



**UNIVERSIDADE FEDERAL DO ESTADO DO CEARÁ
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS MÉDICAS
MESTRADO EM CIÊNCIAS MÉDICAS**

ANTONIO WESLEY ARAÚJO DOS SANTOS

**ANÁLISE DE VARIANTES GENÉTICAS EM SÍTIOS DE *SPLICING* DE GENES DO
MECANISMO DE SÍNTESE DE TRANSLEÇÃO EM PACIENTES COM NEOPLASIA
MIELODISPLÁSICA**

FORTALEZA - CE

2025

ANTONIO WESLEY ARAÚJO DOS SANTOS

**ANÁLISE DE VARIANTES GENÉTICAS EM SÍTIOS DE *SPLICING* DE GENES DO
MECANISMO DE SÍNTESE DE TRANSLEÇÃO EM PACIENTES COM NEOPLASIA
MIELODISPLÁSICA**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Médicas, do Departamento de Medicina Clínica da Universidade Federal do Ceará como requisito final para obtenção do título de Mestre em Ciências Médicas.

Orientador: Prof. Dr. Ronald Feitosa Pinheiro

FORTALEZA - CE

2025

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)

- S233a Santos, Antonio Wesley Araújo.
Análise de variantes genéticas em sítios de splicing de genes do mecanismo de síntese de transcrição em
pacientes com neoplasia mielodisplásica / Antonio Wesley Araújo Santos. – 2025.
74 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Medicina, Programa de Pós-
Graduação em Ciências Médicas, Fortaleza, 2025.
Orientação: Prof. Dr. Ronald Feitosa Pinheiro.
1. Neoplasia Mielodisplásica. 2. Splicing. 3. Next Generation Sequencing. I. Título.

CDD 610

ANTONIO WESLEY ARAÚJO DOS SANTOS

ANÁLISE DE VARIANTES GENÉTICAS EM SÍTIOS DE *SPLICING* DE GENES DO
MECANISMO DE SÍNTESE DE TRANSLAÇÃO EM PACIENTES COM NEOPLASIA
MIELODISPLÁSICA

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Médicas, do Departamento de Medicina Clínica da Universidade Federal do Ceará como requisito final para obtenção do título de Mestre em Ciências Médicas.

Orientador: Prof. Dr. Ronald Feitosa Pinheiro

BANCA EXAMINADORA

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

Roberta Taiane Germano de Oliveira
Universidade Federal do Ceará (UFC)

Silvia Maria Meira Magalhaes
Universidade Federal do Ceará (UFC)

Marcela Cavalcante de Andrade Silva
Universidade Federal de Alagoas (UFAL)

FORTALEZA - CE

2025

Dedico este trabalho à minha Mãe, Francisca
Aquele que me trouxe aqui e me apresentou
a sua forma de ver o mundo. Te amo infinitamente!

AGRADECIMENTOS

Gostaria de utilizar este momento para trazer uma enxurrada de sentimentos acumulados que hora me emocionam e hora me fazem rir bastante. Um momento de resiliência e muita compaixão comigo mesmo e com todos aqueles que influenciaram na construção deste trabalho, ou até mesmo na motivação para continuar persistindo para finalizá-lo com o maior êxito possível.

Primeiramente, agradeço a **Deus e todos os meus guias espirituais**, uma vez que, dentro de crenças variadas neste mundo tão miscigenado me trouxe uma pitada de amor pela vida e pela continuidade dela. Sem eles eu não seria metade do que sou hoje.

Agradeço a minha família carnal, daquela aonde eu vim, especialmente **minha mãe Francisca e meu Pai Ancelmo**, que cultivaram a semente do amor para que eu pudesse vir a esta terra com muita saúde, garra e mais algumas características de personalidade que eles foram me apresentando como consequências da vida que eu iria conhecer pouco a pouco.

Aos meus irmãos **Wellington, Lidiane e Alessandra**, que da maneira como cada um aprendeu a viver, trouxe seus conselhos, sermões e discursos de motivação para que eu pudesse sempre vencer na vida, mesmo quando eu achava que não tinha chance alguma, mesmo quando eu mesmo tinha pensamentos de desistência.

Aos meus sobrinhos **Yara Maria, Lucas Lessa, Michael Douglas** e o mais novinho sobrinho neto, **Benício**. A felicidade que trouxeram e trazer a família é um pilar essencial para fortificar mais ainda nossos sentimentos fraternos.

Enfatizo aqui o meu agradecimento novamente a Deus por manter nossa família sempre unida, pois nesse caminho, as coisas da vida parecem fluir de maneira mais simples e justas.

Agradeço ao **Wesley que mora nos meus pensamentos**, aquele que me faz refletir sobre qual caminho seguir e muito me deixa na dúvida, mas nunca parado. Agradeço por ser aquele que equilibra o excesso de carga e a falta de inspiração própria que pode acontecer nos nossos momentos mais fracos.

Agradeço ao meu **Prof. Dr. Orientador Ronald Pinheiro**, por trazer sua forte bagagem científica, acadêmica e humana para me orientar no caminho certo da história acadêmica. Acredito que tenha sido uma das figuras essenciais para que eu pudesse evoluir não só de maneira acadêmica, pois sendo médico, é impossível renunciar ao seu lado humano, mesmo quando utiliza de pressão e outras táticas eficazes de sua experiência que, de primeira vista, me deixem estressado ou desmotivado por algo. Mas refletindo bem, é um ser humano que está no

meu caminho da evolução nesta vida que me ensina que sem dor na batalha não há ganho real, muito menos valorização.

Meu agradecimento a **Dra. Juliana Cordeiro**, pois na época que decidir trilhar uma carreira acadêmica, foi ela quem me ofereceu a oportunidade de conhecer este mundo tão cheio de caminhos e tão fascinante.

Aos meus amigos e companheiros de toda hora **Daniela Borges, Gabriel Frota e Roberta Taiane**. Sem dúvidas, parceiros que me levantam de se me veem cair, e se me deixam cair, é porque algum ensinamento também precisam me dar. Obrigado por todo apoio e todo carinho, certamente jamais esquecerei deste os primeiros momentos enquanto eu era um simples estudante de faculdade até hoje, um estudante de mestrado que me tomou um rumo bastante interessante e importante na minha vida. Bastante individual, mas certamente cheio de muitas nuances que despertam a nostalgia.

Agradeço a **Letícia, João Victor, Leonardo, Germana, Cinthya, Rinna, Izabelle, Marília, Bruno e mais alguns outros** que compartilharam suas experiências comigo para que contribuísse com meu crescimento acadêmico e pessoal.

Por fim, agradeço ao **CNPQ, à Universidade Federal do Ceará (UFC) e ao Programa de Pós-Graduação em Ciências Médicas** pelo fomento investido neste trabalho e orientações acadêmicas para que ele pudesse ser realizado e concluído.

Pedi a Deus flores,
Ele me deu a chuva!

RESUMO

A Neoplasia mielodisplásica é uma doença caracterizada por citopenias no sangue periférico em indivíduos predominantemente idosos. A SMD é uma doença de lesões no DNA e na maquinária do *splicing*. Existem cerca de oito polimerases que atuam no mecanismo de reparo do DNA com atividade de translesão. Elas estão agrupadas em quatro famílias: família Y (*POLH*, *POLI*, *POLK* e *REVI*), família B (*POLZ*, cuja subunidade catalítica é *REV3L*), família A (*POLQ* e *POLN*) e família X (*POLL*). O estudo buscou analisar a presença de variantes patogênicas nas regiões canônicas de *splicing* de genes de reparo do DNA com atividade de translesão. Foram analisados 30 pacientes com SMD, divididos entre 13 portadores de SMD-EB e 17 portadores de SMD com Sideroblastos em Anel. Foi realizada sequenciamento de nova geração (NGS) da região CDS dos genes com cobertura de 400x conforme orientações do fabricante (Illumina). Após a fase de chamadas de variantes, as análises foram feitas pelo software SpliceAI, sendo avaliadas as variantes com delta escore $\geq 0,8$. Resultados: Foram incluídos no estudo 30 pacientes portadores de SMD. Deste total, 20 pacientes apresentaram 11 variantes em regiões canônicas do *splicing*. O maior número de casos foi detectado no gene *POLI* que apresentou duas variantes, uma na região aceptora e outra na região doadora com variante de frequência alélica (VAF) superior a 36%. As variantes identificadas de *POLI* foram mais observadas em SMD-EB (80%) do que na SMD-SA. O gene *POLN* apresentou três variantes em região aceptora, sendo em dois casos de SMD-EB e um caso de SMD-SA. O gene *POLQ* apresentou uma variante em região doadora na SMD-EB. Já o gene *REVI* apresentou duas variantes na região aceptora também na SMD-EB. Por outro lado, tanto o gene *POLL* quanto *POLM* apresentaram variantes na SMD-SA, sendo duas variantes na região aceptora de *POLL* e uma variante na região doadora de *POLM*. Conclusão: Apesar do pequeno número de casos avaliados, esses resultados de *POLI* não demonstram, no primeiro momento, capacidade de identificar um prognóstico diferente entre os casos de anemia refratária com excesso de blastos, isto é, ter ou não variantes em regiões aceptoras ou doadoras da região de *splicing* no gene *POLI* na SMD-EB não adiciona nenhum efeito prognóstico. No entanto, apesar de bastante preliminar, os dados das variantes em casos de SMD com sideroblastos em anel (uma doença caracterizada em análise molecular, por ter forte relação com mutações em gene de *splicing*) parecem ser condizentes com uma doença que apresenta uma menor profundidade de citopenias com hemoglobinas maiores (consequente menos necessidade transfusional) e plaquetas maiores (acima de 50.000/microlitro), sugerindo que a presença de variantes em regiões aceptoras ou doadoras da região de *splicing* no gene *POLI* nos casos de SMD com sideroblastos em anel possuem um possível efeito protetor.

ABSTRACT

Myelodysplastic Syndrome (MDS) is a disease characterized by cytopenia in peripheral blood from old patients. MDS is a disease with DNA lesions and disturbs on splicing mechanism. There are eight polymerases with translesion activity. They are amount into four protein families: Family Y: POH, POLI, POLK and REV1, Family B (POLZ, with catalytic activity region REV3L), Family A (POLQ and POLN) and Family X (POLL). Our claim was analyse pathogenic variants on canonical splicing region of DNA repair genes with translesion activity. 30 MDS patients were included, 17 whom were MDS-EB and 13 were MDS-RS, but only 20 patients showed splicing variants. We did next Generation sequence (NGS) from CDS gene with coverage of 400x by Illumina. After variants calling, the variants were submitted to SpliceAI software to evaluate your delta score, were considered only delta escore $\geq 0,8$ results. From all 20 cases, 11 variants were detected in canonical splice region for all genes studied. POLI showed two variants, each one in acceptor and donor regions. Was the gene with highest number of cases with VAF $>36\%$ predominately in MDS-EB (80%) than MDS-RS and showed cases with higher cytogenetic disorders. POLN showed three variants on acceptor region on two cases of MDS-EB and another one in MDS-RS. POLQ showed only on variant on donor region of MDS-EB patient. REV1 showed two variants on acceptor region on MDS-EB patient. On the other hand, POLL and POLM showed variants only to MDS-RS, two variants on acceptor region of POLL and one on donor region of POLM. Conclusion: Despite of small number of cases evaluated, these POLI results do not demonstrate, for while, the ability to identify a different prognosis among cases of myelodysplastic neoplasia with excess blasts, that is, having or not variants in acceptor or donor regions of the splicing region in the POLI gene in EB-MDS does not add any prognostic effect. However, despite being quite preliminary, the data on variants in cases of MDS with ringed sideroblasts (a disease characterized in molecular analysis as having a strong relationship with mutations in the splicing gene) seem to be consistent with a disease that presents a shallower depth of cytopenia with higher hemoglobin levels (consequently less need for transfusion) and larger platelets (above 50,000/microliter), suggesting that the presence of variants in acceptor or donor regions of the splicing region in the POLI gene in cases of MDS with ring sideroblasts have a possible protective effect.

LISTA DE TABELAS

Tabela 01. Grupo de RNA auxiliares participantes da maquinaria do spliceossomo, seus variados tamanhos moleculares com respectivas funções nesta cadeia bioquímica.	23
Tabela 02. Tecidos onde já se foi documentado alterações nos genes do splicing e suas respectivas alterações encontradas.	31
Tabela 03. Regiões sequenciadas dos genes REV3L, POLQ, POLH, REV1, POLI, POLL, POLM e POLN e o número de CDS de cada gene.....	41
Tabela 04. Descrição de número de variantes e indivíduos para os genes POLI, POLN, POLQ, REV1, POLL e POLM.	47
Tabela 5. Descrição dos casos: sexo, idade, cariótipo, classificação da doença, classificação prognóstica IPSS-R, identificação da variante, VAF, éxon, região canônica e escore SpliceAI.	48
Tabela 6. Valores de DeltaDetal CT para expressões gênicas dos genes POLI, POLL, POLK, POLN, POQ e REV1. Destacado em negrito que os dois casos com variantes de POLI tiveram maior expressão frente aos outros casos que não possuem variantes de POLI.	52

LISTA DE FIGURAS

Figura 01. Mecanismo de excisão de um íntron e junção de dois éxons.....	24
Figura 02. snRNPs isoladas e em complexo, participantes do mecanismo de splicing.	24
Figura 03. Representação esquemática do mecanismo de splicing. Excisão de um íntron para união de dois éxons através das snRNPs e proteínas complementares.	25
Figura 04. Esfregaço de amostra de medula óssea corado pelo método de coloração de Perls.	34
Figura 05. Variantes detectadas em cada um dos casos analisados..	49

SUMÁRIO

1. INTRODUÇÃO	15
1.1. NEOPLASIA MIELODISPLÁSICA	15
1.2. MECANISMO DE SÍNTESE DE TRANSLEÇÃO	16
1.3. MECANISMOS DE <i>SPLICING</i>	21
1.4. MUTAÇÕES DOS GENES DE SPLICING EM NEOPLASIAS MIELÓIDES	29
1.5. IMPACTO DAS VARIANTES EM SÍTIOS DE SPLICING.....	35
2. OBJETIVOS.....	39
2.1. OBJETIVO GERAL.....	39
2.2. OBJETIVOS ESPECÍFICOS	39
3. PACIENTES, MATERIAIS E MÉTODOS	40
3.1. ASPECTOS ÉTICOS DA PESQUISA	40
3.2. OBTENÇÃO DAS AMOSTRAS DE CÉLULAS DE MEDULA ÓSSEA.....	40
3.2.1. Extração de DNA Genômico.....	40
3.2.2. Customização do Pannel Genético	41
3.2.3. Preparo da Biblioteca	41
3.2.4. 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> e <i>POLN</i>	42
4. RESULTADOS	46
4.1. VARIANTES NO GENE <i>POLI</i>	50
4.2. VARIANTES NO GENE <i>POLN</i>	52
4.3. VARIANTES NO GENE <i>POLQ</i>	53
4.4. VARIANTES NO GENE <i>REV1</i>	53
4.5. VARIANTES NO GENE <i>POLL</i>	54
4.6. VARIANTES NO GENE <i>POLM</i>	55
5. DISCUSSÃO.....	56
6. CONCLUSÕES	63
7. REFERÊNCIAS BIBLIOGRÁFICAS	64
APÊNDICE: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	73

1. INTRODUÇÃO

1.1. NEOPLASIA MIELODISPLÁSICA

A Neoplasia mielodisplásica (SMD), atualmente assim denominada, é uma doença oncohematológica descrita inicialmente em meados de 1982 pelo grupo Franco-Americano-Britânico (FAB) como Síndrome Mielodisplásica. Naquela época, sua classificação era baseada brevemente em dados clínicos epidemiológicos e análise morfológica de células na medula óssea, seja em aspirado medular ou em biópsia. A suspeita inicial se dava através da notação de citopenias periféricas persistentes em quaisquer um dos setores hematológicos visualizados no hemograma. Uma vez admitido o paciente, se realizava um diagnóstico diferencial dentre outros distúrbios mieloides para se então finalizá-lo. Algumas destas condutas permanecem até os dias de hoje, mas foram adaptadas com o passar dos anos pela Organização Mundial de Saúde (OMS) conforme os estudos científicos vieram avançando (BENNET ET AL., 1976; BELLI ET AL., 2015; MAGALHÃES ET AL., 2018).

Trata-se de um fenômeno cancerígeno que ocorre em células proliferativas mieloides na medula óssea de indivíduos acima de 70 anos de idade, em sua maioria, caracterizado por distúrbios de diferenciação celular notado por displasias em pelo menos uma das três linhagens celulares hematológicas. Como consequência deste fenômeno, a produção de células maduras que deverão ser transportadas ao sangue periférico fica alterada e são notados no hemograma completo a presença de citopenias periféricas, também em pelo menos um dos setores hematológicos. A citopenia mais comum descrita e percebida é a do setor eritróide, marcada pela anemia. De acordo com a evolução natural da doença é descrito que cerca de 30 a 40% dos casos evoluam para um quadro de Leucemia Mielóide Aguda (LMA) (ARBER ET AL., 2016; VARDIMAN ET AL., 2008; BEJAR, 2014).

Além dos exames de hemograma, aspirado medular e biópsia de medula óssea, se realiza o exame de citogenética clássica ou molecular para se rastrear alterações no DNA importantes para o prognóstico e diagnóstico da doença. As alterações cromossômicas mais descritas na neoplasia mielóide são as deleções dos cromossomos 5 e 7, trissomia do cromossomo 8 e nulissomia do cromossomo Y (GREENBERG ET AL., 1997; OLNEY & LE BEAU, 2002).

Somente a partir de 1988, foi-se descrito e associada a patogênese desta doença a mutações em genes específicos, tais como *FMS* e *TP53*. Com o advento do sequenciamento de nova geração (*NGS – Next Generation Sequencing*) e a congregação de estudiosos em neoplasia mielodisplásica para publicação do IPSS-M (*International Prognostic Scoring System-*

Molecular), foram mais bem descritas as variantes mutagênicas definidas com seus respectivos prognósticos da doença. Atualmente, se tem descrito alguns genes importantes a serem analisados quando conveniente na evolução clínica do paciente, sendo eles *ASLX1*, *TET2*, *DNMT3A*, *NRAS*, *KRAS*, *EZH2*, *SF3B1*, *SRSF2*, *U2AF1* e *ZRSR2* (BEJAR ET AL, 2011; PAPAEMMANUIL ET AL, 2013; HAFERLACH ET AL, 2014; BERNARD ET A., 2022).

Sabe-se que o seu desenvolvimento está fortemente associado à exposição de diversos fatores de riscos ambientais tais como o benzeno, formol e seus derivados, pesticidas, herbicidas, além de radiações (Chamseddine et al., 2016; Sekeres, 2010). É importante destacar que o grupo que realizou este estudo já analisou exposições crônicas de agrotóxicos em trabalhadores rurais onde foram investigadas alterações cromossômicas, alterações de expressão em genes de reparo do DNA e distúrbios epigenéticos até o presente momento (CAVALCANTE ET AL., 2022; COSTA ET AL., 2020).

1.2. MECANISMO DE SÍNTESE DE TRANSLEÇÃO

À medida que ocorre a divisão celular de maneira fisiológica, o DNA é duplicado pela maquinaria de replicação localizada em seu núcleo. Estima-se que cerca de 10.000 lesões de fitas simples no DNA sejam geradas numa média diária. Essas lesões têm sido descritas como fatores causais provindos do ambiente aos quais o organismo é exposto, como por exemplo a luz ultravioleta, radiação ionizante, metais pesados, fármacos, formol e seus derivados e uma série extensa de agentes tóxicos nocivos ao DNA. Essas lesões precisam ser reparadas pela maquinaria de reparo do DNA, contudo, quando não reparadas, geram mutações pontuais ou até mesmo alterações cromossômicas que podem causar sérios danos celulares como alterações na replicação, transcrição gênica e morte celular. Esses efeitos persistentes a longo prazo resultam em doenças associadas à senescência, neurodegeneração e carcinogênese (CHATTERJEE & WALKER, 2017; LANGIE ET AL., 2015; GUILLIAM & YEELES, 2020).

Diversos mecanismos de monitoramento, reparo e tolerância aos danos no DNA têm sido estudados e descritos e sabe-se que eles existem para impedir consequências deletérias, uma vez que sejam capazes de detectar e sinalizar lesões em locais específicos das fitas de DNA e promover o reparo com as corretas sinalizações de enzimas e moléculas envolvidas nestes mecanismos. Sendo assim, as respostas aos danos no DNA podem resultar em reparo ou tolerância aos danos, sendo o principal objetivo a manutenção da estabilidade genômica (LINDAHL & WOOD, 1999; SHILKIN ET AL., 2020).

O foco de estudo é o mecanismo de reparo com atividade de translesão, sendo assim, as informações introdutórias a serem trazidas manterão esta linha de pesquisa de forma a melhor orientar o conhecimento e organização dos dados publicados na literatura a respeito.

Durante a fase S do ciclo celular, na qual o DNA está sendo replicado, podem haver lesões extensas que inviabilizam as DNA polimerases de alta fidelidade de ligarem, uma vez que os sítios ativos destas enzimas são incompatíveis com esses tipos de lesões, a replicação permanece imobilizada e a forquilha de replicação é reiniciada por um mecanismo de reparo conhecido como *DNA Damage Tolerance* (DDT), no qual enzimas de baixa fidelidade promovem a continuidade da replicação através de um molde alterado, permitindo que o reparo possa ocorrer em um momento posterior. Dessa forma, a célula consegue manter a divisão celular e evitar colapso da forquilha de replicação e instabilidade genômica (POWERS & WASHINGTON, 2018).

Essas DNA polimerases de baixa fidelidade possuem papel importante no mecanismo de reparo do DNA uma vez que tenham maior flexibilidade conformacional para de adaptar a lesões volumosas na forquilha de replicação. Contudo, elas promovem uma replicação com DNA danificado de uma maneira livre de erros, quando o pareamento de bases é biologicamente correto, e com DNA danificado propenso a erros, quando o pareamento de bases foi incorreto; ou elas permitem que a lesão seja transpassada sem que haja remoção de bases e continuam a replicar o DNA (CYBULLA & VINDIGNI, 2023).

Sendo assim, a ativação e regulação do mecanismo de translesão é didaticamente dividida em 3 fases: (1) o acesso ao DNA é restrito ao local da lesão sinalizada; (2) é realizada a seleção de uma das polimerases específicas para a lesão; (3) pode ou não haver troca de enzimas no momento da extensão da fita. Esse controle é mediado por modificações do Antígeno Nuclear de Proliferação Celular (PCNA) (DAIGAKU, DAVIES & ULRICH, 2010; FOX, LEE & MYUNG, 2011; KANAO & MASUTANI, 2017).

No momento de detecção da lesão volumosa, a polimerase de alta fidelidade é trocada por alguma das DNA polimerases de baixa fidelidade para que seja possível a replicação sobre a lesão. Em seguimento, pode haver troca por outra DNA polimerase de baixa fidelidade ou não. Essa etapa permite que a lesão escape do mecanismo de revisão exonucleásica da DNA polimerase replicativa. Em seguida, ocorre sinalização e troca de DNA polimerase para a de alta fidelidade e a replicação do DNA é retomada (MA, TANG & GUO, 2017).

Além disso, devido à diversidade química das lesões no DNA, cada uma das polimerases possui melhor atividade catalítica a depender do tipo de lesão. Essas lesões são descritas como

cognatas. Ou seja, se o dano for uma lesão cognata a polimerase escolhida, haverá uma precisão e eficiência da enzima relativamente alta, ao passo que o inverso apresenta menor precisão e eficiência genômica (POWERS & WASHINGTON, 2018).

Existem cerca de oito polimerases que atuam no mecanismo de translesão. Elas estão agrupadas em quatro famílias: família Y (*POLH*, *POLI*, *POLK* e *REVI*), família B (*POLZ*, cuja subunidade catalítica é *REV3L*), família A (*POLQ* e *POLN*) e família X (*POLL*). Cada uma delas tem especificidades de substratos catalíticos diferentes de acordo com as lesões no DNA (POWERS & WASHINGTON, 2018; GHOSAL & CHEN, 2013).

A *POLQ* é localizada no braço longo do cromossomo 3 e codifica uma proteína com 2.592 aminoácidos composta por um domínio polimerase e outro helicase, essa estrutura bioquímica lhe confere multifuncionalidade. O domínio polimerase permite atuação na síntese de translesão e o domínio helicase permite participação no mecanismo de reparo do DNA (SEKI, 2004; SEKI E WOOD, 2003).

POLQ tem atividade através de sítios lesionados por espécies reativas de oxigênio, como de timina glicol (Tg) e pode inserir nucleotídeos pareado a fita molde. Contudo, quando as lesões são causadas por radiação UV, *POLQ* não tem capacidade catalítica de realizar incorporação de bases, nesse momento, há troca por *POLI* para que haja a extensão (SEKI, 2004; SEKI E WOOD, 2003).

A literatura mostra alterações de expressão do gene *POLQ* em câncer de mama, pulmão, estômago, cólon e carcinoma de células escamosas bucal e estão fortemente associados à péssimo prognóstico com diminuição da sobrevida nesses pacientes. Lemée e cols. (2010) descreveu níveis de expressão mais elevados em pacientes com tumores de câncer de mama do tipo triplo-negativo, bem como altos níveis de expressão de *POLQ* levam ao entrave de forquilha de replicação, aumentam a quantidade de alterações cromossômicas e de mutações somáticas pontuais em câncer de mama.

No entanto, existem poucos relatos de alterações neste gene documentados na literatura científica, sejam de expressão ou de mutações em neoplasias mieloides.

A polimerase kappa (*POLK*) é uma enzima nuclear com atividade de translesão (TLS) que atua no sistema de reparo de DNA. Esta enzima é codificada pelo gene humano *DINB1*, que pertence à família Y de translesão de DNA polimerase e é única em três domínios biológicos: *Eukarya*, *Archaea* e *Bacteria*, enquanto os outros estão presentes apenas em *Eukarya* ou *Bacteria/Archaea*. Esta família de proteínas compartilha múltiplos sítios comuns

em sua sequência primária, o que a diferencia das outras famílias conhecidas de DNA polimerases: A, B, C e X. (OGI et al., 2002).

O núcleo catalítico POLK consiste nos subdomínios *finger*, *palm* e *thumb*, mas os *fingers* e *thumbs* são menores do que nas polimerases de alta fidelidade. É composto por 870 resíduos de aminoácidos, cuja massa molecular parece ser de 99 kDa e seu gene está localizado no cromossomo 5. Esta enzima não possui atividade exonuclease 3'-5' e apresenta atividade ótima de DNA polimerase a 37 °C e pH 6,5-7,5 na presença de cátions Mg²⁺ e de NaCl em níveis ≤ 50 mM. De acordo com o banco de dados Uniprot, esta enzima possui 8 isoformas produzidas por *splicing* alternativo. (GERLACH et al., 2001).

A expressão genética do *DINB1* é diferente entre os tecidos humanos, mas os estudos têm mostrado maior expressão nos testículos, baço, próstata e ovário quando comparados aos rins, cólon, cérebro, coração, fígado, pulmão, placenta, pâncreas e leucócitos do sangue periférico. (GERLACH ET AL., 1999; OGI ET AL., 1999).

Uma das características bioquímicas da POLK é catalizar adutos de benzopireno-N2-dGuanina que podem regular a fidelidade do *bypass* de lesões aromáticas policíclicas e outras estruturas que também podem alterar essa fidelidade, como os adutos AAF-guanina e etenodesoxi-adenosina, mas com baixa fidelidade relativa. Portanto, modificações nas estruturas da *POLK* podem eventualmente resultar em respostas oncogênicas, especialmente em células expostas a moléculas de benzopireno. Experimentos *in vitro* demonstraram que a *POLK* humana também é capaz de contornar lesões causadas por cisplatina. (YADAV ET AL., 2015; GERLACH ET AL., 2001).

Culturas de células de camundongo com expressão transitória do gene *DINB1* de camundongo levam a um aumento de mutações de substituição de base e de mudança de quadro, o que sugere um papel importante do *POLK* na replicação mutagênica do DNA não danificado. (ZHANG et al., 2000).

Um estudo relatou variantes oncogênicas em *POLK* que podem causar câncer de próstata, glioma e outros. Além disso, deleções e mutações no gene *POLK* no câncer de próstata ocorrem a uma taxa de 14,8% e 1,6%, respectivamente. (GRASSO et al., 2012).

Até o momento, existem estudos *in vitro* que abordam a bioquímica da *POLK*. Para investigar a função biológica da *POLK*, um estudo avaliou células-tronco embrionárias de camundongos *knockout POLK* que codifica a enzima e mostrou que as células mutantes são altamente sensíveis tanto à morte quanto à mutagênese induzidas por benzopireno, enquanto expostas à radiação UV e raios X são apenas ligeiramente afetadas. Assim, os resultados deste

estudo indicam que a *POLK* desempenha um papel importante na redução de mutações em lesões de DNA geradas por compostos de benzopireno. (OGI ET AL., 2002).

A literatura afirma que se fazem necessários mais estudos aprofundados sobre os mecanismos desta enzima *POLK*. Mesmo *POLK* tendo mostrado uma grande capacidade de contornar muitas lesões de DNA *in vitro*, seu substrato congênito ainda não foi determinado. (MASUMURA et al., 2017).

A *POLN* está localizada no braço curto do cromossomo 4 e codifica uma proteína de 900 aminoácidos com massa molecular de 100kDa. Estudos *in vitro* demonstram que esta enzima tem baixa eficiência e tem atividade propensa a erros, contudo, sua eficiência maior é em sítios Tg. Em câncer de mama, é descrita mutação neste gene em cerca de 50% dos casos (MOLDOVAN et al, 2010).

POLZ é uma enzima com subunidade catalítica *REV3L* sendo a primeira enzima com atividade de translesão identificada. Esta enzima não tem atividade exonucleásica. O gene que codifica *REV3L* fica localizado no braço curto do cromossomo 6, sendo uma região frágil e já relatada em diversos tumores sólidos e tipos de leucemias como uma região deletada (MAKAROVA & BURGERS, 2015).

Na medula óssea de camundongos com supressão do gene *REV3L* foi observado desenvolvimento de linfoma de células T e aparecimento de tumores mamários. Além disso, alterações pré-neoplásicas em tecidos glandulares adjacentes aos tumores mamários em camundongos com silenciamento do gene *REV3L*, isso pode sugerir que sua deleção ocasiona carcinogênese. Em humanos, a baixa fidelidade da *POLZ* tem como consequência a instabilidade cromossômica (LANGE ET AL., 2011; WITTSCHIEBEN ET AL., 2010).

O gene *POLH* é localizado no cromossomo 6 e codifica uma proteína de 713 resíduos de aminoácidos. Sua estrutura permite interação com *REV1* e tem um domínio de localização nuclear para interação com PCNA (KNOBEL & MARTI, 2011; UCHIYAMA, TERUNUMA & HANAOKA, 2015).

Esta polimerase é descrita na doença Xeroderma Pigmentoso Variante (XP-V), uma vez que haja expressão mutante do gene *POLH*. Representam cerca de 20% dos pacientes portadores desta doença, pois é sabido que vários outros genes envolvidos no mecanismo de reparo do DNA estão diretamente associados ao desenvolvimento desta doença, como *XPA*, *XPB*, *XPC* entre outros. Para os pacientes portadores do gene *POLH* mutado, é relatado maior índice de câncer de pele (JOHNSON et al., 1999; MASUTANI et al., 1999).

O gene *POLI* é localizado no cromossomo 18 e codifica um polipeptídeo de 715 resíduos de aminoácidos. A função de *POLI*, até o momento, não foi bem elucidada, contudo, muitos estudos têm demonstrado que a atividade da mesma está relacionada a proteção contra danos oxidativos, uma vez que variantes encontradas em sua subunidade catalítica provocaram redução da atividade enzimática e os indivíduos ficaram mais susceptíveis à atividade de agentes tóxicos. Além disso, mutações nesse gene têm sido relatadas em carcinoma de células escamosas e adenocarcinoma pulmonar (LI ET AL., 2023)

O gene *REVI* é localizado no cromossomo 2 e codifica uma proteína de 1251 resíduos de aminoácidos. Essa proteína possui dois domínios de ligação à ubiquitina (UMB1 e UMB2), um domínio C-terminal e um domínio BRCT. Esta proteína possui uma atividade de dCMP transferase na qual realiza inserção de citosina oposta a uma guanina. Embora essa enzima tenha esse papel, foi relatado que sua maior atividade está ligada na síntese de translesão reguladora e não catalítica. Essa função reguladora lhe confere a alocação de polimerases TLS nas regiões de dano no DNA e modular a mudança da polimerase no local da lesão (GUO ET AL., 2003; SASATANI, 2017; MALIK ET AL., 2024).

Além disso, mutações pontuais na região do gene REV1 foram relacionadas a risco aumentado de câncer cervical (HE ET AL., 2008).

No contexto de variantes em genes de reparo do DNA, o Xeroderma Pigmentoso (XP) é um modelo importante, especialmente em pacientes com mutações no gene *XPC*, é uma condição clínica fundamental para o estudo dos mecanismos de *splicing* gênico. O gene *XPC* codifica uma proteína essencial no sistema de reparo por excisão de nucleotídeos (NER), que corrige danos no DNA causados por radiação ultravioleta. Alterações neste gene, além de comprometerem o reparo do DNA, podem afetar diretamente o processamento do pré-mRNA, especialmente no que se refere à regulação do *splicing* alternativo. Estudos em indivíduos com XP permitem identificar como mutações específicas, incluindo aquelas em regiões não codificantes, afetam o reconhecimento de sítios de *splicing*, levando à exclusão ou inclusão aberrante de éxons. Dessa forma, o XP torna-se um modelo valioso para investigar como falhas no *splicing* contribuem para a patogênese da doença e para a instabilidade genômica associada (MARTENS, EMMERT & BOECKMANN, 2021).

1.3. MECANISMOS DE *SPLICING*

O mecanismo de *splicing* é um processo bioquímico bastante complexo e muito necessário para o bom funcionamento do metabolismo celular e da manutenção de seus

mecanismos moleculares. Análises realizadas em 2012 pelo Projeto ENCODE relataram que o maior desafio para se obter com acurácia os mecanismos que regulam e reconhecem os sítios da maquinaria de *splicing* é o fato de que os éxons compõem somente cerca de 3% do genoma humano (DUNHAM ET AL., 2012)

Além da participação na produção de mRNA maduros, a maquinaria do *splicing* tem forte atividade em outras etapas da expressão gênica, como exportação nuclear, tradução das moléculas de mRNA e controle de qualidade das moléculas de mRNA geradas através da via *nonsense-mediated decay* (NMD) (BRADLEY & ANCZUKÓW, 2023).

A maioria dos genes humanos são expressos como moléculas precursoras de RNA (pre-mRNA), também chamadas de transcrito primário, que serão posteriormente convertidos a RNA mensageiro (mRNA) através de uma etapa essencial da expressão gênica chamada *splicing*. Nesta etapa, os íntrons são removidos da molécula de pre-mRNA e os éxons são ligados através de enzimas e outras moléculas que fazem parte desta maquinaria (WILL & LÜHRMANN, 2011).

A via NDM é bastante conservada evolutivamente como um mecanismo de regulação pós-transcricional, uma de suas principais ações é a degradação de mRNA com códons de terminação prematuro (PTCs – *Premature Termination Codons*), geralmente derivados de genes mutados, ou mutações *nonsense* para redução de ruídos no transcriptoma (HAN ET AL., 2018).

Sabe-se que cerca de 12% das mutações localizadas em nucleotídeo único encontradas em humanos irão gerar mRNA com códons de terminação prematuro, sendo associadas a algumas doenças já descritas na literatura, como β -Talassemia e Distrofia Muscular de Duchenne. Por outro lado, mutações localizadas em genes da maquinaria NMD implicam em distúrbios neurológicos, doenças autoimunes e até mesmo câncer (HAN ET AL., 2018).

Os mecanismos que envolvem a produção do transcrito primário, como produto da etapa de transcrição gênica, necessitam de uma maquinaria específica para o processamento pós-transcricional onde é recrutada para realizar os ajustes finais na molécula de pré-mRNA. Esta maquinaria é formada por um complexo de proteínas e RNA auxiliares chamadas RNA nucleares pequenos (snRNP – *small nuclear ribonucleoprotein*) os quais interagem entre si em torno do transcrito primário formando o chamado spliceossomo (ROGALSKA, VIVORI & VALCÁRCEL, 2023).

Na tabela 01 a seguir, estão descritos alguns destes snRNP (U1, U2, U4, U5 e U6), cuja função é ajudar na alocação dos agrupamentos reativos dentro da molécula do transcrito

primário, de maneira que os íntrons possam ser retirados, os éxons agrupados e organizados de maneira apropriada para formação do mRNA, esses substratos, além de variados, possuem atuação *trans* em referência à molécula de mRNA (LEE & RIO, 2015).

É importante destacar que as snRNP destacadas acima possuem uma importância maior visto que está presente em 99% da regulação do splicing, enquanto que U11, U12, U4atac, e U6atac são responsáveis por uma menor parcela do *splicing* (BONNER & LEE, 2023).

Tabela 01. Grupo de RNA auxiliares participantes da maquinaria do spliceossomo, seus variados tamanhos moleculares com respectivas funções nesta cadeia bioquímica.

snRNA	Tamanho	Função
U1	165nt	Liga-se ao limite 5' éxon/íntron
U2	185nt	Liga-se ao sítio de ramificação no íntron
U4	145nt	Ajuda a montar o spliceossomo
U5	116nt	Liga-se ao limite 3' íntron/éxon
U6	106nt	Desloca U1 após o primeiro reaquecimento

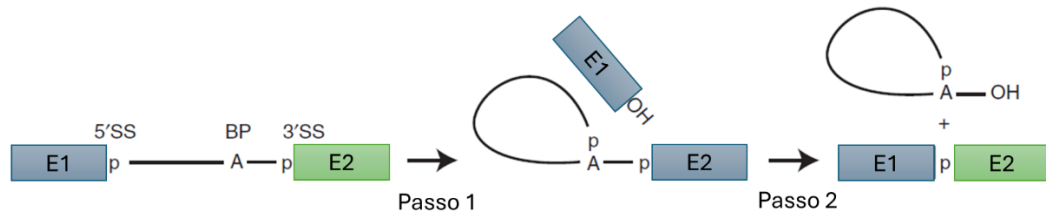
As ligações que ocorrem neste processo bioquímico são mediadas por interações de pareamento de bases em que os snRNAs se ligam aos sítios no pré-mRNA que representam limites entre íntrons e éxons. As moléculas de proteínas participantes deste processo são responsáveis pela manutenção da reação química proporcionando estabilidade nas ligações realizadas (LEE & RIO, 2015).

Considera-se que o mecanismo de *splicing* é dividido em duas etapas bioquímicas. A primeira etapa compreende uma reação de transesterificação para a clivagem da ligação fosfodiéster no limite intron/éxon na extremidade 5' (dinucleotídeo GU) do íntron. Esta reação ocorre entre a porção 2'-OH de um nucleotídeo de adenosina, usualmente encontrada cerca de 30nt da extremidade 3' (dinucleotídeo AG) do íntron, e o fosfato da ligação fosfodiéster de um resíduo de guanosina localizado na extremidade 5' do íntron (Figura 01). Esta reação quebra a cadeia nucleotídica e produz uma estrutura ramificada na extremidade 3' na qual a adenina possui grupamentos fosfato nos carbonos 2', 3' e 5'. Como resultado final desta etapa, o íntron forma uma estrutura curvada (LEE & RIO, 2015; BARALLE & BARALLE, 2005; DUNHAM ET AL., 2012).

É importante destacar que a molécula de pre-mRNA contribui para a definição molecular de um íntron que é limitada e conservada nos sítios 5'ss (splice site), 3'ss e sítio de ramificação (BS – *Branch site*). O sítio de ramificação geralmente é localizado cerca de 18-40 nucleotídeos antes do sítio 3'ss e em eucariotos superiores é seguido de um trato de polipirimidina (PPT – polypyrimidine tract) (WILL & LÜHRMANN, 2011).

Figura 01. Mecanismo de excisão de um íntron e junção de dois éxons. Cada retângulo colorido representa um éxon composto no transcrito primário, à esquerda, ou mRNA, à direita. As reações químicas ocorrem em etapas onde os elementos do *splicing* se ligam às regiões 5' e 3' para catalisar as reações.

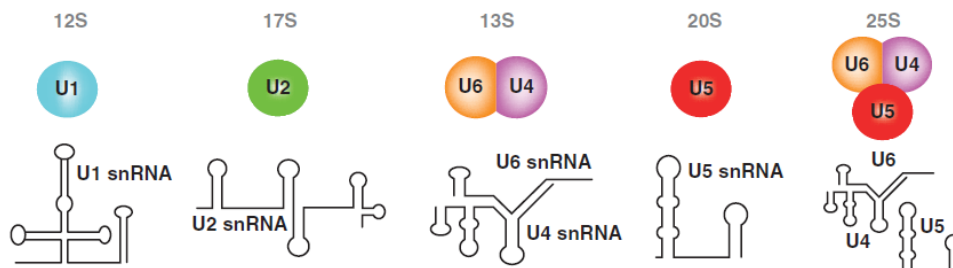
Fonte: Adaptado de Will & Lührmann, 2011.



Devido ao fato de que nenhuma das snRNPs possui sítio ativo já formado, a atividade catalítica de U2 depende da montagem de U1, U2, U5 e U4/U6, além da participação das proteínas não-snRNP, que por sua vez são remodeladas durante o processo de *splicing*. Sendo assim, cada snRNP consiste em um ou dois snRNA agrupado a pelo menos uma das 7 proteínas Sm (B/B', D3, D2, D1, E, F e G) e a variadas proteínas partícuo-específicas (Figura 02) (BRAUNSCHWEIG, 2013).

Figura 02. snRNPs isoladas e em complexo, participantes do mecanismo de *splicing*.

Fonte: Will & Lührmann, 2011.



Caso o íntron não exceda o tamanho de 200-250nts aproximadamente, a maquinaria do spliceossomo se monta através deste próprio íntron. No complexo do *splicing* inicial do íntron (*E complex*), a snRNP U1 é recrutada para a posição 5'ss e fatores não-snRNPs como SF1/mBBP e U2AF interagem com as regiões BS e PPT, respectivamente. Posteriormente, U2 é associado a região BS e formam o complexo A, também chamado de pre-spliceossomo. U4/U6.U5 (tri-snRNP) é pre-formado com a montagem de U5 a U4/U6 que, juntos, formam o complexo pre-catalítico B. É importante destacar que a ligação destas snRNPs entre o transcrito primário formam o centro catalítico do spliceossomo e as proteínas U5 snRNP Prp8 e RNA helicase Brr2 possuem um papel fundamental durante a ativação da catálise (BERTRAM et al., 2017).

A proteína Prp8 funciona como uma base de interação entre Brr2 e Snu114 que formam um meio bioquímico com RNA catalítico e espliceossomo ativado. A Brr2 inicia a ativação do espliceossomo ao desenrolar as hélices U4/U6 (BERTRAM ET AL., 2017).

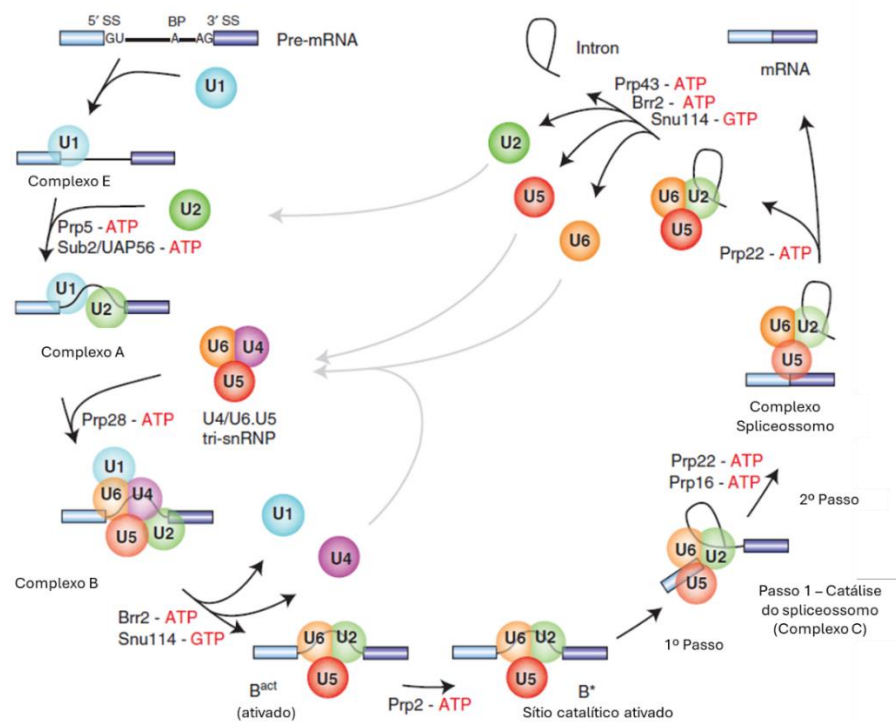
Este complexo é clivado e ativado pela RNA helicase DEAH-box Prp2, que por sua vez catalisa a primeira etapa do splicing. Seguindo a mesma cadeia de reações, esta reação anterior catalisa a produção do complexo C, que consequentemente realiza a catálise da segunda etapa (Figura 03) (WILL & LÜHRMANN, 2011).

Além disso, uma atuação *cis* dos elementos da maquinaria na molécula de pre-mRNA incluem intensificadores (enhancers – ESEs e ISEs) ou silenciadores (silencers – ESSs e ISSs), são caracteristicamente curtos e variados nas sequências de pares de bases que podem modular o splicing alternativo ou o constitutivo uma vez que se ligam às proteínas estimulantes ou silenciadoras da montagem do complexo de *splicing* (WILL & LÜHRMANN, 2011).

A segunda etapa envolve a clivagem da ligação fosfodiéster na extremidade 3' do íntron, que libera a estrutura em forma de laço do complexo. O *splicing* é completado pela junção da extremidade 3' de um éxon com a extremidade 5' do próximo éxon, pela formação de uma ligação fosfodiéster 5'-3' normal. Geralmente, a extremidade 3' de um éxon será rearranjada à extremidade 5' do éxon mais próximo, produzindo um transcrito que exibe todos os éxons na ordem que foram transcritos. Após isso, o espliceossomo então se dissocia e os snRNPs são liberados para as próximas reações bioquímicas (BARALLE & BARALLE, 2005).

Figura 03. Representação esquemática do mecanismo de splicing. Excisão de um íntron para união de dois éxons através das snRNPs e proteínas complementares.

Fonte: Adaptado de Will & Lührmann, 2011.



Sendo assim, pode-se notar que a clivagem do RNA e reações de ligação necessárias para remoção dos íntrons em mRNAs ocorrem em uma proteína RNP (ribonucleoproteína) que faz parte do spliceossomo e que o centro catalítico do spliceossomo é também composto de RNA, desta forma pode-se definir que ele é uma ribozima, como o ribossomo (WAHL, WILL & LÜHRMANN, 2009; WILL & LÜHRMANN, 2011).

Adicionalmente a via de *splicing* constitutiva, se tem também descrito o *splicing* alternativo, que por sua vez é bastante prevalente em eucariotos, onde mais de 95% dos genes humanos têm a maquinaria necessária para executar esta via de expressão, com isso, possibilita a sua elevada variedade de formas e funções proteicas devido ao aumento do número de proteínas expressas a partir de um único gene (NILSEN & GRAVELEY, 2010; LEE & RIO, 2015).

O mecanismo de *splicing* alternativo é definido por um processo pelo qual diferentes complexos de *splicing* são selecionados em uma transcrição pré-mRNA para produzir isoformas diferentes de mRNA e proteína, consecutivamente. Além disso, conhecer profundamente os seus mecanismos é de suma importância pois estão envolvidos nos principais mecanismos celulares como crescimento celular, morte celular, pluripotência, diferenciação celular, desenvolvimento celular, ritmo circadiano, respostas a interações ambientais, exposição a patógenos e doenças (DUNHAM ET AL., 2012).

As vias alternativas de splicing ocorrem nas fases iniciais de montagem do spliciossomo. Nesse contexto, se sabe que a maioria dos pre-mRNAs de mamíferos contém múltiplos íntrons que variam em tamanho de centenas a milhares de nucleotídeos, enquanto os éxons têm aproximadamente 120nt em média. Sendo assim, quando o comprimento do íntron excede 200-250nt (evento comum em eucariotos superiores; em humanos, estão entre 50 e 300 nucleotídeos), o spliciossomo se forma primeiro na região de um éxon, um processo chamado de definição de éxon (KELEMEN ET AL., 2013).

Durante o processo de definição do éxon, U1 se liga ao 5'ss do éxon e promove conexão direta de *U2AF* com o trato de polipirimidina/3'ss (PPT) adicionalmente. Consequentemente, U2 é recrutado para região BS do éxon. Os intensificadores (ESEs) de splicing recrutam proteínas que estabelecem uma rede de interações proteína-proteína através do éxon que estabilizam o complexo definido pelo éxon (BRAUNSCHWEIG, 2013).

Como as etapas bioquímicas do splicing ocorrem em um íntron, subsequentemente à definição do Exon, o 3'ss deve ser pareado no íntron adjacente com um 5'ss a montante. Acredita-se que as interações entre éxon são inicialmente interrompidas e o complexo ligado ao éxon é então convertido a um complexo A ligado ao íntron, onde uma ponte molecular se forma entre U2 e U1 ligada a um 5'ss. Esta etapa é importante para entender qual éxon 5' e 3' será finalmente unido (WILL & LÜHRMANN, 2011).

A maquinaria do splicing possui alta fidelidade para evitar a introdução de erros durante a expressão gênica. Estima-se que a taxa de erros de splicing estejam entre 1 em 100 até 1 em 100.000 erros e estes contribuem para o desenvolvimento de câncer e outras doenças (MAYERLE & GUTHRIE, 2017).

Sendo assim, é importante destacar que os mecanismos de *splicing* são estritamente necessários para uma elevada quantidade de mecanismos celulares vitais e necessários para a evolução humana, pois está diretamente ligado ao funcionamento correto do genoma e proteoma (KELEMEN ET AL., 2013).

Foram relatados alguns distúrbios hematológicos com forte relação a alterações no mecanismo de *splicing*, um exemplo é a Anemia de Fanconi (AF), uma doença genética causada por variantes patogênicas em qualquer um dos 33 genes designados *FANCA-FANCW*. No estudo realizado por Harada e Cols. (2024) analisaram que mutações no gene *CCAR1* foram ocasionadas por falha do mecanismo de *splicing* alternativo em que a expressão deste gene produziu isoformas mutantes que contribuíram para o desenvolvimento da doença.

A Ataxia-telangiectasia (AT) é uma doença neurodegenerativa de caráter autossômico recessivo caracterizada por imunodeficiência, sensibilidade a radiação ionizante, diabetes e aumento do risco de malignidades, principalmente de origem linfóide. Ela é causada por mutações bialélicas no gene *ATM* localizado no cromossomo 11q22.3. Mutações monoalélicas deste gene são reponsáveis pela predisposição ao câncer gástrico, mama, linfoma e outros. Este gene contém 64 éxons que codificam uma proteína de 3050 resíduos de aminoácidos. É sabido que 18% das mutações no gene *ATM* são provocadas por substituições provocadas pelos mecanismos de *splicing*. Essas variantes são decorrentes de adições nos limites éxon-íntro ou ativação de éxons crípticos como resultado de alterações nas regiões canônicas aceptoras e doadoras (ATES ET AL., 2022).

A Hemofilia B é uma doença caracterizada pela deficiência do fator IX da coagulação. Um dos principais sintomas decorrentes desta deficiência é o sangramento prolongado. Este fator é dependente de vitamina K e atua intrinsecamente na cascata de coagulação. Já é relatado em literatura que mutações no gene *F9* tem forte influência dos mecanismos de *splicing* alternativo, uma vez que gerem isoformas não funcionais e causem o desenvolvimento deste distúrbio (GOODEVE, 2015).

A regulação genética do *splicing* canônico é um processo altamente controlado que garante a remoção precisa dos íntrons e a união correta dos éxons durante a maturação do transcrito primário. Essa regulação depende da interação coordenada entre sinais *cis* (sequências no próprio RNA) e fatores *trans* (proteínas reguladoras). Os sinais *cis* incluem os sítios de *splicing* 5' (GU), 3' (AG), o ponto de ramificação e a região polipirimidínica. Esses elementos são reconhecidos pelo spliceossomo, mas sua eficácia pode ser modulada por sequências auxiliares chamadas *enhancers* e *silencers* de *splicing*, localizados nos éxons ou íntrons (ROGALSKA, VIVORI & VALCÁRCEL, 2023).

Os fatores *trans* são proteínas que se ligam a essas sequências regulatórias para ativar ou reprimir o uso de determinados sítios de *splicing*. Entre os principais grupos dessas proteínas estão os fatores da família SR (*Serine/Arginine-rich*), que geralmente promovem o *splicing*, e os hnRNPs (proteínas nucleares heterogêneas ribonucleicas), que costumam agir como repressores. A interação dessas proteínas com o pré-mRNA influencia a montagem do spliceossomo e o reconhecimento dos sítios canônicos, ajustando a eficiência e a especificidade do processo (ROGALSKA, VIVORI & VALCÁRCEL, 2023).

Além dos fatores proteicos, o *splicing* é influenciado por modificações epigenéticas e pelo estado da cromatina. Marcas como a metilação de histonas e a estrutura do DNA podem

afetar a velocidade da transcrição pela RNA polimerase II, impactando indiretamente a exposição e o reconhecimento dos sítios de *splicing* (GURULE ET AL., 2023).

Outro nível importante de regulação ocorre por meio de sinais externos, como hormônios, estresse celular e sinalização extracelular. Esses estímulos podem alterar a expressão ou atividade dos fatores reguladores de *splicing*, levando a mudanças no padrão de combinação de éxons. Embora o *splicing* alternativo seja mais associado a essa flexibilidade, o *splicing* canônico também pode ser ajustado para manter a fidelidade do processamento do RNA em diferentes condições fisiológicas ou ambientais (GURULE ET AL., 2023).

Por fim, a regulação do *splicing* canônico é fundamental para a homeostase celular e o desenvolvimento adequado dos organismos. Alterações nesse controle podem levar à inclusão ou exclusão incorreta de éxons, resultando em proteínas defeituosas ou perda da função gênica. Assim, a célula mantém mecanismos de vigilância, como o *nonsense-mediated decay* (NMD), para eliminar transcritos aberrantes. A compreensão detalhada desses mecanismos tem permitido avanços em terapias gênicas, que visam corrigir erros de *splicing* em doenças genéticas por meio de oligonucleotídeos e outros moduladores (GURULE ET AL., 2023).

1.4. MUTAÇÕES DOS GENES DE SPLICING EM NEOPLASIAS MIELÓIDES

Mutações nos genes de *splicing* são descritas como eventos precoces que ocorrem em clones fundadores e contribuem para o fenótipo displásico das neoplasias hematológicas, neurodegeneração, doenças autoimunes e câncer de uma maneira geral. Em neoplasias mielodisplásicas, compõem cerca de 60% do espectro de alterações genéticas importantes (BIERNACKI ET AL., 2023; BONNER & LEE, 2023; BEJAR ET AL., 2012; PAPAEMMANUIL ET AL., 2013).

Essas mutações no fator de *splicing* ocorrem como mutações pontuais heterozigóticas, mutuamente exclusivas e em resíduos específicos. A presença de um dos fatores de *splicing* selvagem é vital para a sobrevivência das células que expressam um fator de *splicing* mutante. Além disso, estas células alteradas ficam sensibilizadas a fenômenos adicionais na maquinaria de *splicing* por inibidores de *splicing* de pequenas moléculas (LIU ET AL., 2024)

Erros no funcionamento da maquinaria do *splicing* podem ocasionar progressão tumorigênica. Durante as últimas décadas, têm-se associado algumas mutações em genes da

maquinaria do *splicing* ao desenvolvimento de diversos cânceres sólidos como bexiga, cérebro, entre outros, além de neoplasias mieloides como leucemia mieloide aguda, neoplasia mielodisplásica, leucemia mieloide crônica e neoplasia mieloproliferativa (ZHANG ET AL., 2021).

Para alguns tipos de doenças na linhagem mieloide, como Neoplasia Mielodisplásica (SMD) e Leucemia Linfocítica Crônica (LLC), mutações na terminação 3' do sítio de *splicing* dos fatores de reconhecimento, tais como *U2AF* e *SF3b* estão fortemente relacionados ao desenvolvimento de câncer nos pacientes portadores destas alterações genéticas. (COOPER, WAN & DREYFUSS, 2009; PADGETT, 2012; OGAWA, 2012; AUBOEUF ET AL., 2012).

A tabela 02 apresenta um apanhado de exemplos de neoplasias e suas alterações relacionadas a alguns genes envolvidos no mecanismo de *splicing*.

Tabela 02. Tecidos onde já foram documentadas alterações nos genes do *splicing* e suas respectivas alterações encontradas.

Fonte: Adaptado de BONNER & LEE, 2023

Órgão	Fator do <i>Splicing</i>	Tipos de Alterações
Cérebro	<i>SRSF1, SRSF3, HNRNPA1, HNRNPA2, HNRNPHK</i>	Aumento da expressão gênica
Mamas	<i>SRSF1, SRSF3, SRSF4, SRSF5, SRSF6, TRAB2B, HNRNPA1, HNRNPI, RBM5, RBFOX2, HNRNPK, SF3B1</i>	Aumento da expressão gênica Diminuição da expressão gênica Mutação somática
Bexiga	<i>SRSF1, SRSF3</i>	Aumento da expressão gênica
Cólon	<i>SRSF1, SRSF3, SRSF6, SRSF10, TRAB2B, HNRNPK, RBFOX2</i>	Aumento da expressão gênica Diminuição da expressão gênica
Intestino	<i>SRSF1</i>	Aumento da expressão gênica
Rins	<i>SRSF1, SRSF3</i>	Aumento da expressão gênica
Fígado	<i>SRSF3</i>	Aumento da expressão gênica
Pulmão	<i>SRSF1, SRSF3, SRSF5, SRSF6, TRAB2B, RBM5, QKI, RBM10, U2AF1</i>	Aumento da expressão gênica Diminuição da expressão gênica Mutação somática
Pele	<i>SRSF3, HNRNPK, SF3B1, SRSF2</i>	Aumento da expressão gênica Diminuição da expressão gênica Mutação somática
Tireoide	<i>SRSF1, SRSF3, RBM10</i>	Aumento da expressão gênica Mutação somática
Leucemias mieloides	<i>SF3B1, SRSF2, U2AF1, ZRSR2</i>	Mutação somática
Leucemia linfocítica crônica	<i>SF3B1</i>	Mutação somática

Além disso, vários estudos relatam uma série de mutações somáticas recorrentes na região 3' dos sítios de reconhecimento das proteínas e componentes de repetição serina-arginina (SR) dos fatores de *splicing*, nomeados *U2AF1*, *SRSF2*, *SF3B1* e *ZRSR2* (PADGETT, 2012; OGAWA, 2012; MAKISHIMA ET AL., 2012; HAHN & SCOTT, 2012; PRZYCHODZEN ET AL., 2013; BROOKS ET AL., 2014).

Conforme demonstrado na tabela 02, nas leucemias mielóides, a classe de genes mais comumente mutados são os que regulam a maquinaria do *splicing*, estando presentes em aproximadamente metade dos casos de doença relatados. Dentre os genes estudados na maquinaria do *splicing*, mutações nos genes *SF3B1*, *U2AF1*, *ZRSR2* e *SRSF2* são as mais prevalentes. As mutações no spliceossomo provêm ganho de função ou geram mutações neomórficas que induzem alterações significativas na expressão gênica do pre-mRNA em escala genômica. Por exemplo, a inibição de *SF3B1*, *U2AF1*, *ZRSR2* e *SRSF2* devido a mutações nestes genes provocam redução dos mecanismos inflamatórios pela alteração do *splicing* dos genes da via de sinalização TLR; em contraste, quando ocorrem mutações neomórficas elas provocam o aumento de fatores pro-inflamatórios (GURULE ET AL., 2023).

Além disso, Bonner e Lee (2023) afirmam que mutações nos genes *SF3B1*, *SRSF2* e *U2AF1* ocorrem exclusivamente como mutações *missense* de caráter heterozigótico em

regiões *hotspot*, enquanto mutações no gene *ZRSR2* ocorrem de maneira aleatória no gene e predispõe a perda de função gênica (*Loss-of-function - LoF*).

Frequências elevadas de mutações no gene *U2AF1* (*U2 small nuclear RNA auxiliary factor 1*) em pacientes portadores de neoplasia mielodisplásica ou leucemia mieloide aguda corroboram com distúrbios hematopoiéticos ineficazes e progressão do câncer (ZHAO ET AL., 2022)

U2AF1 em conjunto com *U2AF2* formam o complexo *U2AF* que reconhece o sítio de *splicing* 3' (3'SS) dos íntrons e recruta a snRNP *U2*. O domínio de *U2AF1* inclui um domínio homólogo (UHM) a *U2AF* que media a heterodimerização com *U2AF2*, e duas proteínas dedo de zinco bastante conservadas *CCCH* (*ZnF1* e *ZnF2*) que cooperam na ligação ao sítio 3'SS nos limites íntro-éxon. Além disso, mutações heterozigóticas pontuais que afetam os resíduos *S34* e *Q157* com os domínios de *ZnF1* e *ZnF2*, respectivamente, resultam em arranjos de *splicing* incorretos, dependentes da sequência do transcrito primário, que resultam em alterações importantes nos genes críticos para a hematopoiese (BIANCON ET AL., 2022).

Sendo assim, é descrito na literatura que os *hotspots S34* e *Q157* mutados são críticos uma vez que perturbam as interações químicas de *U2AF1* e o pre-mRNA. Mutações em *S34* são mais fortemente descritas em neoplasias hematológicas, mas ambas estão presentes nesta classe de doenças. Além disso, essas duas mutações contribuem bastante nos processos biológicos tumorigênicos, tais como ciclo celular, reparo do DNA, transcrição e processos de tradução e processamento de RNA (BIANCON ET AL., 2022).

O gene *SF3B1* codifica a ribonucleoproteína *U2*, uma das principais envolvidas no processo de *splicing* responsável pelo reconhecimento e seleção do ponto de ramificação (*Branchpoint*) do pre-mRNA de maneira independente, bem como o reconhecimento da maioria das regiões 3'ss, promovendo estabilidade nesta ligação, além disso, é sabido que este é o gene mais mutado nos pacientes portadores de neoplasia mielodisplásica (20-28% de todos os pacientes), sendo normalmente alterações heterozigóticas (CIESLA ET AL., 2023).

Estudos sugeriram através de análises estruturais que mutações no gene *SF3B1* podem alterar a carga e a forma de acordo com os resíduos de aminoácidos, que por sua vez interrompem a interação com o transcrito primário por cerca de 10-30 nucleotídeos, dessa forma, dificultando o processo de *splicing* inicializar (Bonner & Lee, 2023). Por exemplo, mutações no resíduo de arginina 625 são fortemente associadas a melanomas uveal, enquanto mutações na posição *K700E* são bastante presentes em leucemias mieloides e cânceres epiteliais (LIU et al., 2020).

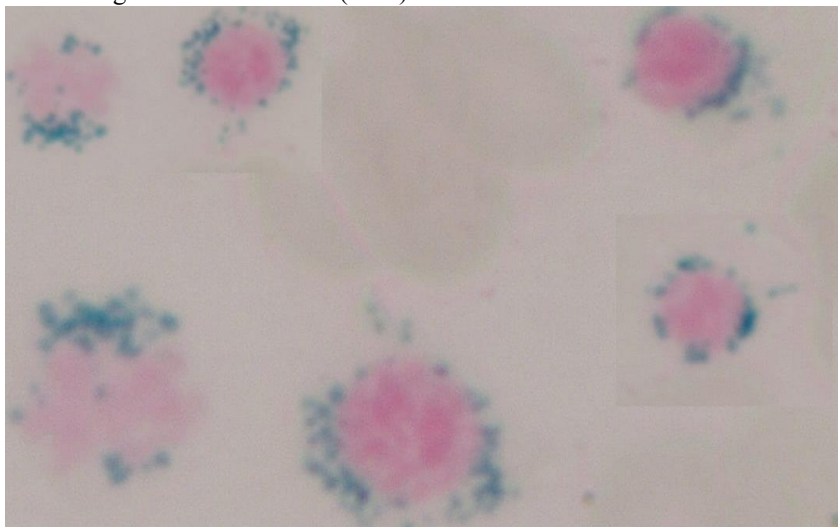
Até hoje, milhares de transcritos aberrantes têm sido identificados em diversos tipos de cânceres, demonstrando um papel importante deste gene na maquinaria do *splicing*. Contudo, poucos alvos têm sido fortemente associados a fenótipos de doenças incluindo *ABCB7*, *TMEM14c*, *ALAS2* na biossíntese do heme, *BRD9* na iniciação e manutenção de tumores sólidos e *PPP2R5A* na regulação do gene *MYC* (BONNER & LEE, 2023).

Por sua vez, já se sabe que existe uma grande correlação entre mutações no gene *SF3B1* e a presença de sideroblastos em anel nos exames morfológicos, principalmente mielograma e biópsia de medula óssea. Esse achado morfológico se dá pelo acúmulo de ferro mitocondrial nos precursores eritróides que, durante a coloração de *Perls* são evidenciados ao redor do núcleo (Figura 03). Um estudo de Malcovati e Cols. (2011) evidenciou um valor preditivo positivo de 97,7% quando analisada a mutação do *SF3B1* e o fenótipo morfológico descrito acima (GAJZER & CECILIA, 2023).

Mutações no gene *SF3B1* implicam em alteração do metabolismo de formação do heme que contribui para o acúmulo de ferro mitocondrial descrito acima, uma vez que os genes *TMEM14c*, *PPOX* e *ABCB7* sofrem alteração de *splicing* (CLOUGH. ET AL., 2022).

De acordo com as classificações da SMD pela Organização Mundial de Saúde, em 2008, para se ter o diagnóstico de Anemia Refratária com Sideroblastos em Anel (ARSA), se era necessária a presença de 15% ou mais de sideroblastos em anel no aspirado de medula óssea (Figura 04) e 5% ou menos de blastos. Em 2016, a Organização aprimorou esta classificação com dados genéticos onde se inferiu que na presença de mutação no gene *SF3B1*, 5% de sideroblastos em anel é suficiente para se realizar o diagnóstico de ARSA, além disso a nomenclatura foi modificada para SMD-SA. É importante destacar que nos casos em que não se possa realizar detecção da mutação, é possível realizar exame morfológico, contudo, há perda de aproximadamente 17% do diagnóstico correto da doença. Já em 2022, a mesma doença passa a ser chamada de Neoplasia Mielóide com mutação no *SF3B1* (VARDIMAN ET AL., 2008; ARBER ET AL., 2016).

Figura 04. Esfregaço de amostra de medula óssea corado pelo método de coloração de Perls.
Fonte: Laboratório Citogenômica do Câncer (LCC).



Mutações no gene *SF3B1* (*Splicing Factor 3b subunit 1*) são bastante associadas a um bom prognóstico e ótima resposta as terapias padrões, enquanto mutações nos outros genes da maquinaria do *splicing* geralmente apresentam um prognóstico mais desfavorável, como os genes *U2AF1* e *SRSF2* (Biernacki et al., 2023). Neste mesmo sentido, mutações *hotspots* no *SF3B1* e a ausência de mutações *frameshift* e *nonsense* em pacientes portadores de SMD geralmente são mutações com ganho de função. Por outro lado, estudos em modelos de camundongos *knockout* sugeriram que mutações heterozigóticas apresentam desenvolvimento de sideroblastos em anel, já outros estudos não obtiveram este mesmo resultado. Logo, pode-se inferir que é necessário determinar se a mutação no gene *SF3B1* acarretará ganho ou perda de função (MATSUNAWA ET AL., 2014; WANG ET AL., 2014; GAJZER & CECILIA, 2023).

Além disso, é importante destacar que mutações no gene *TP53* e/ou deleções no braço longo do cromossomo 5 (del.5q) mesmo com a presença de mutações no gene *SF3B1* no mesmo clone predizem prognóstico bastante desfavorável da doença (GAJZER & CECILIA, 2023).

O gene *SRSF2* (*Serine and Arginine-rich splicing factor 2*) é um membro da família de proteínas ricas em serina-arginina que estão fisiologicamente relacionadas ao funcionamento da maquinaria do *splicing* constitutivo e alternativo ao se ligar a sequências intensificadoras de *splicing* exônico dentro do pre-mRNA por meio de seu domínio de reconhecimento de RNA (*RRM – recognition motif domain*). Tem sido descrita uma mutação bastante importante neste gene envolvendo resíduos de prolina 95 (*SRSF2P95*) em pacientes portadores de diversos distúrbios hematológicos tais como neoplasia mielodisplásica, leucemia mielomonocítica crônica e leucemia mieloide aguda (YANG ET AL., 2023).

Mutações no gene *SRSF2* compõem cerca de 20% das alterações descritas na maquinaria do *splicing* em pacientes portadores de neoplasia mielodisplásica. Essas mutações prejudicam o ciclo de diferenciação hematopoiética *in vivo* mas não relacionada a perda de função do gene. Essas alterações dificultam a atividade normal de ligação do RNA específico da sequência do *SRSF2*, alterando assim o reconhecimento de intensificadores de *splicing* (*enhancer motifs*) exônicos específicos para induzir *splicing* incorreto recorrente de reguladores hematopoiéticos essenciais. Um exemplo interessante é o *splicing* do gene *EZH2* dependente da mutação *SRSF2 P95H* que desencadeia decaimento mediado por uma mutação nonsense que resulta em diferenciação hematopoiética prejudicada pela redução dos níveis de *EZH2* (HOSONO, 2019; KIM ET AL., 2015).

Além disso, este gene mutado está associado com mutações *JAK2V617F* em pacientes portadores de mielofibrose, apresentando um prognóstico ruim para estes pacientes (YANG ET AL., 2023).

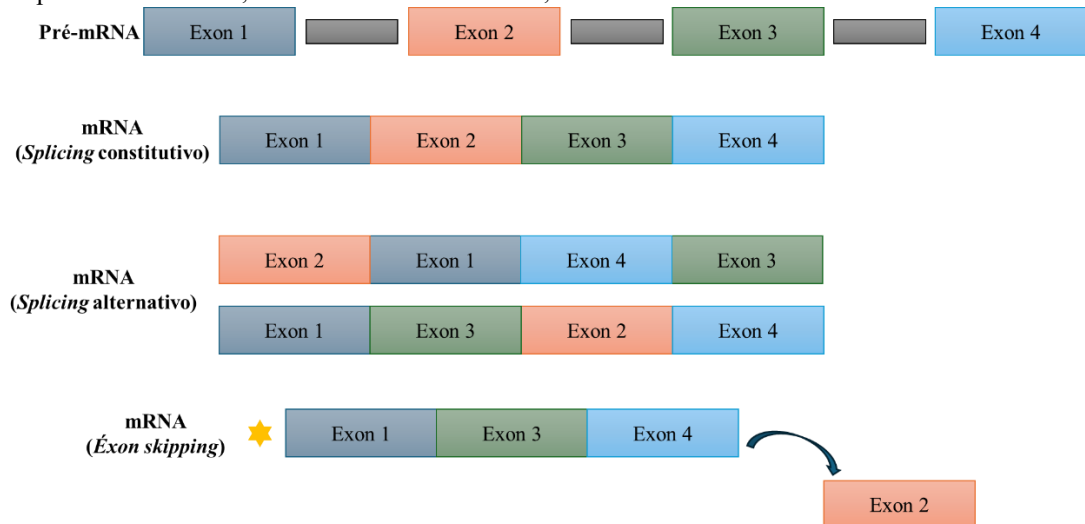
Um estudo realizado por Kim e cols. (2015) revelou que a mutação *SRSF2 P95H* é distinta da mutação *loss-of-function SRSF2*. Além disso, possui papel chave no entendimento da fisiopatologia da neoplasia mielodisplásica uma vez que aumenta produção de *HSPCs* (*Hematopoietic Stem and Progenitor Cells*) em camundongos com mutação *Srsf2P95H*, altera a cinética do ciclo celular e aumento de apoptose com o desenvolvimento de citopenias periféricas e displasias morfológicas, sendo assim, este gene demonstra ser importante para a hematopoiese.

1.5. IMPACTO DAS VARIANTES EM SÍTIOS DE SPLICING

Distúrbios causados por variantes em sítios de *splicing* ocorrem quando alterações genéticas afetam o processo de remoção de íntrons e união de éxons no RNA mensageiro (mRNA). Essas variantes podem ocorrer nos sítios canônicos de *splicing* ou em regiões regulatórias próximas. Quando elas ocorrem, o maquinário de *splicing* pode falhar em reconhecer corretamente os éxons e íntrons, resultando em eventos anômalos como o *exon skipping* (salto de éxon), inserções, deleções ou retenção de íntrons (Figura 05). Tais alterações muitas vezes comprometem a leitura do mRNA e levam à produção de proteínas defeituosas ou não funcionais, contribuindo para uma variedade de doenças genéticas (CARTEGINI, CHEW & KRAINER, 2002).

Figura 05. Mecanismos de splicing constitutivo e alternativo a partir do transcrito primário. Cada retângulo representa visualmente um éxon, as cores diferenciam cada éxon dentro de uma mesma fita de mRNA, já os retângulos coloridos em cinza representam íntrons. A estrela amarela demonstra uma das alterações que ocorre no mecanismo de splicing, o chamado éxon skipping. Neste fenômeno ocorre desacoplamento de éxons do mRNA final após a finalização do mecanismo de splicing, no exemplo da imagem o éxon dois é desacoplado da fita de mRNA em referência ao transcrito primário (Pré-mRNA).

Fonte: Adaptado de TAKEDA, CLEMENS & HOFFMAN, 2021.



As regiões canônicas do *splicing* são sequências altamente conservadas localizadas nas extremidades dos íntrons: o sítio doador (no início do íntron, geralmente com a sequência GT) e o sítio aceptor (no final do íntron, geralmente AG). Essas regiões são fundamentais para o correto reconhecimento e remoção dos íntrons durante o processamento do transcrito primário. Variantes que afetam essas regiões podem impedir o reconhecimento adequado desses sítios pelo maquinário de *splicing*, levando a eventos aberrantes como o éxon *skipping*, a retenção de íntrons ou o uso de sítios de *splicing* alternativos, como *splicing* crípticos. Como consequência, o mRNA resultante pode estar truncado, conter sequências extras ou perder informações essenciais, comprometendo a estrutura e a função da proteína codificada (XU ET AL, 2022).

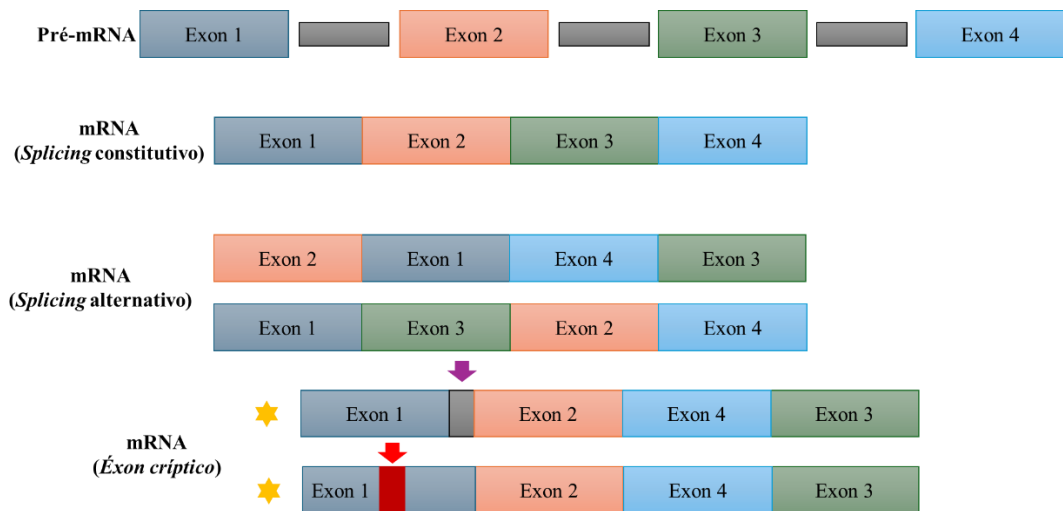
O éxon *skipping* é o fenômeno mais comum descrito, nele um éxon essencial é ignorado durante o processamento do mRNA. Isso pode resultar em proteínas truncadas, instáveis ou com função comprometida. Por exemplo, na distrofia muscular de Duchenne, mutações que afetam os sítios de *splicing* do gene da distrofina levam frequentemente à exclusão de éxons importantes, interrompendo a produção da proteína. Outro exemplo é a fibrose cística, em que mutações no gene *CFTR* podem afetar o *splicing* normal e resultar em perda de função da proteína reguladora do canal de cloro (HAKIM ET AL., 2017).

Além disso, mutações podem levar à ativação de éxons crípticos, que são regiões do pré-mRNA normalmente não incluídas no transcrito maduro, mas que passam a ser reconhecidas como éxons devido a alterações nos sinais de *splicing*. A inclusão de éxons crípticos pode introduzir códon de parada prematuros, alterar domínios funcionais da proteína ou

desestabilizar o RNA por mecanismos como o NMD (decaimento mediado por nonsense) (Figura 06). Esses eventos têm sido associados a diversas doenças genéticas e cânceres, nos quais a expressão gênica é amplamente desregulada (QIU ET AL., 2023).

Figura 06. Mecanismos de splicing constitutivo e alternativo a partir do transcrito primário. Cada retângulo representa visualmente um éxon, as cores diferenciam cada éxon dentro de uma mesma fita de mRNA, já os retângulos coloridos em cinza representam íntrons. As estrelas amarelas demonstram alterações que ocorrem no mecanismo de splicing denominadas éxon críptico. Neste fenômeno, estruturas de íntrons chamadas pseudoéxons são incluídas entre éxons (seta roxa) ou dentro de um éxon (seta vermelha).

Fonte: Adaptado de ULE & BLENCOWE, 2019.



O éxon críptico é uma sequência do pré-mRNA que, em condições normais, não é reconhecida pelo maquinário de *splicing* como um éxon funcional. No entanto, mutações genéticas, tanto na própria sequência críptica quanto em elementos regulatórios, podem alterar o mecanismo do *splicing* e levar à ativação dessas regiões, fazendo com que sejam incluídas no mRNA maduro. A ativação de um éxon críptico geralmente ocorre quando uma mutação fortalece um sítio de *splicing* ou quando um sítio canônico adjacente é enfraquecido, levando o *splicing* a “procurar” alternativas nas proximidades. Essas alterações são particularmente críticas em genes essenciais, pois podem interromper a leitura ou adicionar sequências aberrantes à proteína codificada (ULE & BLENCOWE, 2019).

A inclusão de éxons crípticos no mRNA pode ter diversas consequências patogênicas. O éxon críptico introduz domínios peptídicos não funcionais ou interferentes, alterando a estrutura e a função da proteína final. Diversas doenças genéticas e cânceres têm sido associados à ativação de éxons crípticos. Um exemplo conhecido ocorre no gene *SMN2*, relacionado à atrofia muscular espinhal (AME), em que uma única mutação silenciosa influencia o *splicing* e a ativação de um éxon críptico, resultando em níveis reduzidos da proteína funcional. Isso

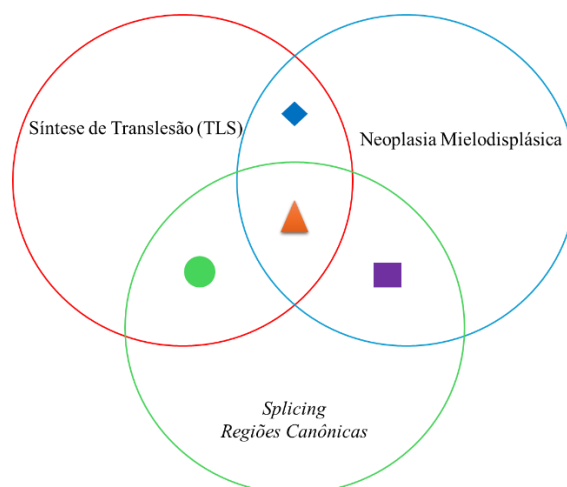
demonstra como alterações aparentemente pequenas podem ter um impacto profundo na regulação do *splicing* e na fisiologia celular (ULE & BLENCOWE, 2019).

Uma vez que a variabilidade de transcritos gênicos provindos de mecanismos de *splicing* alternativos possam variar dentre os tecidos e que genes que tem forte influência na atividade celular, como por exemplo, o mecanismo de síntese de translesão, se faz importante analisar se as regiões canônicas de *splicing* dos genes *POLI*, *POLQ*, *POLL*, *POLN*, *POLH*, *REV1*, *POLK* e *REV3L* estão preservadas como deveriam, ou se existem variantes importantes para a neoplasia mielodisplásica.

Este estudo busca avaliar a interseção entre os achados clínicos da neoplasia mielodisplásica, o metabolismo do mecanismo de síntese de translesão e as alterações que podem ocorrer nos sítios de *splicing* dos genes *POLI*, *POLQ*, *POLL*, *POLN*, *POLH*, *REV1*, *POLK* e *REV3L*. A figura 07 representa visualmente o desenho deste estudo, uma vez que haja na literatura estudos que analisem interseções entre neoplasia mielodisplásica e mecanismos de *splicing*, mecanismo de síntese de translesão e a neoplasia mielodisplásica, bem como os mecanismos de *splicing* e os mecanismos de síntese de translesão.

Figura 7. Desenho do estudo que realiza a interseção entre os dados da neoplasia mielodisplásica, mecanismo de síntese de translesão e mecanismos de *splicing* genética com ênfase em regiões canônicas dos genes *POLI*, *POLQ*, *POLN*, *POLL*, *POLM*, *REV1*, *REV3L* e *POLH*. Cada circunferência representa uma área de estudo específica (Neoplasia mielodisplásica em azul, mecanismos de síntese de translesão em vermelho e mecanismos de *splicing* em ênfase em regiões canônicas em verde). O losango azul representa estudos que avaliaram a interseção entre os mecanismos de translesão e a neoplasia mielodisplásica. O quadrado roxo representa a interseção entre a neoplasia mielodisplásica e os mecanismos de *splicing* com ênfase em regiões canônicas, o círculo verde representa a interseção de estudos que avaliaram os mecanismos de *splicing* e os mecanismos de síntese de translesão. Já o triângulo laranja representa de maneira ampla o objetivo deste estudo que avalia a interseção entre estas três grandes áreas de pesquisa.

Fonte: Próprio autor.



2. OBJETIVOS

2.1. OBJETIVO GERAL

Descrever variantes genéticas de genes pertencentes ao mecanismo de reparo do DNA com tolerância a danos e síntese de translesão e quais predições patogênicas dessas variantes nos mecanismos de *splicing* destes genes na neoplasia mielodisplásica.

2.2. OBJETIVOS ESPECÍFICOS

- Descrever variantes de *splicing* somáticas nos genes *POLQ*, *POLN*, *POLL*, *POLM*, *REV1*, *REV3L*, *POLI* e *POLH* e qual seu impacto patogênico na neoplasia mielodisplásica;
- Analisar variantes de *splicing in silico* dos genes *POLQ*, *POLN*, *POLL*, *POLM*, *REV1*, *REV3L*, *POLI* e *POLH* através da ferramenta SpliceAI para extrair seus escores de patogenicidade;
- Associar as variantes de *splicing* dos genes *POLQ*, *POLN*, *POLL*, *POLM*, *REV1*, *REV3L*, *POLI* e *POLH* com delta score superior a 0.8 com a patogênese e prognóstico da neoplasia mielodisplásica.

3. PACIENTES, MATERIAIS E MÉTODOS

3.1. ASPECTOS ÉTICOS DA PESQUISA

O presente estudo foi submetido e aprovado 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.

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 8 genes em 30 amostras de *pool* celular de medula óssea de pacientes diagnosticados com SMD, dos subtipos SMD com sideroblastos em anel (SMD-SA) e SMD com excesso de blastos (SMD-EB), subtipos de baixo e alto risco de evolução para LMA, respectivamente. Estes pacientes foram 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.

3.2.1. 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 30 pacientes diagnosticados com Neoplasia Mielodisplásica incluídos neste estudo.

3.2.2. 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 8 genes codificadores de DNA Polimerases com atividade TLS (*REV3L*, *POLQ*, *POLH*, *REV1*, *POLI*, *POLL*, *POLM* e *POLN*), com uma cobertura total de 98%. Abaixo são apresentadas as regiões que foram avaliadas e o número de CDS de cada gene (Tabela 03).

Tabela 03. Regiões sequenciadas dos genes *REV3L*, *POLQ*, *POLH*, *REV1*, *POLI*, *POLL*, *POLM* e *POLN* 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

3.2.3. 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:

- a) **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*).
- b) **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.
- c) **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.
- d) **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, foi acompanhado o *Score*
- e) Qualidade e alinhamento foi realizada no sistema Miseq repórter.

3.2.4. Análise das variantes obtidas nos genes *POLQ*, *REV3L*, *POLH*, *REV1*, *POLI*, *POLL*, *POLM* e *POLN*

3.2.4.1. 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.

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 escore 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*.

3.2.4.2. Análise Secundária: Chamada de variantes

Para a chamada de variantes (*variant calling*) foi utilizados os softwares IGV, 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.

Para análise de variantes das regiões de splicing dos genes com atividade TLS, utilizamos inicialmente a ferramenta *SpliceAI* (JAGANATHAN ET. AL, 2019).

3.2.4.3. Análise terciária: patogenicidade pela ferramenta SpliceAI

O SpliceAI é uma ferramenta de inteligência artificial baseada em *deep learning* que realiza a predição do transcrito primário nas regiões canônicas, acceptora ou doadora, ou em nenhuma delas. Neste software, a rede neural aprende os mecanismos determinantes do *splicing* diretamente do DNA genômico avaliando 10.000 pares de bases para prever a função deste mecanismo de cada posição do transcrito primário. É importante destacar que a rede neural realiza teste utilizando posições nucleossômicas posicionadas em regiões de sítios de splicing, o estudo recomenda que para uma ótima análise de regiões doadoras ou acceptoras, a visualização de 150 nucleotídeos do genoma de interesse para predição de alteração de nucleotídeo através de inclusão ou exclusão de éxon. Os dados de notação do projeto

GENCODE foram base para o treinamento e estratificação dos parâmetros de aprendizado do software.

Essa rede de aprendizado realiza avaliação de variantes genéticas para a função do *splicing*, realizando predição dos limites éxon-íntron tanto para a sequência transcrita de pre-mRNA referência quanto para a sequência transcrita alternativa contendo a variante a ser analisada, deste modo o software realiza uma análise matemática da diferença entre pontuações medidas pelo delta score.

Variantes de *splicing* críptico apresentam pontuação ≥ 0.5 (delta score) e são validadas em RNA-seq apresentando 71% de sensibilidade quando a variante está próxima aos éxons. Quando a variante está na sequência de íntrons, apresenta sensibilidade de 41%. Variantes validadas com baixa pontuação ≤ 0.5 são transcritos aberrantes e normais nos dados de RNA-seq resultantes de *splicing* alternativo.

Em seguimento de análises e atualizações do aprendizado da ferramenta, foi observado que 35% das variantes estudadas de *splicing* críptico com pontuações intermediárias entre 0.35 e 0.8 apresentam relações importantes nas transcrições normais e aberrantes produzidas nos diferentes tecidos. É importante destacar que novas isoformas de *splicing* criada por variantes de *splicing* críptico pode variar substancialmente entre tecidos.

Esse dado tem efeito oposto quando observadas variantes com pontuações altas >0.8 , que apresentaram menor probabilidade de produzir efeitos específicos no tecido, uma vez que essas variantes estão sob forte seleção negativa. Esses achados corroboram a observação de que éxons com *splicing* alternativo tendem a apresentar pontuações intermediárias, em comparação com éxons de *splicing* constitutivo.

Com base na fração de mutações de *splicing* críptico que estão sob seleção negativa, a média em humanos é de 5 mutações funcionais raras, representando uma frequência alélica $<0,1\%$, em comparação com 11 variantes raras que apresentam proteínas truncadas. É notado que uma fração significativa das variantes de *splicing* críptico podem não anular totalmente a função do gene, mas podem produzir alterações em seu quadro e por deslocamento incompleto do *splicing* para isoforma aberrante.

Variantes de *splicing* críptico de pontuação forte provavelmente irão provocar deslocamento do *splicing* da isoforma normal para uma alterada, independente de fenômenos

epigenéticos, ao passo que variantes de pontuação mais baixa aproximam a seleção da junção de *splicing*, resultando no uso alternativo da junção em diferentes tecidos e contextos celulares.

No geral, foi analisada uma depleção de 56% de mutações crípticas comuns de *splicing* (delta escore >0.8) a uma distância de pelo menos 50 nucleotídeos de uma junção exon-ítron, consistente com maior dificuldade em prever o impacto de variantes intrônicas profundas, como observado nos dados de RNA-seq.

Para este estudo, foram consideradas as variantes que apresentaram delta escore acima ou igual a 0.8 com objetivo de manter a máxima acurácia possível dos resultados conforme demonstrado por Jaganathan e cols. (2019).

4. RESULTADOS

Após realização da nomenclatura das variantes localizadas em região de *splicing*, elas foram consultadas no software online de *Deep Learning SpliceAI* (<https://spliceailookup.broadinstitute.org/>). Essa ferramenta permite predição da posição no transcrito primário, se a variante está nas regiões doadora, acceptora ou ausente nestas regiões. O resultado fornecido pela ferramenta é dado através de um delta score que varia em ≤ 0.2 , ≥ 0.5 e ≥ 0.8 , onde variantes com delta score ≥ 0.8 predizem variantes de *splicing* críptico onde pode haver importante alteração do mRNA a ser gerado demonstrando spliciogenicidade. Já variantes com score < 0.5 apresentam isoformas diferentes do transcrito original, resultantes geralmente de *splicing* alternativo, contudo, apresentam transcritos normais com ausência de spliciogenicidade.

Em seguida, foram filtradas as variantes que ocorrem nas regiões doadora e acceptora dos genes, pois de acordo com Walker e cols. (2023), variantes localizadas nessa região tem forte impacto de função gênica levando a perda de função (*LoF variants*) e consequente desenvolvimento de doenças.

Foram incluídos no estudo, inicialmente, 30 pacientes portadores de neoplasia mielodisplásica. Deste total, 20 pacientes apresentaram 11 variantes em regiões canônicas do *splicing* na somatória dos genes estudados (Tabela 05).

Para o gene *REV3L* não foram notadas variantes em regiões canônicas do *splicing*. Já para o gene *POLH*, não foram detectadas variantes em quaisquer regiões de *splicing*. Para o gene *POLQ*, foi observado uma variante na região canônica doadora. Para o gene *REV1*, foram notadas duas variantes em região canônica acceptoras. O gene *POL1* apresentou duas variantes em região canônica, sendo uma na região acceptora e outra na região doadora. Para o gene *POLL*, foram notadas duas variantes canônicas em região acceptora. O gene *POLN* apresentou três variantes canônicas em região acceptora. Por fim, o gene *POLM* apresentou somente uma variante na região canônica doadora.

A tabela 04 apresenta as variantes detectadas nos casos analisados, dividindo entre SMD-EB e SMD-SA.

Tabela 04. Descrição de número de variantes e indivíduos para os genes *POLI*, *POLN*, *POLQ*, *REV1*, *POLL* e *POLM*.

NEOPLASIA MIELODISPLÁSICA COM EXCESSO DE BLASTOS (SMD-EB)				
Gene	Nº de variantes	Nº de indivíduos	Variante região aceptora	Variante região doadora
POLI	2	12 (60%)	POLI(NM_007195.3):c.1199-2A>G	POLI(NM_007195.3):c.1067+2T>G
POLN	3	2 (10%)	POLN(NM_181808.4):c.1309+2T>A POLN(NM_181808.4):c.1179+2T>G	Ausente
POLQ	1	5 (25%)	Ausente	POLQ(NM_199420.4):c.5774-1G>A
REV1	2	2 (10%)	REV1(NM_016316.4):c.1321+1G>A REV1(NM_016316.4):c.54+2T>G	Ausente
NEOPLASIA MIELODISPLÁSICA COM SIDEROBLASTOS EM ANEL (SMD-SA)				
POLI	2	3 (15%)	POLI(NM_007195.3):c.1199-2A>G	POLI(NM_007195.3):c.1067+2T>G
POLN	1	1 (5%)	POLN(NM_181808.4):c.2197+2 T>A	Ausente
POLL	2	2 (10%)	POLL(NM_001174084.2):c.1363+2T>A POLL(NM_001174084.2):c.1363+1G>A	Ausente
POLM	1	1 (5%)	Ausente	POLM(NM_013284.4):c.715-1G>T

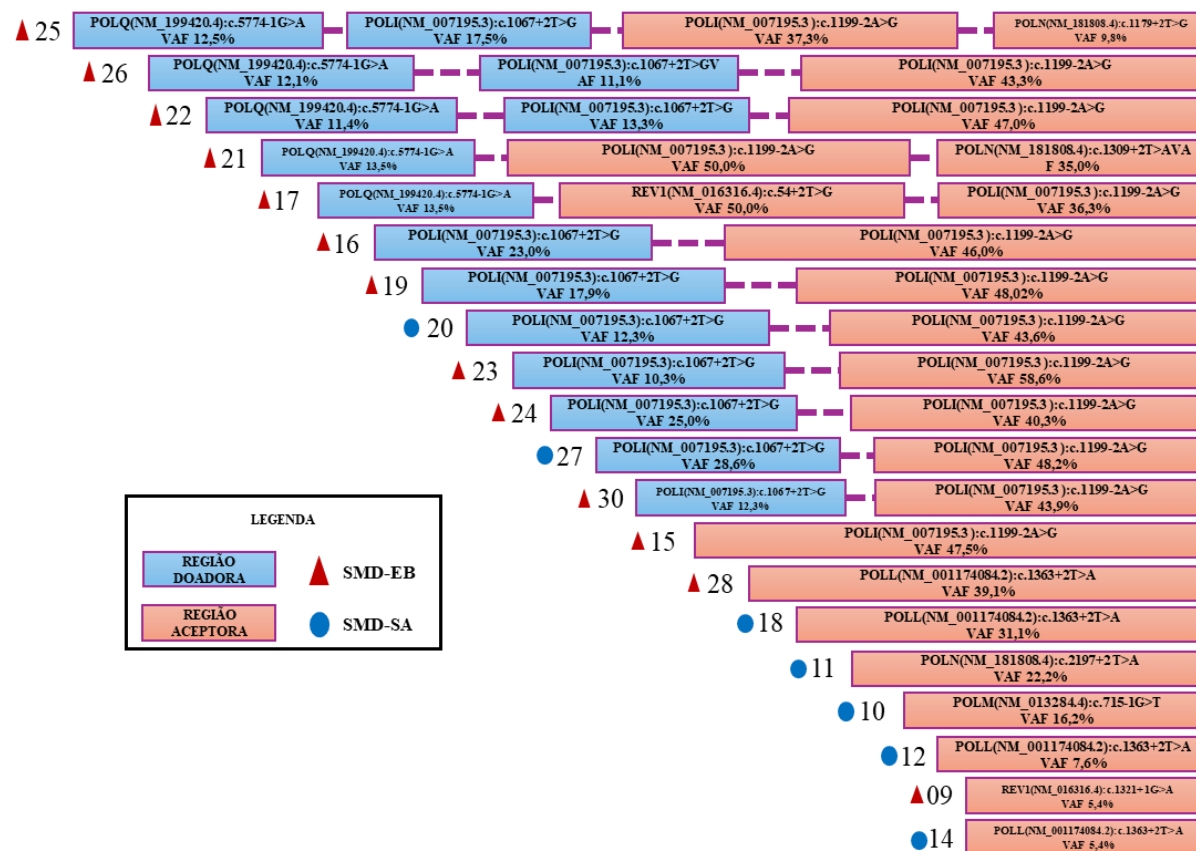
É importante destacar que dos 20 casos analisados, cinco (25%) apresentaram variantes em mais de um gene, sendo todos em indivíduos diagnosticados com SMD-EB (Figura 05).

Tabela 5. Descrição dos casos: sexo, idade, cariótipo, classificação da doença, classificação prognóstica IPSS-R, identificação da variante, VAF, éxon, região canônica e escore SpliceAI.

ID Paciente	Sexo	Idade	Cariótipo	Classificação Doença	IPSS-R	Variante	VAF(%)	Éxon	Região Canônica	Escore SpliceAI	Classificação Oncogenicidade Franklin
9	Masculino	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	REV1(NM_016316.4):c.1321+1G>A	5,4	7	Acceptora	0.93	Variante de Significado Indeterminado (VUS)
10	Masculino	72	45,X,-Y[18]/46,XY[7]	SMD-SA	Muito Baixo	POLM(NM_013284.4):c.715-1G>T	16,2	6	Doadora	0.99	Variante de Significado Indeterminado (VUS)
11	Feminino	70	Ausência de Metáfases	SMD-SA	Intermediário	POLN(NM_181808.4):c.2197+2 T>A	22,2	21	Acceptora	0.87	Variante de Significado Indeterminado (VUS)
12	Feminino	87	Ausência de Metáfases	SMD-SA	Baixo	POLL(NM_001174084.2):c.1363+2T>A	7,6	8	Acceptora	0.96	Variante de Significado Indeterminado (VUS)
14	Feminino	56	46, XX [11]	SMD-SA	Muito Baixo	POLL(NM_001174084.2):c.1363+1G>A	5,4	8	Acceptora	0.96	Variante de Significado Indeterminado (VUS)
15	Masculino	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	POLI(NM_007195.3):c.1199-2A>G	47,5	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
16	Feminino	72	46,XX,inv(3)(q21q26)[16]/46,XX[4]	SMD-EB1	Muito Alto	POLI(NM_007195.3):c.1067+2T>G	23	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
17	Masculino	73	47,XY,+8[3]/48,XY,+8,+12,dup(1)(q21q44)[5]/46,XY[12]	SMD-EB2	Alto	POLI(NM_007195.3):c.1199-2A>G	46	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
						POLQ(NM_199420.4):c.5774-1G>A	13,5	18	Doadora	0.8	Variante de Significado Indeterminado (VUS)
						REV1(NM_016316.4):c.54+2T>G	50	2	Acceptora	0.97	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1199-2A>G	36,3	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
18	Feminino	77	46,XX[20]	SMD-SA	Baixo	POLI(NM_007195.3):c.1199-2A>G	31,1	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
19	Masculino	55	45,XY,-7[15]/46,XY,-7,+mar[5]	SMD-EB2	Muito Alto	POLI(NM_007195.3):c.1067+2T>G	17,9	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
20	Masculino	92	46,XY,del(5)(q32)[13]/46,XY[17]	SMD-SA	Baixo	POLI(NM_007195.3):c.1199-2A>G	48,02	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	12,3	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
21	Masculino	58	46,XY[20]	SMD-EB1	Intermediário	POLI(NM_007195.3):c.1199-2A>G	43,6	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
						POLQ(NM_199420.4):c.5774-1G>A	13,5	18	Doadora	0.8	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	50	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
22	Masculino	79	46,XY[20]	SMD-EB1	Intermediário	POLN(NM_181808.4):c.1309+2T>A	35	10	Acceptor	0.96	Variante de Significado Indeterminado (VUS)
						POLQ(NM_199420.4):c.5774-1G>A	11,4	18	Doadora	0.8	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	13,3	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1199-2A>G	47	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
23	Masculino	64	46,XY,del(11)(q23)[3]/46,XY[12]	SMD-EB1	Intermediário	POLI(NM_007195.3):c.1067+2T>G	10,3	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
24	Masculino	84	46,XY[20]	SMD-EB1	Alto	POLI(NM_007195.3):c.1199-2A>G	58,6	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	25	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
25	Feminino	84	46,XY[20]	SMD-EB1	Alto	POLI(NM_007195.3):c.1199-2A>G	40,3	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
						POLQ(NM_199420.4):c.5774-1G>A	12,5	18	Doadora	0.8	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	17,5	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1199-2A>G	37,3	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
26	Feminino	81	46,XY[20]	SMD-EB1	Alto	POLN(NM_181808.4):c.1179+2T>G	9,8	8	Acceptor	0.96	Variante de Significado Indeterminado (VUS)
						POLQ(NM_199420.4):c.5774-1G>A	12,1	18	Doadora	0.8	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	11,1	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1199-2A>G	43,3	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
27	Feminino	82	47,XX,+8[9]/47,XX,+8,del(20)(q12)[5]/46,XX[6]	SMD-SA	Intermediário	POLI(NM_007195.3):c.1067+2T>G	28,6	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
28	Feminino	71	46,XX[20]	SMD-EB1	Baixo	POLI(NM_007195.3):c.1199-2A>G	48,2	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	39,1	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)
30	Masculino	55	45,X,-Y[15]/45,X,-Y,del(5)(q32)[3]/46,XY[2]	SMD-EB2	Muito Alto	POLI(NM_007195.3):c.1199-2A>G	12,3	7	Doadora	0.98	Variante de Significado Indeterminado (VUS)
						POLI(NM_007195.3):c.1067+2T>G	43,9	9	Acceptora	0.98	Variante de Significado Indeterminado (VUS)

Figura 08. Variantes detectadas em cada um dos casos analisados. Cada número externo aos retângulos representa a identificação dos casos estudados. As formas retangulares representam a(s) variante(s) detectada(s) em cada amostra analisada. Cada forma retangular possui um tamanho diferente quando analisado individualmente caso a caso. Quando a amostra possui somente uma variante detectada, a proporção de tamanhos se dá de maneira regressiva do ponto mais superior ao inferior da figura. Esses tamanhos representam a VAF (*Variant Allele Frequency* – Variante de Frequência Alélica). A coloração laranja dos retângulos representa variantes detectadas na região aceptora do *splicing*, ao passo que retângulos na coloração azul representam variantes detectadas na região doadora do *splicing*. A forma triangular no início de cada linha da figura representa a identificação de paciente portador de SMD-EB (Neoplasia Mielodisplásica com Excesso de Blastos) e a forma circular, paciente portador de SMD-SA (Neoplasia Mielodisplásica com Sideroblastos em anel).

Fonte: Próprio autor



4.1. VARIANTES NO GENE POLI

O gene *POLI* demonstrou uma frequência maior na casuística estudada, sendo observado em 75% do total de casos estudados (20 casos). Foram observadas 2 variantes para este gene, sendo uma variante na região aceptora do éxon nove (*POLI*(NM_007195.3):c.1199-2A>G) e outra na região doadora do éxon 7 (*POLI*(NM_007195.3):c.1067+2T>G). Do total de pacientes estudados, 11 (36,6%) apresentaram as duas variantes simultaneamente na mesma amostra e quatro (13,3%) apenas na região aceptora.

As mesmas variantes encontradas na região canônica de *splicing* do gene *POLI* em SMD-EB foram também observadas em SMD-SA. Contudo, a frequência foi maior na SMD-EB (80%), ao passo que em SMD-SA foi de 20%.

Dentre os indivíduos com SMD-EB (12 casos) com variantes na região canônica de *splicing* do gene *POLI* houve prevalência de 73,3% de indivíduos do sexo masculino.

Foi observado que houve maior número de pacientes com idade inferior a 70 anos (25%) com variantes na região canônica de *splicing*.

Dentre os indivíduos com SMD-EB com variantes de *POLI* em sua região de *splicing* houve uma proporção de 50% entre KT normal e alterado, sendo que os alterados apresentam as seguintes alterações desfavoráveis: monossomia do cromossomo 7, deleção do cromossomo 11q, inversão do cromossomo 3 e cariótipo complexo. Além disso, houve prevalência de 53% de pacientes de maior risco prognóstico de acordo com o IPSS-R.

Analisando exposição a tóxicos, houve uma leve prevalência de indivíduos não expostos (53,3%).

Dentre os indivíduos com SMD-EB com variantes de *POLI* em sua região de *splicing*, 12 pacientes apresentaram a variante na região aceptora e 8 na região doadora. Contudo, 3 pacientes (25%) apresentaram a mesma variante isolada na região aceptora e 8 pacientes (75%) apresentaram as duas variantes tanto na região canônica aceptora, quanto na região canônica doadora.

Analisando os parâmetros hematológicos dos indivíduos com SMD-EB com variantes de *POLI* em sua região de *splicing*, houve maior prevalência (58,3%) de pacientes com hemoglobina inferior a 8g/dL, ao passo que houve menor prevalência (33,3%) dos casos com menos de 50.000 plaquetas por microlitro de sangue total.

Dentre os indivíduos com SMD-EB com variantes de *POLI* em sua região de *splicing* houve uma maior prevalência (66,6%) de dependentes transfusionais.

Do total de indivíduos com SMD-EB estudados (13), 46,1% (6) deles não apresentaram variantes na região canônica de *splicing* do gene *POLI*. Ainda assim, houve maior prevalência do sexo masculino (66,6%).

Ainda dos pacientes SMD-EB sem variantes no gene *POLI*, metade apresentou hemoglobina inferior a 8 g/dL e 33% destes mesmos pacientes, apresentaram número de plaquetas inferior a 50.000/microlitro.

Dentre os indivíduos com SMD-SA com variantes de *POLI* em sua região de *splicing* houve maior prevalência do sexo feminino (66,6%).

Os indivíduos com SMD-SA com variantes de *POLI* em sua região de *splicing* demonstraram maior prevalência de cariótipo alterado (50%), sendo as principais alterações encontradas a del5q, trissomia do cromossomo 8 e del20q. Todos apresentaram prognóstico favorável segundo o IPSS-R.

Em relação à exposição a tóxicos, houve maior prevalência de indivíduos expostos (66,6%).

Dentre os indivíduos com SMD-SA com variantes de *POLI* em sua região de *splicing*, 3 casos apresentaram variante na região aceptora (100%) e 2 casos na região doadora (66,6%). Destes 3 casos, 2 apresentaram variantes simultâneas na região aceptora e doadora na mesma amostra (66,6%) e apenas 1 somente na região aceptora (33,3%).

Analisando os parâmetros hematológicos, os indivíduos com SMD-SA com variantes de *POLI* em sua região de *splicing*, demonstraram menor prevalência (33,3%) de indivíduos com hemoglobina inferior a 8 g/dL, ao passo que não foram observados pacientes com plaquetas inferiores a 50.000/microlitro. Além disso, houve maior prevalência (66,6%) de indivíduos não dependentes transfusionais.

Apesar do pequeno número de casos avaliados, esses resultados de *POLI* não demonstram, no primeiro momento, capacidade de identificar um prognóstico diferente entre os casos de anemia refratária com excesso de blastos, isto é, ter ou não variantes em regiões aceptoras ou doadoras da região de *splicing* no gene *POLI* na SMD-EB pode não adicionar nenhum efeito prognóstico. No entanto, apesar de bastante preliminar, os dados das variantes em casos de SMD com sideroblastos em anel (uma doença caracterizada em análise molecular, por ter forte relação com mutações em gene de *splicing*) parecem ser condizentes com uma doença que apresenta uma menor profundidade de citopenias com hemoglobinas maiores (consequente menos necessidade transfusional) e plaquetas maiores (acima de 50.000/microlitro), sugerindo que a presença de variantes em regiões aceptoras ou doadoras da

região de *splicing* no gene *POLI* nas SMD-SA adicionam um possível efeito de melhor prognóstico.

O nosso grupo realizou um estudo prévio de expressão gênica, ainda não publicado, dos genes *POLI*, *POLL*, *POLN*, *POLK*, *POLQ*, e *REVI*. Realizando a extração destes dados de expressão e conectando com as amostras estudadas no presente estudo, obteve-se um número de casos final pareado de sete pacientes (Tabela 06). Para efeitos sinérgicos de dados, observou-se que dois pacientes com variantes de *POLI*, um SMD-EB e outro SMD-SA tiveram maior valores de $2^{\Delta\Delta Ct}$, o que poderia significar uma tendência ao aumento expressão gênica (Wierschke et al., 2010; Xie, Wang & Zhang, 2023).

Tabela 6. Valores de DeltaDetal CT para expressões gênicas dos genes *POLI*, *POLL*, *POLK*, *POLN*, *POQ* e *REVI*. Destacado em negrito que os dois casos com variantes de *POLI* tiveram maior expressão frente aos outros casos que não possuem variantes de *POLI*.

PACIENTE ID	OMS 2016	Nomenclatura	Região canônica	Exon	POLI	POLL	POLK	POLN	POLQ	REVI
9	SMD-EB1	REVI(NM_016316.4):c.1321+1G>A	Acceptora	7	0,0018503	0,1520154	0,1315056	0,0012469	0,125602003	0,143929484
15	SMD-EB1	POLI(NM_007195.3):c.1199-2A>G	Acceptora	9	0,0440707	0,04481	0,3553393	0,051466	0,510751455	0,740632591
20	SMD-SA-DM	POLI(NM_007195.3):c.1067+2T>G	Doadora	7	0,1160570	0,1562956	0,359991	0,0195274	0,547720252	0,555400828
20	SMD-SA-DM	POLI(NM_007195.3):c.1199-2A>G	Acceptora	9	0,1160570	0,1562956	0,359991	0,0195274	0,547720252	0,555400828
12	SMD-SA	POLL(NM_001174084.2):c.1363+2T>A	Acceptor	8		0,7819992	0,6517546	0,0198811	0,292720284	0,0041602
14	SMD-SA	POLL(NM_001174084.2):c.1363+1G>A	Acceptor	8	0,0071113	0,2489687	0,0089132	0,0004043	0,035760154	0,012564158
11	SMD-SA	POLN(NM_181808.4):c.2197+2 T>A	Acceptor	21	0,0047222	0,085397	0,1466101	0,0029583	0,135252718	0,148440363
10	SMD-SA	POLM(NM_013284.4):c.715-1G>T	Donor	6	0,0169999	0,4680119	0,4826288	0,0091959	0,71848459	0,62661572

4.2. VARIANTES NO GENE POLN

O gene *POLN* apresentou menor frequência de variantes por número de casos (15%) quando comparado com a frequência de *POLI* (75%). Para essas variantes identificadas, todas ocorreram em regiões canônicas aceptoras de *splicing* do gene *POLN*, contudo, a variante do éxon 21 *POLN*(NM_181808.4):c.2197+2 T>A foi identificada na SMD-SA, ao passo que as outras duas identificadas do éxon 10 *POLN*(NM_181808.4):c.1309+2T>A e do éxon oito *POLN*(NM_181808.4):c.1179+2T>G foram descritas na SMD-EB. Havendo assim, prevalência em casos de SMD-EB (66,6%).

Devido ao número de casos pequeno, somente 3 indivíduos, as análises descritas abaixo foram agrupadas para SMD-EB e SMD-AS e descritas individualmente em cada parágrafo a seguir.

Dentre os indivíduos com variantes na região canônica de *splicing* do gene *POLN*, houve maior prevalência (66,6%) de casos do sexo feminino. Em relação ao cariótipo, não foram observadas alterações cromossômicas.

Em relação ao risco de escore prognóstico, somente 33,3% apresentou risco prognóstico alto (IPSS-R), sendo um dos dois casos de SMD-EB. Em relação aos expostos a agentes tóxicos, somente 33,3% dos casos tiveram exposição, sendo um paciente de SMD-EB e outro de SMD-SA.

Dentre os indivíduos com variantes na região canônica de *splicing* do gene *POLN*, houve menor prevalência (33,3%) de pacientes com hemoglobina inferior a 8g/dL e plaquetas inferior a 50.000/microlitro, sendo um caso de SMD-EB. Contudo, houve maior prevalência (66,6%) de pacientes dependentes de transfusão, sendo um paciente de SMD-EB e outro de SMD-SA.

4.3. VARIANTES NO GENE POLQ

O gene *POLQ* apresentou uma única variante identificada no éxon dezoito na região canônica doadora de *splicing* em 25% dos casos (POLQ(NM_199420.4):c.5774-1G>A). Para este gene, a variante identificada foi observada somente em pacientes SMD-EB com prevalência do sexo masculino (60%).

Dentre os indivíduos com variantes na região canônica de *splicing* do gene *POLQ*, 80% apresentaram cariótipo normal, mas os 20% alterados apresentaram as alterações: trissomia do cromossomo 8 e duplicação do cromossomo 1. Em relação ao risco prognóstico, 60% dos pacientes apresentaram risco alto segundo o IPSS-R.

Em relação à exposição a agentes tóxicos, somente 20% dos pacientes tiveram exposição relatada.

Analisando os parâmetros do hemograma, 40% dos pacientes apresentaram hemoglobina inferior a 8g/dL e plaquetas inferiores a 50.000/microlitro.

Por fim, 60% dos pacientes com variantes na região canônica de *splicing* do gene *POLQ* apresentaram dependência transfusional.

Apesar do pequeno número de casos avaliados, esses resultados de *POLQ* não demonstram, no primeiro momento, capacidade de diferenciar prognóstico entre os casos de neoplasia mielodisplásica com excesso de blastos, isto é, ter ou não variantes em regiões aceptoras ou doadoras da região de *splicing* no gene *POLQ* na SMD-EB não adiciona nenhum efeito prognóstico.

4.4. VARIANTES NO GENE REV1

O gene *REVI* apresentou 2 variantes na região canônica aceptora de *splicing*, sendo elas: no éxon sete REV1(NM_016316.4):c.1321+1G>A e no éxon dois REV1(NM_016316.4):c.54+2T>G. Além disso, foram observadas em apenas 10% dos casos.

As variantes foram observadas somente em pacientes SMD-EB. E todos os casos (2 indivíduos) demonstraram alto risco segundo o IPSS-R.

Todas as variantes foram identificadas somente no sexo masculino. Além disso, todos os casos apresentaram cariótipo alterado, sendo um complexo e outro com monossomias.

Em relação à análise de exposição a tóxicos, 50% dos casos foram expostos.

Analisando os parâmetros do hemograma, 100% dos casos tiveram hemoglobina inferior a 8g/dL, ao passo que nenhum dos pacientes apresentou plaquetas com contagem inferior a 50.000/microlitro. Por fim, 100% dos casos apresentaram dependência transfusional.

Apesar do pequeno número de casos avaliados, esses resultados de *REVI* não demonstram, no primeiro momento, capacidade de diferenciar prognóstico entre os casos de neoplasia mielodisplásica com excesso de blastos, isto é, ter ou não variantes em regiões aceptoras ou doadoras da região de *splicing* no gene *REVI* na SMD-EB não adiciona nenhum efeito prognóstico.

4.5. VARIANTES NO GENE POLL

O gene *POLL* apresentou duas variantes no éxon oito na região canônica aceptora de *splicing*, sendo elas: POLL(NM_001174084.2):c.1363+2T>A e POLL(NM_001174084.2):c.1363+1G>A. E assim como REV1, foi observado em 10% dos casos. Diferente do observado em *REVI*, os casos com variantes na região canônica aceptora de *splicing* do gene *POLL* foram observados somente em SMD-SA.

Para este gene, só foi observado pacientes do sexo feminino e não visualizado risco alto de prognóstico, segundo IPSS-R. Para análise de cariótipo, foi visualizado somente cariótipo normal. Em relação a exposição aos tóxicos, somente 50% dos casos foi observada exposição.

Analisando os parâmetros do hemograma, não foram observados pacientes com hemoglobina inferior a 8g/dL ou plaquetas com contagem inferior a 50.000/microlitro. Por fim, 100% dos casos demonstraram não ser dependentes de transfusão.

Os dados das variantes de *POLL* em casos de SMD com sideroblastos em anel (uma doença por si só de um gene de *splicing*) parecem ser condizentes com uma doença que apresenta uma menor profundidade de citopenias com hemoglobinas maiores (consequente menos necessidade transfusional) e plaquetas maiores (acima de 50.000/microlitro), sugerindo

que a presença de variantes em regiões aceptoras ou doadoras da região de *splicing* no gene *POLL* nas SMD-SA adicionam um possível efeito de melhor prognóstico.

4.6. VARIANTES NO GENE POLM

O gene *POLM* apresentou uma única variante em 5% dos casos estudados, sendo ela *POLM* (NM_013284.4):c.715-1G>T presente na região canônica doadora de *splicing* no éxon seis.

Esta variante foi observada em um paciente do sexo masculino com nulissomia do cromossomo Y portador de SMD-SA de risco prognóstico muito baixo. Foi um paciente exposto a produtos tóxicos como solventes e tintas e nos parâmetros hematológicos, não apresentou hemoglobina inferior a 8g/dL nem plaquetas com contagem inferior a 50.000/microlitro. O paciente não apresentava dependência transfusional

5. DISCUSSÃO

Este é o primeiro trabalho a estudar variantes nas regiões canônicas de *splicing* dos genes *POLI*, *POLQ*, *POLL*, *POLN*, *POLM*, *REV1*, *POLH* e *REV3L*.

É comum que um único transcrito primário possa conter diferentes regiões que são reconhecidas pelos mecanismos de *splicing* alternativo, eles resultam em efeitos combinatórios provindos da seleção destes diferentes sítios de *splicing* do transcrito primário, gerando entre 10 e 100 isoformas diferentes do mesmo gene (CARTEGINI, CHEW & KRAINER, 2002).

Todas as variantes encontradas de *POLI*, *POLL*, *POLM*, *POLN*, *REV1* e *POLQ* demonstraram frequência extremamente baixa em banco de dados populacionais do gnomAD, estando ausente do mesmo. Além disso, não foram encontradas evidências clínicas de caráter somático no software Franklin.

Nesta mesma plataforma, se observou uma variação de número de transcritos para cada uma das variantes analisadas: *POLI*(NM_007195.3):c.1199-2A>G apresenta 25 transcritos; *POLI*(NM_007195.3):c.1067+2T>G apresenta 29 transcritos; *POLL*(NM_001174084.2):c.1363+1G>A e *POLL*(NM_001174084.2):c.1363+1G>A apresentam 19 transcritos; *POLM*(NM_013284.4):c.715-1G>T apresenta 22 transcritos; *POLN*(NM_181808.4):c.1309+2T>A e *POLN*(NM_181808.4):c.1179+2T>G apresentam sete transcritos, já a variante *POLN*(NM_181808.4):c.2197+2 T>A apresentou 11 transcritos; *POLQ*(NM_199420.4):c.5774-1G>A apresenta dois transcritos; *REV1*(NM_016316.4):c.1321+1G>A apresenta 19 transcritos e a variante *REV1*(NM_016316.4):c.54+2T>G apresenta 14 transcritos. Dessa forma, sugere-se que estes genes possuem alta atividade dos mecanismos de *splicing* documentadas na literatura, embora ainda não tenham sido evidenciadas associações de oncogenicidade e patogenicidade para eles.

Analisando o prognóstico (IPSS-R) dos pacientes que possuem duas variantes de *POLI* na mesma amostra, observou-se que dois casos foram SMD-SA e apresentaram prognósticos intermediário e baixo. Já para os casos de SMD-EB ainda com duas variantes de *POLI* na mesma amostra, observou-se que houve um aumento do risco prognóstico, uma vez que um paciente apresentou risco intermediário, um paciente com risco alto e três com risco muito alto. Esse achado pode sugerir que a presença da variante de *POLI* na SMD sabidamente de alto risco, pode influenciar negativamente no

prognóstico da doença, pois provavelmente aumenta o risco patogênico na evolução natural da doença. Além disso, um estudo realizado por Oliveira (2023) observou nos dados ainda não publicados que houve aumento de expressão gênica nos dois casos que tem a variante de *POLI*, pode-se inferir de maneira leve que há um aumento de expressão desta variante podendo contribuir para algum processo evolutivo da doença disforme na produção transcritos que possam levar a produção de proteínas com funções alteradas que influenciem diretamente na atividade metabólica do mecanismo de *splicing*.

Além disso, foi encontrado a mesma quantidade de pacientes com SMD-EB portadores de variantes de *POLI*, *POLQ*, *POLN* e *REVI*, na mesma amostra, onde dois apresentam risco intermediário e três apresentaram risco alto. É importante destacar que nas cinco amostras, sempre há a presença de pelo menos uma variante de *POLI*. Observando que houve uma redução, embora leve, da avaliação do risco prognóstico da doença, este achado pode sugerir que as variantes de *POLQ*, *POLN* e *REVI*, quando detectadas em conjunto com *POLI*, influenciam na redução do risco prognóstico de pacientes com SMD-EB.

Comparando os pacientes portadores de SMD-EB com variantes isoladas de *POLI* (somente na região aceptora do *splicing* e no éxon nove) e com as duas variantes detectadas neste estudo (uma em região aceptora e outra na região doadora do *splicing* e no éxon sete) observou-se que é mais prevalente a presença de casos com as duas variantes simultâneas (dois casos com variantes isoladas e cinco casos com variantes simultâneas). Além disso, observou-se que quando a variante de *POLI* se apresenta isoladamente, o risco prognóstico varia entre baixo e muito alto, já quando se apresentam de maneira simultânea, houve a prevalência de risco muito alto (três casos), além de um paciente com risco alto e outro com risco intermediário. Este achado pode sugerir que quando as variantes se apresentam de maneira simultânea, há um aumento do risco prognóstico da doença. Analisando que cada uma das variantes foi detectada em éxons diferentes, pode-se inferir que há uma maior possibilidade de produção de proteínas truncadas e com alteração de função, visto que provavelmente foram resultado do fenômeno de éxon críptico, em concordância com a análise *in silico* provinda do software SpliceAI (JAGANATHAN et al., 2019).

Analisando o cariótipo dos pacientes portadores de SMD-EB e nos quais foram encontradas pelo menos uma variante de *POLI*, observou-se alterações citogenéticas graves como cariótipo complexo, monossomia do cromossomo sete e deleção do cromossomo 11q. Foi observada também deleção do cromossomo 5q, para esta alteração,

é importante destacar que quando o paciente é portador de SMD-EB esta alteração se demonstra associada de nulissomia do cromossomo Y no mesmo clone, ao passo que quando o paciente é portador de SMD-SA, a del(5q) se apresenta de maneira isolada. Estes achados sugerem e contribuem com os achados do índice prognóstico IPSS-R, ou seja, as variantes de *POLQ* podem estar associadas a uma doença de maior risco prognóstico.

É descrito que a atividade de *POLQ* pode alterar sequências intrônicas e exônicas, que por sua vez, podem potencializar ou inativar elementos regulatórios do *splicing* (CECCALDI et al., 2015). Associando que este gene foi o único, dentre os estudados, que apresentou menor números de transcritos, pode ser um indicativo de forte seleção natural neste gene.

O número de transcritos documentados para um gene nem sempre reflete diretamente a ação da seleção natural, pois pode depender do nível de estudo e interesse científico. No entanto, quando múltiplos transcritos são funcionais e conservados entre espécies, isso pode indicar ação da seleção natural, seja positiva, ao diversificar isoformas proteicas, ao manter variantes essenciais. Genes com maior diversidade transcricional tendem a estar associados a processos adaptativos complexos, como resposta imune ou regulação celular, sugerindo que essa diversidade pode oferecer vantagem evolutiva e ser mantida por pressões seletivas (ARRIBAS et al., 2024).

Além disso, a presença de *splicing* alternativo e a geração de múltiplas isoformas podem permitir que um único gene atue em diferentes contextos celulares, temporais ou ambientais, aumentando a complexidade regulatória sem necessidade de novos genes. Esse mecanismo é particularmente vantajoso em organismos mais complexos, onde a diversidade funcional precisa ser alcançada com um número limitado de genes. Assim, quando essa diversidade transcricional é mantida ao longo da evolução e apresenta expressão específica ou regulada, reforça-se a hipótese de que está sob influência da seleção natural, contribuindo para a adaptação e sobrevivência da espécie (VERTA & JACOBS, 2021).

Ceccaldi e cols (2015) e Junior e cols (2019) sugerem que a inibição de *POLQ* em células com os genes pertencentes à maquinaria do *splicing*, podem levar ao fenômeno genético conhecido como “*Synthetic Lethality*”, esse fenômeno é descrito quando há uma interrupção mútua de dois ou mais genes que levam à morte celular, enquanto a mutação de um único gene isolado não causaria este dano celular. Esse fenômeno demonstra uma

relação compensatória onde as células são dependentes da atividade de *POLQ* quando o *splicing* é defeituoso e vice-versa.

Um estudo *in vivo* realizado no Canadá, procurou descrever o papel dos eventos de *splicing* e o seu impacto na instabilidade genômica em células tronco embrionárias (*ESCs – Embryonic Stem Cells*) de camundongos. Neste experimento, eles realizaram sequenciamento de RNA para realizar análise comparativa de expressão dos genes reguladores da resposta de danos ao DNA (*DDR – DNA Damage Response*) em células tronco embrionárias comparando com fibroblastos embrionários (*MEF – Mouse Embryonic Fibroblasts*) (JIANG ET AL., 2023).

Eles avaliaram 805 genes estudados que apresentaram diferenças significativas nos eventos de *splicing* entre células tronco embrionárias e células somáticas. Um dos resultados observados foi que 20% dos genes apresentaram alteração no padrão do *splicing* alternativo (JIANG ET AL., 2023).

Eles identificaram uma célula tronco embrionária específica, chamada RBP DPPA5A, que garante a estabilidade genômica ao regular o *splicing* do éxon críptico dos principais genes de replicação e reparo do DNA *REVI* e *POLQ* (JIANG ET AL., 2023).

Dentre os eventos de *splicing* alternativo, o *splicing* de éxon críptico, também chamados de pseudoéxons, tem impacto importante na expressão proteica. Eles são normalmente sequências intrônicas que quando incluídas em transcritos de mRNAs maduros geram alterações frameshift e/ou a geração de códons de parada prematuros. Esses eventos geralmente promovem redução da expressão proteica. Nos genes que apresentaram alteração no padrão de *splicing* alternativo (20%, 162 genes) 15 deles foi visualizado inclusão de pseudoéxons, os genes *POLQ* e *REVI* mostraram uma alta taxa de inclusão em células tronco embrionárias comparadas com fibroblastos embrionários (JIANG ET AL., 2023).

Eles concluíram que a alta taxa de inclusão de pseudoéxons nos genes *REVI* e *POLQ* reduzem a expressão proteica, suprimem a replicação mutagênica e as vias de reparo do DNA. Contudo, é importante destacar que a regulação da expressão destes genes é mediada a nível pós-transcricional, pela inclusão de pseudoéxons (JIANG ET AL., 2023).

Sendo assim, a variante de *POLQ* neste estudo foi encontrada somente em pacientes com SMD-EB e ela sempre está acompanhadas de variantes nos genes *POLII*, *POLN* e *REVI*. Este achado pode sugerir que este gene está associado à alta instabilidade genômica acompanhada de variantes associadas a outros genes do mecanismos de síntese

de translesão. É importante destacar que a SMD-EB é uma doença de alto risco prognóstico que geralmente apresentam maior número de displasias na medula óssea e com contagem mais alta de blastos relativo às outras formas da doença. Corroborando este dado, embora os pacientes não tenham apresentado hemoglobina inferior a 8g/dL e contagem de plaquetas inferior a 50.000, houve prevalência de pacientes dependentes transfusionais. Isso significa que a atividade medular para produção de células funcionais para o sangue periférico está prejudicada e que é necessária reposição delas como forma de tratamento paliativo.

Corroborando o estudo de Jiang e cols (2023), pode-se inferir que a alta taxa de pseudoéxons nos genes *POLQ* e *REVI* descrita pode estar associada com essa instabilidade metabólica a nível pós-transcricional destes genes, podendo evidenciar como um dos fenótipos importantes na doença a dependência transfusional.

É importante destacar também que os cinco pacientes portadores de múltiplas variantes apresentaram variante de *POLQ*, contudo, somente três pacientes eram dependentes transfusionais. Os dois pacientes que não eram dependentes transfusionais não demonstraram nenhum dado diferente entre os cinco casos estudados que pudesse justificar a disparidade do dado de dependência transfusional total para *POLQ*.

Embora haja estudos que demonstram algumas atividades do gene *REVI* relativas aos mecanismos de *splicing*, essas interações ainda não são bem elucidadas, uma vez que os mecanismos de reparo do DNA com atividade de translesão e os mecanismos de *splicing* sejam vias diferentes (JIANG ET AL., 2023).

Para este estudo, observou-se a presença de variantes de *REVI* somente na SMD de alto risco (SMD-EB), apresentando duas variantes em éxons diferentes, uma no éxon sete e outra no éxon 2, mas em pacientes diferentes. Adicionalmente, todos os pacientes apresentaram hemoglobina inferior a 8g/dL e foram dependentes transfusionais. O cariótipo dos pacientes apresentou alterações graves como cariótipo complexo e monossomias. Mutações no gene *REVI* foram associadas a tumores sólidos como câncer de ovário e pulmão, embora não tenham sido estudadas variantes na região de *splicing* deste gene é importante explicitar que este gene já tem transcritos importantes para evolução oncogênica (ZHU ET AL., 2022).

Sendo assim, observou-se que os casos que apresentaram variantes de *POLL*, *POLQ*, *POLN* e *REVI* têm forte associação à dependência transfusional, demonstrando um alto risco de doença e prognóstico preocupante para a evolução da doença.

No estudo de Cleaver, Collins e Volik (2003), conduziram uma análise detalhada das regiões genômicas dos genes *POLH* e *POLI* tanto de humanos como de camundongos. Essa análise permitiu observar a eficiência da maquinaria do *splicing* nas regiões de cada junção exon-intron para o *POLH* em humanos. O objetivo central foi entender melhor as consequências das duplicações gênicas e causas de exon *skipping*

Para o gene *POLI* foi observado em um ensaio de modelos murinos *knockout* cepa 129/J através de RT-PCR que o gene apresenta uma mutação no códon de parada do éxon 2 resultando em uma proteína truncada, pois há perda deste éxon na formação spliciogênica do mRNA. Esse éxon corresponde aos resíduos 14-55 ou 34-75 em duas variantes de *POLI* conhecidas (D34 e Y59) com sítio de iniciação da tradução alternativos. Os autores especularam que estas variantes alteram interações com outras proteínas que por consequência se pode obter alterações regulatórias bioquímicas (KAZACHENKO ET AL., 2017).

Para o gene *POLN*, o fenômeno é similar aos demais já descritos acima, contudo, só se tem estudos *in vitro* documentados até o momento. O estudo realizou cultura de linhagem celular K562 e 833K e analisou o cDNA de 20 clones humanos que foram sequenciados. Desta casuística, 14 apresentaram variações no *splicing* alternativo com capacidade de produzir proteínas inativas causadas por exon *skipping* ou inserção de éxons alternativos (MARINI ET AL., 2003).

Dentre os casos analisados neste estudo, as variantes de *POLN* foram detectadas em pacientes portadores de SMD-SA e SMD-EB, sendo a maior prevalência na SMD de alto risco. Embora o cariótipo e hemograma não tenham trazido alterações importantes, observou-se dependência transfusional em maior número de pacientes.

Para os genes *POLL* e *POLM*, não foram encontrados registros na literatura se há alguma associação funcional biológica com os mecanismos de *splicing*. Para *POLM*, um único estudo apenas levanta a hipótese de que é possível que o *splicing* alternativo parece fazer parte da regulação de expressão deste gene durante resposta imune (LUCAS ET AL., 2005).

Para os casos analisados neste estudo com variantes de *POLL* e *POLM*, observou-se que foram detectadas somente em pacientes portadores da doença de menor risco (SMD-SA). Em todos os casos não houve dependência transfusional ou quaisquer alteração grave no cariótipo.

Sendo assim, se fazem necessárias maiores investigações acerca dos mecanismos envolvidos no *splicing* do funcionamento das DNA polimerases com atividade de

translesão, uma vez que apresentam importante papel no reparo do DNA, bem como possuem alto potencial para desenvolvimento carcinogênico.

6. CONCLUSÕES

Este foi o primeiro trabalho a estudar variantes nas regiões canônicas dos genes *POLQ*, *POLN*, *POLL*, *POLM*, *REVI* e *POLI*. Devido a este fato, houve poucas evidências na literatura que esclarecessem e corroborassem os dados encontrados neste estudo. Com exceção dos outros genes, não foram encontradas na casuística variantes para os genes *POLH* e *REV3L*. Além disso, dentre as variantes detectadas, houve maior destaque para os genes *POLI*, *POLQ* e *POLN*, nas quais foram mais prevalentes na SMD-EB. Um achado importante foi observado quando há dependência transfusional, há uma prevalência destas variantes, sugerindo que elas podem ter algum papel importante neste aspecto, sendo necessários outros ensaios mais direcionados para confirmar a hipótese.

Além disso, houve uma maior prevalência de variantes no gene *POLI* dentre os casos estudados e uma vez extraído o dado da ferramenta SpliceAI como variante patogênica, verifica-se que este gene, dentre os estudados foi associado a cariótipo complexo, prognóstico de alto risco e pacientes dependentes transfusionais. A união destas características confirma que a variante detectada neste gene tem associação com forma de alto risco da doença.

O presente trabalho buscou descrever variantes nos genes estudados e correlacionar com algumas variáveis clínicas importantes para prognóstico da SMD. Estes dados poderão contribuir para projetos futuros que objetivem classificar essas variantes detectadas quanto a sua oncogenicidade.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ARBER D. A, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. **Blood**. 127(20):2391-405. 2016.
- ARBER, Daniel A. et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. **Blood**. 127(20): 2391-405. 2016.
- ARRIBAS, Yago A. et al. Transposable element exonization generates a reservoir of evolving and functional protein isoforms. **Cell**. 187(26):7603-7620. 2024.
- ATES, Esra Arslan et al. A Novel ATM Gene Mutation Affecting Splicing in an Ataxia-Telangiectasia Patient. **Mol Syndromol**. 13:80–84. 2022.
- AUBOEUF D, Carno-Fonseca M, Valcárcel J, Biamonti G. Alternative splicing and cancer. **J Nucleic Acids**. 2012;2012:363809
- BARALLE, D;Baralle, M. Splicing in action: assessing disease causing sequence changes. **J Med Genet**. 42: 737-748. 2005.
- BEJAR R. Clinical and genetic predictors of prognosis in myelodysplastic syndromes. **Haematologica**. 2014; 99(6):956-964.
- BEJAR R. et al. Clinical effect of point mutations in myelodysplastic syndromes. **N Engl J Med**. 364(26):2496-506. 2011.
- BELLI, Carolina B. et al. Myelodysplastic syndromes in South America: a multinational study of 1080 patients. **Am J Hematol**. 90(10): 851-8. 2015.
- BENNETT, J. M. et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. **Br J Haematol**. 33(4): 451-8.1976.
- BERNARD, Elsa et al. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. **NEJM Evid**. 1(7). 2022.
- BERTRAM, Karl. et al. Cryo-EM Structure of a pre-catalytic human spliceosome primed for activation. **Cell**. 170:701-713. 2017.
- BIANCON, Giulia et al. Precision analysis of mutant U2AF1 activity reveals deployment of stress granules in myeloid malignancies. **Mol Cell**. 82(6): 1107-1122.2022.
- BIERNACKI, Melinda Ann et al. Discovery of U2AF1 neoantigens in myeloid neoplasms. **Journal for ImmunoTherapy of Cancer**. 11.2023.
- BONNER, Elizabeth A., Lee, Stanley C. Therapeutic Targeting of RNA Splicing in Cancer. **Genes**. 14 (1378). 2023.

- BRADLEY, Robert K., Anczuków, Olga. RNA splicing dysregulation and the hallmarks of cancer. **Nature reviews cancer**. 23: 135-155. 2023.
- BRAUNSCHWEIG, Ulrich et al. Dynamic integration of splicing within gene regulatory pathways. **Cell**. 152 (6): 1252-1269. 2013.
- BROOKS AN, Choi PS, de Waal L, Sharifnia T, Imielinski M, et al. A pan-cancer analysis of transcriptome changes associated with somatic mutations in *U2AF1* reveals commonly altered splicing events. **PLOS ONE**. 2014;9:e87361.
- CARTEGINI, Luca; Chew, Shern L.; Krainer, Adrian. Listening to silence and understanding nonsense: exonic mutation that affect splicing. **Nature reviews**. 3: 285-298. 2002.
- CAVALCANTE, Gabrielle M. et al. Tissue methylation and demethylation influence translesion synthesis DNA polymerases (TLS) contributing to the genesis of chromosomal abnormalities in myelodysplastic syndrome. **J Clin Pathol**. 75(2): 85-93. 2022.
- CECCALDI, Raphael et al. Homologous-recombination-deficient tumors are dependent on POLQ-mediated repair. **Nature**. 518(7538): 258-62. 2015.
- CHAMSEDDINE An et al. Unraveling Myelodysplastic Syndromes: Current Knowledge and Future Directions. **Curr Oncol Rep**. 2016; 18(4): 1-11.
- CHATTERJEE, Nimrat; Walker, Graham C. Mechanisms of DNA damage, repair, and mutagenesis. **Environ Mol Mutagen**. 58(5):235-263. 2017.
- CIESLA, Maciej. et al. m6A-driven SF3B1 translation control steers splicing to direct genome integrity and leukemogenesis. **Molecular Cell**. 83: 1165-1179. 2023.
- CLEAVER, J.E.; Collins, C.; Volik, S. Genome sequence and splice site analysis of low-fidelity DNA polymerases H and I involved in replication of damaged DNA. **Genomics**. 82(5):561-70. 2003.
- CLOUGH, Courtnee. et al. Coordinated missplicing of TMEM14c and ABCB7 causes ring sideroblasts formation in SF3B1-mutant myelodysplastic syndrome. **Blood**. 139(13): 2038-2049. 2022.
- COOPER TA, Wan L, Dreyfuss G. RNA and disease. **Cell**. 2009; 136:777–93.
- COSTA, Marilia B. et al. Chromosomal abnormalities and dysregulated DNA repair gene expression in farmers exposed to pesticides. **Environ Toxicol Pharmacol**. 2021.
- CYBULLA, Emily; Vindigni, Alessandro. Leveraging the replication stress response to optimize cancer therapy. **Nat Rev Cancer**. 23(1):6-24. 2023.

- DAIGAKU Y.; Davies AA.; Ulrich HD. Ubiquitin-dependent DNA damage bypass is separable from genome replication. **Nature**. 465(7300):951-5. 2010.
- DUNHAM I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Fietze S, HARROW J, et al. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. **Nature**. 2012; 489:57–74.
- FLANAGAN, Annabelle M. et al. The human POLH gene is no mutated, and is expressed in a cohort of patients with basal or squamous cell carcinoma of the skin. **Int J Mol Med**. 19(4): 589-96. 2007.
- FOX, JT.; Lee, KY.; Myung K. Dynamic regulation of PCNA ubiquitylation/deubiquitylation. **FEBS Lett**. 585(18):2780-5. 2011.
- GERLACH et al. Human and mouse homologs of Escherichia coli DinB (DNA polymerase IV), members of the UmuCyDinB superfamily. **Proc Natl Acad Sci U S A(PNAS)**. 96(21): 11922-7. 1999.
- GERLACH et al. Purification and Characterization of polk, a DNA Polymerase Encoded by the Human DINB1 Gene. **The Journal of Biological Chemistry**. 276 (1): 92-98. 2001.
- GHOSAL G.; Chen J. DNA damage tolerance: a double-edged sword guarding the genome. **Transl Cancer Res**. 2013;2(3):107-129.
- GOODEVE, A. C. Hemophilia B: molecular pathogenesis and mutation analysis. **Journal of Thrombosis and Haemostasis**. 13: 1184–1195. 2015.
- GRASSO et al. The mutational landscape of lethal castration-resistant prostate cancer. **Nature**. 487(7406):239-243. 2012.
- GREENBERG P, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. **Blood**, v. 89, n. 6, p. 2079-88, 1997.
- GUILLIAM, Thomas A; Yeeles, Joseph T. P. Reconstitution of translesion synthesis reveals a mechanism of eukaryotic DNA replication restart. **Nat Struct Mol Biol**. 27(5):450-460. 2020.
- GUO, C et al. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. **EMBO J**. 22:6621–6630. 2003.
- GURULE, Natalia J. et al. Myelodysplastic neoplasm-associated U2AF1 mutations induce
- Hahn CN, Scott HS. Spliceosome mutations in hematopoietic malignancies. **Nat Genet**. 2012;44:9–10

- HAKIM, Nor Hakimah Ab et al. Neuron-specific splicing. *Biosci Trends*. 11(1):16-22. 2017.
- HAN, Xin. et al. Nonsense-mediated mRNA decay: a ‘nonsense’ pathway makes sense in **stem cell biology**.
- HARADA, Naoya et al. The splicing factor CCAR1 regulates the Fanconi anemia/BRCA pathway. **Molecular Cell**. 84(14): 2618-2633. 2024.
- HE, X. et al., REV1 genetic variants associated with the risk of cervical carcinoma. **European journal of epidemiology**. 23(6): 403–409. 2008.
- HO, The V.; Schärer, Orlando D. Translesion DNA synthesis polymerases in DNA interstrand crosslink repair. **Environ Mol Mutagen**. 51(6): 552-66. 2010.
- HOEIJMAKERS, J. H. DNA damage, aging, and cancer. **N Engl J Med**. 361(15):1475-85. 2009.
- HOSONO, Naoko. Genetic abnormalities and pathophysiology of MDS. **International Journal of Clinical Oncology**. 24: 885-892. 2019.
- host defense defects by compromising neutrophil chemotaxis. **Leukemia**. 37: 2115-2124. 2023.
- HUANG, Min et al. RNA-splicing factor SART3 regulates translesion DNA synthesis. **Nucleic Acids Research**. 46(9). 2018.
- IYAMA T.; Wilson D.M. DNA repair mechanisms in dividing and non-dividing cells. **DNA Repair**. 12(8):620-36. 2013.
- JAGANATHAN, K. et al. Predicting Splicing from Primary Sequence with Deep Learning. **Cell**. 176:535–548. 2019.
- JIANG, Fangjie et al. DPPA5A suppresses the mutagenic TLS and MMEJ pathways by modulating the cryptic splicing of REV1 and POLQ in mouse embryonic stem cells. **Proc Natl Acad Sci U.S.A.** 120(30). 2023.
- 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.
- JUNIOR, Howard Lopes Ribeiro et al. Can synthetic lethality approach be used with DNA repair genes for primary and secondary MDS?. **Med Oncol**. 36(12):99. 2019.
- KANAO, R.; Masutani, C. Regulation of DNA damage tolerance in mammalian cells by post-translational modifications of PCNA. **Mutat Res**. 803-805:82-88. 2017.

- KAZACHENKO, Konstantin Y. et al. Alternative splicing at exon 2 results in the loss of the catalytic activity of mouse DNA polymerase ι in vitro. **DNA repair**.50:77-82. 2017.
- KELEMEN, Olga et al. Function of alternative splicing. **Gene**. 514 (1): 1-30. 2013.
- KIM, Eunhee et al. SRSF2 mutations contribute to myelodysplasia through mutant-specific effects on exon recognition. **Cancer Cell**. 27(5): 617-630. 2015.
- KNOBEL, PA.; Marti, T. M. Translesion DNA synthesis in the context of cancer research. **Cancer Cell Int**.11:39. 2011.
- LANGIE, S. A. et al. Causes of genome instability: the effect of low dose chemical exposures in modern society. **Carcinogenesis**. 36(1):61-88. 2015
- LEE, Yeon; Rio, Donald C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. **Annu Rev Biochem**. 84: 291-323.2015.
- LI, Xiaoqing et al. Polymerase ι (POLI) confers radioresistance of esophageal squamous cell carcinoma by regulating RAD51 stability and facilitating homologous recombination. **Cell Death Discov**. 9(1):291. 2023.
- LINDAHL T.; Wood R.D. Quality control by DNA repair. **Science**. 286(5446):1897-905. 1999.
- LIU, Xiaolei et al. A mitochondrial surveillance mechanism activated by SRSF2 mutations in hematologic malignancies. **The Journal of Clinical Investigation**. 134(12). 2024.
- LIU, Zhaoqi. et al. Mutations in the RNA splicing factor SF3B1 promote tumorigenesis through MYC stabilization. **Cancer discov**. 10(6):806-821. 2020.
- LUCAS, Daniel et al. Polymerase ι is up-regulated during the T cell-dependent immune response and its deficiency alters developmental dynamics of spleen centroblasts. **Eur. J. Immunology**. 35:1601-1611. 2005.
- MA, Xiaolu; Tang, Tie-Shan; Guo, Caixia. Regulation of Translesion DNA synthesis in mammalian cells. **Environ Mol Mutagen**. 61(7): 680-692. 2020.
- MACIEJEWSKI, Jaroslaw P.; Padgett, Richard A. Defects oin spliceosomal machinery: a new pathway of leukemogenesis. **Br. J Haematol**. 158(2): 165-173. 2012.
- MAGALHÃES, Silvia M. et al. Guidelines on myelodysplastic syndromes: Associação Brasileira de Hematologia, Hemoterapia e Terapia Celular. **Hematol Transfus Cell Ther**. 40(3): 255-261. 2018.
- Makarova AV, Burgers PM. Eukaryotic DNA polymerase ζ . **DNA Repair (Amst)**. 29:47-55. 2015.

- MAKISHIMA H, Visconte V, Sakaguchi H, Jankowska AM, Abu Kar S, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. **Blood**. 2012;119:3203–10
- MALCOVATI L. et al; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium and of the Associazione Italiana per la Ricerca sul Cancro Gruppo Italiano Malattie Mieloproliferative. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. **Blood**. 118(24):6239-6246. 2011.
- MALIK, Radhika et al. Cryo-EM structure of the Rev1-Pol ζ holocomplex reveals the mechanism of their cooperativity in translesion DNA synthesis. **Nat Struct Mol Biol**. 31(9):1394-1403. 2024.
- MARINI, Federica; Kim, Nayun; Schuffert, Anthony and Wood, Richard. POLN, a nuclear POLA Family DNA polymerase homologous to the DNA cross-link sensitivity protein Mus308. **The Journal of Biological Chemistry**. 278(34):32014-19. 2003.
- MARTENS, Marie C.; Emmert, Steffen; Boeckmann, Lars. Xeroderma Pigmentosum: Gene Variants and Splice Variants. **Genes (Basel)**. 29;12(8):1173. 2021.
- MARTIN, Sara K.; Wood, Richard D. DNA Polymerase in DNA replication and repair. **Nucleic Acids Res**. 47(16):8348-8361. 2019.
- MASUMURA et al. Limited ability of DNA Polymerase Kappa to suppress Benzo[a]pyrene-induced genotoxicity in vivo. **Environmental and molecular mutagenesis**. 58: 644-653. 2017.
- MASUTANI, C. et al. Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. **EMBO J**. 18(12):3491-501. 1999.
- MATSUNAWA M, Yamamoto R, Sanada M, et al. Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia. **Leukemia** 2014; 28(9):1844–50.
- MAYERLE, Megan. Guthrie, Christine. Genetics and biochemistry remain essential in the structural era of the spliceosome. **Methods**. 125:3-9. 2017.
- MOLDOVAN, G. L. et al. DNA polymerase POLN participates in cross-link repair and homologous recombination. **Mol Cell Biol**. 30(4):1088-96. 2010.
- MORELLI, C. et al. Alternative splicing, genomic structure, and fine chromosome localization of REV3L. **Cytogenet Cell Genet**. 83:18-20. 1998.

- MUNFORD, V et al. A genetic cluster of xeroderma pigmentosum variant patients with two different founder mutations. **Br J Dermatol**. 176(5):1270-1278. 2017.
- NILSEN TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. **Nature**. 463: 457–463. 2010.
- OGAWA . Splicing factor mutations in myelodysplasia. **Int J Hematol**. 2012;96:438–42
- OGI et al. Mutation enhancement by DINB1, a mammalian homologue of the Escherichia coli mutagenesis protein DinB. **Genes to cells**. 4(11): 607-618. 1999.
- OGI et al. PolK protects mammalian cells Against the lethal and mutagenic effects of benzo[a]pyrene. 99(24): 15548-15553. 2002.
- OGIET al., Polk protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. **Proc Natl Acad Sci U S A**. 99(24): 15548–15553. 2002.
- 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.
- OZTURK, S.; Demir, N. DNA repair mechanisms in mammalian germ cells. **Histol Histopathol**. 26(4):505-17. 2011.
- PADGETT RA. New connections between splicing and human disease. **Trends Genet**. 2012; 28:147–54.
- PADGETT RA. New connections between splicing and human disease. **Trends Genet**. 2012;28:147–54
- PAPAEEMMANUIL, E. et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. **Blood**. 122(22):3616-27. 2013.
- POWERS KT, Washington MT. Eukaryotic translesion synthesis: Choosing the right tool for the job. **DNA Repair**. 71:127-134. 2018.
- PRZYCHODZEN B, Jerez A, Guinta K, Sekeres MA, Padgett R, et al. Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. **Blood**. 2013;122:999–1006.
- QIU, Han et al. Efficient exon skipping by base-editor-mediated abrogation of exonic splicing enhancers. **Cell Rep**. 42(11):113340. 2023.
- ROGALSKA, Malgorzata Ewa; Vivori, Claudia; Juan, Valcárcel. Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects. **Natures Review Genetic**. 24 (4): 251-269. 2023.

- SASATANI, M. et al. Overexpression of Rev1 promotes the development of carcinogeninduced intestinal adenomas via accumulation of point mutation and suppression of apoptosis proportionally to the Rev1 expression level. **Carcinogenesis**, 38(5): 570–578. 2017.
- 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.** 23(22): 4484-94. 2004.
- SEKI, M.; Marini, F.; Wood, RD. POLQ (Pol theta), a DNA polymerase and DNA dependent ATPase in human cells. **Nucleic Acids Res.** 31(21): 6117-26. 2003.
- SHILKIN, E.S. Translesion DNA Synthesis and Carcinogenesis. **Biochemistry.** 85(4): 425-435. 2020.
- THAKUR, M et al., DNA polymerase H undergoes alternative splicing, protects against UV sensitivity and apoptosis, and suppresses Mre11- dependent recombination. 32: 222–235. 2001.
- TOMAS-ROCA, Laura et al. De Novo mutations in PLXND1 and REV3L cause Möbius syndrome. **Nature communications.** 7199. 2015.
- 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.** 10(7). 2015.
- ULE, Jernej; Blencowe, Benjamin J. Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. **Mol Cell.** 76(2): 329-345. 2019.
- 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.** 114(5): 937-951. 2009.
- VARDIMAN, James W. et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. **Blood.** 114 (5): 937-951. 2008.
- VERTA, Jukka-Pekka; Jacobs, Arne. The role of alternative splicing in adaptation and Evolution. **Trends Ecol Evol.** 37(4): 299-308. 2021
- WAHL MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. **Cell.** 136: 701–18. 2009.

- WANG C, Sashida G, Saraya A, et al. Depletion of Sf3b1 impairs proliferative capacity of hematopoietic stem cells but is not sufficient to induce myelodysplasia. **Blood** 2014;123(21):3336–43.
- WILL CL, Lührmann R. Spliceosome structure and function. **Cold Spring Harb Perspect Biol.** 3: 3707. 2011.
- XU, Shaohai et al. SRRM2 organizes splicing condensates to regulate alternative splicing. **Nucleic Acids Res.** 50(15):8599-8614. 2022.
- YADAV et al. Somatic mutations in catalytic core of POLK reported in prostate cancer alter translesion DNA synthesis. **Hum. Mutat.** 36(9): 873-880. 2015.
- YANG, Yue et al. SRSF2 mutation reduces polycythemia and impairs hematopoietic progenitor functions in JAKV617F-driven myeloproliferative neoplasm. **Blood cancer journal.** 13(1): 171. 2023.
- ZHANG et al. Culture mouse cells with transient expression of the mouse DINB1 gene leads to an increase of base substitution and frameshifts mutations. **Nucleic Acids Research.** 28(21): 4138-4146. 2000.
- ZHANG, Yuanjiao et al. Alternative splicing and cancer: a systematic review. **Signal Transduction and Targeted Therapy – Nature.** 6:78. 2021.
- ZHU, Ning et al. REV1: A novel biomarker and potential therapeutic target for various cancers. **Front Genet.** 2022.
- XIE, Fuliang; Wang, Jianyu; Zhang, Baohong. RefFinder: a web-based tool for comprehensively analyzing and identifying reference genes. **Funct Integr Genomics.** 23(2): 125. 2023.
- WIERSCHKE, Stephan et al. Evaluating reference genes to normalize gene expression in human epileptogenic brain tissues. **Biochem Biophys Res Commun.** 403(3-4):385-90. 2010.

APÊNDICE: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO



UNIVERSIDADE FEDERAL DO ESTADO DO CEARÁ FACULDADE DE MEDICINA DEPARTAMENTO DE MEDICINA CLÍNICA

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO PROJETO: ANÁLISE DE VARIANTES GENÉTICAS E SEU IMPACTO NO MECANISMO DE SPLICING EM GENES DE REPARO DO DNA COM ATIVIDADE DE 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 arroxeadas) 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) órgãos(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: Antonio Wesley Araújo dos Santos (085996641058) ou com seu Orientador acadêmico: Prof. Dr. Ronald Feitosa Pinheiro (085981881972) 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