectivamente (tabela 7).

Duas das variantes já foram descritas anteriormente neste trabalho e classificadas como claramente patogênicas por se tratar de variantes com perda de função (*stop gained* e *splice acceptor variant*) (tabela 7).

A variante POLQ(NM\_199420.4):c.7271A>T (p.Asn2424Ile), relatada pelo COSM1693249, teve sua predição de patogenicidade indicada como deletéria em duas das quatro ferramentas utilizadas, garantindo a esta variante a pontuação OP1 (*Oncogenic Supporting-1*). Apesar de ser considerada deletéria pelas ferramentas de predição, as bases de dados indicam que 100% das variantes *missense* relatadas na literatura para *POLQ* são variantes benignas, o que confere a pontuação SBS2 (*Somatic Benign Strong-2*). Após somatória das pontuações adquiridas, a variante foi classificada como Provavelmente Benigna (tabela 7).

As variantes do tipo SNPs que já haviam sido depositadas em bancos de dados genômicos foram detectadas em um número relevante dentre os pacientes incluídos no estudo, contudo, o escore total obtido para esses polimorfismos revelou um baixo potencial deletério, apesar dos relatos existentes na literatura demonstrando suas associações a outros tipos de câncer. A caracterização e classificação de oncogenicidade dessas variantes polimórficas identificadas no gene *POLQ* estão descritas na tabela abaixo (tabela 7).

**Tabela 7.** Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene *POLQ*

POLIMORFISMOS (SNPs) NO GENE <i>POLQ</i> IDENTIFICADOS NAS AMOSTRAS ANALISADAS						
POLIMORFISMO	VARIANTE ALTERAÇÃO NO AMINOÁCIDO	CONSEQUÊNCIA DA VARIANTE	Nº DE PACIENTES COM A VARIANTE	SOP (ACMG/AMP/VICC)	PACIENTE ID/VAF	SIGNIFICADO CLÍNICO
rs702018	c.1437C>T p.Gly479=	Sinônima	14	Benigna	<b>1</b> (100,0%), <b>2</b> (100,0%), <b>3</b> (100,0%), <b>4</b> (100,0%), <b>5</b> (100,0%), <b>6</b> (99,8%), <b>7</b> (100,0%), <b>8</b> (100,0%), <b>9</b> (100,0%), <b>10</b> (100,0%), <b>11</b> (100,0%), <b>12</b> (99,9%), <b>13</b> (99,9%), <b>14</b> (100,0%)	Ependimoma, Câncer Colorretal
rs702017	c.197G>T p.Arg66Ile	Missense	14	Benigna	<b>1</b> (100,0%), <b>2</b> (100,0%), <b>3</b> (100,0%), <b>4</b> (99,9%), <b>5</b> (100,0%), <b>6</b> (100,0%), <b>7</b> (100,0%), <b>8</b> (100,0%), <b>9</b> (99,9%), <b>10</b> (99,9%), <b>11</b> (100,0%), <b>12</b> (99,9%), <b>13</b> (100,0%), <b>14</b> (99,9%)	Ependimoma, Câncer Colorretal
rs757793630	c.5637A>C p.Ser1879=	Sinônima	13	-	<b>16</b> (33,0%), <b>17</b> (32,2%), <b>18</b> (25,9%), <b>19</b> (30,0%), <b>21</b> (31,7%), <b>22</b> (34,7%), <b>23</b> (31,1%), <b>24</b> (29,0%), <b>25</b> (28,6%), <b>26</b> (33,0%), <b>27</b> (31,5%), <b>28</b> (25,1%), <b>30</b> (28,0%)	NR
rs768860607	c.6719-22T>C p.?	Intrônica	13	-	<b>17</b> (14,4%), <b>18</b> (14,1%), <b>19</b> (13,3%), <b>20</b> (14,0%), <b>21</b> (7,6%), <b>22</b> (7,6%), <b>23</b> (21,6%), <b>24</b> (13,7%), <b>25</b> (9,3%), <b>26</b> (9,0%), <b>27</b> (10,0%), <b>28</b> (16,3%), <b>30</b> (12,0%)	NR
rs367591282	c.5653A>T p.Ile1885Phe	Codificante	13	-	<b>17</b> (34,4%), <b>18</b> (30,5%), <b>19</b> (32,3%), <b>20</b> (48,8%), <b>21</b> (50,6%), <b>22</b> (46,2%), <b>23</b> (36,9%), <b>24</b> (40,0%), <b>25</b> (46,2%), <b>26</b> (47,1%), <b>27</b> (44,3%), <b>28</b> (35,2%), <b>30</b> (25,9%)	NR

rs1381057	c.7538A>G p.Gln2513Arg	Missense	11	Benigna	<b>1</b> (100,0%), <b>2</b> (99,9%), <b>3</b> (73,8%), <b>4</b> (99,9%), <b>5</b> (51,6%), <b>7</b> (99,8%), <b>8</b> (99,9%), <b>9</b> (65,0%), <b>12</b> (100,0%), <b>13</b> (99,8%), <b>4</b> (48,6%)	Leucemia Linfoblástica B, Leucemia Mieloide Aguda, Ependimoma, Câncer Colorretal
rs2030531	c.7390-76G>A p.?	Intrônica	11	-	<b>1</b> (99,9%), <b>2</b> (47,0%), <b>3</b> (72,9%), <b>4</b> (99,9%), <b>5</b> (57,5%), <b>7</b> (99,9%), <b>8</b> (99,9%), <b>9</b> (21,5%), <b>12</b> (99,9%), <b>13</b> (99,9%), <b>14</b> (45,3%)	NR
rs13059229	c.1109-62G>C p.?	Intrônica	11	Benigna	<b>1</b> (97,9%), <b>2</b> (70,2%), <b>3</b> (77,9%), <b>4</b> (100,0%), <b>5</b> (42,8%), <b>7</b> (99,7%), <b>8</b> (100,0%), <b>9</b> (47,2%), <b>12</b> (100,0%), <b>13</b> (99,5%), <b>14</b> (48,5%)	Câncer Colorretal, Câncer de cabeça e pescoço
rs1381058	c.6846-56T>C p.?	Intrônica	10	-	<b>1</b> (100,0%), <b>2</b> (28,3%), <b>3</b> (89,6%), <b>4</b> (100,0%), <b>5</b> (51,5%), <b>7</b> (100,0%), <b>8</b> (100,0%), <b>12</b> (100,0%), <b>13</b> (99,5%), <b>14</b> (43,2%)	NR
rs3218649	c.2945C>G p.Thr982Arg	Missense	10	Benigna	<b>1</b> (99,9%), <b>2</b> (59,5%), <b>3</b> (81,4%), <b>4</b> (100,0%), <b>5</b> (47,1%), <b>7</b> (100,0%), <b>8</b> (99,9%), <b>12</b> (100,0%), <b>13</b> (99,8%), <b>14</b> (42,8%)	Leucemia Linfoblástica B, Leucemia Mieloide Aguda, Ependimoma, Câncer Colorretal
rs3218651	c.3602A>G p.His1201Arg	Missense	5	Benigna	<b>3</b> (36,0%), <b>4</b> (50,3%), <b>8</b> (99,9%), <b>12</b> (99,3%), <b>13</b> (53,5%)	Câncer Colorretal, Câncer de esôfago, Câncer de Pulmão
rs761787398	c.5774-1G>A p.?	Região de Splicing	5	VUS	<b>17</b> (22,2%), <b>21</b> (13,5%), <b>22</b> (11,4%), <b>25</b> (12,5%), <b>26</b> (12,1%)	NR
rs3218636	c.2884A>G p.Ser962Gly	Missense	4	-	<b>3</b> (24,8%), <b>4</b> (58,8%), <b>8</b> (99,8%), <b>12</b> (100,0%)	NR
rs41544013	c.1959+8C>G p.?	Região de Splicing; Intrônica	2	-	<b>2</b> (46,4%), <b>4</b> (27,5%)	NR

rs3218634	c.7612C>G p.Leu2538Val	Missense	2	Benigna	<b>6</b> (45,7%), <b>10</b> (45,1%)	Câncer de Pulmão de Pequenas Células
rs532411	c.6911C>T p.Ala2304Val	Missense	2	Benigna	<b>6</b> (45,7%), <b>10</b> (53,8%)	Câncer de Pulmão de Pequenas Células, Adenocarcinoma de Colon, Câncer de Tireoide
rs487848	c.1742C>T p.Ala581Val	Missense	2	Benigna	<b>6</b> (51,1%), <b>10</b> (50,1%)	Câncer de Pulmão de Pequenas Células, Câncer Colorretal
rs377717694	c.7248C>T p.Phe2416=	Sinônima	1	Provavelmente Benigna	<b>1</b> (52,2%)	Doenças Genéticas Congênitas
rs3218647	c.6406-16C>T p.?	Intrônica	1	-	<b>2</b> (53,2%), <b>22</b> (15,0%)	NR
rs3218642	c.4141C>A p.Pro1381Thr	Missense	1	-	<b>2</b> (56,8%)	NR
rs35766343	c.1323G>A p.Ala441=	Sinônima	1	-	<b>2</b> (47,8%)	NR
rs36065146	c.543C>T p.Phe181=	Sinônima	1	Benigna	<b>2</b> (55,4%)	Câncer de Pulmão de Pequenas Células, Linfoma Linfoblástico de Células T
rs375451342	c.1598G>A p.Arg533Gln	Missense	1	Provavelmente Benigna	<b>3</b> (5,8%)	Câncer Colorretal, Câncer de Endométrio
rs41540016	c.7390G>A p.Ala2464Thr	Missense;Região de Splicing	1	-	<b>5</b> (42,3%)	NR
rs3218650	c.7153-71C>T p.?	Intrônica	1	NR	<b>5</b> (43,4%)	NR
rs61734794	c.5350G>T p.Asp1784Tyr	Codificante	1	-	<b>17</b> (20,0%)	NR

rs751970101	c.4685A>T p.Asp1562Val	Missense	1	-	<b>17</b> (12,5%)	NR
rs761966065	c.6406-35T>C p.?	Intrônica	1	-	<b>26</b> (9,8%)	NR
rs192758267	c.5408G>A p.Ser1803Asn	Missense	1	-	<b>30</b> (20,0%)	NR

**Legenda.** RS: número de referência da variante no banco de dados dbSNP do NCBI; Variante/alteração no aminoácido: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; NR: variante não relatada nos bancos de dados; Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora

#### 4.2.2. Caracterização, análise *in silico* da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene *REV3L*

Ao analisarmos o gene *REV3L*, identificamos 284 variantes, das quais 221 foram exônicas e 63 intrônicas. Ao classificarmos as variantes de acordo com a consequência no transcrito ou na proteína, obtivemos o seguinte resultado: 145 variantes *missense*, 50 variantes *synonymous*, 17 variantes *frameshift*, 1 variante *inframe*, 8 variantes *stop gained* e 1 variante do tipo *Splice Donor Variant*.

##### 4.2.2.1. Variantes com Potencial Perda de Função (LoF)

Durante a avaliação de patogenicidade, foram detectadas variantes cujas consequências são seriamente prejudiciais para a proteína resultante (variantes potencialmente *loss of function*, perda de função). São elas: 17 variantes *frameshift*, 1 variante *inframe*, 8 variantes do tipo *stop gained*, e 1 *Splice Donor Variant* (tabela 9). Todas essas variantes são consideradas como Claramente Patogênicas devido às consequências de alto potencial deletério, as quais comumente comprometem processos transcricionais e traducionais, gerando proteínas truncadas ou não funcionais (tabela 8)

A maioria das variantes identificadas no gene *REV3L* demonstraram-se **Provavelmente Oncogênicas** pela classificação SOP. Isso se dá, provavelmente, porque já são descritas na literatura três variantes LoF, patogênicas, que causam perda de função neste gene, e que são causadoras de NMD (do inglês, *nonsense-mediated decay*), fenômeno celular que causa a degradação de mRNA em genes que contêm códons de parada prematuros. A tradução desses mRNAs aberrantes poderia, em alguns casos, levar a ganhos de funções deletérias ou atividade dominante negativa das proteínas resultantes. Essa característica de mutações com perda de função causadoras de NMD no gene *REV3L* garante a pontuação de oncogenicidade atribuída a esta variante.

Dentre as variantes identificadas, **cinco** variantes de consequência *frameshift* foram detectadas em múltiplos pacientes e classificadas como **Provavelmente Oncogênicas** pelos critérios SOP. A frequência dessas variantes não foi encontrada no banco de dados populacional gnomAD (critério OP4) e todas as ferramentas computacionais suportam a patogenicidade desta variante (OP1), além disso, são variantes causadoras de NDM e este foi o principal critério para a classificação destas

variantes como **Provavelmente Oncogênicas**. A maioria dos pacientes com estas variantes foram diagnosticados com SMD com excesso de blastos 1 ou 2 (tabela 8).

O gene *REV3L* apresentou **sete variantes** com perda de função que foram classificadas como **variantes de significado incerto – VUS** seguindo os critérios SOP. Quatro das seis variantes foram detectadas em um pequeno número de pacientes, enquanto as variantes *REV3L*(NM\_002912.4):c.7485del (p.Phe2496LeufsTer36) e *REV3L*(NM\_001372078.1):c.7804G>T (p.Glu2602Ter), foram identificadas em múltiplos pacientes. A frequência destas variantes não foi encontrada no gnomAD (critério OP4), e, apesar de ser uma variante *nonsense*, que causa perda de função e garantiria a pontuação (OP1), a alteração não ocorreu em uma região possivelmente causadora de NMD e por este motivo, classificamos estas variantes como de significado incerto (tabela 8)

Foram identificadas, em pacientes individuais deste estudo, **12 mutações** distintas classificadas como **Provavelmente Oncogênicas** pelos critérios SOP, que apresentavam valores de VAF menores (de 2,7 a 16%) do que os valores apresentados por outras variantes no gene *REV3L*. Valores menores de VAF são geralmente atribuídos a mutações passageiras subclonais que contribuem com a instabilidade genômica da célula (tabela 8).

**Tabela 8.** Análise in silico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene *REV3L*

VARIANTE	PACIENTE ID/VAF	SOP (ACMG/AMP/VICC)	MUTATION TASTER	SIFT	PROVEAN	FATHH	ESCORE TOTAL
REV3L(NM_002912.4):c.784del (p.Gln262ArgfsTer20)	2(6,0%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.3078del (p.Val1027PhefsTer14)	15(6%),17(8,5%),18(9,6%),19(8,4%),20(4,9%),21(12,7%),22(7,3%),24(9,2%),25(9,6%),27(9,3%),28(6,4%),30(8,7%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.3070del (p.Trp1024GlyfsTer17)	15(7,6%),17(4%),18(12,7%),19(4,9%),20(12,7%),21(8,8%),22(3,2%),23(11,8%),24(4,1%),25(6,9%),26(10,7%),27(9%),30(4,8%),30(9,2%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.3081del (p.Asp1029ThrfsTer12)	17(2,7%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.3074del (p.Pro1025GlnfsTer16)	17(2,7%),22(4,2%),24(4,1%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.2494_2495del (p.Leu832GlufsTer2)	17(3,6%),18(8,1%),21(6,7%),24(5,8%),26(5,7%),27(5,6%),30(4,8%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001286432.1):c.2259del (p.Lys753AsnfsTer2)	17(10,3%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.2166dup (p.Gly723ArgfsTer3)	17(12,9%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.441del (p.Asn147LysfsTer26)	17(10,5%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.7485del (p.Phe2496LeufsTer36)	18(20,6%),22(8,5%),26(12,9%),28(13,2%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.6491del (p.Ile2164LysfsTer17)	18(9,3%),26(9,1%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.3641del (p.Thr1214LysfsTer18)	19(13,2%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-

REV3L(NM_002912.4):c.8818del (p.Arg2940AlafsTer27)	22(11,4%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.1927del (p.Glu643ArgfsTer70)	30(7,5%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.2489del (p.Leu830Ter)	15(5,6%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.1926_19 31del (p.Lys642_Glu643del)	7(43,6%)	VUS	Polimorfismo	Tolerada	Neutra	Tolerada	0
REV3L:g.103735G>A	9(11,6%)	-	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.4279C> T (p.Gln1427Ter)	9(11,1%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.3114T> A (p.Tyr1038Ter)	15(38%),16(35%),17(36,4%)18 (29,3%),19(50,8%)20(30%),21( 38,3%),22(35,2%),23(35%),24( 35,3%),25(38,5%),26(39,3%),2 7(35,5%),28(27,9%),30(21,7%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.2781T> G (p.Tyr927Ter)	15(11,9%),18(10,7%),21(14,6% ,24(9,4%),25(10,3%),28(15,3% ,30(18,5%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.3105T>A (p.Tyr1035Ter)	17(5%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.3468T> A (p.Tyr1156Ter)	28(7,9%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.7804G> T (p.Glu2602Ter)	1(18%),2(15,2%),3(15,5%),4(1 6,5%),5(15,8%),6(15,4%),7(15, 2%),8(15,4%),9(17,7%),10(17 ,11(16,8%),12(14%),13(15,4 ,14(16,3%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.8659A> T (p.Lys2887Ter)	4(5,3%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_001372078.1):c.676G>T (p.Glu226Ter)	11(16,8%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV3L(NM_002912.4):c.9253-6T>C COSM5424379/5424380	3(16%)	VUS	intron	intron	intron	intron	-

REV3L(NM_001372078.1):c.2482T> C (p.Ser828Pro) COSM5011659,COSM5011660	15(48,1%),17(51,1%),18(56,5%),19(56,7%),20(45,8%),21(46,3%),22(50,6%),23(47,6%),24(46,4%),25(46,2%),26(48,3%),27(52,1%),28(52,2%),30(50,7%)	Provavelmente Benigna	Causadora de doença	Prejudicial	Neutra	Tolerada	2
REV3L(NM_001372078.1):c.2787T> C(p.Thr929=) COSM3381403,COSM3381404	21(34,8%),26(30%),28(11,9%)	Provavelmente Benigna	Causadora de doença	Neutra	Neutra	Tolerada	1
REV3L(NM_001372078.1):c.1954G> A (p.Val652Ile) COSM3619574,COSM3619575	21(7,1%)	Provavelmente Benigna	Polimorfismo	Prejudicial	Neutra	Tolerada	1

Variante: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; VUS: variante de significado incerto; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; PROVEAN/SIFT/MutationTaster/MutPred: resultado da análise de patogenicidade *in silico* das variantes de acordo com cada ferramenta. Escore total: número total de ferramentas que concordaram entre si acerca do potencial patogênico da variante. Publicações: número de referências encontradas para determinada variante na literatura. Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora.

#### 4.2.2.2. Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)

Seguindo os critérios idealizados neste estudo, realizamos o levantamento nos bancos de dados dbSNP e COSMIC. Dentre as variantes identificadas, 28 foram registradas no banco de dados dbSNP e classificadas como variantes polimórficas e duas foram registradas no COSMIC e classificadas como variantes com potencial oncogênico.

A **primeira variante** com potencial oncogênico, REV3L(NM\_002912.4):c.9253-6T>C (COSM5424379/5424380), é um SNV intrônico que ocorreu em uma região de splicing e se caracteriza por uma troca de A por G. Não foi possível realizar a análise de predição de patogenicidade, pois a variante ocorre em uma região de intron, no entanto, não encontramos sua frequência na base de dados gnomAD e a literatura reportou a variante como Benigna Moderada (algoritmo de predição DANN), recebendo a classificação de oncogenicidade como **VUS** (tabela 8).

A **segunda variante**, REV3L(NM\_001372078.1):c.2482T>C(p.Ser828Pro) (COSM5011659, COSM5011660), é uma troca de serina por prolina, do tipo *missense*. Ao realizar a análise de predição, a variante foi prevista como deletéria por 2 dos quatro algoritmos utilizados, conseguindo a pontuação. No entanto, a variante já foi descrita como oncogênica em um paciente com Melanoma. A classificação de oncogenicidade indicou a variante como **Provavelmente Benigna** e a mesma já foi relatada em um paciente com Carcinoma papilífero de células renais (tabela 8).

A **terceira variante**, REV3L(NM\_001372078.1):c.2787T>C(p.Thr929=), apesar de ser apontada pelo COSMIC como potencialmente oncogênica, é uma variante sinônima e não tem como consequência a troca de aminoácido. Ao realizarmos a análise de predição, a variante foi prevista como não prejudicial por 3 dos quatro algoritmos utilizados. Foi classificada como **Provavelmente Benigna** pelos critérios de oncogenicidade (SOP-ACMG/AMP) e já foi relatada em carcinoma ductal de pâncreas (tabela 8).

A **quarta variante** com potencial oncogênico, REV3L(NM\_001372078.1):c.1954G>A(p.Val652Ile)(COSM3619574, COSM3619575), é uma troca de C por T, do tipo *missense*. Tem como consequência a troca do aminoácido Valina por Isoleucina (p.Val652Ile). Ao realizar a análise de predição, a variante foi prevista como deletéria por apenas um dos quatro algoritmos utilizados. A classificação de oncogenicidade indicou a variante como **Provavelmente Benigna**, seguindo,

principalmente, o critério que garante benignidade de suporte a variantes *missense* ocorrendo em um gene para o qual variantes *nonsense* são conhecidas por causar doenças. No entanto a mesma variante já foi descrita como oncogênica em um paciente com Melanoma (tabela 8).

A caracterização e classificação das variantes polimórficas identificadas no gene *REV3L* nas amostras analisadas estão descritas na tabela abaixo (tabela 9).

**Tabela 9.** Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene *REV3L*

POLIMORFISMOS (SNPs) NO GENE <i>REV3L</i> IDENTIFICADOS NAS AMOSTRAS ANALISADAS						
POLIMORFISMO	VARIANTE ALTERAÇÃO NO AMINOÁCIDO	CONSEQUÊNCIA DA VARIANTE	Nº DE PACIENTES COM A VARIANTE	SOP (ACMG/AMP/VICC)	PACIENTE ID/VAF	SIGNIFICADO CLÍNICO
rs461646	c.3301C>T p.Leu1101=	Codificante	14	Benigna	<b>1</b> (25,8%), <b>2</b> (100,0%), <b>3</b> (100,0%), <b>4</b> (50,6%), <b>5</b> (43,6%), <b>6</b> (99,8%), <b>7</b> (99,9%), <b>8</b> (48,9%), <b>9</b> (98,7%), <b>10</b> (99,9%), <b>11</b> (51,4%), <b>12</b> (100,0%), <b>13</b> (99,8%), <b>14</b> (99,9%)	Leucemia Mieloide Aguda; Câncer Colorretal; Câncer de Pulmão
rs462779	c.3671C>T p.T1224I	Missense	14	Benigna	<b>1</b> (100,0%), <b>2</b> (100,0%), <b>3</b> (100,0%), <b>4</b> (50,5%), <b>5</b> (50,5%), <b>6</b> (100,0%), <b>7</b> (99,9%), <b>8</b> (45,4%), <b>9</b> (99,8%), <b>10</b> (100,0%), <b>11</b> (50,7%), <b>12</b> (100,0%), <b>13</b> (99,9%), <b>14</b> (99,9%), <b>18</b> (31,4%), <b>21</b> (30,0%), <b>27</b> (43,5%)	Leucemia Mieloide Aguda; Carcinoma Hepatocelular, Ependimoma
rs455645	c.4485G>A p.Ser1495=	Sinônima	14	Benigna	<b>1</b> (57,5%), <b>2</b> (100,0%), <b>3</b> (100,0%), <b>4</b> (52,8%), <b>5</b> (45,3%), <b>6</b> (100,0%), <b>7</b> (99,9%), <b>8</b> (59,3%), <b>9</b> (99,5%), <b>10</b> (99,9%), <b>11</b> (49,8%), <b>12</b> (100,0%), <b>13</b> (99,9%), <b>14</b> (100,0%)	Leucemia Mieloide Aguda; Câncer Colorretal
rs458486	c.2706A>G p.Gly902=	Sinônima	11	Benigna	<b>1</b> (16,9%), <b>3</b> (53,3%), <b>4</b> (49,5%), <b>5</b> (48,2%), <b>8</b> (54,1%), <b>9</b> (99,4%), <b>10</b> (96,3%), <b>11</b> (47,3%), <b>12</b> (99,9%), <b>13</b> (48,6%), <b>14</b> (100,0%)	Leucemia Linfoblástica B, Neoplasia Linfoide, Ependimoma
rs455732	c.4290G>A p.Val1430=	Sinônima	11	Benigna	<b>1</b> (15,7%), <b>3</b> (52,0%), <b>4</b> (44,8%), <b>5</b> (46,3%), <b>8</b> (54,2%), <b>9</b> (99,4%), <b>10</b> (98,9%), <b>11</b> (49,0%), <b>12</b>	Leucemia Linfoblástica B, Câncer de Pulmão de Pequenas Células, Ependimoma

						(100,0%), <b>13</b> (54,7%), <b>14</b> (99,9%)	
rs1623806	c.7538+57A>G p.?	Missense	11	Benigna	<b>1</b> (44,3%), <b>3</b> (46,9%), <b>4</b> (57,1%), <b>5</b> (57,0%), <b>8</b> (47,5%), <b>9</b> (99,7%), <b>10</b> (99,2%), <b>11</b> (48,7%), <b>12</b> (100,0%), <b>13</b> (43,7%), <b>14</b> (100,0%)	Leucemia Mieloide Aguda, Câncer de Pulmão de Pequenas Células	
rs3204954	c.9219G>C p.Arg3073=	Sinônima	4	Benigna	<b>1</b> (52,2%), <b>2</b> (59,5%), <b>6</b> (56,0%), <b>7</b> (56,6%)	Adenocarcinoma Esofágico	
rs3218591	c.3429A>T p.Ala1143=	Sinônima	4	Benigna	<b>1</b> (15,7%), <b>2</b> (45,2%), <b>6</b> (31,1%), <b>7</b> (99,9%)	Adenocarcinoma Esofágico	
rs3204953	c.9190G>A p.Val3064II	Missense	4	NR	<b>1</b> (8,7%), <b>2</b> (59,6%), <b>6</b> (55,9%), <b>7</b> (56,7%)	Adenocarcinoma Esofágico	
rs78491998	c.6754-4A>T p.?	Região de Splicing, Intrônica	3	-	<b>17</b> (9,8%), <b>20</b> (12,2%), <b>27</b> (14,3%)	NR	
rs770452846	c.6754-14G>T p.?	Intrônica	3	Benigna	<b>17</b> (8,6%), <b>22</b> (12,2%), <b>30</b> (11,1%)	Câncer de Esôfago	
rs769634899	c.4701T>A p.His1567Gln	Missense	3	-	<b>17</b> (6,5%), <b>20</b> (8,0%), <b>24</b> (8,0%)	NR	
rs761329594	c.1097-32C>T p.?	Intrônica	3	-	<b>18</b> (16,7%), <b>22</b> (9,8%), <b>24</b> (22,7%)	NR	
rs748063363	c.8352-30A>T p.?	Intrônica	2	-	<b>17</b> (20,0%), <b>18</b> (23,5%), <b>22</b> (20,8%), <b>24</b> (25,0%), <b>28</b> (32,3%)	NR	
rs572989019	c.139+9G>A p.?	Intronica/Upstream	1	-	<b>2</b> (55,9%)	NR	
rs458017	c.3467A>G p.Tyr1156Cys	Missense	1	Benigna	<b>2</b> (48,5%)	Câncer de Esôfago	

rs3218574	c.663-70A>G p.?	Intronica	1	Benigna	3(13,6%)	NR
rs3218606	c.5909G>A p.Arg1970His	Missense	1	NR	3(9,3%)	Câncer Colorretal
rs41289890	c.2191G>A p.Ala731Thr	Missense	1	-	6(6,4%)	NR
rs141702431	c.3559G>A p.Val1187Ile	Missense	1	-	6(54,9%)	NR
rs535670392	c.7073-117T>C p.?	Intronica	1	-	8(62,3%)	NR
rs770402288	c.180C>T p.Tyr60=	Sinônima	1	-	9(5,6%)	NR
rs192998811	c.948-19A>G p.?	Intrônica	1	-	11(47,9%)	NR
rs747822274	c.1097-19del p.?	Intrônica	1	NR	12(54,5%)	Câncer de Próstata, Câncer Colorretal
rs3218587	c.6064C>T p.Leu2022=	Sinônima	1	Benigna	6(73,7%)	Neoplasia Linfoide - Mucose Fungoide
rs139348785	c.6338A>G p.Tyr2113Cy	Missense	1	-	6(32,9%)	NR
rs779047506	c.1452G>A	Sinônima	1	-	17(30,0%)	NR
rs773259666	c.4698G>A p.Glu1566=	Sinônima	1	-	30(7,8%)	NR

**Legenda.** RS: número de referência da variante no banco de dados dbSNP do NCBI; Variante/alteração no aminoácido: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; NR: variante não relatada nos bancos de dados; Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora.

### 4.2.3. Caracterização, análise in silico da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene *POLH*

O sequenciamento completo do gene *POLH* apresentou baixa cobertura nas amostras analisadas, tendo como consequência a ausência de leitura de 09 amostras (IDs 15,16,19,20,21,23,24,26,27,) e baixa leitura nas demais amostras.

Nas amostras em que foi possível a completa ou parcial leitura, foram identificadas 22 variantes, das quais 16 localizadas em éxons e 6 em introns. Ao classificarmos as variantes de acordo com a consequência da mesma para o transcrito ou na proteína, obtivemos o seguinte resultado: 11 variantes *missense*, 4 variantes *synonymous*, 0 variantes *frameshift*, 0 variante *inframe*, 0 variantes *stop gained* e 1 variante do tipo *3\_prime\_UTR\_variant*. Esta última variante ocorre na região 3'UTR do RNA mensageiro (mRNA), a qual é seguida pelo códon de terminação da tradução. A região 3'-UTR geralmente contém regiões de regulação da expressão gênica. Por ocorrer em uma região não codificante, seu impacto para o transcrito ou a proteína é de difícil predição.

#### 4.2.3.1. Variantes com Potencial Perda de Função (LoF)

O gene *POLH* não apresentou variantes do tipo LoF, claramente patogênicas, com potencial perda de função para a proteína, possivelmente devido à baixa cobertura de sequenciamento deste gene.

#### 4.2.3.2. Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)

Dentre as variantes identificadas, 06 foram anotadas no banco de dados dbSNP e classificadas como variantes polimórficas. Não houve variantes anotadas no COSMIC Database.

As variantes do tipo SNPs que já haviam sido depositadas em bancos de dados genômicos obtiveram um escore total de baixo potencial deletério, apesar dos relatos existentes na literatura demonstrando suas associações a outros tipos de câncer, que estão demonstrados na tabela e serão discutidos posteriormente.

Destacamos duas variantes polimórficas identificadas no mesmo paciente (ID19), que já foram relatadas na literatura em Leucemia Mieloide Aguda. A primeira variante, *POLH*(NM\_006502.3):c.1434G>A (p.Thr478=), é uma troca de G por A,

anotada como uma variante polimórfica (rs3734690) e ocorre na região (ch6)43581586. Sendo uma variante do tipo Sinônima, não há substituição do aminoácido e, por consequência, não há perda da função da proteína. A segunda variante, POLH(NM\_006502.3):c.1939A>T (p.Met647Leu), caracterizada pela troca de uma base A por T, é uma variante do tipo *missense*, que leva à substituição do aminoácido metionina (Met) por leucina (Leu) (tabela 10)

As duas variantes possuem registro no banco de dados Clinvar (VCV000259987.12; VCV000259989.13) e, ao realizarmos a análise de predição utilizando os quatro algoritmos propostos, ambas foram classificadas como Polimorfismo pelo *Mutation Taster* e Neutras pelos outros três algoritmos propostos. A pontuação de oncogenicidade permitiu classificá-las como benignas (tabela 10)

A terceira variante com destaque neste estudo, POLH(NM\_006502.3):c.698A>G (p.Asn233Ser), ocorre na região (chr6)43568762. É uma variante polimórfica (rs61756403), do tipo *missense*, caracterizada pela troca de A por G. Essa substituição de base única na região indicada do gene *POLH* leva à modificação do aminoácido Asparagina (Asn) pelo aminoácido Serina (Ser) na posição 233 da cadeia polipeptídica e foi classificada como causadora de doença pelo algoritmo *Mutation Taster*, deletéria pelos algoritmos *Sift* e *Provean* e tolerada pelo algoritmo *Mutpred*. Apesar da predição ter demonstrado patogenicidade da variante, após avaliação dos critérios de oncogenicidade, a variante foi classificada como benigna (tabela 10)

A caracterização e classificação de oncogenicidade dessas variantes polimórficas identificadas no gene *POLH* estão descritas na tabela abaixo (tabela 10)

**Tabela 10.** Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene *POLH*

POLIMORFISMOS (SNPs) NO GENE <i>POLH</i> IDENTIFICADOS NAS AMOSTRAS ANALISADAS						
POLIMORFISMO	VARIANTE ALTERAÇÃO NO AMINOÁCIDO	CONSEQUÊNCIA DA VARIANTE	Nº DE PACIENTES COM A VARIANTE	SOP (ACMG/AMP/VICC)	PACIENTE ID/VAF	SIGNIFICADO CLÍNICO
rs142864589	c.1473G>A p.Gln491=	Sinônima	2	Provavelmente Benigna	<b>18</b> (17,4%), <b>30</b> (26,9%)	Xeroderma Pigmentoso
rs1064260	c.*50A>G p.?	Variante 3'-UTR	1	Provavelmente Benigna	<b>3</b> (24,9%)	Xeroderma Pigmentoso
rs2307465	-	Intrônica	1	-	<b>3</b> (5,4%)	NR
rs3734690	c.1434G>A p.Thr478=	Sinônima	1	Benigna	<b>5</b> (52,2%)	Leucemia Mieloide Aguda, Adenocarcinoma de Colon, Xeroderma Pigmentoso
rs6941583	c.1939A>T p.Met647Leu	Missense	1	Benigna	<b>5</b> (57,6%)	Leucemia Mieloide Aguda, Câncer Colorretal, Câncer de Pulmão
rs61756403	c.698A>G p.Asn233Ser	Missense	1	Benigna	<b>8</b> (49,1%)	Xeroderma Pigmentoso
rs552285614	c.1144C>T p.Arg382Cys	Missense	1	-	<b>9</b> (5,9%)	NR

**Legenda.** RS: número de referência da variante no banco de dados dbSNP do NCBI; Variante/alteração no aminoácido: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; NR: variante não relatada nos bancos de dados; Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora

#### 4.2.4. Caracterização, análise in silico da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene *POLI*

Foram identificadas, no total, 58 variantes no gene *POLI*. Das quais, 35 foram localizadas em éxons e 23 em regiões intrônicas. A classificação das variantes quanto à consequência para o transcrito ou proteína, levou ao seguinte resultado: 22 variantes *missense*, 7 variantes *Synonymous*, 3 variantes *frameshift*, 1 variante *inframe*, 1 variantes *stop gained* e 3 variantes *Splice Acceptor Variant* e 1 *Splice Donor Variant*. Variantes em regiões doadoras ou receptoras de splice podem causar truncamento da proteína ou perda de função, por este motivo são classificadas como de alto impacto pela classificação de severidade do *Ensembl Database*.

##### 4.2.4.1. Variantes com Potencial Perda de Função (LoF)

A maioria das variantes **LoF** identificadas no gene *POLI* demonstraram-se **Provavelmente Oncogênicas** pela classificação SOP. Diferentemente de *REV3L*, as variantes nesta região de *POLI*, mesmo classificadas como perda de função, não são conhecidas como causadoras de NMD, fenômeno celular que causa a degradação de mRNA em genes que contêm códons de parada prematuros. No entanto, as variantes destacadas podem afetar o domínio funcional da proteína. Essa característica das mutações em *POLI* com perda de função aqui destacadas garante a pontuação de oncogenicidade atribuída a estas variantes.

A **primeira variante** do tipo *frameshift*, *POLI*(NM\_007195.3):c.349del(p.Val117TyrfsTer2), identificada em apenas um paciente, é caracterizada pela substituição do aminoácido valina por um códon de parada, o que afeta o domínio funcional da proteína, garantindo a pontuação OVS1 (*Oncogenic Very Strong-1*) e não foi identificada no banco de dados gnomAD (OP4 - *Oncogenic Supporting*), o que, após o somatório das pontuações, a classifica como **Provavelmente Oncogênica** pelos critérios de oncogenicidade (SOP-ACMG/AMP) (tabela 11).

A **segunda variante** do tipo *frameshift*, *POLI*(NM\_007195.3):c.965\_966insT(p.Pro323ThrfsTer6), detectada em apenas um paciente, é caracterizada pela substituição de uma prolina por um códon de terminação, como consequência da inserção de uma timina na posição (chr18)51818206. A única pontuação que conseguimos atribuir a esta variante foi a (OP4 - *Oncogenic Supporting*),

por estar ausente no banco de dados gnomAD. Após o somatório das pontuações, a variante foi classificada como **VUS** pelos critérios de oncogenicidade (tabela 11)

A **terceira variante** do tipo *frameshift*, POLI(NM\_007195.3):c.1185del (p.Lys396AsnfsTer2), detectada em apenas um paciente, é caracterizada pela substituição de uma lisina por um códon de terminação, como consequência de uma deleção na posição (chr18)51818206. Esta variante afeta diretamente o domínio funcional da proteína, garantindo a pontuação OVS1 (*Oncogenic Very Strong-1*), e não foi identificada no banco de dados gnomAD (OP4 - *Oncogenic Supporting*), o que, após o somatório das pontuações, a classifica como **Provavelmente Oncogênica** pelos critérios de oncogenicidade (SOP-ACMG/AMP) (tabela 11).

A **quarta variante do tipo LoF identificada no gene POLI**, POLI(NM\_007195.3):c.51\_53del (p.Asp17del) tem como consequência uma deleção do tipo *inframe*, foi identificada em 10 dos 30 pacientes e está registrada na base de dados dbSNP como polimorfismo. A análise de predição classificou a variante como não prejudicial nos quatro algoritmos utilizados, garantindo a pontuação SBP1 (*SBP1: Somatic Benign Supporting-1*), além disso a variante foi reportada no gnomAD com uma frequência alélica de 70,5%, o que confere a pontuação SBVS1 (*Somatic Benign Very Strong-1*). O somatório das pontuações permitiu classificar esta variante como **benigna** pelos critérios de oncogenicidade (SOP-ACMG/AMP). A variante já foi relatada em câncer de próstata, câncer de mama, câncer colorretal carcinoma adrenocortical e carcinoma de células escamosas (tabela 11).

A variante do tipo *stop gained*, POLI(NM\_007195.3):c.1618C>T (p.Gln540Ter), identificada no gene *POLI* em apenas um paciente (ID14), tem como consequência a substituição de uma glicina por um códon de terminação. Esta variante afeta diretamente o domínio funcional da proteína, garantindo a pontuação OVS1 (*Oncogenic Very Strong-1*), e não foi identificada no banco de dados gnomAD (OP4 - *Oncogenic Supporting*), o que, após o somatório das pontuações, a classifica como **Provavelmente Oncogênica** pelos critérios de oncogenicidade (SOP-ACMG/AMP) (tabela 11).

As demais variantes do tipo LoF foram identificadas em um número considerável de pacientes, no entanto, por estarem localizadas em regiões intrônicas não foram realizadas avaliações de predição da patogenicidade e classificação de oncogenicidade das variantes (tabela 11).

**Tabela 11.** Análise in silico de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene *POLI*

VARIANTE	PACIENTE ID/VAF	SOP (ACMG/AMP/VICC)	MUTATION TASTER	SIFT	PROVEAN	MUTPRED	ESCORE TOTAL
POLI(NM_007195.3):c.349del (p.Val117TyrfsTer2)	12(7,3%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLI(NM_007195.3):c.965_966ins T (p.Pro323ThrfsTer6)	18(5,4%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLI(NM_007195.3):c.1185del (p.Lys396AsnfsTer2)	22(9,5%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLI(NM_007195.3):c.51_53del (p.Asp17del)	1,2,3,4,6,7,8,10,12,14*	Benigna	Polimorfismo	Tolerada	Neutra	Tolerada	0
POLI(NM_007195.3):c.1618C>T (p.Gln540Ter) COSM4275745, COSM4275746	14(6,4%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
POLI:g.22428_22428delA	17(10,5%),18(11,3%),19(7,4%),20(515,9%),21(13,5%),22(12,4%),24(17%),25(14,9%),26(17,7%),27(17,3%),28(17,1%),30(12,7%)	-	intron	intron	intron	intron	-
POLI:g.22428A>G,A>T	15(47,5%),16(46%),17(36,3%),17(5,7%),18(31,1%),19(48,2%),20(43,6%),21(50%),22(47%),23(58,6%),24(40,3%),25(37,3%),26(43,3%),27(48,2%),28(39,1%),30(43,9%)	-	intron	intron	intron	intron	-
POLI(chr18):51818197ATTTA>T TTT, ATTTA>TTTTG	17(4,7%),19(10%),21(5,4%),28(4,7%)	-	intron	intron	intron	intron	-
POLI(chr18)51818198ATTA>TTT	18(4,7%)	-	intron	intron	intron	intron	-

POLI:g.142039T>C	16(23%),19(17,9%),20(12,3%),22(13,3%),23(10,3%),24(25%),25(17,5%),26(11,1%),27(28,6%),30(12,3%)	-	intron	intron	intron	intron	-
POLI(NM_007195.3):c.1185G>A (p.Gln395=)	28(12,5%)	Provavelmente Benigna	Causadora de doença	Tolerada	Neutra	Neutra	1
COSM4072584,COSM4072585							

**Legenda.** Variante: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; VUS: variante de significado incerto; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; PROVEAN/SIFT/MutationTaster/MutPred: resultado da análise de patogenicidade *in silico* das variantes de acordo com cada ferramenta. Escore total: número total de ferramentas que concordaram entre si acerca do potencial patogênico da variante. Publicações: número de referências encontradas para determinada variante na literatura. Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Células contendo “\*” referem-se a amostras nas quais não foi possível determinar o VAF. Fonte: Elaborada pela autora.

#### 4.2.4.2. Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)

Dentre as variantes identificadas, 13 foram anotadas no banco de dados dbSNP e classificadas como variantes polimórficas e 02 são variantes anotadas no COSMIC.

Destacamos a variante potencialmente oncogênica, anotada no banco de dados COSMIC (COSM4275745, COSM4275746), uma variante LoF, claramente patogênica, que foi descrita na sessão anterior (tabela 11)

A **segunda variante**, POLI(NM\_007195.3):c.1185G>A (p.Gln395=) (COSM4072584, COSM4072585), caracterizada pela substituição de uma base G por A, é uma variante do tipo *synonymous*, que não acarreta alteração no aminoácido e, por consequência, na função da proteína (indicada pelo símbolo =). A análise de predição da variante a classificou como causadora de doença por apenas um dos quatro algoritmos utilizados. Ao realizamos a classificação de oncogenicidade, a variante foi classificada como **Provavelmente Benigna** e já foi descrita em câncer esofágico (tabela 11).

A classificação das variantes polimórficas no gene *POLI* identificadas nas amostras analisadas estão descritas na tabela abaixo (tabela 12)

**Tabela 12.** Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene *POLH*

POLIMORFISMOS (SNPs) NO GENE <i>POLH</i> IDENTIFICADOS NAS AMOSTRAS ANALISADAS						
POLIMORFISMO	VARIANTE ALTERAÇÃO NO AMINOÁCIDO	CONSEQUÊNCIA DA VARIANTE	Nº DE PACIENTES COM A VARIANTE	SOP (ACMG/AMP/VICC)	PACIENTE ID/VAF	SIGNIFICADO CLÍNICO
rs8305	c.2191G>A (p.Ala731Thr)	Missense	05	-	1,2(48,5%),7(99,9%), 10(99,8%),14(53,6%)	NR
rs660793	c.1405-30G>C	Intrônica	06	-	1(11,6%),2(54,1%),7 (45,4%),9(47,4%),12 (45,2%),14(49,7%)	NR
rs610281	c.559+34C>T	Intrônica		Benigna	1(52,9%),2(53,6%),7 (53,3%),9(30,7%),12 (52,8%),14(54,1%)	Leucemia Mieloide Aguda, Carcinoma de Pulmão
rs3730688	c.241+93A>G	Intrônica	04	-	1(19%),2(49,3%),14( 40,1%)	NR
rs10584411; rs3729509; rs763063557	c.51_53del (p.Asp17del)	Inframe	08	Benigna	1,2,3,4,6,7,8,10	Cancer de Próstata, Câncer de Mama, Câncer Colorretal, Carcinoma Adrenocortical, Carcinoma de endométrio, Astrocitoma
rs202045769	c.74C>G (p.Ala25Gly)	Missense	01	-	4(100%)	NR
rs3730669	c.87G>A (p.Ala29=)	Sinônima	01	Benigna	6(37,2%)	Melanoma, Carcinoma de Intestino
rs3730809	c.1404+12_1404+13 del	Intrônica	02	-	<b>7(41,2%),12(51,4%)</b>	NR

rs34362008	c.1404+41_1404+42 del	Intrônica	02	-	<b>7(43%),12(51,4%)</b>	NR
rs770802279	c.149C>T (p.Ala50Val)	Missense	01	-	<b>9(10,9%)</b>	NR
rs3730727	c.559+92A>G	Intrônica	01	-	<b>9(30,8%)</b>	NR
rs3730689	c.241+96G>A	Intrônica	01		<b>14(59,8%)</b>	
rs771168096	c.1447G>A (p.Glu483Lys)	Missense	08	-	<b>16(61%),17(63,6%), 20(56,7%),22(43,8%) ,24(60%),25(59,3%) ,26(57,1%),28(76,9 %)</b>	NR

**Legenda.** RS: número de referência da variante no banco de dados dbSNP do NCBI; Variante/alteração no aminoácido: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; NR: variante não relatada nos bancos de dados; Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora

#### 4.2.5. Caracterização, análise in silico da patogenicidade e classificação de oncogenicidade de variantes identificadas no gene *REV1*

Foram identificadas 108 variantes no gene *REV1*, das quais 62 foram em éxons e 46 em introns. Dentre aquelas, detectadas 39 variantes *missense*, 12 variantes *synonymous*, 05 variantes *frameshift*, 01 variante do tipo *stop gained* e 02 variantes do tipo *Splice Donor Variant*. Variantes em regiões doadoras ou receptoras de splice podem causar truncamento da proteína ou perda de função, por este motivo são classificadas como de alto impacto pela classificação de severidade do *Ensembl Database*

##### 4.2.5.1. Variantes com Potencial Perda de Função (LoF)

Durante a avaliação de patogenicidade, foram detectadas variantes cujas consequências são seriamente prejudiciais para a proteína resultante (variantes potencialmente *loss of function*, perda de função). São elas: 05 variantes *frameshift*, 01 variante do tipo *stop gained* e 02 variantes do tipo *Splice Donor Variant*. Para a análise de predição in silico, todas essas variantes são consideradas como Claramente Patogênicas devido às consequências de alto potencial deletério, as quais comumente comprometem processos transcricionais e traducionais, gerando proteínas truncadas ou não funcionais (tabela 13).

A **primeira variante** de consequência ***frameshift***, *REV1(NM\_001321455.2):c.2423del(p.Leu808ArgfsTer50)*, é caracterizada pela deleção de uma timina na posição (ch2)100024512, e foi detectada em apenas um paciente dos 30 pacientes avaliados (ID03). A frequência da variante não foi encontrada no gnomAD (OP4 - *Oncogenic Supporting*), e, após avaliação dos critérios de oncogenicidade (SOP-ACMG/AMP), a variante pôde ser classificada como VUS (tabela 13).

A **segunda variante** de consequência ***frameshift***, *REV1(NM\_001321458.2):c.1648del (p.Gln550LysfsTer25)*, detectada em 01 paciente dos 30 pacientes avaliados (ID11), é caracterizada pela deleção de uma guanina que tem como consequência a troca do aminoácido lisina por um códon de terminação que afeta o domínio funcional da proteína, garantindo a pontuação OVS1 (*Oncogenic Very Strong-I*), e não foi identificada no banco de dados gnomAD (OP4 - *Oncogenic Supporting*), o que, após o somatório das pontuações, a classifica como **Provavelmente Oncogênica** pelos critérios de oncogenicidade (SOP-ACMG/AMP) (tabela 13).

A **terceira variante *frameshift***, REV1(NM\_001321454.2):c.2057del (p.Pro686GlnfsTer25), identificada no pacientes de ID 18 e 21, teve como consequência a substituição do aminoácido prolina por um códon de terminação (Ter). A única pontuação que conseguimos atribuir à esta variante foi a OP4 - *Oncogenic Supporting*, por não ter sido identificada no gnomAD. Após avaliação dos critérios de oncogenicidade (SOP-ACMG/AMP), a variante pôde ser classificada como VUS (tabela 13).

A **quarta variante com perda de função**, REV1(NM\_016316.4):c.42\_43del (p.Trp15GlyfsTer35), foi identificada em dois dos 30 pacientes avaliados (ID18 e ID27) e foi classificada como Variante de Significado Incerto (**VUS**) após pontuação seguindo os critérios SOP-ACMG/AMP (não encontrada no gnomAD) (tabela 13).

Para a **quinta variante** identificada na região REV1(chr2)100027185CT>T, não foi possível realizar a classificação e predição de patogenicidade e oncogenicidade, pois o software não foi capaz de identificar a região da variante (tabela 13).

Para a variante do tipo *Splice Donor* não foi possível realizar a classificação e predição de patogenicidade e oncogenicidade, pois a variante se encontra em uma região intrônica (tabela 13).

A variante do tipo *Stop Gained*, REV1(NM\_001321454.2):c.3547G>T (p.Asp1183Tyr), identificada em 14 dos 30 pacientes avaliados, não teve a frequência alélica identificada no gnomAD, garantindo a pontuação OP4 (*Oncogenic Supportin - 4*) a esta variante e classificando-a como VUS (tabela 13).

**Tabela 13.** Análise *in silico* de predição de patogenicidade e classificação de oncogenicidade das variantes selecionadas para o gene *REV1*

VARIANTE	PACIENTE ID/VAF	SOP (ACMG/AMP/VICC)	MUTATION TASTER	SIFT	PROVEAN	FATHH	ESCORE TOTAL
REV1(NM_001321455.2):c.2423del (p.Leu808ArgfsTer50)	3(5,7%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV1(NM_001321458.2):c.1648del (p.Gln550LysfsTer25)	11(5,5%)	Provavelmente Oncogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV1(NM_001321454.2):c.2057del (p.Pro686GlnfsTer25)	18(6,3%),21(10%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV1(NM_016316.4):c.42_43del (p.Trp15GlyfsTer35)	18(11,1%),27(19%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV1(chr2)100027185CT>T	25(14,3%),27(8,6%),28(10%)	-	error	error	error	error	-
REV1:g.54203G>A	9(5,4%),17(13,3%)	-	intron	intron	intron	intron	-
REV1(NM_001321454.2):c.3547G>T (p.Asp1183Tyr)	1(27,7%),2(25%),3(29,2%),4(28, 7%),5(27,6%),6(26,7%),7(28,3%), 8(26,4%),9(28,4%),10(27,6%),1 1(26,4%),12(28,7%),13(28,1%),1 4(25,9%)	VUS	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	Claramente Patogênica	-
REV1(NM_016316.4):c.2023A>C COSM4083698	17(10%),18(11%),22(8,8%),24(9, 3%),28(9,9%)	VUS	Polimorfismo	Deletéria	Neutra	Tolerada	1

**Legenda.** Variante: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; VUS: variante de significado incerto; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; PROVEAN/SIFT/MutationTaster/MutPred: resultado da análise de patogenicidade *in silico* das variantes de acordo com cada ferramenta. Escore total: número total de ferramentas que concordaram entre si acerca do potencial patogênico da variante. Publicações: número de referências encontradas para determinada variante na literatura. Células contendo “-” referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Células contendo “\*” referem-se a amostras nas quais não foi possível determinar o VAF. Fonte: Elaborada pela autora.

#### 4.2.5.2. Variantes Potencialmente Oncogênicas (COSMIC Database) e polimórficas (dbSNP)

Dentre as variantes identificadas, 15 foram registradas no banco de dados *Single Nucleotide Polymorphism* (dbSNP) e 01 variante identificada no banco de dados *COSMIC Database*.

A variante REV1(NM\_016316.4):c.2023A>C (COSM4083698), anotada no banco de dados COSMIC, foi identificada em 05 pacientes. É um SNV do tipo *missense*, e se caracteriza pela substituição da base T por G, que tem como consequência a troca do aminoácido treonina por prolina (p.T675P). Ao realizarmos a análise de predição, a variante foi prevista como deletéria por apenas um dos 4 algoritmos utilizados e a frequência alélica da variante não foi descrita no gnmAD, o que permitiu classificar esta variante como VUS (tabela 13).

As variantes do tipo SNPs que já haviam sido depositadas em bancos de dados genômicos foram detectadas em um número relevante dentre os pacientes incluídos no estudo e, em sua maioria, relatadas em neoplasias mieloides ou linfoides, além de tumores sólidos, contudo, o escore total obtido para esses polimorfismos revelou um baixo potencial deletério. A caracterização e classificação de oncogenicidade dessas variantes polimórficas identificadas no gene *POLQ* estão descritas na tabela abaixo (tabela 14).

**Tabela 14.** Caracterização e classificação de oncogenicidade das variantes polimórficas detectadas no gene *REVI*

POLIMORFISMOS (SNPs) NO GENE <i>POLH</i> IDENTIFICADOS NAS AMOSTRAS ANALISADAS						
POLIMORFISMO	VARIANTE ALTERAÇÃO NO AMINOÁCIDO	CONSEQUÊNCIA DA VARIANTE	Nº DE PACIENTES COM A VARIANTE	SOP (ACMG/AMP/VICC)	PACIENTE ID/VAF	SIGNIFICADO CLÍNICO
rs2305354	c.3240T>C (p.Cys1080=)	Sinônima	10	Provavelmente Benigna	1(100%),2(39,6%),3(49,4%),4(56,7%),7(43,9%),8(100%)9(100%),12(30,1%),13(64,2%),14(37,2%)	Leucemia Mieloide Aguda, Câncer de mama, Adenocarcinoma
rs717454	c.2614+15A>G p.?	Intrônica	9	Benigna	1(100%),3(31,1%),4(36,7%),7(41,9%),8(99,9%),9(14,2%),12(46,8%),13(49,7%),14(39,3%)	Leucemia Mieloide Aguda, Câncer de mama, Câncer de pulmão, Carcinoma de cólon
rs3087386	c.770T>C (p.Phe257Ser)	Missense	9	Benigna	1(100%),2(55,5%),3(34,4%),4(54,2%),7(50,3%),8(100%),12(55,8%),13(55%),14(52,1%)	Leucemia Mieloide Aguda, Câncer de mama, Câncer de pulmão, Carcinoma de esôfago, Carcinoma de pâncreas, Neoplasia de tireoide
rs3087403	c.412G>A (p.Val138Met)	Missense	3	Benigna	1(51,1%),7(46,2%),8(99,8%)	Leucemia Mieloide Aguda, Adenocarcinoma de Colon, Osteossarcoma
rs28369944	c.412G>A (p.Val138Met)	Intrônica	3	-	1(54,8%),12(41,9%),14(41,7%)	NR
rs28369943	c.412G>A (p.Val138Met)	Intrônica	10	-	1(99,9%),2(48,1%),3(46,8%),4(54,5%),7(47%),8(99,9%),9(100)	NR

						%12(41,8%),13(51,1%),14(41,6%)	
rs13394927	c.2615-49C>T	Intrônica	3	-		<b>2(55,3%),3(20,4%),9(43,1%)</b>	NR
rs13409359	c.1438+54T>G	Intrônica	2	-		<b>2(58,6%),3(30,1%)</b>	NR
rs10175852	c.1213+30G>T	Intrônica	2	-		<b>2(46,5%),3(10,3%)</b>	NR
rs3087399	c.1118A>G (p.Asn373Ser)	Missense	2	Benigna		<b>2(56,5%),3(16,2%)</b>	Adenocarcinoma de Colon
rs3087401	c.3007G>A (p.Ala1003Thr)	Missense	1	Benigna		<b>8(50,7%)</b>	Linfoma Difuso de Células B
rs201548583	c.1832-34_1832-33del	Intrônica	2	-		<b>9(74,9%),28(45,8%)</b>	NR
rs138292917	c.3046-14G>A	Intergênica	1	-		<b>11(47%)</b>	NR
rs370682495	c.216G>A (p.Met72Ile)	Missense		-		<b>13(51,4%)</b>	NR
rs779212144	c.504-38T>G	Intrônica	1	-		<b>28(19%)</b>	NR

**Legenda.** RS: número de referência da variante no banco de dados dbSNP do NCBI; Variante/alteração no aminoácido: nomenclatura da variante identificada seguindo os critérios do HGVS; ID: número de identificação do paciente; VAF: Variant allele frequency - frequência do alelo variante; SOP(ACMG/AMP): resultado do procedimento operacional padrão de classificação de oncogenicidade; NR: variante não relatada nos bancos de dados; Células contendo “-“ referem-se a informações que não constam nos bancos de dados ou análises que não foram passíveis de realização. Fonte: Elaborada pela autora

## 5. DISCUSSÃO

O advento do IPSS Molecular, publicado em 2022, demonstrou que até 94% dos pacientes com Neoplasia Mielodisplásica apresentam pelo menos uma anormalidade molecular, demonstrando um alto grau de instabilidade genômica nesta doença. A instabilidade genômica, por sua vez, pode gerar mutações aleatórias, indo de mutações pontuais no genoma, até grandes rearranjos cromossômicos, como translocações, deleções e monossomias, características compartilhadas por 40-60% dos pacientes com SMD (BERNARD et al, 2022)

Diante disso, as evidências de que genes participantes das maquinarias de reparo do DNA estão, em sua maioria, alterados na Neoplasia Mielodisplásica, tem sido crescentes. No mesmo paciente, o número de alterações no DNA (mutações ou anormalidades cromossômicas) geralmente aumenta à medida que a doença progride, potencializando o risco de evolução para Leucemia Mieloide Aguda. Isso demonstra que a célula da SMD pode se tornar incapaz de lidar com a grande quantidade de lesões conforme a doença progride (ZHOU et al, 2015)

Alguns dos processos que proporcionam a integridade do genoma são os mecanismos de resposta a danos no DNA que as células desenvolveram para remover a maioria das lesões. Apesar da proteção proporcionada pelos mecanismos de reparo, eles podem não funcionar da forma correta, devido a alterações funcionais em suas proteínas. Isso pode ocasionar o bloqueio da forquilha de replicação do DNA e gerar alterações cromossômicas. Como uma segunda opção para lidar com o dano ao DNA no caso de falha dos mecanismos de reparo, as células desenvolveram mecanismos adicionais para sintetizar através das lesões ou corrigir lesões passadas, os mecanismos de síntese translesão, os quais nos dedicamos a analisar neste estudo (RIBEZZO, 2016; VAISMAN, 2017; POWERS, 2018).

No entanto, devido à falta de acurácia das polimerases TLS, sua atuação precisa ser cuidadosamente regulada, pois alterações que causem aumento ou diminuição de seus processos de síntese translesão ou que induzam o fenótipo TLS propenso a erros, como alterações de expressão ou mutações, podem levar à indução de mutações no genoma e acelerar o processo de carcinogênese. Evidências epidemiológicas sugerem que até dois terços das mutações no câncer são causadas por erros durante a replicação do DNA e que muitas das mutações pontuais encontradas no genoma de células saudáveis e, em maior número, nas neoplásicas, sejam em virtude do mecanismo TLS, reforçando a importância da síntese translesão e do sistema de reparo do DNA. (BIELAS, 2006; WATERS, 2009; TOMASETTI et

al, 2017). A SMD se torna um modelo interessante de estudos das polimerases TLS devido ao alto número de mutações presentes nos pacientes diagnosticados.

Reportamos e discutimos aqui os resultados do primeiro estudo científico que avaliou a presença de variantes nos genes *REV3L*, *POLQ*, *POLH*, *REVI* E *POLI*, DNA Polimerases com atividade TLS, em pacientes com Neoplasia Mielodisplásica.

### **Alta incidência de variantes *LoF* nos genes *POLQ* e *REV3L* representam perda do mecanismo de Síntese Translesão na Neoplasia Mielodisplásica**

Neste estudo, identificamos um alto índice de variantes do tipo *LoF* (perda de função) nos genes *POLQ* e *REV3L*. Foram 15 variantes para o gene *POLQ* e 25 variantes para o gene *REV3L*, todas consideradas como Claramente Patogênicas devido às consequências de alto potencial deletério, as quais comumente comprometem processos transcricionais e traducionais, gerando proteínas truncadas ou não funcionais.

Datam de 2014 os primeiros estudos a demonstrarem que mutações que resultam em perda de função ou eficiência reduzida da proteína *POLQ* podem levar ao colapso de forquilhas de replicação e controle anormal do tempo de replicação, levando a célula à instabilidade cromossômica (FERNANDEZ-VIDAL, 2014).

Sabendo que os subtipos mais avançados da SMD, com excesso de blastos, exibem maior instabilidade genômica e alto risco de progressão para leucemia mieloide aguda, comparamos a incidência dessas mutações nos dois subtipos avaliados neste estudo, SMD-EB e SMD-SA. Para os dois genes avaliados, dentre os pacientes incluídos neste estudo, tivemos predominância de mutações *LoF* em pacientes com subtipos de doença avançada, SMD-EB1 e SMD-EB2 (*POLQ*=15; *REV3L*=24) comparados a pacientes com o subtipo de menor risco (SMD com sideroblastos em anel / *POLQ*=10; *REV3L*=12).

Além disso, também tivemos predominância das demais variantes caracterizadas (*missense* e *synonymous*). Em *POLQ*, 157 variantes foram identificadas em pacientes dos subtipos SMD-EB e 116 variantes identificadas em pacientes dos subtipos SMD-AS. Para o gene *REV3L* identificamos 246 variantes nos subtipos SMD-EB e 173 nos subtipos SMD-SA.

Sabemos que a polimerase theta (*POLQ*) é uma DNA polimerase de baixa fidelidade que não possui função exonucleásica 3'-5', estando envolvida tanto na via de síntese translesão, quanto na via alternativa de junções por extremidades não-homólogas, denominada theta-mediated-joining ou alt-NHEJ, que é um mecanismo de backup do reparo de quebra de DNA

de fita dupla. As vias mediadas por POLQ predominam nas células cancerígenas quando outras vias de reparo do DNA estão ausentes ou não funcionam da maneira correta (CARACCILO, 2021; CECCALDI, 2015)

A perda de POLQ, o que supomos ser o caso deste estudo pelo alto índice de mutações que conferem perda de função, sensibiliza as células à radiação ionizante e estudos mostram que camundongos com deficiência desta enzima exibem alterações cromossômicas que aumentam a instabilidade do genoma, fato que sugere um papel importante para POLQ na manutenção da integridade da célula (YOUSEFZADEH et al., 2013). Seria esta enzima uma possível nova guardiã do genoma?

Não seria surpresa determinarmos que POLQ desempenha a função de guardiã do genoma, pois os estudos demonstram que tanto o aumento da expressão quanto a perda de função de POLQ aumentam as frequências de mutações. Além disso, POLQ atua ativamente na síntese translesão e no reparo do DNA, com suas subunidades estruturais interagindo com o DNA e com demais enzimas de reparo, como RAD51 e BRCA1. Interessantemente, diversos estudos têm demonstrado que POLQ forma um complexo com o PARP-1 em uma via de letalidade sintética interagindo com BRCA1 e RAD51, o que torna esta enzima importante alvo terapêutico, com estudos clínicos já em andamento (ZATREANU et al, 2021; DRZEWIECKA, 2022)

Recentemente, publicamos um artigo em que avaliamos a expressão de genes das vias de reparo de quebra de fita simples (*XPA*, *XPC*, *XPG/ERCC5*, *CSA/ERCC8* e *CSB/ERCC6*) e quebra de fita dupla (*ATM*, *BRCA1*, *BRCA2*, *RAD51*, *XRCC5*, *XRCC6*, *LIG4*) em 92 pacientes com Neoplasia Mielodisplásica. Demonstramos a diminuição de expressão do eixo BRCA1-BRCA2-RAD51 em pacientes com SMD e corroboramos este achado com uma análise *in silico* no banco de dados GEPIA Database, onde demonstramos um padrão de diminuição da expressão desses genes em 30 tipos de câncer. Neste estudo sugerimos que a abordagem de letalidade sintética pode ser testada com genes de reparo de DNA (além do status *BRCA1/2*) para Neoplasia Mielodisplásica e encorajamos ensaios clínicos avaliando o uso de inibidores de PARP1 na SMD (RIBEIRO-JUNIOR et al, 2019)

Diante dos resultados aqui apresentados, com perda de função da proteína POLQ devido à mutações LoF e diante de estudos que demonstram a atuação desta enzima em vias de letalidade sintética, interagindo com BRCA1/2 e PARP-1, nos questionamos se essas mutações que garantem perda de função à POLQ criam a oportunidade de se empregar a abordagem de

letalidade sintética nos pacientes com Neoplasia Mielodisplásica, assim como foi aplicada, e ainda é até hoje, com terapias aprovadas, em células cancerígenas que abrigam mutações em *BRCA1* e *BRCA2* usando inibidores de PARP1 (PARPi) (FARMER et al, 2005)

Encontramos relatos na literatura de alterações na função de *POLQ*, seja como consequência de mutações ou expressão gênica e proteica em diversos tumores sólidos, tais como câncer de pulmão, estômago e reto, sendo o aumento de expressão associado a pior sobrevida nestes pacientes (KAWAMURA et al, 2004; LEONCINI et al, 2014; LEMEE, 2010). Esse aumento de expressão é ainda mais característico em neoplasias com deficiência dos mecanismos de recombinação homóloga. Para câncer de mama, por exemplo, é estimado que, aproximadamente, 70% dos tumores sejam caracterizados pelo aumento da expressão de *POLQ* (HIGGINS et al, 2010).

Nas neoplasias mieloides, no entanto, há poucos estudos sendo conduzidos que objetivam avaliar a função desta enzima no processo de carcinogênese destas doenças. Não encontramos, na literatura, nenhum relato de mutações no gene *POLQ* em pacientes com Neoplasia Mielodisplásica, sendo este estudo o primeiro a relatar e caracterizar as variantes presentes no gene em pacientes com SMD.

A proteína REV3L (subunidade catalítica de POLZ) é reconhecida como a principal polimerase de síntese translesão, pois é capaz de contornar um amplo espectro de lesões. No entanto, não possui atividade exonucleásica de revisão, o que pode contribuir para uma síntese translesão propensa a erros. POLZ/REV3L tem sido associado à carcinogênese em cânceres de mama, pulmão, gliomas e estômago e na modulação da sensibilidade à cisplatina (MAKAROVA et al, 2015; VARADI et al, 2011). No que concerne a presença de mutações, em neoplasias mieloides, um estudo realizado em pacientes com leucemia promielocítica aguda identificou mutações no gene *REVL3* em dois de 41 pacientes avaliados (RIVA et al, 2013), diferentes das mutações que identificamos no presente estudo.

O único relato de mutações neste gene identificada em paciente com Neoplasia Mielodisplásica foi publicado pelo nosso grupo de pesquisa como um relato de caso para a mutação REV3L(NM\_002912.4):c.9253-6T>C, identificada no paciente ID03. O paciente apresentava hemoglobina muito baixa, número aumentado de blastos, cariótipo:47,XY,+8[6]/47,XY,del(7)(q32),+8[7] e sem resposta à terapia hipometilante (decitabina). A variante teve VAF de 16% e foi considerada patogênica pela análise funcional realizada à época da publicação (Análise Funcional através de Modelos ocultos de Markov –

FATHMM). Ao realizarmos a classificação de oncogenicidade desta variante, por estar ausente no gnomAD e ter sido classificada como patogênica pelos algoritmos utilizados, foi possível considerar as pontuações OP1 (*Oncogenic Supporting-1*) e OP4 (*Oncogenic Supporting-4*) para esta variante, classificando-a como variante de significado incerto (VUS). Para as demais neoplasias mieloides, um estudo realizado em pacientes com leucemia promielocítica aguda identificou mutações no gene *REVL3* em dois de 41 pacientes avaliados (OLIVEIRA et al, 2020)

Corroborando com nosso estudo, uma pesquisa conduzida por Huang e colaboradores, em 2020, publicada na revista BMC Medical Genetics, realizou uma abordagem de bioinformática para identificar mutações em polimerases no banco de dados TCGA. A análise feita pelos pesquisadores demonstrou uma maior frequência de mutações em *POLQ* e *REV3L/POLZ* em comparação com outras polimerases. (HUANG et al, 2020). Um segundo estudo publicado por Huang e demais pesquisadores, em 2016, sequenciou o exoma de pacientes com carcinoma espinocelular de cabeça e pescoço e identificou predominância de mutações do tipo perda de função no gene *REV3L*, as quais conferiam aos pacientes sensibilidade ao tratamento com a droga dacomitinib (HUANG et al, 2016).

No presente estudo, a alta incidência de mutações que conferem perda de função às proteínas *POLQ* e *REV3L* evidencia o importante papel destas enzimas na manutenção da estabilidade do genoma, uma vez que a instabilidade genômica é um passo importante para o início, progressão e evolução da Neoplasia Mielodisplásica. Nos questionamos, no entanto, sobre a possibilidade de os mecanismos de tolerância ao dano estarem sendo afetados pela instabilidade genômica intrínseca à doença, o que representaria um grave problema, pois nessa conjuntura, ambas as vias, reparo e tolerância aos danos, que protegem as células de lesões no DNA, se encontrariam insuficientes em pacientes com SMD, predispondo às conhecidas alterações cromossômicas, de grande impacto prognóstico para a SMD.

As análises realizadas neste estudo nos permitem hipotetizar e propor que *POLQ* e *REV3L* podem cooperar na geração de instabilidade genômica nos pacientes com Neoplasia Mielodisplásica por falha no mecanismo de síntese translesão devido ao alto índice de mutações que conferem perda de função a estas proteínas. Mecanismo este o mais potencialmente patogênico, dentre os diversos sistemas de reparo do DNA, e que, até então, não havia sido descrito na Neoplasia Mielodisplásica.

### **A alta taxa de variantes em *REVI* e *POLI* pode representar um biomarcador da instabilidade genômica na SMD?**

Os genes *REVI* e *POLI* não apresentaram alta incidência de mutações do tipo perda de função, como os genes *REV3L* e *POLQ*. No entanto, a presença de altas taxas mutacionais em múltiplos pacientes e, especificamente, de mutações do tipo *missense* (134 para o gene *REVI* e 14 para o gene *POLI*), as quais alteram o aminoácido codificado, pode representar um aumento da instabilidade do genoma de pacientes com Neoplasia Mielodisplásica, refletindo, inclusive, a falha dos mecanismos de reparo e a ausência do mecanismo TLS mediado por *REV3L* e *POLQ*. Além disso, uma vez ausente a síntese TLS de *POLQ* e *REV3L*, a célula recorrerá a outras polimerases que desempenhem a mesma atividade, como *REVI* e *POLI*, as quais podem ter suas funções prejudicadas como consequência de mutações *missense*, ou, pior ainda, poderiam estar atuando através do mecanismo de síntese translesão propenso a erros.

Nossos achados demonstram que *REV1* e *POLI* podem desempenhar um papel fundamental na integridade do sistema hematopoiético. Alguns estudos na literatura mostram a importância desses genes para a manutenção da integridade genômica nessas células. Para relatar a importância de *REV1* para a manutenção do sistema hematopoiético, citamos aqui um estudo colaborativo publicado por Martín-Pardillos e colaboradores (2017) que questionou se as polimerases TLS, mais especificamente *REV1*, poderiam desempenhar um papel protetor no sistema hematopoiético. Os pesquisadores realizaram um transplante de HSCs em camundongos irradiados com *knockout* de *REV1*, para investigar como a depleção de *Rev1* pode prejudicar o sistema hematopoiético, limitando sua capacidade funcional e induzindo alterações genômicas que predisõem a neoplasias hematológicas. (MARTIN-PARDILLOS et al, 2017). Os camundongos desenvolveram falência progressiva da medula e apresentaram citopenias leves em todas as linhagens. Além disso, o nocaute combinado de *REV1* e *XPC*, outra enzima de reparo do DNA, resultou em deficiência do reparo do genoma global por excisão de nucleotídeos (ggNER) com aplasia de medula óssea, aumento do estresse oxidativo e, finalmente, levou os camundongos à anemia aplásica entre três e quatro meses de idade (MARTIN-PARDILLOS et al, 2017). Adicionalmente, outros estudos funcionais demonstraram que linhagens *knockout* para genes da via da Anemia de Fanconi (FA) e genes das vias de reparo, exibiram acentuados níveis de Pol  $\iota$  (enzima referente ao gene *POLI*) como uma tática de escape contra a instabilidade genética que poderia levar a célula à apoptose.

Nosso estudo ratifica a importância dessas enzimas na estabilidade do genoma de células da medula óssea. Diferentemente de *REV3L* e *POLQ*, não foi identificada

predominância de variantes em algum dos subtipos de SMD em detrimento do outro. Para *REV1*, pacientes dos subtipos SMD-EB apresentaram 81 variantes e pacientes diagnosticados com o subtipo SMD-SA apresentaram 73 variantes. Para o gene *POL1*, foram identificadas 46 variantes nos subtipos SMD-EB e 41 variantes nos subtipos SDM-SA.

Dentre os resultados obtidos, destaca-se a presença da variante REV1(NM\_001321454.2):c.3547G>T (p.Asp1183Tyr), variante do tipo *stop gained*, que confere perda de função à proteína. Esta variante foi identificada em 14 pacientes de ambos os subtipos incluídos em nosso estudo, com valores de VAF que variavam de 25-28,7% e, após análise dos critérios de oncogenicidade, foi classificada como VUS – variante de significado incerto. Considerando a ausência de referências na literatura a respeito da variante encontrada, seu elevado potencial patogênico previsto e uma variante de frequência alélica relativamente alta em todos os pacientes nas quais foi identificada, acreditamos que esta variante pode ser importante candidata para estudos adicionais na Neoplasia Mielodisplásica.

Atualmente, a polimerase TLS REV1 é mais notavelmente reconhecida pela sua atividade regulatória. Mais especificamente, essa polimerase atua como uma proteína de *scaffold* para as demais enzimas envolvidas na síntese translesão e possui capacidade de interagir simultaneamente com outros componentes do complexo diretamente envolvidos no processo de ultrapassagem do dano, como Rev7 (subunidade estrutural de POLZ), Pol η e Pol κ, tendo assim um papel indispensável no agrupamento do complexo multi-proteico constituinte do mecanismo TLS, denominado de mutassoma. Tais evidências sugerem que ela facilita não só o recrutamento, como também a regulação das polimerases TLS, de forma similar à PCNA (RIZZO et al, 2019)

### **Acúmulo de mutações LoF e ausência de polimorfismos benignos nos genes *POLQ* e *REV3L* em paciente com cariótipo complexo**

Alterações cromossômicas são observadas em até 60% dos pacientes com Neoplasia Mielodisplásica, sendo mais frequentes as alterações de perda ou ganho, como deleções nos cromossomos 5 e 7, trissomia do cromossomo 8 e nulissomia do cromossomo Y, em oposição às translocações balanceadas mais comuns às leucemias agudas. Pacientes com cariótipos complexos, por exemplo, apresentam péssimo prognóstico, com sobrevida inferior a 6 meses. Se as polimerases com atividade translesão são mediadoras de um mecanismo que se propõe a evitar alterações cromossômicas por paradas na forquilha de replicação, quais seriam as

associações destas alterações citogenéticas, tão comuns à SMD, com as mutações nas DNA Polimerases TLS?

Identificamos neste estudo que o paciente **ID15**, diagnosticado com SMD-EB1, portador de cariótipo complexo, incluindo deleção e monossomia dos cromossomos 7 e 5, classificado como de muito alto risco pelo IPSS-R e exposto a agrotóxicos há mais de 10 anos, apresentou acúmulo de mutações do tipo LoF nos genes *POLQ* (4 *frameshifts*) e *REV3L* (5 *frameshifts*). Três das quatro mutações em *POLQ* foram classificadas como VUS e as cinco mutações com perda de função identificadas neste paciente foram classificadas como Potencialmente Oncogênicas pelos critérios SOP.

Como explicitado anteriormente, mutações LoF são caracterizadas pelo alto potencial patogênico, gerando proteínas truncadas e causando perda de sua função. Os achados para este paciente nos fazem hipotetizar que a função protetora da via de tolerância ao dano na prevenção de eventos mutacionais mais graves, como aberrações cromossômicas desbalanceadas, deleções e rearranjos, mediada por *POLQ* ou *REV3L* estaria totalmente disfuncional neste paciente. Nesse sentido, é possível inferir que a alta frequência de variantes claramente patogênicas identificadas em um paciente de alto risco e prognóstico desfavorável, representam fenômenos que possivelmente estariam comprometendo o mecanismo TLS como um todo e induzindo eventos como o colapso da forquilha de replicação, que acentuaria a instabilidade genômica nas células e permitiria a geração das alterações cromossômicas observadas no cariótipo em estudo, como a deleção no cromossomo 7, monossomia do cromossomo 5 e presença de material adicional no cromossomo 11.

Interessantemente, não identificamos neste paciente nenhum SNP de classificação oncogênica benigna, os quais estariam associados a uma função protetora em Leucemia Mieloide Aguda e Leucemia Linfoblástica do tipo B, bem como em diversos tumores sólidos, como Câncer Colorretal, Câncer de Pulmão, Carcinoma Hepatocelular, Câncer de Esôfago e Câncer de Próstata, conforme caracterizamos nas tabelas de número 06 e 08. Acreditamos que isso denote a perda do mecanismo de síntese translesão neste paciente, tanto pela alta incidência de mutações do tipo perda de função, quanto pela ausência de polimorfismos benignos que teriam uma função protetora para a célula e estariam associados com neoplasias hematológicas e tumores sólidos.

Os polimorfismos benignos são variações genéticas que não causam danos ao indivíduo e alguns estudos demonstram que certos polimorfismos benignos podem ter uma função

protetora contra o câncer. Por exemplo, Alharbi et al. (2021) reportaram que certas variantes genéticas em *TPCN2* e *P2RX4* foram significativamente associadas a uma diminuição do risco de desenvolver várias neoplasias, como melanoma, câncer de próstata e de tecidos moles (ALHARBI et al., 2021). Sobre este tema, nosso grupo tem empregado esforços no estudo de novas variantes polimórficas de genes de reparo do DNA com efeito protetor para a instabilidade genômica na SMD. Ribeiro et al. (2013), reportaram que o genótipo heterogigoto A/G do polimorfismo rs3835, no gene *XRCC5* está associado com uma diminuição de risco para o acometimento de SMD e, também, à variáveis de bom prognóstico da doença, denotando um caráter protetor destes polimorfismos na doença (ex., baixo risco para o IPSS) (RIBEIRO et al., 2013).

### **Pacientes dos subtipos de baixo risco com cariótipo normal apresentam menor frequência de variantes LoF e maior frequência de variantes *missense* para os genes *REVI* e *POLI***

Foi possível identificar que os pacientes diagnosticados com subtipos de baixo risco (SMD-AS, SMD-AS-DU e SMD-SA-DM) apresentando cariótipos sem alterações cromossômicas possuíram menor incidência de mutações LoF, bem como maior frequência de mutações do tipo *missense*, além do acúmulo de mutações em regiões intrônicas no gene *REVI* e *POLI*. Esses resultados evidenciam que o perfil molecular obtido está em concordância com o prognóstico favorável dos pacientes, uma vez que seus quadros clínicos se configuram como quadros de baixo risco, apresentando número reduzido de citopenias, poucas displasias na medula óssea, ausência de blastos e ausência de alterações cromossômicas.

A presença de maior quantidade de variantes em regiões não codificadoras dos genes *REVI* e *POLI* e a maior incidência de variantes *missense*, assim como baixa incidência de variantes LoF nestes pacientes pode demonstrar que a maquinaria responsável pela via de tolerância ao dano não está completamente comprometida nesses casos, de modo que, mesmo às custas da possível inserção de mutações passageiras, as forquilhas de replicação não sofrem colapso, permitindo que o processo replicativo seja retomado diante de lesões volumosas e protegendo a célula de grandes deleções. Tal hipótese é corroborada pela ausência de alterações genéticas nos cariótipos destes pacientes, como aneuploidias, deleções e translocações. Nesse sentido, para os pacientes avaliados neste estudo, compreende-se que a instabilidade genômica mediada pelos mecanismos de síntese translesão não é suficientemente acentuada em casos de SMD com cariótipo normal cujo prognóstico é favorável. Possivelmente, estes pacientes possuem falhas nos mecanismos de reparo, no entanto, com atuação dos mecanismos de síntese translesão.

### **Patogenicidade x Oncogenicidade: qual seria o melhor critério de classificação?**

Este estudo de *screening* genético nos permitiu obter conhecimento acerca da predição de patogenicidade de variantes, bem como iniciarmos estudos de classificação da oncogenicidade dessas mesmas variantes.

Sempre foi um grande desafio prever as consequências funcionais ou impactos clínicos de variantes genéticas no câncer. Desde a popularização do NGS, um número crescente de variantes genéticas somáticas foi descoberto e documentado em bancos de dados públicos, como o COSMIC, no entanto, a grande maioria destas variantes não possui anotações funcionais ou clínicas, fato que foi impulsionador para realizarmos a classificação de oncogenicidade nestas variantes. Exceto por mutações em *hotspots* ou mutações recorrentes já incluídas em painéis genéticos, a interpretação clínica de variantes recentemente descobertas continua sendo um desafio.

Existem mais de 10 milhões de variantes com curadoria no banco de dados COSMIC. No entanto, observamos apenas a anotação da variante, o que não nos garante o conhecimento se ela é ou não oncogênica. Por outro lado, alguns bancos de dados, como o CIViC contêm variantes rotuladas manualmente como oncogênicas, mas esses bancos de dados tendem a ser relativamente pequenos, porque a curadoria manual é cara e não tão eficaz quanto os softwares de predição (TATE et al, 2019; GRIFFITH et al, 2017).

A literatura nos apresenta numerosas ferramentas para prever patogenicidade ou oncogenicidade de variantes, no entanto, normalmente elas não incorporam evidências clínicas ou estudos funcionais como recursos preditivos, mas apenas previsões computacionais sobre os efeitos da mutação na sequência, estrutura e conservação da proteína, sendo sua previsão mais focada em priorizar mutações funcionalmente importantes, o que as tornaria patogênicas, em vez de mutações com significado clínico oncogênico.

Com o intuito de minimizar a variabilidade de predições entre laboratórios e instituições clínicas, os grupos de trabalho do *Clinical Genome Resource* (ClinGen), *Cancer Genomics Consortium* (CGC) e *Variant Interpretation for Cancer Consortium* (VICC) se uniram no intuito de criar um procedimento operacional padrão (SOP) para classificação de oncogenicidade de variantes inspirado nas diretrizes de patogenicidade da linhagem germinativa do *American College of Medical Genetics and Genomics* (ACMG) (HORAK et al, 2022; RICHARDS et al, 2015)

Como explanado na metodologia, com base neste procedimento operacional padrão (SOP), variantes somáticas de nucleotídeo único (SNPs) e pequenas inserções/deleções, tanto as já bem estabelecidas quanto as desconhecidas, podem ser atribuídas as seguintes categorias: oncogênica, provavelmente oncogênica, variante de significado incerto (VUS), provavelmente benigna e benigna, auxiliando sua posterior interpretação clínica (HORAK et al, 2022)

O uso da classificação SOP, como chamamos neste estudo, não exclui a importância da predição de patogenicidade das variantes somáticas. Os resultados das ferramentas de predição *in silico* de patogenicidade são fundamentais para os critérios de classificação de oncogenicidade do SOP, os quais incluem: dados populacionais, dados de anotação de variantes no câncer, dados de estudos funcionais e dados de predição de patogenicidade fornecidos por tais ferramentas. No entanto, esclarecemos que predição de patogenicidade é diferente de predição da oncogenicidade, sendo a primeira relacionada a ferramentas de predição *in silico*, as quais se detém a prever o impacto da alteração na proteína e em sua conservação, e a segunda, relacionada ao impacto da variante na prática clínica.

Alguns de nossos resultados demonstram a fragilidade de se utilizar apenas ferramentas de predição para avaliar a oncogenicidade de variantes. O gene *POLQ*, por exemplo, apresentou 10 variantes classificadas como claramente patogênicas pelos algoritmos de predição, que, no entanto, se apresentaram como variantes de significado incerto após classificação de oncogenicidade pelo SOP. A variante REV3L(NM\_001372078.1):c.2482T>C (p.Ser828Pro), anotada no (COSM5011659,COSM5011660), foi dita como causadora de doença e prejudicial por duas das quatro ferramentas de predição utilizadas, mas foi classificada como provavelmente benigna após pontuação pelo SOP. Da mesma forma, a variante POLQ(NM\_199420.4):c.7271A>T (p.Asn2424Ile), registrada no COSMIC (COSM1693249), foi indicada como patogênica por dois dos quatro algoritmos utilizados, no entanto, foi classificada como Potencialmente Oncogênica pelos critérios SOP.

Dessa forma, acreditamos que é de extrema importância a incorporação da classificação SOP, recentemente preconizada, na classificação de variantes de importância clínica na oncologia, pois demonstramos que variantes claramente patogênicas não são, necessariamente, oncogênicas.

## 6. CONCLUSÕES

1. A alta incidência de variantes LoF nos genes *POLQ* e *REV3L* em pacientes em estágio avançado da SMD pode representar o aumento de instabilidade genômica por ausência do mecanismo de Síntese Translesão mediado por estas enzimas na Neoplasia Mielodisplásica, fato que possivelmente contribuiu para a progressão da doença nestes pacientes;
2. A alta taxa de variantes em *REVI* e *POLI*, com poucas variantes com perda função, pode representar um biomarcador da instabilidade genômica na Neoplasia Mielodisplásica;
3. Pacientes com cariótipos apresentando múltiplas alterações cromossômicas (cariótipo complexo), acumulam mutações de perda de função (LoF) em genes de síntese translesão, o que pode explicar a grande instabilidade genética observada nestes indivíduos.
4. As ferramentas de predição *in silico* não são suficientes para classificação de variantes oncogênicas, sendo cada vez mais necessária a incorporação da classificação SOP na prática clínica oncológica;
5. Estes resultados suportam a importância do mecanismo de síntese translesão na estabilidade genômica das células tronco hematopoiéticas, provendo informações importantes para o melhor entendimento da fisiopatologia e progressão da Neoplasia Mielodisplásica.

## 6. LIMITAÇÕES E PERSPECTIVAS FUTURAS

Estamos cientes das limitações advindas com este estudo, como a baixa cobertura observada em algumas variantes e amostras analisadas ou a ausência de análise estatística de associação ou correlação que permitisse traçar uma associação das variantes encontradas aos diferentes fenótipos da SMD, em razão do pequeno número de amostras para as quais foi possível a realização das análises de bioinformática. Ademais dessas limitações, estudos de expressão gênica foram conduzidos por nosso grupo e utilizaremos os resultados aqui apresentados para traçar uma análise de associação entre as variantes identificadas e a expressão dos genes de síntese translesão nos pacientes com SMD. Esperamos inter-relacionar os resultados obtidos a partir dessas duas metodologias para conseguir traçar um perfil da atuação do mecanismo de síntese translesão na Neoplasia Mielodisplásica e, possivelmente, partir para estudos funcionais com linhagens celulares e/ou camundongos *knockout*.

Este é um estudo guarda-chuva, que se propôs a implementar o Sequenciamento Genético por NGS no Laboratório de Citogenômica do Câncer, bem como estabelecer linhas de análise de dados baseadas em uma análise de bioinformática criteriosa, com o estudo de ferramentas de predição *in silico* e utilizando o mais novo critério de classificação de variantes oncogênicas, o SOP-ACMG-AMP-VICC.

Somos o primeiro grupo de pesquisa que se propôs a estudar o mecanismo de síntese translesão na SMD, e estudar um mecanismo complexo como o TLS não foi fácil. Por conta da sua natureza ambígua “livre de erros” ou “propensa a erros” muitas podem ser as possibilidades de atuação dessas enzimas no processo de carcinogênese, atuando como genes supressores de tumor ou como oncogenes. Estamos em um caminho longo e árduo e teremos muitas dificuldades no percurso, especialmente pela natureza multifacetada das Polimerases TLS e heterogênea da SMD, bem como pelas dificuldades e limitações do NGS. No entanto, nosso anseio por biomarcadores prognósticos que possam auxiliar no manejo clínico e tratamento dos pacientes com SMD torna essa jornada mais leve.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ADAM, S. S.; CHRISTOPHER, J. G.; BENJAMIN L. E.; The genetics of myelodysplastic syndrome: from clonal hematopoiesis to secondary leukemia. *Nat Rev Cancer*. 2017 Jan; 17(1): 5–19.

ADAR, S.; IZHAR, L.; HENDEL, A.; GEACINTOV, N.; LIVNEH, Z. Repair of gaps opposite lesions by homologous recombination in mammalian cells. *Nucleic Acids Research*, vol. 37, no. 17, pp. 5737–5748, 2009.

ADÈS, L.; ITZYKSON, R.; FENAUX, P. Myelodysplastic syndromes. *The Lancet*, v. 383, n. 9936, p. 2239 - 2252, 2014.

AGGARWAL, S., VAN DE LOOSDRECHT, A. A., ALHAN, C., OSSENKOPPELE, G. J., WESTERS, T. M., & BONTKES, H. J. (2011). Role of immune responses in the pathogenesis of low-risk MDS and high-risk MDS: implications for immunotherapy. *British Journal of Haematology*, 153(5), 568–581.

ALHARBI AF, PARRINGTON J. The role of genetic polymorphisms in endolysosomal ion channels TPC2 and P2RX4 in cancer pathogenesis, prognosis, and diagnosis: a genetic association in the UK Biobank. *NPJ Genom Med*. 2021 Jul 12;6(1):58.

ALISON J.; RATTRAY AND JEFFREY N. STRATHERN. ERROR-PRONE DNA POLYMERASES: When Making a Mistake is the Only Way to Get Ahead. *Annu Rev Genet*. 2003;37:31-66.

ALLERA-MOREAU, C. et al. DNA replication stress response involving PLK1, CDC6, POLQ, RAD51 and CLASPIN upregulation prognoses the outcome of early/mid-stage non-small cell lung cancer patients. *Oncogenesis*, v. 1, p. e30, Oct 2012.

ALTER, P. B. Fanconi anemia and the development of leukemia. *Best Pract Res Clin Haematol*. 2014 ;

ANDERSEN PL.; XU F.; XIAO W. Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA. *Cell Res*. 2008 Jan;18(1):162-73.

ARANA, ME.; TAKATA, K.; GARCIA-DIAZ, M.; WOOD, RD.; KUNKEL, TA. A unique error signature for human DNA polymerase  $\eta$ . *DNA Repair (Amst)*. 2007 Feb 4;6(2):213-23.

ARBER, D.A.; HASSERJIAN, R.P. Reclassifying myelodysplastic syndromes: what's where in the new WHO and why. *Hematology*, p. 294–298, 2015.

ARBER DA, ORAZI A, HASSERJIAN R, THIELE J, BOROWITZ MJ, LE BEAU MM, BLOOMFIELD CD, CAZZOLA M, VARDIMAN JW. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016 May 19;127(20):2391-405. doi: 10.1182/blood-2016-03-643544. Epub 2016 Apr 11. PMID: 27069254.

AUL, C.; GIAGOUNIDIS, A.; GERMING, U.; Epidemiological features of myelodysplastic syndromes: results from regional cancer surveys and hospital-based statistics. *Int J Hematol*. 2001 Jun;73(4):405-410.

1000 GENOMES PROJECT CONSORTIUM; AUTON A, BROOKS LD, DURBIN RM, GARRISON EP, KANG HM, KORBEL JO, MARCHINI JL, MCCARTHY S, MCVEAN GA, ABECASIS GR. A global reference for human genetic variation. *Nature*. 2015 Oct 1;526(7571):68-74. doi: 10.1038/nature15393. PMID: 26432245; PMCID: PMC4750478.

BAVOUX, C.; LEOPOLDINO, A. M.; BERGOGLIO, V. Up-Regulation of the Error-Prone DNA Polymerase  $\kappa$  Promotes Pleiotropic Genetic Alterations and Tumorigenesis. *Cancer Res* 2005 65; 325.

BAYNTON, K.; BRESSON-ROY, A.; FUCHS, RP.; Analysis of damage tolerance pathways in *Saccharomyces cerevisiae*: a requirement for Rev3 DNA polymerase in translesion synthesis. *Mol Cell Biol*. 1998 Feb; 18(2):960-6.

BEAGAN, K. et al. *Drosophila* DNA polymerase theta utilizes both helicase-like and polymerase domains during microhomology-mediated end joining and interstrand crosslink repair. *PLoS Genet*. 2017. May, v. 13, n. 5, p. e1006813.

BEJAR R, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med*, 2011; 364(26):2496-506.

BEJAR R. Clinical and genetic predictors of prognosis in myelodysplastic syndromes. *Haematologica*. 2014; 99(6):956-964.

BELLI, C. B. et al. Myelodysplastic syndromes in South America: A multinational study of 1080 patients. *Am J Hematol*, Jun 2015.

BENNARDO, N.; CHENG, A.; HUANG, N.; STARK, JM. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet*. 2008 Jun 27;4(6):e1000110

BENNETT, J. M. et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*, v. 33, n. 4, p. 451-8, Aug 1976.

BENNETT, J. M. Proposals for the classification of the myelodysplastic syndrome. *Br. J. Hematol.*, v.51, p.189-199, 1982.

BERGOGLIO, V.; BAVOUX, C.; VERBIEST, V.; HOFFMANN, JS.; CAZAUX, C. Localisation of human DNA polymerase kappa to replication foci. *Journal of Cell Science*, 2002. 115(Pt 23), 4413-8.

BERNARD, E.; TUECHLER, H.; GREENBERG, P.L.; HASSERJIAN, R.P.; OSSA, J.E.A.; NANNYA, Y.; DEVLIN, S.M.; CREIGNOU, M.; PINEL, P.; MONNIER, L.; ET AL. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *NEJM Evid*. 2022, 1, EVIDoa2200008

BERNASCONI, P. et al. Validation of the new comprehensive cytogenetic scoring system (NCCSS) on 630 consecutive de novo MDS patients from a single institution. *Am J Hematol*, v. 88, n. 2, p. 120-9, Feb 2013.

BERTOLIN, A.P.; MANSILLA, S.F.; GOTTIFREDI, V. The identification of translesion DNA synthesis regulators: Inhibitors in the spotlight. *DNA Repair (Amst)*. 2015 Aug;32:158-64

Bielas, J. H. (2006). Non-transcribed strand repair revealed in quiescent cells. *Mutagenesis*, 21(1), 49–53.

BIELAS, J. H. et al. Human cancers express a mutator phenotype. *Proc Natl Acad Sci U S A*, v. 103, n. 48, p. 18238-42, Nov 2006.

BIENKO M.; GREEN CM.; CROSETTO N.; RUDOLF F.; ZAPART G.; COULL B.; KANNOUCHE P.; WIDER G.; PETER M.; LEHMANN AR.; HOFMANN K.; DIKIC I.

Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science*. 2005 Dec 16;310(5755):1821-4.

BIERTÜMPFEL, C.; ZHAO, Y.; KONDO, Y.; RAMÓN-MAIQUES, S.; GREGORY, M.; LEE, JY. Structure and Mechanism of Human DNA Polymerase  $\eta$ . *Nature*. 2010 Jun 24; 465(7301): 1044–1048.

BLACK, SJ.; KASHKINA, E.; KENT, T.; POMERANTZ, RT. DNA Polymerase  $\theta$ : A Unique Multifunctional End-Joining Machine. *Genes (Basel)*. 2016 Sep; 7(9): 67.

BOHGAKI, T.; BOHGAKI, M.; HAKEM, R. DNA double-strand break signaling and human disorders. *Genome Integr*, v. 1, n. 1, p. 15, 2010.

BOSTIAN, AC.; MADDUKURI, L.; REED, MR.; SAVENKA, T.; HARTMAN, JH1.; DAVIS, L. Kynurenine Signaling Increases DNA Polymerase Kappa Expression and Promotes Genomic Instability in Glioblastoma Cells. *Chem Res Toxicol*. 2016 Jan 19;29(1):101-8.

BOWEN, DT. Occupational and environmental etiology of MDS. *Best Pract Res Clin Haematol.*, v. 26, n. 4, p 319-26, 2013.

BRANDALIZE, A. P. C. et al. Genetic variants involved in specialized DNA replication and their relation with breast cancer risk, disease progression and patient prognosis. *BMC Proceedings*, v. 7, n. Suppl 2, p. P7-P7, 04/04 2013.

BRANDALIZE, A. P. et al. A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. *BMC Cancer*, v. 14, p. 850, Nov 2014.

BRECCIA, M., et al. Analysis of prognostic factors in patients with refractory anemia with excess of blasts (RAEB) reclassified according to WHO proposal. *Leuk Res*, 2009. Mar,v.33, n.3, p.391-4,

BRUNNING, R. D.; ORAZI, A.; GERMING, U. Myelodysplastic syndromes/neoplasms, overview. In: SWERDLOW, S. H.; CAMPO, E.; HARRIS, N. L.; JAFFE, E. S.; PILERI, S. A.; STEIN, H.; THIELE, J.; VARDIMAN, J. W. WHO Classification of Tumours of Hematopoietic and Lymphoid Tissues. Lyon: IARC Press, 2008. p. 88-93.

BUDCZIES, J.; KLAUSCHEN, F.; SINN, B.V.; GYÖRFFY, B.; SCHMITT, W.D.; DARB-ESFAHANI, S.; DENKERT, C. Cutoff Finder: a comprehensive and straightforward Web

application enabling rapid biomarker cutoff optimization. *PLoS One*, v. 7, n. 12, p. e51862, 2012.

BOUDSOCQ, F.; LING, H.; YANG, W. Structure-based interpretation of missense mutations in Y-family DNA polymerases and their implications for polymerase function and lesion bypass. *DNA repair*, v. 1, n. 5, p. 343–358, 2002.

BUSTIN, S.A.; BENES, V.; GARSON, J.A.; HELLEMANS, J.; HUGGETT, J.; KUBISTA, M.; et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem.*, v. 55, n. 4, p. 611-22, 2009.

CARACCIOLO D, RIILLO C, DI MARTINO MT, TAGLIAFERRI P, TASSONE P. Alternative Non-Homologous End-Joining: Error-Prone DNA Repair as Cancer's Achilles' Heel. *Cancers (Basel)*. 2021 Mar 19;13(6):1392. doi: 10.3390/cancers13061392. PMID: 33808562; PMCID: PMC8003480.

CAVALCANTE GM, BORGES DP, DE OLIVEIRA RTG, FURTADO CLM, ALVES APNN, SOUSA AM, DE PAULA DS, FILHO FDR, MAGALHÃES SMM, RIBEIRO-JR HL, PINHEIRO RF. Tissue methylation and demethylation influence translesion synthesis DNA polymerases (TLS) contributing to the genesis of chromosomal abnormalities in myelodysplastic syndrome. *J Clin Pathol*. 2022 Feb;75(2):85-93. doi: 10.1136/jclinpath-2020-207131. Epub 2020 Nov 24. PMID: 33234697.

CECCALDI, R. et al. Homologous-recombination-deficient tumours are dependent on Pol $\theta$ -mediated repair. *Nature*. 2015; Feb, v. 518, n. 7538, p. 258-62.

CECCALDI, R., LIU, J., AMUNUGAMA, R. *et al.* Homologous-recombination-deficient tumours are dependent on Pol $\theta$ -mediated repair. *Nature* 518, 258–262 (2015). <https://doi.org/10.1038/nature14184>

CEPPI, P. et al. Polymerase eta mRNA expression predicts survival of non-small cell lung cancer patients treated with platinum-based chemotherapy. *Clin Cancer Res*, v. 15, n. 3, p. 1039-45, Feb 2009.

CGOWAN-JORDAN, J.; SIMONS, A.; SCHMID, M. *ISCN : an international system for human cytogenomic nomenclature* (2016). Basel ; New York : Karger, 2016.

CHAMSEDDINE AN et al. Unraveling Myelodysplastic Syndromes: Current Knowledge and Future Directions. *Curr Oncol Rep*. 2016; 18(4): 1-11.

CHANG DJ, CIMPRICH KA. DNA damage tolerance: when it's OK to make mistakes. *Nat Chem Biol*. 2009;5:82–90.

CHEN, CY.; LIN, LI.; TANG, JL., et al. RUNX1 gene mutation in primary myelodysplastic syndrome -- the mutation can be detected early at diagnosis or acquired during disease progression and is associated with poor outcome. *Br J Haematol* 2007;139:405-414

CHIHARA, D.; ITO, H.; KATANODA, K.; SHIBATA, A.; MATSUDA, T.; SOBUE, T.; MATSUO, K.; Incidence of Myelodysplastic Syndrome in Japan. *J Epidemiol*. 2014;24(6):469-73

CHIU, RK. et al. Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations. *PLoS Genet*, v. 2, n. 7, p. e116, Jul 2006.

CHOI, J.; LIM, S.; KIM, E.; JO, A.; GUENGERICH, FP. Translesion Synthesis Across Abasic Lesions by Human B-family and Y-family DNA Polymerases  $\alpha$ ,  $\delta$ ,  $\eta$ ,  $\iota$ ,  $\kappa$ , and REV1. *J Mol Biol*. 2010 Nov 19; 404(1): 34–44.

CIOC, AM1.; WAGNER, JE.; MACMILLAN, ML.; DEFOR, T.; HIRSCH, B. Diagnosis of myelodysplastic syndrome among a cohort of 119 patients with fanconi anemia: morphologic and cytogenetic characteristics. *Am J Clin Pathol*. 2010 Jan;133(1):92-100.

COGLE, C. R. et al. Incidence of the myelodysplastic syndromes using a novel claims-based algorithm: high number of uncaptured cases by cancer registries. *Blood*, v. 117, n. 26, p. 7121-5, Jun 2011.

COGLE, C. R. Incidence and Burden of the Myelodysplastic Syndromes. *Curr Hematol Malig Rep*, v. 10, n. 3, p. 272-81, Sep 2015.

CORE TEAM. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (Ed.). <http://www.r-project.org>. 2012.

COSTA MB, FARIAS IR, DA SILVA MONTE C, FILHO LIPF, DE PAULA BORGES D, DE OLIVEIRA RTG, RIBEIRO-JUNIOR HL, MAGALHÃES SMM, PINHEIRO RF. Chromosomal abnormalities and dysregulated DNA repair gene expression in farmers exposed

to pesticides. *Environ Toxicol Pharmacol.* 2021 Feb;82:103564. doi: 10.1016/j.etap.2020.103564. Epub 2020 Dec 14. PMID: 33326828.

DAIGAKU Y.; DAVIES AA.; ULRICH HD. Ubiquitin-dependent DNA damage bypass is separable from genome replication. *Nature.* 2010 Jun 17;465(7300):951-5.

DICKER, F.; HAFERLACH, C.; SUNDERMANN, J., et al. Mutation analysis for RUNX1, MLL-PTD, FLT3-ITD, NPM1 and NRAS in 269 patients with MDS or secondary AML. *Leukemia* 2010;24:1528-1532

DRZEWIECKA M, BARSZCZEWSKA-PIETRASZEK G, CZARNY P, SKORSKI T, ŚLIWIŃSKI T. Synthetic Lethality Targeting Pol $\theta$ . *Genes (Basel).* 2022 Jun 20;13(6):1101.

FARMER H, MCCABE N, LORD CJ, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature.* 2005;434(7035):917–21.

FEARON, E. R. Tumor Suppressor Genes. In: VOGELSTEIN, B.; KINZLER, K. W. (Ed.). *The genetics basis of human cancer.* 2nd ed. New York: McGraw-Hill, p. 197-206, 2002

FERNANDEZ-VIDAL A, GUITTON-SERT L, CADORET JC, DRAC M, SCHWOB E, BALDACCI G, CAZAUX C, HOFFMANN JS. A role for DNA polymerase  $\theta$  in the timing of DNA replication. *Nat Commun.* 2014 Jul 3;5:4285.

FOX, JT.; LEE, KY.; MYUNG K. Dynamic regulation of PCNA ubiquitylation/deubiquitylation. *FEBS Lett.* 2011 Sep 16;585(18):2780-5

FRIEDBERG EC.; LEHMANN AR.; FUCHS RP. Trading places: how do DNA polymerases switch during translesion DNA synthesis? *Mol Cell.* 2005;18:499–505

FRIEDBERG, EC. A brief history of the DNA repair field. *Cell research,* 2008; 18(1), 3-7.

FRIEDBERG, EC.; WAGNER, R.; RADMAN, M. Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science.* 2002 May 31;296(5573):1627-30.

GARCIA-MANERO, G. Myelodysplastic syndromes: 2015 Update on diagnosis, risk-stratification and management. *Am J Hematol,* v. 90, n. 9, p. 831-41, Sep 2015.

GHOSAL G.; CHEN J. DNA damage tolerance: a double-edged sword guarding the genome. *Transl Cancer Res.* 2013;2(3):107-129.

GOFF, JP.; SHIELDS, DS.; SEKI, M.; CHOI, S.; EPPERLY, MW.; DIXON, T.; WANG, H et al. Lack of DNA polymerase theta (POLQ) radiosensitizes bone marrow stromal cells in vitro and increases reticulocyte micronuclei after total-body irradiation. *Radiat Res.* 2009 Aug;172(2):165-74.

GREENBERG P, COX C, LEBEAU MM, FENAUX P, MOREL P, SANZ G, SANZ M, VALLESPI T, HAMBLIN T, OSCIER D, OHYASHIKI K, TOYAMA K, AUL C, MUFTI G, BENNETT J. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*, v. 89, n. 6, p. 2079-88, 1997

GREENBERG, P. L.; TUECHLER, H.; SCHANZ, J.; SANZ, G.; GARCIA-MANERO, G.; SOLÉ, F.; BENNETT, J. M.; BOWEN, D.; FENAUX, P.; DREYFUS, F.; KANTARJIAN, H.; KUENDGEN, A.; LEVIS, A.; MALCOVATI, L.; CAZZOLA, M.; CERMAK, J.; FONATSCH, C.; LE BEAU, M. M.; SLOVAK, M.L.; KRIEGER, O.; LUEBBERT, M.; MACIEJEWSKI, J.; MAGALHAES, S. M.; MIYAZAKI, Y.; PFEILSTÖCKER, M.; SEKERES, M.; SPERR, W. R.; STAUDER, R.; TAURO, S.; VALENT, P.; VALLESPI, T.; VAN DE LOOSDRECHT, A. A.; GERMING, U.; HAASE, D. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*, v. 120, n. 12, p. 2454-2465, 2012.

GREENBERG, P.L.; ATTAR, E.; BENNETT, J.M.; BLOOMFIELD, C.D.; BORATE, U.; DE CASTRO, C.M., et al. NCCN Practice Guidelines for myelodysplastic syndromes, Version 1. *J Natl Compr Canc Netw*, v. 480, n. 11, p. 838-874, 2013.

GRIFFITH M, SPIES NC, KRYSIAK K, MCMICHAEL JF, COFFMAN AC, DANOS AM, AINSCOUGH BJ, RAMIREZ CA, RIEKE DT, KUJAN L, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat Genet.* 2017;49(2):170–4.

GUO C, et al. REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. *Mol Cell.* 2006. 23(2):265-71

GUO, C.; FISCHHABER, PL.; LUK-PASZYC, L.; MASUDA, Y.; ZHOU, J.; KAMIYA, K.; KISKER, C.; FRIEDBERG, EC. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *EMBO J.* 2003. 22:6621–6630.

HAASE, D. Cytogenetics features in myelodysplastic syndromes. *Ann. Hematol.*, v. 87, p.515-526, 2008.

HAASE, D.; GERMING, U.; SCHANZ, J.; PFEILSTÖCKER, M.; NÖSSLINGER, T.; HILDEBRANDT, B. et al. New insights into the prognostic impact of the karyotype in MDS and correlation with subtypes: evidence from a core dataset of 2124 patients. *Blood*. 2007 Dec 15; 110(13):4385-95.

HAFERLACH, T., NAGATA, Y., GROSSMANN, V., OKUNO, Y., BACHER, U., NAGAE, G., SCHNITTGER, S., SANADA, M., KON, A., ALPERMANN, T., YOSHIDA, K., ROLLER, A., NADARAJAH, N., SHIRAISHI, Y., SHIOZAWA, Y., CHIBA, K., TANAKA, H., KOEFFLER, H. P., KLEIN, H. U., ... OGAWA, S. (2013). Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* 2014 28:2, 28(2), 241–247.

HANAHAHAN, D.; WEINBERG, AR. Hallmarks of Cancer: The Next Generation. *Cell*. 2011 Mar 4;144(5):646-74.

HARACSKA, L.; PRAKASH, S.; PRAKASH, L. Replication past O(6)-methylguanine by yeast and human DNA polymerase  $\epsilon$ . *Mol Cell Biol*. 2000b Nov;20(21):8001-7.

HARACSKA, L.; YU, SL.; JOHNSON, RE.; PRAKASH, L.; PRAKASH, S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase  $\epsilon$ . *Nat Genet*. 2000a Aug;25(4):458-61.

HARRIS, PV.; MAZINA, OM.; LEONHARDT, EA.; CASE, RB.; BOYD, JB.; BURTIS, KC. Molecular cloning of *Drosophila* mus308, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. *Mol Cell Biol*. 1996 Oct;16(10):5764-71.

HASLE, H.; NIEMEYER, C. M.; CHESSELLS, J. M.; BAUMANN, I.; BENNETT, J. M.; KERNDRUP, G.; HEAD, D. R. A pediatric approach to the WHO classification of myelodysplastic and myeloproliferative diseases. *Leukemia*. 2003 Feb;17(2):277-82.

HASSERJIAN, R. P. Myelodysplastic Syndrome Updated. *Pathobiology*, v. 86, n. 1, p. 7–13, 2019.

HE, X. et al., REV1 genetic variants associated with the risk of cervical carcinoma. *European journal of epidemiology*, v. 23, n. 6, p. 403–409, 2008.

HIGGINS, G.S.; HARRIS, A.L.; PREVO, R.; HELLEDAY, T.; MCKENNA, W.G.; BUFFA, F.M. Overexpression of POLQ Confers a Poor Prognosis in Early Breast Cancer Patients. *Oncotarget* 2010, 1, 175–184.

HIGGINS, GS.; PREVO, R.; LEE, YF.; HELLEDAY, T.; MUSCHEL, RJ.; TAYLOR, S. A Small interfering RNA screen of genes involved in DNA repair identifies tumor-specific radiosensitization by POLQ knockdown. *Cancer Res.* 2010. 70, 2984–2993.

HIRAI H, OKADA M, MIZOGUCHI H, MANO H, KOBAYASHI Y, NISHIDA J, TAKAKU F. Relationship between an activated N-ras oncogene and chromosomal abnormality during leukemic progression from myelodysplastic syndrome. *Blood.* 1988 Jan;71(1):256-8. PMID: 3275473.

HOEIJMAKERS, J. H. DNA damage, aging, and cancer. *N Engl J Med*, v. 361, n. 15, p. 1475-85, Oct 2009.

HOFFMANN, JS.; CAZAUX, C. Aberrant expression of alternative DNA polymerases: a source of mutator phenotype as well as replicative stress in cancer. *Semin Cancer Biol*, v. 20, n. 5, p. 312-9, Oct 2010.

HOLLAENDER, A.; CURTIS, JT. Effect of sublethal doses of monochromatic ultraviolet radiation on bacteria in liquid suspension. *Proceedings of the Society for Experimental Biology and Medicine.* 1935; 33, 61-62.

HORAK, P., GRIFFITH, M., DANOS, A. M., PITEL, B. A., MADHAVAN, S., LIU, X., CHOW, C., WILLIAMS, H., CARMODY, L., BARROW-LAING, L., RIEKE, D., KREUTZFELDT, S., STENZINGER, A., TAMBORERO, D., BENARY, M., RAJAGOPAL, P. S., IDA, C. M., LESMANA, H., SATGUNASEELAN, L., ... SONKIN, D. (2022). Standards for the classification of pathogenicity of somatic variants in cancer (oncogenicity): Joint recommendations of Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC), and Variant Interpretation for Cancer Consortium (VICC). *Genetics in Medicine*, 24(5), 986–998.

HUANG F, TANAKA H, KNUDSEN BS, RUTGERS JK. Mutant POLQ and POLZ/REV3L DNA polymerases may contribute to the favorable survival of patients with tumors with POLE mutations outside the exonuclease domain. *BMC Med Genet.* 2020 Aug 24;21(1):167.

HUANG KK, JANG KW, KIM S, KIM HS, KIM SM, KWON HJ, KIM HR, YUN HJ, AHN MJ, PARK KU, RAMNARAYANAN K, MCPHERSON JR, ZHANG S, RHEE JK, VETTORE AL, DAS K, ISHIMOTO T, KIM JH, KOH YW, KIM SH, CHOI EC, TEH BT, ROZEN SG, KIM TM, TAN P, CHO BC. Exome sequencing reveals recurrent REV3L mutations in cisplatin-resistant squamous cell carcinoma of head and neck. *Sci Rep*. 2016 Jan 21;6:19552.

HWA YUN, B.; GUO, J.; BELLAMRI, M.; TURESKY, RJ. DNA adducts: Formation, biological effects, and new biospecimens for mass spectrometric measurements in humans. *Mass Spectrom Rev*. 2018 Jun 11.

IYAMA T.; WILSON DM. DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair (Amst)*. 2013 Aug;12(8):620-36

JAFFE, E. S., SWERDLOW, S. H., CAMPO, E., HARRIS, N. L., PILERI, S. A., STEIN, H., THIELE, J., VARDIMAN, J. W. WHO classification of tumours: tumours of haematopoietic and lymphoid tissues. Lyon, France, 2001.

JAN, M.; SPERLING, A. S.; EBERT, B. L. Cancer therapies based on targeted protein degradation — lessons learned with lenalidomide. *Nature Reviews Clinical Oncology*, v. 18, n. 7, p. 401–417, jul. 2021.

JANKOWSKA, AM.; GONDEK, LP.; SZPURKA, H.; NEARMAN, ZP.; TIU, RV.; MACIEJEWSKI, JP. Base excision repair dysfunction in a subgroup of patients with myelodysplastic syndrome. *Leukemia*. 2008 Mar;22(3):551-8.

JANSEN, JG.; TSAALBI-SHTYLIK, A.; DE WIND, N. Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. *DNA Repair (Amst)*. 2015 May;29:56-64

JOHNSO, RE.; WASHINGTON, MT.; HARACSKA, L. et al. Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature*. 2000;406:1015–9.

JOHNSON, RE.; KONDRATICK, CM.; PRAKASH, S.; PRAKASH, L. hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science*. 1999 Jul 9;285(5425):263-5.

JOHNSON, RE.; PRAKASH, S.; PRAKASH, L. The human DINB1 gene encodes the DNA polymerase Poltheta. *Proceedings of the National Academy of Sciences of the United States of America*, 2000. 97(8), 3838-43.

JONAS, B. A.; GREENBERG, P. L. MDS prognostic scoring systems – past, present, and future. *Best Pract Res Clin Haematol*, v. 28, n. 1, p. 3-13, Mar 2015.

Junior, H.L.R., de Oliveira, R.T.G., de Paula Borges, D. et al. Can synthetic lethality approach be used with DNA repair genes for primary and secondary MDS?. *Med Oncol* 36, 99 (2019).

KANAO, R.; MASUTANI, C. Regulation of DNA damage tolerance in mammalian cells by post-translational modifications of PCNA. *Mutat Res.* 2017 Oct;803-805:82-88.

KANNOUCHE PL.; LEHMANN AR. Ubiquitination of PCNA and the polymerase switch in human cells. *Cell Cycle*. 2004 Aug;3(8):1011-3.

KAO, J.; ROSENSTEIN, B. S.; PETERS, S.; MILANO, M. T.; KRON, S. J. Cellular response to DNA damage. *Ann. N Y Acad. Sci.*, v. 1066, p. 243-258, 2005.

KAWAMURA, K. et al. DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. *Int J Cancer*, v. 109, n. 1, p. 9-16, Mar 2004.

KAWAMURA, K.; BAHAR, R.; SEIMIYA, M.; CHIYO, M.; WADA, A.; OKADA, S.; HATANO, M.; TOKUHISA, T.; KIMURA, H.; WATANABE, S.; et al. DNA Polymerase? Is Preferentially Expressed in Lymphoid Tissues and Upregulated in Human Cancers. *Int. J. Cancer* 2004, 109, 9–16.

KHOURY JD, SOLARY E, ABLA O, AKKARI Y, ALAGGIO R, APPERLEY JF, BEJAR R, BERTI E, BUSQUE L, CHAN JKC, CHEN W, CHEN X, CHNG WJ, CHOI JK, COLMENERO I, COUPLAND SE, CROSS NCP, DE JONG D, ELGHETANY MT, TAKAHASHI E, EMILE JF, FERRY J, FOGELSTRAND L, FONTENAY M, GERMING U, GUJRAL S, HAFERLACH T, HARRISON C, HODGE JC, HU S, JANSEN JH, KANAGALSHAMANNA R, KANTARJIAN HM, KRATZ CP, LI XQ, LIM MS, LOEB K, LOGHAVI S, MARCOGLIESE A, MESHINCHI S, MICHAELS P, NARESH KN, NATKUNAM Y, NEJATI R, OTT G, PADRON E, PATEL KP, PATKAR N, PICARSIC J, PLATZBECKER U, ROBERTS I, SCHUH A, SEWELL W, SIEBERT R, TEMBHARE P, TYNER J, VERSTOVSEK S, WANG W, WOOD B, XIAO W, YEUNG C, HOCHHAUS A. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours:

Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia*. 2022 Jul;36(7):1703-1719. doi: 10.1038/s41375-022-01613-1. Epub 2022 Jun 22. PMID: 35732831; PMCID: PMC9252913.

KNOBEL, PA.; MARTI, T. M.; Translesion DNA synthesis in the context of cancer research. *Cancer Cell Int*. 2011 Nov 2;11:39.

KRIEG AJ, HAMMOND EM, GIACCIA AJ. Functional analysis of p53 binding under differential stresses. *Mol Cell Biol*. 2006;26:7030-45.

KURAMOTO, K.; BAN, S.; ODA, K.; TANAKA H.; KIMURA, S.; SUZUKI, G. Chromosomal instability and radiosensitivity in myelodysplastic syndrome cells. *Leukemia*. 2002 Nov;16(11):2253-8.

KUSUMOTO, R.; MASUTANI, C.; IWAI, S.; HANAOKA, F. Translesion synthesis by human DNA polymerase eta across thymine glycol lesions. *Biochemistry*. 2002 May 14;41(19):6090-9.

LANGIE, S. A. et al. Causes of genome instability: the effect of low dose chemical exposures in modern society. *Carcinogenesis*, v. 36 Suppl 1, p. S61-88, Jun 2015

LEE, E. J. et al. The evolving field of prognostication and risk stratification in MDS: Recent developments and future directions. *Blood Rev*, Jun 2015.

LEHMANN AR. Translesion synthesis in mammalian cells. *Exp Cell Res*. 2006 Aug 15;312(14):2673-6

LEMÉE, F. et al. DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. *Proc Natl Acad Sci U S A*, v. 107, n. 30, p. 13390-5, Jul 2010.

LEONCINI, E.; RICCIARDI, W.; CADONI, G.; ARZANI, D.; PETRELLI, L.; PALUDETTI, G.; BRENNAN, P.; LUCE, D.; STUCKER, I.; MATSUO, K.; ET AL. Adult Height and Head and Neck Cancer: A Pooled Analysis within the INHANCE Consortium. *Eur. J. Epidemiol*. 2014, 29, 35–48.

LESSA, RC. et al. Identification of upregulated genes in oral squamous cell carcinomas. *Head Neck*, v. 35, n. 10, p. 1475-81, Oct 2013.

LIEW E, OWEN C. Familial myelodysplastic syndromes: a review of the literature. *Haematologica*. 2011 Oct;96(10):1536-42. doi: 10.3324/haematol.2011.043422. Epub 2011 May 23. PMID: 21606161; PMCID: PMC3186316.

LINDAHL T.; WOOD RD. Quality control by DNA repair. *Science*. 1999 Dec 3;286(5446):1897-905.

LINDAHL, T. Instability and decay of the primary structure of DNA. *Nature*, 1993; 362(6422):709-15.

LIVAK, K.J.; SCHMITTGEN, T.D. Analysis of relative gene expression data using real time quantitative PCR and the 2(-Delta DeltaC(T) ) method. *Methods*., v. 25, p. 402-408, 2001.

LIVNEH Z1.; ZIV O.; SHACHAR S. Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle*. 2010 Feb 15;9(4):729-35

LIVNEH, Z. Keeping mammalian mutation load in check. Regulation of the activity of error-prone DNA polymerases by p53 and p21. *Cell Cycle*. 2006; 5:1918-22.

MA, X., et al. Myelodysplastic syndromes: incidence and survival in the United States. *Cancer*, v.109, n.8, p.1536-42, Apr 15 2007.

MAGA, G. et al. DNA polymerase theta purified from human cells is a high-fidelity enzyme. *J Mol Biol*. 2002. May v. 319, n. 2, p. 359-69.

MAGALHÃES, S. M. M. et al. Epidemiological and Clinicopathological Data From the Brazilian Registry of Patients with Myelodysplastic Syndromes and Comparative Analysis between different geographic areas. *Blood*. 116, abstract 1884 2010.

MAKAROVA AV, BURGERS PM. Eukaryotic DNA polymerase  $\zeta$ . *DNA Repair (Amst)*. 2015 May; 29:47-55.

MALCOVATI L, PORTA MG, PASCUTTO C, et al. Prognostic factors and life expectancy in MDS classified according to WHO criteria: a basis for clinical decision making. *J Clin Oncol*., v. 23, p. 7594-603, 2005.

MALCOVATI, L. et al. Impact of the degree of anemia on the outcome of patients with myelodysplastic syndrome and its integration into the WHO classification-based Prognostic Scoring System (WPSS). *Haematologica*, 2011; 96:1433-1440.

MALCOVATI, L.; GERMING, U.; KUENDGEN, A. et al. Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. *J Clin Oncol*. 2007;25(23):3503–10.

MANKE, IA.; LOWERY, DM.; NGUYEN, A.; YAFFE, MB.; BRCT repeats as phosphopeptide-binding modules involved in protein targeting. *Science (New York, N.Y.)*, 2003; 302(5645), 636-9.

MAUL, R. et al. DNA polymerase  $\tau$  functions in the generation of tandem mutations during somatic hypermutation of antibody genes. *The Journal of experimental medicine*, v. 213, n. 9, p. 1675–1683, 2016.

MARTIN, MJ.; JUAREZ, R.; BLANCO, L.; DNA-binding determinants promoting NHEJ by human Pol $\mu$ . *Nucleic Acids Res*. 2012 Dec; 40(22): 11389–11403.

MARTÍN-PARDILLOS, A., TSAALBI-SHTYLIK, A., CHEN, S., LAZARE, S., VAN OS, R. P., DETHMERS-AUSEMA, A., FAKOURI, N. B., BOSSHARD, M., APRIGLIANO, R., VAN LOON, B., SALVATORI, D. C. F., HASHIMOTO, K., VAN DER SPEK, C. D., MORIYA, M., RASMUSSEN, L. J., DE HAAN, G., RAAIJMAKERS, M. H. G. P., & DE WIND, N. (2017). Genomic and functional integrity of the hematopoietic system requires tolerance of oxidative DNA lesions. *Blood*, 130(13), 1523–1534.

MASUTANI, C.; ARAKI, M.; YAMADA, A.; KUSUMOTO, R.; NOGIMORI, T.; MAEKAWA, T. et al. Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *EMBO J*. 1999 Jun 15;18(12):3491-501.

MASUTANI, C.; KUSUMOTO, R.; IWAI, S.; HANAOKA, F. Mechanisms of accurate translesion synthesis by human DNA polymerase  $\eta$ . *EMBO J*. 2000 Jun 15; 19(12): 3100–3109.

MATEOS-GOMEZ, PA.; GONG, F.; NAIR, N.; MILLER, KM.; LAZZERINI-DENCHI, E.; SFEIR, A. Mammalian polymerase  $\theta$  promotes alternative NHEJ and suppresses recombination. *Nature*. 2015 Feb 12;518(7538):254-7.

MCINTYRE, J.; WOODGATE, R. Regulation of translesion DNA synthesis: Posttranslational modification of lysine residues in key proteins. *DNA Repair (Amst)*. 2015 May;29:166-79.

MCQUILTEN ZK.; POLIZZOTTO MN.; WOOD EM.; SUNDARARAJAN V. Myelodysplastic syndrome incidence, transfusion dependence, health care use, and complications: an Australian population-based study 1998 to 2008.

MCVEY, M1.; LEE, SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet*. 2008 Nov;24(11):529-38.

MOLDOVAN, G. L.; MADHAVAN, M. V.; MIRCHANDANI et al. DNA polymerase POLN participates in cross-link repair and homologous recombination. *Mol Cell Biol*. 2010 Feb;30(4):1088-96.

MUFTI, GJ. Chromosomal deletions in the myelodysplastic syndrome. *Leuk Res*. 1992;16(1):35-41.

NAEIM, F.; RAO, P. N.; GRODY, W. W. Myelodysplastic Syndromes. *Hematopathology: morphology, immunophenotype, cytogenetics and molecular approaches*. 1st ed. New York: Academic Press Publication, p. 129-154, 2008.

NATELSON, E. A.; PYATT, D. Acquired myelodysplasia or myelodysplastic syndrome: clearing the fog. *Adv Hematol*, 2013; 309637.

NAZHA A.; NARKHEDE M.; RADIVOYEVITCH T.; SEASTONE DJ.; PATEL BJ.; GERDS AT.; MUKHERJEE S.; KALAYCIO M.; ADVANI A.; PRZYCHODZEN B.; CARRAWAY HE.; MACIEJEWSKI JP.; SEKERES MA. Incorporation of molecular data into the Revised International Prognostic Scoring System in treated patients with myelodysplastic syndromes. *Leukemia*. 2016 Nov;30(11):2214-2220

NEUKIRCHEN, J.; SCHOONEN, W. M.; STRUPP, C.; GATTERMANN, N.; AUL, C.; HAAS, R.; GERMING, U. Incidence and prevalence of myelodysplastic syndromes: data from the Düsseldorf MDS-registry. *Leuk Res*. 2011 Dec;35(12):1591-6.

NIEMEYER, C. M.; BAUMANN, I. Myelodysplastic syndrome in children and adolescents. *Semin Hematol.*, v. 45, n. 1, p. 60-70, 2008.

OGI, T.; SHINKAI, Y.; TANAKA, K et al. Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. *Proc Natl Acad Sci U S A*. 2002;99:15548–53.

OHMORI, H.; FRIEDBERG, EC.; FUCHS, RPP et al. The Y-family of DNA polymerases. *Molecular Cell*, vol. 8, no. 1, pp. 7–8, 2001.

OLIVEIRA RTG, FRANÇA IGF, JUNIOR HLR, RIELLO GBC, BORGES DP, CAVALCANTE GM, MAGALHÃES SMM, PINHEIRO RF. c.9253-6T>c REV3L: A novel marker of poor prognosis in Myelodysplastic syndrome. *Hematol Transfus Cell Ther*. 2021 Jul-Sep;43(3):377-381.

OLNEY, H. J.; LE BEAU, M. M. The Cytogenetics and molecular biology of the Myelodysplastic syndromes. In: STEENSMA, D. P. (Ed.). *Myelodysplastic syndromes: pathobiology and clinical management*. 2nd ed. New York: Informa Healthcare, 2002. p. 49-86.

O-WANG J.; KAJIWARA, K.; KAWAMURA, K.; KIMURA, M.; MIYAGISHIMA, H.; KOSEKI, H.; TAGAWA, M.; An essential role for REV3 in mammalian cell survival: absence of REV3 induces p53-independent embryonic death. *Biochem Biophys Res Commun*. 2002 May 10;293(3):1132-7.

O-WANG, J. et al. DNA polymerase kappa, implicated in spontaneous and DNA damage-induced mutagenesis, is overexpressed in lung cancer. *Cancer Res*, v. 61, n. 14, p. 5366-9, Jul 2001.

OZTURK, S.; DEMIR, N. DNA repair mechanisms in mammalian germ cells. *Histol Histopathol*, 2011; 26(4):505-17.

PAGÈS, V.; FUCHS, RP. How DNA lesions are turned into mutations within cells? *Oncogene*. 2002 Dec 16;21(58):8957-66.

PAPAEMMANUIL, E. et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*, 2013; 122(22):3616-27.

PFEIFFER, P.; GOEDECKE, W.; KUHFITIG-KULLE, S.; OBE, G. Pathways of DNA double-strand break repair and their impact on the prevention and formation of chromosomal aberrations. *Cytogenet Genome Res*. 2004;104(1-4):7-13.

PFEILSTÖCKER, M.; TUECHLER, H.; SANZ, G. et al. Time-dependent changes in mortality and transformation risk in MDS. *Blood*. 2016;128:902–10.

PILIÉ, PG.; TANG, C.; MILLS, GB.; YAP, TA. State-of-the-art strategies for targeting the DNA damage response in cancer. *Nat Rev Clin Oncol*. 2018 Oct 24.

PILLAIRE, MJ. et al. A 'DNA replication' signature of progression and negative outcome in colorectal cancer. *Oncogene*, v. 29, n. 6, p. 876-87, Feb 2010.

PINHEIRO, R. F.; CHAUFFAILLE, M. L. Comparison of I-FISH and G-banding for the detection of chromosomal abnormalities during the evolution of myelodysplastic syndrome. *Braz J Med Biol Res*. 2009 Nov;42(11):1110-2

PLOSKY, BS.; WOODGATE, R.; Switching from high-fidelity replicases to low-fidelity lesion-bypass polymerases. *Curr Opin Genet Dev*. 2004;14:113–9.

POUSADA, G. et al. Mutational and clinical analysis of the ENG gene in patients with pulmonary arterial hypertension. *BMC Genetics*, v.17, n. 1, p. 7–9, 2016.

POWERS KT, WASHINGTON MT. Eukaryotic translesion synthesis: Choosing the right tool for the job. *DNA Repair (Amst)*. 2018 Nov; 71:127-134.

RIBEIRO, H. L. JR.; OLIVEIRA, R. T.; MAIA, A.R.; SOUSA, J.C.; HEREDIA, F.F.; MAGALHÃES, S.M.; PINHEIRO R.F. ATM polymorphism is associated with low risk myelodysplastic syndrome. *DNA Repair (Amst)*. 2013 Feb 1;12(2):87-9.

RIBEIRO, HR. Expressão de genes relacionados às vias de reparo de danos em fita dupla no DNA em pacientes com síndrome mielodisplásica. 2016. Tese (Doutorado em Ciências Médicas) – Faculdade de Medicina, Universidade Federal do Ceará. Fortaleza, 2016.

RIBEZZO, F., SHILOH, Y., & SCHUMACHER, B. (2016). Systemic DNA damage responses in aging and diseases. *Seminars in Cancer Biology*, 37–38, 26–35.

RICHARDS, S., AZIZ, N., BALE, S., BICK, D., DAS, S., GASTIER-FOSTER, J., GRODY, W. W., HEGDE, M., LYON, E., SPECTOR, E., VOELKERDING, K., & REHM, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the

Association for Molecular Pathology. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 17(5), 405–424.

RIVA L, RONCHINI C, BODINI M, LO-COCO F, LAVORGNA S, OTTONE T, MARTINELLI G, IACOBUCCI I, TARELLA C, CIGNETTI A, VOLORIO S, BERNARD L, RUSSO A, MELLONI GE, LUZI L, ALCALAY M, DELLINO GI, PELICCI PG. Acute promyelocytic leukemias share cooperative mutations with other myeloid-leukemia subgroups. *Blood Cancer J*. 2013 Sep 13;3(9):e147.

RIZZO AA, KORZHNEV DM. The Rev1-Pol $\zeta$  translesion synthesis mutasome: Structure, interactions and inhibition. *Enzymes*. 2019;45:139-181.

ROLLISON, D. E. et al. Epidemiology of myelodysplastic syndromes and chronic myeloproliferative disorders in the United States, 2001-2004, using data from the NAACCR and SEER programs. *Blood*, 2008; 112(1):45-52.

ROSSI, DJ1.; BRYDER, D.; SEITA, J.; NUSSENZWEIG, A.; HOEIJMAKERS, J.; WEISSMAN, IL. Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature*. 2007 Jun 7;447(7145):725-9.

SALE, J. E.; LEHMAN, A. R.; WOODGATE, R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol*. 2012; 13(3): 141–152.

SANTINI, V. Clinical Use of Erythropoietic Stimulating Agents in Myelodysplastic Syndromes. *Oncologist*. 2011;16 Suppl 3:35-42.

SASATANI, M., XI, Y., KAJIMURA, J., KAWAMURA, T., PIAO, J., MASUDA, Y., HONDA, H., KUBO, K., MIKAMOTO, T., WATANABE, H., XU, Y., KAWAI, H., SHIMURA, T., NODA, A., HAMASAKI, K., KUSUNOKI, Y., ZAHARIEVA, E. K., & KAMIYA, K. (2017). Overexpression of Rev1 promotes the development of carcinogen-induced intestinal adenomas via accumulation of point mutation and suppression of apoptosis proportionally to the Rev1 expression level. *Carcinogenesis*, 38(5), 570–578.

SAUTA E, ROBIN M, BERSANELLI M, TRAVAGLINO E, MEGGENDORFER M, ZHAO LP, CABALLERO BERROCAL JC, SALA C, MAGGIONI G, BERNARDI M, DI GRAZIA C, VAGO L, RIVOLI G, BORIN L, D'AMICO S, TENTORI CA, UBEZIO M, CAMPAGNA A, RUSSO A, MANNINA D, LANINO L, CHIUSOLO P, GIACCONE L, VOSO MT, RIVA

M, OLIVA EN, ZAMPINI M, RIVA E, NIBOUREL O, BICCHIERI M, BOLLI N, RAMBALDI A, PASSAMONTI F, SAVEVSKI V, SANTORO A, GERMING U, KORDASTI S, SANTINI V, DIEZ-CAMPELO M, SANZ G, SOLE F, KERN W, PLATZBECKER U, ADES L, FENAUX P, HAFERLACH T, CASTELLANI G, DELLA PORTA MG. Real-World Validation of Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *J Clin Oncol*. 2023 Mar 17;JCO2201784. doi: 10.1200/JCO.22.01784. Epub ahead of print. PMID: 36930857.

SCHANZ, J. et al. New comprehensive cytogenetic scoring system for primary myelodysplastic syndromes (MDS) and oligoblastic acute myeloid leukemia after MDS derived from an international database merge. *J Clin Oncol*, 2012; 30(8):820-9.

SCHMITTGEN, T.D.; LIVAK, K.J. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.*, v. 3, n. 6, p. 1101-8, 2008

SCHNITTGER, S.; BACHER, U.; HAFERLACH, C, et al. Characterization of NPM1-mutated AML with a history of myelodysplastic syndromes or myeloproliferative neoplasms. *Leukemia* 2011;25:615-621

SEKERES, M. A. The epidemiology of myelodysplastic syndromes. *Hematol Oncol Clin North Am*, v. 24, n. 2, p. 287-94, Apr 2010.

SEKI, M. et al. High-efficiency bypass of DNA damage by human DNA polymerase Q. *EMBO J*. 2004.v. 23, n. 22, p. 4484-94, Nov.

SEKI, M.; MARINI, F.; WOOD, RD. POLQ (Pol theta), a DNA polymerase and DNA dependent ATPase in human cells. *Nucleic Acids Res*. 2003. Nov v. 31, n. 21, p. 6117-26.

SFEIR, A.; SYMINGTON, LS. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway?. *Trends Biochem Sci*. 2015 Nov;40(11):701-714.

SHADDUCK, R.K., et al. Recent advances in myelodysplastic syndromes. *Exp Hematol*, v.35, n.4 Suppl 1, p.137-43, Apr 2007.

SHAHRABI, S.; et al. Genetics and epigenetics of myelodysplastic syndromes and response to drug therapy: new insights. *Oncology Reviews*. 2016; 10(4): 311.

SHARIEF, FS. et al. Cloning and chromosomal mapping of the human DNA polymerase theta (POLQ), the eighth human DNA polymerase. *Genomics*. v. 59, n.1, p. 90-6, Jul 1999.

SHIH, LY.; HUANG, CF.; WANG, PN. et al. Acquisition of FLT3 or N-ras mutations is frequently associated with progression of myelodysplastic syndrome to acute myeloid leukemia. *Leukemia* 2004;18:466-475

SHIMA, N.; HARTFORD, SA.; DUFFY, T.; WILSON, LA.; SCHIMENTI, KJ.; SCHIMENTI, JC. Phenotype-based identification of mouse chromosome instability mutants. *Genetics*. 2003 Mar;163(3):1031-40.

SHIMA, N.; MUNROE, RJ.; SCHIMENTI, JC.; The mouse genomic instability mutation *chaos1* is an allele of *Polq* that exhibits genetic interaction with *Atm*. *Mol Cell Biol*. 2004 Dec;24(23):10381-9.

SHRIVASTAV, M.; DE HARO, LP.; NICKOLOFF, JA. Regulation of DNA double-strand break repair pathway choice. *Cell Res*. 2008 Jan;18(1):134-47.

SIMONS, A., SHAFFER, L. G., & HASTINGS, R. J. (2013). Cytogenetic Nomenclature: Changes in the ISCN 2013 Compared to the 2009 Edition. *Cytogenetic and Genome Research*, 141(1), 1–6.

SOLÉ, F.; LUÑO, E.; SANZO, C.; ESPINET, B.; SANZ, G. F.; CERVERA, J.; CALASANZ, M. J.; CIGUDOSA, J. C.; MILLÀ, F.; RIBERA, J. M.; BUREO, E.; MARQUEZ, M. L.; ARRANZ, E.; FLORENSA, L. Identification of novel cytogenetic markers with prognostic significance in a series of 928 patients with primary myelodysplastic syndromes. *Haematologica*, v. 90, p.1168-1178, 2005.

SUZUKIA, T.; GRÚZA, P.; HONMAA, M.; ADACHIB, N.; NOHMIA, T. The role of DNA polymerase  $\zeta$  in translesion synthesis across bulkyDNA adducts and cross-links in human cells. *Mutation Research*. 791–792 (2016) 35–41

SWERDLOW, S. H. et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th edition. Lyon, France: IARC Press, 2008.

SWERDLOW, SH; CAMPO, E; PILERI, SA; HARRIS, NL; STEIN, H. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016 May 19;127(20):2375-90.

TAKATA, K.; SHIMIZU, T.; IWAI, S.; WOOD, RD. Human DNA polymerase N (POLN) is a low fidelity enzyme capable of error-free bypass of 5S-thymine glycol. *J Biol Chem*. 2006 Aug 18;281(33):23445-55

Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res*. 2019;47(D1):D941–7.

TERRADOS, G.; CAPP, J. P.; CANITROT, Y. et al. Characterization of a natural mutator variant of human DNA polymerase lambda which promotes chromosomal instability by compromising NHEJ. *PLoS One*. 2009;4:e7290.

THERNEAU, T.; LUMLEY, T. Survival: survival analysis, including penalised likelihood. [http://cran.r-project.org/package = survival](http://cran.r-project.org/package=survival). 2011.

TOMASETTI C, LI L, VOGELSTEIN B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science*. 2017;355(6331):1330–4.

UCHIYAMA, M.; TERUNUMA, J.; HANAOKA, F. The Protein Level of Rev1, a TLS Polymerase in Fission Yeast, Is Strictly Regulated during the Cell Cycle and after DNA Damage. *PLoS One*. 2015 Jul 6;10(7):e0130000.

Vaisman A, Woodgate R. Translesion DNA polymerases in eukaryotes: what makes them tick? *Crit Rev Biochem Mol Biol*. 2017 Jun;52(3):274-303. doi: 10.1080/10409238.2017.1291576. Epub 2017 Mar 9.

VAISMAN, A.; MASUTANI, C.; HANAOKA, F.; CHANEY, SG. Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase eta. *Biochemistry*. 2000 Apr 25;39(16):4575-80.

VALERIA, V.; RAMON, V.; TIU AND HEESUN J. Pathogenesis of myelodysplastic syndromes: an overview of molecular and non-molecular aspects of the disease. *Blood Res*. 2014 Dec; 49(4): 216–227.

VANDESOMPELE, J.; DE PRETER, K.; PATTYN, F.; POPPE, B.; VAN ROY, N.; DE PAEPE, A.; SPELEMAN, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, v. 3, n. 7, p. RESEARCH0034, 2002.

Varadi V, Bevier M, Grzybowska E, Johansson R, Enquist K, Henriksson R, Butkiewicz D, Pamula-Pilat J, Tecza K, Hemminki K, et al. Genetic variation in genes encoding for polymerase zeta subunits associates with breast cancer risk, tumour characteristics and survival. *Breast Cancer Res Treat.* 2011;129(1):235–245.

VARDIMAN, J. W. et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood.* 2009; 114(5): 937-951.

VASSALLO, J., & MAGALHÃES, S. M. M. (2009). Síndromes mielodisplásicas e mielodisplásicas/mieloproliferativas. *Revista Brasileira de Hematologia e Hemoterapia*, 31(4), 267–272.

VISCONTE, V., TIU, R. V., & ROGERS, H. J. (2014). Pathogenesis of myelodysplastic syndromes: an overview of molecular and non-molecular aspects of the disease. *Blood Research*, 49(4), 216.

VISSER O.; TRAMA A.; MAYNADIÉ M.; STILLER C.; MARCOS-GRAGERA R.; DE ANGELIS R.; MALLONE S.; TEREANU C.; ALLEMANI C.; RICARDI U.; SCHOUTEN HC.; RARECARE WORKING GROUP. Incidence, survival and prevalence of myeloid malignancies in Europe. *Eur J Cancer.* 2012 Nov;48(17):3257-66

WALTER, MJ.; SHEN, D.; DING, L. et al. Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med.* 2012;366(12): 1090-1098.

WANG, H. et al. Analysis of specialized DNA polymerases expression in human gliomas: association with prognostic significance. *Neuro Oncol*, v. 12, n. 7, p. 679-86, Jul 2010

WANG, X. et al. Mutational analysis of thirty-two double-strand DNA break repair genes in breast and pancreatic cancers. *Cancer Res*, v. 68, n. 4, p. 971-5, Feb 2008.

WASHINGTON, MT.; JOHNSON, ROBERT E, PRAKASH, LOUISE, PRAKASH, SATYA. Human DINB1-encoded DNA polymerase kappa is a promiscuous extender of mispaired

primer termini. Proceedings of the National Academy of Sciences of the United States of America, 2002. 99(4), 1910-4.

WATERS, L. S., MINESINGER, B. K., WILTROUT, M. E., D'SOUZA, S., WOODRUFF, R. V., & WALKER, G. C. (2009). Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiology and Molecular Biology Reviews : MMBR*, 73(1), 134–154.

WHO. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System. Geneva, World Health Organization, 2011 (WHO/NMH/NHD/MNM/11.1) (<http://www.who.int/vmnis/indicators/haemoglobin.pdf>, acesso em 10/08/2018).

WHO. Nutritional anaemias. Report of a WHO scientific group. *World Health Organ Tech Rep Ser*. 1968;405:5-37.

WITTSCHIEBEN, J. P.; PATIL, V.; GLUSHETS, V. et al. Loss of DNA polymerase zeta enhances spontaneous tumorigenesis. *Cancer Res*. 2010 Apr 1;70(7):2770-8.

WOLFLE, W. T., WASHINGTON, M. T., PRAKASH, LOUISE, & PRAKASH, SATYA. Human DNA polymerase kappa uses template – primer misalignment as a novel means for extending mispaired termini and for generating single-base deletions. *Genes & Development*, 2003. 17(17), 2191-9.

WU F, LIN X, OKUDA T, HOWELL SB. DNA polymerase zeta regulates cisplatin cytotoxicity, mutagenicity, and the rate of development of cisplatin resistance. *Cancer Res*. 2004;64:8029-35.

WYATT, DW.; FENG, W.; CONLIN, MP.; YOUSEFZADEH, MJ.; ROBERTS, SA.; MIECZKOWSKI, P. et al. Essential Roles for Polymerase  $\theta$ -Mediated End Joining in the Repair of Chromosome Breaks. *Mol Cell*. 2016 Aug 18;63(4):662-673.

YADAV S.; MUKHOPADHYAY S.; ANBALAGAN M.; MAKRIDAKIS N. Somatic Mutations in Catalytic Core of POLK Reported in Prostate Cancer Alter Translesion DNA Synthesis. *Hum Mutat*. 2015 Sep;36(9):873-80.

YANG, J. et al. Altered DNA polymerase iota expression in breast cancer cells leads to a reduction in DNA replication fidelity and a higher rate of mutagenesis. *Cancer Res*, v. 64, n. 16, p. 5597-607, Aug 2004.

YAO, Y.; DAI, W. Genomic Instability and Cancer. *J Carcinog Mutagen*. 2014; 5.

YOUSEFZADEH MJ, WOOD RD. DNA polymerase POLQ and cellular defense against DNA damage. *DNA Repair (Amst)* 2013;12(1):1-9.

YOUSEFZADEH, M. J.; WOOD, R. D. DNA polymerase POLQ and cellular defense against DNA damage. *DNA Repair (Amst)*. 2013 Jan 1;12(1):1-9.

YU, X.; CHINI, C. C.; HE, M.; MER, G.; CHEN, J. The BRCT domain is a phospho-protein binding domain. *Science*. 2003;302:639-42

ZATREANU, D., ROBINSON, H.M.R., ALKHATIB, O. et al. Pol $\theta$  inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. *Nat Commun* 12, 3636 (2021).

ZELENSKY, NA.; SCHIMMEL, J.; KOOL, H.; KANAAR, R.; TIJSTERMAN, M. Inactivation of Pol  $\theta$  and C-NHEJ eliminates off-target integration of exogenous DNA. *Nat Commun*. 2017. Jul 7;8(1):66

ZHOU, BL.; PATA, JD.; STEITZ, TA. Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. *Molecular Cell*, 2001; 8(2), 427-37.

ZHOU, T.; CHEN, P.; GU, J., et al. Potential relationship between inadequate response to DNA damage and development of myelodysplastic syndrome. *Int. J. Mol. Sci.*, v. 5, n. 16, p. 966-89, 2015.

ZHOU, W.; CHEN, Y.; LIU, X.; CHU, P.; LORIA, S.; WANG, Y. Expression of DNA Translesion Synthesis Polymerase  $\eta$  in Head and Neck Squamous Cell Cancer Predicts Resistance to Gemcitabine and Cisplatin-Based Chemotherapy. *PLoS One*. 2013; 8(12): e83978

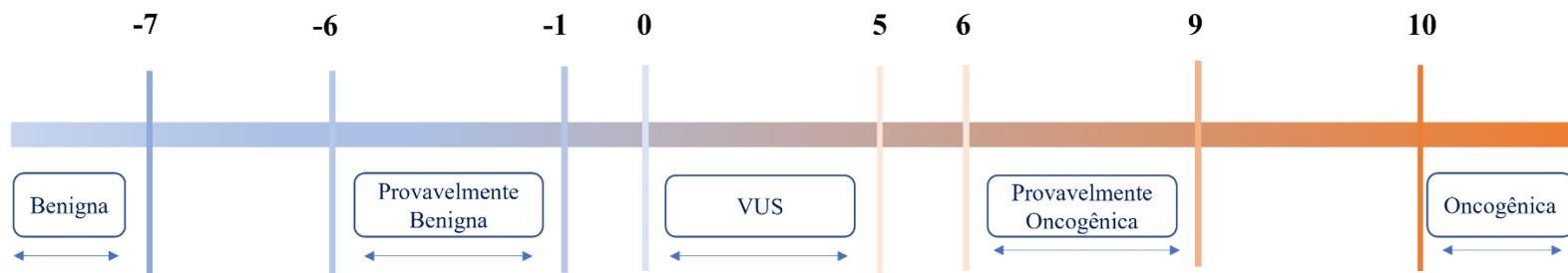
**Anexo A** - Critérios de classificação de variantes de somáticas em câncer proposto pelos grupos de trabalho: *Clinical Genome Resource (ClinGen)*, *Cancer Genomics Consortium (CGC)*, e *Variant Interpretation for Cancer Consortium (VICC)*

CRITÉRIOS PARA EVIDÊNCIA DE ONCOGENICIDADE		
CATEGORIA	EVIDÊNCIA	CRITÉRIOS
<b>Muito forte 8 pontos</b>	OVS1	Variante que comprometa completamente a função proteica ( <i>null</i> ) em um gene cuja a perda de função (LOF) é um mecanismo de doença conhecido - <i>nonsense</i> / <i>frameshift</i> / sítios de <i>splicing</i> $\pm$ 1 ou 2 / códon inicial / deleções de éxons
<b>Forte 4 pontos</b>	OS1	Mesma alteração de aminoácido que uma variante antes descrita como oncogênica, independentemente da troca de nucleotídeos. Exemplo: Val→Leu causada por ambos G>C or G>T no mesmo códon
	OS2	Estudos funcionais ( <i>in vitro</i> / <i>in vivo</i> ) que demonstrem um efeito oncogênico da variante
	OS3	Variante localizada em um <i>hotspot</i> no <a href="http://cancerhotspots.org">cancerhotspots.org</a> com pelo menos 50 amostras com uma variante somática na mesma posição de aminoácidos e a mesma contagem de alteração de aminoácidos em pelo menos 10 amostras
<b>Moderada 2 pontos</b>	OM1	Localizado em uma parte crítica e bem estabelecida de um domínio funcional (por exemplo, sítio ativo de uma enzima)
	OM2	Alterações no comprimento da proteína como resultado de deleções/inserções em um oncogene ou gene supressor de tumor conhecido, ou variantes de <i>stop-loss</i> em um gene supressor de tumor conhecido.
	OM3	Localizado em um dos hotspots em <a href="http://cancerhotspots.org">cancerhotspots.org</a> com menos de 50 amostras com uma variante somática na mesma posição de aminoácidos, e a mesma contagem de alteração de aminoácidos no <a href="http://cancerhotspots.org">cancerhotspots.org</a> é de pelo menos 10.
	OM4	Variante missense em um resíduo de aminoácido onde uma variante missense diferente determinada como oncogênica (usando este padrão) foi documentada. A diferença de aminoácido em relação ao aminoácido de referência deve ser maior ou pelo menos aproximadamente a mesma para a alteração missense determinada como oncogênica.
<b>Suporta 1 ponto</b>	OP1	Todas as linhas de evidência computacional utilizadas suportam um efeito oncogênico da variante (conservação/evolução, impacto de <i>splicing</i> , etc.).
	OP2	Variante somática em um gene em uma malignidade com uma única etiologia genética. Exemplo: o retinoblastoma é causado pela inativação bialélica de RB1.
	OP3	Localizado em um dos hotspots em <a href="http://cancerhotspots.org">cancerhotspots.org</a> com menos de 50 amostras com uma variante somática na mesma posição de aminoácidos, e a mesma contagem de alteração de aminoácidos no <a href="http://cancerhotspots.org">cancerhotspots.org</a> é de pelo menos 10.
	OP4	Ausente de controles (ou em uma frequência extremamente baixa) no banco de dados de agregação de genoma (gnomAD).

Legenda: OVS1: Oncogenic Very Strong-1, OS1: Oncogenic Strong-1, OS2: Oncogenic Strong-2, OS3: Oncogenic Strong-3, OM:1 Oncogenic Moderate-1, OM2: Oncogenic Moderate-2, OM3: Oncogenic Moderate-3, OM4: Oncogenic Moderate-4, OP1: Oncogenic Supporting-1, OP2: Oncogenic Supporting-1, OP3: Oncogenic Supporting-3, OP4: Oncogenic Supporting-4

CRITÉRIOS PARA EVIDÊNCIA DE BENIGNIDADE		
CATEGORIA	EVIDÊNCIA	CRITÉRIOS
<b>Muito forte -8 pontos</b>	SBVS1	Frequência alélica > 5% no banco de dados de agregação do genoma (gnomAD) em qualquer uma das 5 populações continentais gerais: africana, leste da Ásia, europeia (não finlandesa), latina e sul da Ásia.
<b>Forte -4 pontos</b>	SBS1	Frequência alélica > 1% no banco de dados de agregação do genoma (gnomAD) em qualquer uma das 5 populações continentais gerais: africana, leste da Ásia, europeia (não finlandesa), latina e sul da Ásia.
	SBS2	Estudos funcionais bem estabelecidos <i>in vitro</i> ou <i>in vivo</i> não mostram efeitos oncogênicos.
<b>Suporta -1 pontos</b>	SBP1	Todas as linhas de evidência computacional utilizadas não sugerem impacto da variante (conservação/evolução, impacto de splicing, etc.).
	SBP2	Variante sinônima (silenciosa) para a qual os algoritmos de previsão de splicing não prevêem nenhum impacto na sequência de consenso de splicing nem a criação de um novo local de splicing e o nucleotídeo não é altamente conservado.

Legenda: SBVS1: Somatic Benign Very Strong-1, SBS1: Somatic Benign Strong-1, SBS2: Somatic Benign Strong-2, SBP1: Somatic Benign Supporting-1, SBP2: Somatic Benign Supporting-



Legenda: Régua utilizada para avaliação de oncogenicidade após somatório das categorias de evidência

## **APÊNDICE: Termo de Consentimento Livre e Esclarecido**



### **UNIVERSIDADE FEDERAL DO CEARÁ FACULDADE DE MEDICINA DEPARTAMENTO DE MEDICINA CLÍNICA**

#### **TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

#### **PROJETO: RASTREAMENTO DE VARIANTES GENÉTICAS EM DNA POLIMERASES COM ATIVIDADE TRANSLEÇÃO NA NEOPLASIA MIELODISPLÁSICA**

Você está sendo convidado a participar de um projeto de pesquisa. Sua participação é importante, porém, você não deve participar contra a sua vontade. Leia atentamente as informações abaixo e faça qualquer pergunta que desejar.

O abaixo assinado, \_\_\_\_\_,  
\_\_\_\_ anos, RG nº \_\_\_\_\_, declara que é de livre e espontânea vontade que está participando como voluntário do projeto de pesquisa supracitado, de responsabilidade do pesquisador Roberta Taiane Germano de Oliveira / Prof. Dr. Ronald Feitosa Pinheiro. O abaixo assinado está ciente de que:

#### **NATUREZA E PROPÓSITO DO ESTUDO**

O objetivo da pesquisa é estudar o material genético de participantes portadores de Neoplasia Mielodisplásica, gerando dados que favoreçam a uma melhor compreensão do surgimento e evolução clínica desta doença.

#### **PROCEDIMENTOS A SEREM REALIZADOS E RESPONSABILIDADES**

A amostra biológica utilizada na presente pesquisa corresponde a medula óssea do participante. A medula óssea corresponde a um tecido líquido-gelatinoso que ocupa o interior dos ossos, sendo conhecida popularmente por 'tutano'. Na medula óssea são produzidos

os componentes do sangue: as hemácias (glóbulos vermelhos), os leucócitos (glóbulos brancos) e as plaquetas.

A coleta da medula óssea será realizada por médico hematologista experiente com agulha de mielograma mediante punção esternal. O osso do esterno é um osso chato, plano e ímpar localizado no centro do tórax. O osso do esterno é um importante osso hematopoético, ou seja, de produção das células sanguíneas.

Serão coletadas somente 3mL de amostra de medula óssea com o uso de uma agulha específica para aspiração da medula óssea. Todo o procedimento de coleta da medula óssea é realizado mediante administração de anestésico local com duração máxima de 15 minutos. O procedimento de coleta esternal da medula óssea, pode, raramente, determinar uma equimose (mancha arroxeadada) ao redor do local de onde foi retirado, desaparecendo em poucos dias, e poderá ocorrer dor discreta e de fácil alívio, podendo ser, ocasionalmente, dor de maior intensidade. Excepcionalmente, poderá ocorrer sangramento local. Raramente, pode ocorrer infecção local. Para participantes com maior sensibilidade dolorosa existe a possibilidade de realizar o procedimento sob anestesia geral. A coleta deste material é necessária, em poucos casos, pela amostra ser insuficiente ou inadequada para análise.

Após o procedimento, serão coletados dados pessoais de sexo e idade, apresentação clínica e checagem dos exames laboratoriais indicados para esclarecimento diagnóstico, tais como: hemograma completo, citologia e histologia (análise microscópica das células) da medula óssea, avaliação para depósitos de ferro medular, bem como outros exames que são feitos mesmo para os participantes que não participam de nenhuma pesquisa.

É de sua responsabilidade: comparecer nas datas e horários informados e submeter-se aos procedimentos de rotina do serviço.

## **PARTICIPAÇÃO VOLUNTÁRIA**

Sua participação é voluntária, você não receberá nenhum ressarcimento monetário por participar da pesquisa e tem a liberdade de desistir ou interromper a participação neste estudo no momento em que desejar. Neste caso, você deve informar imediatamente sua decisão ao pesquisador responsável ou a qualquer um membro de sua equipe, sem necessidade de qualquer explicação e sem que isto venha interferir no seu atendimento nesta instituição.

Independentemente de seu desejo e consentimento, sua participação no estudo poderá ser interrompida, em função da ocorrência de qualquer doença que, a critério médico, prejudique a continuação de sua participação no estudo, do não cumprimento das normas

estabelecidas, de qualquer outro motivo que, a critério da pesquisadora, seja do interesse de seu próprio bem-estar ou dos demais participantes e, por fim, da suspensão do estudo como um todo.

O Laboratório de Citogenômica do Câncer o manterá informado, em tempo oportuno, sempre que houver alguma informação adicional que possa influenciar seu desejo de continuar participando no estudo e prestará qualquer tipo de esclarecimento em relação ao progresso da pesquisa, conforme sua solicitação.

### **DIVULGAÇÃO DE INFORMAÇÕES QUANTO À PARTICIPAÇÃO NO ESTUDO**

Os registros que possam identificar sua identidade serão mantidos em sigilo, a não ser que haja obrigação legal de divulgação. Você não será identificado por ocasião da publicação dos resultados obtidos.

Contudo, o(s) monitor(es) do Estudo, auditor(es), membros do Comitê de Ética em Pesquisa com Seres Humanos, ou autoridades do(s) órgão(s) regulamentar(es) envolvido(s) terão direito de ter acesso aos registros originais de dados clínicos de sua pessoa, coletados durante a pesquisa, na extensão em que for permitido pela Lei e regulamentações aplicáveis, como o propósito de verificar os procedimentos e dados do estudo, sem, no entanto, violar a condição de que tais informações são confidenciais.

### **CONTATOS E PERGUNTAS**

Caso surja algum imprevisto ou dúvidas, você deverá entrar em contato solicitar contato direto com o pesquisador responsável pelo estudo: **Roberta Taiane Germano de Oliveira (99662327)** ou com seu **Orientador acadêmico: Prof. Dr. Ronald Feitosa Pinheiro (81881972)** ou entre em contato com o **Comitê de Ética em Pesquisa da UFC/PROPESQ** - Rua Coronel Nunes de Melo, 1000 - Rodolfo Teófilo / Telefone: 3366-8344/46. (Horário: 08:00-12:00 horas de segunda a sexta-feira). O CEP/UFC/PROPESQ é a instância da Universidade Federal do Ceará responsável pela avaliação e acompanhamento dos aspectos éticos de todas as pesquisas envolvendo seres humanos.

Somente assine este termo se você tiver a certeza de que recebeu todos os esclarecimentos e informações para decidir conscientemente sobre a sua participação neste estudo.

## ASSINATURAS

Autorizo o acesso às minhas informações de saúde aos membros da equipe de pesquisadores, nas condições estabelecidas descritas nos itens acima.

Não renunciei qualquer direito legal que eu venha a ter ao participar deste Estudo.

Não receberei nenhum ressarcimento monetário por participar da pesquisa.

Eu, por fim, declaro que li cuidadosamente todo este documento denominado **Termo de Consentimento Livre e Esclarecido** e que, após a assinatura, tive oportunidade de fazer perguntas sobre o conteúdo do mesmo e também sobre o referido estudo, recebendo explicações que responderam por completo minhas dúvidas e reafirmando estar livre e espontaneamente decidido a participar do estudo, ficando munido de uma via do documento assinado pelo pesquisador responsável.

\_\_\_\_/\_\_\_\_/\_\_\_\_

**Data**

\_\_\_\_\_

**Assinatura do participante da pesquisa**

\_\_\_\_/\_\_\_\_/\_\_\_\_

**Data**

\_\_\_\_\_

**Assinatura do pesquisador responsável**

\_\_\_\_/\_\_\_\_/\_\_\_\_

**Data**

\_\_\_\_\_

**Assinatura do responsável pela aplicação do TCLE**